

FUNDAMENTAL EXPLORATION OF SOOT FORMATION
AND MORPHOLOGY FROM A MOLECULAR
MODELING PERSPECTIVE

by

Khaled Mosharraf Mukut

A Dissertation submitted to the Faculty of the Graduate School,
Marquette University,
in Partial Fulfillment of the Requirements for
the Degree of Ph.D.

Milwaukee, Wisconsin

August 2025

ABSTRACT
FUNDAMENTAL EXPLORATION OF SOOT FORMATION
AND MORPHOLOGY FROM A MOLECULAR
MODELING PERSPECTIVE

Khaled Mosharraf Mukut

Marquette University, 2025

Soot formation remains one of the least understood phenomena in combustion science, posing significant challenges due to its complex physicochemical nature and considerable environmental and health impacts. This dissertation presents a comprehensive molecular-level investigation into the fundamental mechanisms governing soot inception, particle growth, and morphological evolution through state-of-the-art reactive molecular dynamics (RMD) simulations of acetylene pyrolysis. A novel computational analysis utility, Molecular Arrangement and Fringe Identification and Analysis from Molecular Dynamics (MAFIA-MD), was developed to accurately characterize soot particle formation, providing detailed insights into chemical composition, internal structure, and surface characteristics. Critical physicochemical markers defining the boundary between gas-phase species and particulate soot were rigorously identified, including specific primary particle mass and number of carbon atoms. Extensive analyses of internal structural evolution revealed essential insights into particle density variations, radial distribution of internal features, and structural complexity during soot nucleation and growth. Furthermore, detailed characterization of soot particle pore structures and surface morphology established quantitative relationships among morphological features, fractal dimensions, and particle reactivity, significantly enhancing predictive capabilities of engineering scale models. This dissertation effectively addresses critical gaps in current knowledge by providing robust methodologies and extensive molecular-level insights that substantially advance our understanding of soot formation processes. The outcomes of this research offer essential contributions towards refining predictive combustion models and developing targeted strategies for emissions control, thereby supporting cleaner combustion technologies, improved public health, and environmental sustainability.

ACKNOWLEDGMENTS

Khaled Mosharraf Mukut

I would like to express my deepest gratitude to my advisor, Dr. Somesh Roy. His invaluable guidance, patience, and mentorship were the cornerstone of my doctoral research. I am especially thankful for the intellectual freedom he gave me to explore my own ideas, and for teaching me how to think critically and approach complex problems with confidence. His insightful feedback was instrumental in shaping this dissertation. I am also deeply grateful to my dissertation committee members for their time and expertise. Their thoughtful critiques and challenging questions significantly improved the clarity and rigor of this work.

My sincere thanks extend to my collaborators, Dr. Georgios A. Kelesidis and Dr. Eirini Goudeli, whose expertise and insightful discussions were crucial to this research. I particularly acknowledge Akaash Sharma and Anindya Ganguly for their dedicated support with the molecular dynamics simulations. This work was also enriched by the engaging and collaborative spirit of my peers and research group members; their support and camaraderie made this journey both memorable and enjoyable. On a special note, I am deeply grateful to my friend Dr. Hasanul Aziz for his invaluable help with Python, and to my roommate Dr. Rafee Mahbub, who was always a willing sounding board for new ideas.

This research would not have been possible without the generous financial support from the National Science Foundation (NSF) grant No. 2144290. Their support facilitated the extensive simulations and analyses essential to this work.

On a deeply personal note, I dedicate my heartfelt thanks to my family. To my father, Abdul Mannan; my mother, Khaleida Mannan; and my siblings, Dr. Asadullah Hill Galib and Akib Al Mueyed, your unwavering support and encouragement were my foundation.

And finally, to my wife, Dr. Farha Binte Karim. I am profoundly grateful for your constant love, patience, and understanding throughout this demanding journey. Your belief in me was my greatest motivation, and this achievement is as much yours as it is mine.

DEDICATION

To my beloved late grandparents, whose enduring wisdom, love, and sacrifices have deeply inspired and shaped my life's journey. Their memories live within me, guiding my steps and reminding me of the values they instilled. This work is a humble tribute to their cherished legacy.

Thank You

TABLE OF CONTENTS

ACKNOWLEDGMENTS	i
DEDICATION	ii
LIST OF TABLES	viii
LIST OF FIGURES	ix
1 INTRODUCTION AND BACKGROUND	1
1.1 Background and Significance of Soot Formation	1
1.2 Fundamental Processes in Soot Formation	2
1.3 Review of Existing Soot Models	3
1.4 Gaps in Current Knowledge and Motivation for this Research	4
1.5 Objectives and Research Questions	5
1.6 Methodological Innovations and Approach	6
1.7 Outline of the Dissertation	7
2 MOLECULAR ARRANGEMENT, FRINGE IDENTIFICATION AND ANALYSIS FROM MOLECULAR DYNAMICS (MAFIA-MD): A POST-PROCESSING TOOL FOR SOOT STUDY	9
2.1 Introduction	9
2.2 Theory and Algorithm	14
2.2.1 Identification of the cyclic structures	14
2.2.2 Chemical characterization of atom clusters	19
2.2.3 Identification and analysis of molecular fringes	20
2.3 Workflow	23
2.3.1 Code structure	23
2.3.2 Program deployment	23

2.3.3	Analysis modes	27
2.3.4	Program input	28
2.3.5	Numerical example	30
2.4	Future works	39
2.5	Summary	39
3	ACETYLENE PYROLYSIS CASE-I: BOUNDARY BETWEEN GAS PHASE SPECIES AND SOOT	41
3.1	Introduction	41
3.2	Numerical Methodology	43
3.3	Chemical structure of soot clusters	43
3.4	Results and discussion	44
3.5	Conclusion and future works	49
4	INTERNAL STRUCTURE OF INCIPIENT SOOT PARTICLES FROM REACTIVE MOLECULAR DYNAMICS SIMULATIONS	51
4.1	Introduction	51
4.2	Numerical Methodology	54
4.2.1	Simulation configurations	54
4.2.2	Workflow	55
4.2.3	Extraction of physicochemical properties	56
4.2.4	Characterization of internal structure	57
4.3	Results and Discussion	58
4.3.1	Formation of incipient soot particles in RMD	58
4.3.2	Classification of incipient soot particles	60
4.3.3	Comparison with experimental data	63

4.3.4	Internal structure of incipient soot particles	63
4.3.5	The boundary between the core and shell	69
4.4	Conclusion	71
5	PHYSICOCHEMICAL EVOLUTION OF INCIPIENT SOOT PARTICLES IN ACETYLENE PYROLYSIS: A REACTIVE MOLECULAR DYNAMICS STUDY	73
5.1	Introduction	73
5.2	Numerical Methodology	77
5.2.1	Simulation configurations	77
5.2.2	Extraction of physicochemical properties	79
5.3	Results and Discussion	81
5.3.1	Inception of soot and classification of incipient soot particles .	81
5.3.2	Physicochemical features of incipient particles	83
5.3.3	Evolution of ring structures in incipient soot particles	95
5.3.4	Summary statistics	97
5.4	Conclusion	100
6	ELUCIDATING PORE AND SURFACE FEATURES OF SOOT NANOPARTICLES USING MOLECULAR DYNAMICS SIMULATIONS	103
6.1	Introduction	103
6.2	Methodology	105
6.2.1	Simulation of incipient soot particles	106
6.2.2	Extraction of surface and pore information	107
6.3	Results and Discussion	108
6.3.1	Bulk morphological properties of incipient particles	109
6.3.2	Bulk morphological properties of pores	116

6.3.3	Fractal characteristics of incipient particles and cavities	117
6.3.4	Pore size distribution inside incipient particles	121
6.3.5	Correlations for pore area and pore volume	124
6.3.6	Summary statistics	126
6.4	Conclusion	126
6.5	Acknowledgments	129
7	SUMMARY AND FUTURE WORKS	130
7.1	Summary	130
7.2	Future Works	131
7.2.1	Expansion of Reactive Molecular Dynamics (RMD) Simulations	131
7.2.2	Temperature and Pressure Effects on Soot Evolution	132
7.2.3	Modeling Soot Aggregate and Coagulation	132
7.2.4	Further Investigation into Reaction Pathways	133
7.2.5	Integration RMD Data into Engineering-Scale Model	133
	BIBLIOGRAPHY	135
A	SYMBOLS, NOMENCLATURE, AND DEFINITIONS USED IN CHAPTER 4 .	161
A.1	Symbols and nomenclature	161
A.2	Expressions for physical properties of soot particles	162
A.3	Physicochemical data used and analyzed in this study	163
A.3.1	Feature set	163
A.3.2	Sample data	164
B	SYMBOLS AND DEFINITIONS USED IN CHAPTER 6	166
B.1	Mathematical Equations	166

B.2	List of Symbols and Definitions	167
B.3	Definitions and Key Parameters	169
C	SUPPLEMENTARY DATA FOR CHAPTER 6	172
D	DEVELOPED SOFTWARE AND TOOLS	182
D.1	MAFIA-MD: Molecular Arrangements and Fringe Identification and Analysis from Molecular Dynamics	182
D.2	StereoFractAnalyzer: 2D and 3D Fractal Analyzer	182
D.3	PyPack: A Molecular System Builder for LAMMPS	182
E	EVIDENCE FOR THE RESTRUCTURING OF NON-HEXAGONAL RINGS .	184
E.1	Thermodynamic Driving Force: Ring Stability	184
E.2	The Stone-Wales Transformation	185
E.3	Collision and Conversion of 5- and 7-Membered Rings	185
E.4	Evidence from Nanotube Growth Simulations	186
E.5	Summary	187

LIST OF TABLES

2.1	Example file format for XYZ files	14
2.2	Computational cost breakdown of MAFIA-MD	39
3.1	Characterization of soot particle evolution with time	48
5.1	Equation of linear curves fitted to the data for different temperatures shown in Figs. 5.3 and 5.4. V is the particle volume in \AA^3 , A is the particle surface area in \AA^2 , and N_C is the total number of carbon atoms.	88
5.2	Mean, Standard Error of the Mean (SEM), and Standard Deviation (SD) of the physicochemical features of soot clusters.	99
6.1	Mean, Standard Error of the Mean (SEM), and Standard Deviation (SD) of the surface and pore features of incipient particles.	127
B.1	List of symbols and their definitions used in the main text.	168

LIST OF FIGURES

2.1	Different domain decomposition strategies (2.1(a) and 2.1(b)) and minimum amount of overlap required for a 6-carbon ring (2.1(c)).	16
2.2	Examples and characterization of optical fringes obtained from a hypothetical HRTEM image of a soot particle.	22
2.3	Code structure and relevant computational loads	24
2.4	Callgraph of MAFIA-MD	25
2.5	Program graphical user interface	28
2.6	Execution and output of MAFIA-MD	36
2.7	Fringe spacing calculation mode outputs.	38
3.1	Evolution of incipient soot inside the domain from acetylene molecules	45
3.2	Comparison of (a) C/H ratio and (b) particle diameter (d_p) of the soot particles with results reported in literature.	46
3.3	Distribution of (a) atomic fractal dimension and (b) packing density of the soot particles with molecular weight.	47
3.4	Population of hydrocarbon molecules present in the simulation domain at four different times based on (a-d) number of molecules and (e-h) molecular weight (kg/kmol).	50
4.1	Schematic representation of the calculation of radial distribution of internal features in the soot particle.	57
4.2	A general representation of steps during the formation and evolution of incipient soot cluster during acetylene pyrolysis (from a simulation performed at 1650 K).	59
4.3	C/H ratio vs. molar mass of soot clusters at different temperatures.	61
4.4	Classification of incipient soot particles	62

4.5	Radial distribution of (a,c) non-cyclic and (b,d) cyclic carbon atoms ($N_{\text{C}}''(r^*)$ and $N_{\text{O}}''(r^*)$) in (a,b) type 1 and (c,d) type 2 soot particles as a function of the normalized radial distance ($r^* = \frac{r}{R_g}$) from the center of mass. The blue vertical line is at a radial distance equivalent to R_g	64
4.6	Radial distribution of the median fraction of cyclic carbon.	65
4.7	Radial distribution of normalized C/H ratio ($\theta_{\text{C/H}}^*(r^*) = \frac{\theta_{\text{C/H}}(r^*)}{\Theta_{\text{C/H}}}$) in incipient particles as a function of normalized radial distance ($r^* = \frac{r}{R_g}$) from the center of mass. The blue vertical line is at a radial distance equivalent to R_g . The $\frac{C}{H}$ ratio of the strip and the particle are the same along the red horizontal line.	67
4.8	Radial distribution of normalized local density ($\rho^*(r^*) = \frac{\rho(r^*)}{\rho_s}$) inside incipient particles as a function of normalized radial distance ($r^* = \frac{r}{R_g}$) from the center of mass. The blue vertical line is at a radial distance equivalent to R_g . The density of the strip and the particle are the same along the red horizontal line.	68
4.9	Identification of core and shell based on the radial distribution of normalized medians (a) density, (b) C/H ratio and radial distribution of the number of (c) cyclic and (d) non-cyclic carbon atoms per unit area in incipient particles as a function of normalized radial distance ($r^* = \frac{r}{R_g}$) from the center of mass.	70
4.10	Schematic of core and shell structure of incipient particles.	71
5.1	Overview of the workflow employed in this study. The blocks numbered (1) to (4) have been discussed in more detail in our previous work [110]. This work focuses on the block (5).	78
5.2	One example particle from each class obtained from RMD simulations.	82
5.3	Evolution of particle volume with the number of carbon atoms at different temperatures. Top row: Type 1 particles, bottom row: Type 2 particles.	84
5.4	Evolution of particle surface area with the number of carbon atoms at different temperatures. Top row: Type 1 particles, bottom row: Type 2 particles.	85
5.5	Evolution of fraction of carbon atoms in ring structures ($N_{\text{O}}/N_{\text{C}}$) with molar mass.	86
5.6	Relationship between molar mass and (a) volume, (b) surface area, (c) radius of gyration, and (d) total number of rings in incipient soot particles	90
5.7	Relationship between radius of gyration and (a) mass, (b) volume, (c) surface area, and (d) total number of rings in incipient soot particles.	91
5.8	Evolution of atomic fractal dimension (D_f) with molar mass	94

5.9	Evolution of radius of gyration (R_g) and volume equivalent radius (R_{eq}): Comparison between volume equivalent radius and radius of gyration for all particles (a); ratio of radius of gyration and volume equivalent radius vs number of carbon atoms for type 1 particles (b) and for type 2 particles (c).	95
5.10	Evolution of cyclic structures in incipient soot particles.	97
5.11	Kendall's rank correlation coefficient between different physicochemical features of type 1 (lower-left triangular matrix) and type 2 (upper-right triangular matrix) incipient soot particles.	100
6.1	Schematic representation of different types of cavities inside an irregularly shaped particle.	105
6.2	Overview of the workflow utilized in this work.	106
6.3	One example particle obtained from RMD simulations along with examples of surface renderings and identified cavities.	108
6.4	Distribution of sphericity (Ψ) of incipient particles.	109
6.5	Average circularity ($\bar{\sigma}$) of incipient particles at four different temperatures.	111
6.6	Distribution (a) porosity, Φ and (b) specific pore volume [cm^3/g] of incipient particles.	112
6.7	Distribution of specific surface area [m^2/g] of incipient particles.	114
6.8	Contribution of different cavities to specific surface area [m^2/g] of incipient particles.	115
6.9	(a) Average pore sphericity, $\bar{\psi}$ and (b) average specific surface area of cavities [$1/\text{\AA}$] in incipient particles.	116
6.10	Distribution of (a) surface fractal dimension (D_S), and (b) volume fractal dimension (D_V) of incipient particles.	119
6.11	Distribution of (a) surface fractal dimension (D_{SC}), and (b) volume fractal dimension (D_{VC}) of the cavities within incipient particles.	120
6.12	Cavity size (d_p) distribution of all cavities across all the particles.	122
6.13	Distribution of average fraction of void volume occupied by cavities of different sizes within incipient particles.	123

6.14 Average pore size distribution (PSD) of different cavities within incipient soot primary particles.	124
A.1 Two sample soot particles and their attributes investigated in this study	165
B.1 Types of cavities in soot particles.	169
B.2 Material and Bulk Volumes of a Soot Particle.	169
B.3 External and Exterior Surface Areas of a Soot Particle.	170
C.1 Distribution of particle sphericity at different temperatures.	172
C.2 Distribution of average circularity of particles at different temperatures.	172
C.3 Distribution of porosity, Φ of particles at different temperatures.	173
C.4 Distribution of specific pore volume [cm^3/g] of particles at different temperatures.	173
C.5 Distribution of specific surface area of incipient particles at four different temperatures.	174
C.6 Contribution of different cavities to specific surface area [m^2/g] of incipient particles.	174
C.7 (a) Average pore sphericity, $\bar{\psi}$ and (b) average specific surface area of cavities [$1/\text{\AA}$] in incipient particles at different temperatures.	175
C.8 Distribution of surface fractal dimension (D_S) of incipient particles at four different temperatures.	176
C.9 Distribution of volume fractal dimension (D_V) of incipient particles at four different temperatures.	176
C.10 Distribution of surface fractal dimension (D_{SC}) of cavities at four different temperatures.	177
C.11 Distribution of volume fractal dimension (D_{VC}) of cavities at four different temperatures.	177
C.12 Distribution of average cavity size (\bar{d}_p) distribution of different cavities within an incipient soot primary particle.	178

C.13 Pore size distribution (PSD) of the entire population of cavities from all incipient soot primary particles.	178
C.14 Correlation between total pore volume and pore surface area within incipient particles at different temperatures. The equation for linear fit to the data also added to the plot. An exponential and a quadratic fit were also explored, but no significant improvement of correlation coefficient was observed.	179
C.15 Correlation between particle volume and pore volume within incipient particles at different temperatures. The equation for linear fit to the data also added to the plot. An exponential and a quadratic fit were also explored, but no significant improvement of correlation coefficient was observed.	180
C.16 Correlation between individual pore volume and pore surface area for different cavity types.	181
E.1 Illustration of the “collision” of edge seven- and five-membered rings transforming into two edge six-membered rings (top), and the migration of a seven-membered ring (bottom). This demonstrates a direct pathway for the conversion of non-hexagonal rings into more stable hexagonal structures. (Adapted from Frenklach and Mebel [251].)	185
E.2 Recreation of the key data illustrating the ring formation pathways during carbon growth. The figure shows (a) the schematics of the direct (Path 1) and indirect (Path 2) pathways, (b) the significantly lower energy barrier for Path 2, (c) the statistical dominance of Path 2 for hexagon formation. Data adapted from Wang et al. [252].	186

CHAPTER 1

INTRODUCTION AND BACKGROUND

1.1 Background and Significance of Soot Formation

Soot, also known as black carbon, consists of tiny nanoparticles primarily formed during incomplete combustion of carbon-based fuels. These particles are typically less than 100 nanometers in diameter [1] and are composed of elemental carbon, along with various organic compounds [2]. Soot formation is a complex process that occurs in numerous combustion systems, including internal combustion engines, coal-fired power plants, wildfires, and residential heating systems. The incomplete combustion of hydrocarbons leads to the production of soot particles [3], which can have significant environmental and health implications, affecting both human wellbeing [4] and climate dynamics [5, 6].

Exposure to soot particles has been linked with numerous negative health outcomes, including respiratory problems [7], cardiovascular diseases [8], and increased mortality rates, due to their penetration deep into the human respiratory tract [9]. Due to their nanoscale size, soot particles easily penetrate deep into human lungs, exacerbating respiratory issues and significantly increasing health risks, particularly in densely populated areas.

Beyond soot's direct health effects, soot significantly impacts the global environment. It contributes substantially to atmospheric heating through radiative forcing, directly absorbing sunlight and altering climate patterns [10]. Additionally, soot deposition accelerates the melting of polar ice by reducing surface albedo, posing significant threats to sensitive ecosystems and contributing further to global climate change.

The process of soot formation is inherently complex, characterized by intricate chemical pathways, short reaction timescales, and high molecular variability. Despite extensive research, significant gaps persist in fully understanding how soot forms at the molecular level, particularly the initial transition from gas-phase species to solid nanoparticles. Addressing this knowledge gap is crucial for developing effective strategies for emission control and cleaner combustion technologies, thus safeguarding public health

and environmental quality.

1.2 Fundamental Processes in Soot Formation

The formation of soot is generally understood to occur in several interconnected stages: first, the formation of gas-phase precursor molecules including polycyclic aromatic hydrocarbons (PAHs); second, the inception or nucleation of these precursors into initial solid soot particles; third, subsequent particle surface growth by adsorption and chemical reactions; fourth, coagulation and aggregation of smaller particles into larger soot aggregates; and finally, particle oxidation and fragmentation [11, 12, 13, 14, 15, 16, 17]. Each of these stages is governed by complex interactions between chemical kinetics and physical transformations.

Efforts to predict and control soot have led to the development of a wide range of computational models, from empirical and semi-empirical correlations to detailed phenomenological frameworks [18, 19, 20]. These engineering-level models are built upon our understanding of the dominant underlying processes governing soot particle evolution, including inception, growth, coagulation, and oxidation. The theories describing polycyclic aromatic hydrocarbon (PAH) formation and growth have been particularly central to this effort.

Despite substantial advancements, major gaps remain in our understanding of soot inception—the critical and least understood step [14]. Existing numerical models commonly rely on assumptions, such as spherical incipient particles, fixed particle densities, and simplified chemical reactions, that overlook the intricate chemistry and evolving morphology of early soot particles. Furthermore, the exact chemical pathways and precise morphological transitions during particle formation and growth remain unclear, partly due to limitations in experimental diagnostics and the inherent complexity of the processes [21]. Addressing these unresolved aspects is crucial for accurate modeling, control, and mitigation of soot emissions.

This dissertation addresses these critical knowledge gaps by employing reactive molecular dynamics (RMD) simulations to systematically investigate the fundamental

chemical and physical transitions occurring during soot inception. Specifically, this research develops robust methodologies to identify the boundary between gas-phase precursors and incipient soot particles, characterize internal structural evolution, and quantify surface and morphological properties during early soot growth. Leveraging the computational analysis tool MAFIA-MD [22], detailed insights into molecular structures—including cyclic and aliphatic hydrocarbons—are obtained, significantly enhancing the fundamental understanding of soot formation mechanisms and improving upon current simplified soot models.

1.3 Review of Existing Soot Models

Various approaches have been employed to model soot formation, generally categorized as empirical, semi-empirical, and detailed models. Empirical models rely purely on experimental correlations and typically provide limited quantitative insights into underlying soot formation processes [23, 24, 25, 26]. Semi-empirical models integrate some theoretical foundations but still depend significantly on experimental data to derive critical parameters [18, 27, 28, 29]. Detailed models, including discrete sectional methods (DSM) [30], methods of moments [31, 13, 32, 31], and stochastic soot models [33, 34], aim to provide deeper insights by solving particle size distribution functions (PSDF) and capturing detailed chemical and physical interactions during soot formation [18].

To track the evolution of the soot particle size distribution, population balance models are commonly employed. A widely used approach is the method of moments (MOM), with implementations like the method of moments with interpolative closure (MOMIC) [31], which solves the population balance equation efficiently. Discrete sectional method (DSM) [30, 35, 36] is another approach to do this, where the entire particle size distribution (PSD) is discretized into bins or sections, and evolution of soot within each section is tracked. These mathematical frameworks are agnostic to the underlying chemistry and physics, allowing it to incorporate various sub-models for different processes such as nucleation, coagulation, oxidation, and growth. For instance, particle surface growth can be described using mechanisms like the hydrogen-abstraction-carbon-addition (HACA)

pathway [14, 11], nucleation can be described either by PAH dimerization [37] or by lumped acetylene reactions [28] while other modules describe nucleation, coagulation, and oxidation.

However, these existing soot models consistently rely on simplifying assumptions that restrict their predictive capabilities. Common assumptions include spherical soot particles [38], fixed primary particle size [39], densities [38], and simplified surface growth mechanisms [11, 40, 12], which neglect the complex molecular structures and reactions identified in recent molecular-level studies [39, 41]. Additionally, many models assume pre-defined initial soot nuclei and omit variations in particle morphology and internal structure during soot inception and early growth [42]. These limitations underscore the need for more detailed, accurate, and fundamentally sound modeling approaches capable of capturing complex morphological and chemical evolution processes.

1.4 Gaps in Current Knowledge and Motivation for this Research

Although significant advancements have been made in understanding soot formation mechanisms, several crucial gaps remain, particularly surrounding the nucleation and initial growth stages of soot particles. The current gaps can be summarized as follows:

- **Model-level simplifications:** Even detailed models still invoke idealized primary particles (fixed density, diameter, sphericity) and single-path surface-growth schemes such as HACA to simplify the complex chemistry of soot formation and growth. These simplifications mask the roles of non-HACA radical chemistry, mixed aliphatic-aromatic bonding, and evolving particle density revealed by recent RMD studies.
- **Diagnostic limitations:** Ultrafast, nanoscale transformations during inception occur on sub-microsecond timescales, well below the temporal and spatial resolution of most flame diagnostics. Key intermediate species (five- and seven-membered rings, resonance-stabilized radicals (RSR), curved fringe planes) are therefore inferred only indirectly, limiting our understanding of the earliest soot formation processes.
- **Morphological uncertainty:** TEM reconstructions show nascent particles that are neither compact spheres nor loose PAH stacks but heterogeneous entities with

internal voids and curved graphitic layers [42]. Current models cannot accommodate this structural richness, limiting predictive fidelity.

These unresolved issues propagate into *system-level* uncertainties leading to inaccurate number-mass correlations, unreliable growth/oxidation rates, and large scatter in predicted optical properties. Filling these gaps is essential not only for basic science but also for engine design, emissions legislation, and climate modelling.

Addressing these critical gaps in our fundamental understanding is thus of utmost importance for advancing both theoretical and practical aspects of combustion science. Enhanced insights at the molecular level, including precise characterization of the transition from gas-phase precursors to solid particulate matter, internal particle structure, and surface morphological properties, will enable more accurate and predictive soot formation models. Achieving this level of detailed understanding has substantial practical implications, facilitating the design and optimization of cleaner combustion technologies, more efficient emissions control strategies, and targeted regulatory frameworks for environmental protection. Motivated by these significant needs, this dissertation employs advanced reactive molecular dynamics (RMD) simulations, combined with sophisticated computational tools and analyses, to rigorously investigate the intricate physicochemical transformations during soot inception and early particle growth. By developing and applying methodologies capable of precisely characterizing soot formation processes, this research directly contributes to bridging these existing knowledge gaps and advancing the state-of-the-art understanding in soot formation and morphology.

1.5 Objectives and Research Questions

Given the significant gaps and limitations identified above, this dissertation aims to provide a molecular-level characterization of incipient soot. The primary objective is to investigate the evolving chemical composition and physical properties—such as internal structure, bonding, and density—of incipient soot particles during their inception and early growth. This research moves beyond the idealized particle representations common in

engineering models by focusing on the complex nature of the particles themselves.

Specifically, this work seeks to:

1. Characterize the fundamental transition from gas-phase precursors to incipient soot by identifying key physicochemical markers (e.g., molecular size, density, connectivity) that define the gas-particle boundary at the molecular level.
2. Develop robust computational methodologies for accurately characterizing the internal structural evolution, surface properties, and morphological transformations of soot particles during their inception and subsequent growth phases.
3. Investigate the role of molecular structures, particularly cyclic and aliphatic hydrocarbons, in the soot nucleation and early growth stages, providing insights into their significance and interaction mechanisms.
4. Establish quantitative relationships among morphological features, internal structures, and chemical compositions to enhance predictive capabilities and improve upon existing simplified soot formation models.

These objectives collectively aim to address the previously discussed knowledge gaps and limitations, thereby contributing substantial advancements to the field of combustion science and improving strategies for controlling soot emissions.

1.6 Methodological Innovations and Approach

Given the inherent complexity and limitations of existing experimental and modeling approaches, novel methodological strategies are essential for advancing our understanding of soot formation at a fundamental level. To address these needs, this dissertation leverages state-of-the-art reactive molecular dynamics (RMD) simulations capable of accurately capturing chemical reactions and structural transformations at the atomic scale.

A major methodological innovation of this research is the development and utilization of an advanced computational analysis tool, Molecular Arrangement and Fringe Identification and Analysis from Molecular Dynamics (MAFIA-MD). This novel post-processing utility provides an efficient, comprehensive means of extracting and

analyzing detailed molecular structures, including cyclic and aliphatic hydrocarbons, internal structural configurations, and morphological characteristics directly from large-scale RMD simulation outputs. By precisely identifying ring structures, carbon-hydrogen ratios, and fringe patterns, MAFIA-MD substantially enhances the ability to characterize soot particle formation and growth accurately.

Additionally, this research establishes robust analytical frameworks for quantifying internal structural evolution, surface features, and morphological transformations of soot particles. These innovative methodologies allow for precise differentiation between gas-phase precursors and particulate matter, elucidating clear chemical and morphological boundaries. Consequently, this methodological approach provides unprecedented molecular-level resolution, significantly surpassing existing simplified modeling approaches and offering critical insights to inform future experimental validations and soot modeling practices.

1.7 Outline of the Dissertation

The remainder of this dissertation is structured as follows:

- Chapter 2 introduces and describes in detail the development and implementation of the novel computational analysis utility MAFIA-MD, specifically designed to characterize molecular structures and morphological features from reactive molecular dynamics (RMD) simulations of soot.
- Chapter 3 presents initial investigations into the molecular-level boundary between gas-phase precursor molecules and solid soot particles during acetylene pyrolysis. This chapter identifies distinct physicochemical markers that differentiate particulate soot from gas-phase species.
- Chapter 4 provides an extensive analysis of the internal structural evolution of incipient soot particles formed during acetylene pyrolysis, employing the methodologies developed in Chapter 2 to characterize key morphological and chemical transitions.

- Chapter 5 systematically explores the physicochemical properties of incipient soot particles, analyzing critical features such as size, mass distribution, and structural complexity during the early stages of soot formation.
- Chapter 6 presents detailed insights into the evolution of pore structures, surface characteristics, and fractal morphology of soot particles. This chapter highlights the importance of these features in understanding soot reactivity and maturity.
- Chapter 7 synthesizes key findings and contributions of this dissertation, discusses their broader implications for combustion science, and outlines recommended avenues for future research to further extend this work.

CHAPTER 2

MOLECULAR ARRANGEMENT, FRINGE IDENTIFICATION AND ANALYSIS FROM MOLECULAR DYNAMICS (MAFIA-MD): A POST-PROCESSING TOOL FOR SOOT STUDY

This chapter is a reproduction of the article already published as: "Mukut, K. M., Roy, S., & Goudeli, E. (2022). Molecular arrangement and fringe identification and analysis from molecular dynamics (MAFIA-MD): A tool for analyzing the molecular structures formed during reactive molecular dynamics simulation of hydrocarbons. *Comput. Phys. Commun.*, 276, 108325. doi: 10.1016/j.cpc.2022.108325" [22]

2.1 Introduction

Reactive molecular dynamics (RMD) simulation is a method of analyzing both the physical and chemical changes of atoms and molecules. During the simulation, individual atoms and molecules are allowed to interact with each other based on the chemical and inter-atomic potentials along with the Newton's equations of motion. With the recent advancement of computational resources, reactive molecular dynamics (RMD) has become more and more practical for performing the first principal analysis of unknown physics, especially in the field of biomolecular chemistry [43].

Several molecular dynamics potentials have been developed in recent years to capture the chemical and structural changes in reactive atomic clusters containing carbon [44]. Among them, quantum molecular dynamics (QMD) [45] and empirical models such as Tersoff [46, 47], REBO [48, 49], AIREBO [50], Reactive force-field potential (ReaxFF) [51], Environment Dependent Interaction [52], and Charge Optimized Many-Body [53] potentials have been widely used in recent studies [25, 54, 55, 56, e.g.,]. Among these, for comprehensive studies focusing on the chemical changes of complex hydrocarbon species (e.g. soot formation), ReaxFF potential [51] has gained popularity among the contemporary researchers [57, 58, 59, 60, 61, 62, 63]

Soot is primarily carbon nano-particles formed during incomplete combustion of hydrocarbon fuels [3]. Under certain conditions (usually in absence of enough oxidizing species) during combustion, mostly gaseous hydrocarbon molecules react with each other to create large molecular clusters which can no longer be treated as gaseous species. These solid or liquid-like clusters are incipient soot and the process of this gas-to-particle transition is called soot inception or nucleation. The exact mechanism of this transformation from gas phase species to solid particles is arguably the least understood phenomenon during soot formation [21]. The widely accepted theories of soot inception indicates that soot inception starts with the formation of polycyclic aromatic hydrocarbons (PAHs) [64, 65, 21], particularly with 5- and 6-membered rings. The initial gas-phase PAH molecules grow and combine together via physical and chemical interactions to form the first soot particles. But the exact physico-chemical processes are yet to be confirmed. The complexity of hydrocarbon systems and the time and length scales of the process makes it very difficult to study the inception process experimentally. This has led to multiple competing and complimentary hypotheses of soot inception [64, 65, 21, 61]. The reactive molecular dynamics (RMD) simulation can be very helpful in unraveling the physico-chemical processes of soot inception and can help bridge the gap between theoretical hypotheses and experimental observation at different scales. The use of RMD can also extend to the later stages of soot evolution, where the incipient soot particle grows via physical and chemical interaction with other soot and gas phase species. In order to properly interpret and validate RMD results, it is important to identify, differentiate, and analyze various physical and chemical structures such as rings and fringes that are formed in soot particles. These characteristics often are related to soot reactivity and maturity [66, 67]. The study of temporal evolution of these features will help resolve the mystery of soot inception and evolution. The post-processing utility “Molecular Arrangement and Fringe Identification and Analysis from Molecular Dynamics” or (MAFIA-MD) presented in this manuscript provides practitioners an ability to study these features with ease.

In usual practice, RMD simulations of soot related studies start with a set of hydrocarbon molecules in a confined domain [57]. The chemical interactions are then

captured using an appropriate molecular dynamics potential as time progresses. The bond-related information (bond orders) is stored in a bond information file (also known as bonds file). The time-resolved snapshot and atomic trajectories in the simulation domain are also available in “.XYZ” [68] format. These “.XYZ” files (also referred as trajectory files) contain the coordinates of individual atoms at different times. The size of the bonds and trajectory files depends on how frequently the data is saved and the interconnectivity of the atoms. In high-temperature application such as combustion, where the chemical reactions are fast, it is important to save these files frequently (e.g. every few picosecond) in order to capture the fast chemical evolution. Furthermore, the atomic interconnectivity of hydrocarbon systems can grow large quite fast. For example, a 0.1 ns of acetylene combustion simulation consisting of 6000 atoms at 1500 K can lead to a bond information file of approximately 5 GB and a corresponding trajectory file of 300 MB. Because of its smaller size, the capability of extracting key features of atomic redistribution and molecular restructuring only from the trajectory file lead to easier visualization and faster analysis.

The goal of MAFIA-MD is to use the trajectory files individually and extract the chemical and structural information, primarily by detecting the presence of 5-, 6-, and 7-membered alicyclic and aromatic ring structures in an atom cluster. The identification of ring structure is important because the stability of such structures, particularly of aromatic rings, is thought to be a key for inception and growth of soot. Therefore, quantitative information on cyclic structures in soot particles can be helpful for identifying the important chemical pathways in soot formation and growth. Additionally, the planarity and curvature of the cyclic structures play important roles in dictating the stability of individual molecules and morphology of an incipient soot nucleus. The identification of these features can thus be utilized to analyze the stability of molecular clusters. The presence of planar and curved surfaces inside molecular clusters creates the optical fringes observed in high-resolution transmission electron microscope (HRTEM) images of soot particles. The quantification of these surfaces provides an accurate and meaningful pathway for direct experimental validation of RMD simulations with HRTEM images.

MAFIA-MD extracts this information from RMD simulations using an efficient algorithm implemented using **Python**. For the sake of completeness, the chemical bond information is calculated using an algorithm proposed by Kim and Kim [69]. This enables **MAFIA-MD** to capture the chemical information of the relevant soot forming molecules and export them in chemically understandable and usable formats such as simplified molecular-input line-entry system (SMILES) [70] and spatial data file (SDF) [71]. In short, **MAFIA-MD** captures the number of different alicyclic and aromatic structures present in the simulation domain, the percentage of alicyclic/aromatic and aliphatic carbons, the chemical representations of the constituent molecules existing in the domain, and provide a way to calculate the fringe spacing statistics of soot cluster obtained from RMD simulations. A graphical user interface (GUI) is developed for the easy management of input parameters. The main functions of the utility can be divided into following three segments:

1. Identification of cyclic structures

- C/H ratio
- Number/percentage of alicyclic and aromatic carbons
- Number/percentage of aliphatic carbons
- Statistics of 5-, 6-, and 7-membered ring structures

2. Chemical characterization of atom clusters

- SMILE string for easier vizualization of molecules
- Export into a molar file format (SDF)

3. Identification and analysis of molecular fringes in soot for validation with HRTEM images

It should be noted here that the **MAFIA-MD** does not strictly check for all four classical aromaticity conditions as per Huckel’s rule. It is impossible to extract the exact electronic structure from the trajectory files, therefore conditions such as the presence of $(4n + 2) \pi$ electrons or perpendicularity of p -orbitals are not possible to check. Furthermore, due to the nature of the MD simulations the atoms of an aromatic ring may not always lie in a single plane at an instantaneous timestamp. A strict planarity check will rule out some

of these aromatics which may have a single out-of-plane carbon in an instantaneous time. A planarity check criterion is included and implemented in **MAFIA-MD** for future development in checking the planarity condition for aromaticity. However, in the current version we only focus on bond distance and closed nature of the ring structures. In the chemical characterization segment of **MAFIA-MD**, the bond order is calculated from the trajectory files. This information can be used to differentiate the aromatics from the alicyclic hydrocarbons.

Finally, some, not all, of the outputs provided by **MAFIA-MD** can be obtained from the bonds information file. However, as mentioned earlier, the bond information files can be orders of magnitude larger in size than the trajectory (**.XYZ** files). The capability of post-processing without requiring these large bond information files gives more flexibility to the user during analysis and sharing data. Since the bonds file is not needed to be saved during runtime in order to extract the chemical information, the memory and I/O requirement for the RMD simulation can be made leaner with the help of **MAFIA-MD**. This can lead to faster runtime, easier collaboration and sharing of data (by only sharing small trajectory files). The ability to do batch processing on a number of molecular dynamics trajectories and get a lot of soot-relevant qualitative and quantitative information at the same time also makes **MAFIA-MD** very useful. There are some open-source tools available in the literature that can detect cyclic structures in MD simulation results [72, 73, 74], but these tools were not designed specifically for hydrocarbon systems and cannot analyze the fringe statistics, which is important for soot-relevant physics. To the best of our knowledge, **MAFIA-MD** is the first utility to do these large-scale analyses automatically and directly from the trajectory files. This makes **MAFIA-MD** unique and very useful to practitioners for analyzing the RMD simulations of hydrocarbons.

2.2 Theory and Algorithm

2.2.1 Identification of the cyclic structures

Read the Input Trajectory File

Table 2.1: Example file format for XYZ files

<i>testFile.xyz</i>				
Content	Description			
N				Number of Atom, N
Timestep:	t			Timestamp, t
C	x	y	z	Atom identifier and coordinates of individual atoms
C	x	y	z	
H	x	y	z	
C	x	y	z	
H	x	y	z	
C	x	y	z	
⋮	⋮	⋮	⋮	⋮
H	x	y	z	
C	x	y	z	

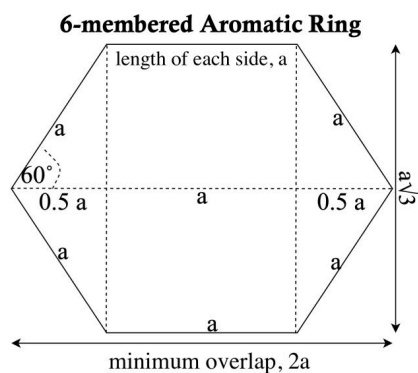
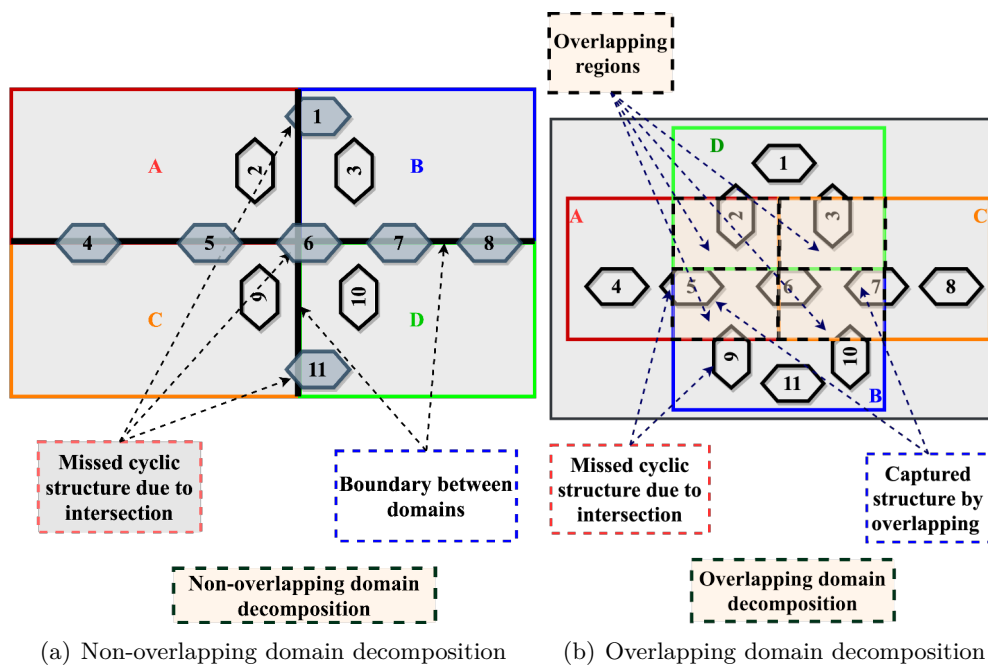
MAFIA-MD takes individual snapshot/trajectory files (in “.XYZ” format) from molecular dynamics simulation as the input. The “.XYZ” files contain coordinates of individual atoms existing in the domain as well as the total number of atoms and the timestamp. An example of the “.XYZ” file structure is presented in Table 2.1. In the current version, the ring-structure detection is carried out solely based on the carbon atoms in the domain. Therefore, once the “.XYZ” file is read, the hydrogen atoms are disregarded. For bookkeeping purpose, the timestamp and the total number of atom in the domain is also read.

Pre-Processing of Input Data

The input coordinates of the carbon atoms are sorted first. The computational complexity of the overall system depends on the compactness (the degree of interconnectivity between the atoms) of the cluster of carbon atoms. If too many atoms are interconnected with each other (very compact cluster), the execution time will be inconveniently long. To remedy this situation, a divide and conquer approach [75] is implemented to reduce the computational complexity of the problem. In this approach, the computational domain is divided into a number of small spatially overlapping subdomains, each of which has significantly lower computational complexity than the entire domain. These overlapping subdomains are then analyzed sequentially.

The division or splitting of the computational domain in overlapping subdomains can be referred to as *overlapped domain decomposition*. The overlapping is required as some of the ring structures can be shared between multiple subdomains. If not properly accounted for, these shared rings will not be counted. This problem is shown in Fig. 2.1(a) which presents non-overlapping (i.e., conventional) domain decomposition. It is evident from Fig. 2.1(a) that, even though the sub-domains A, B, C and D encircle all the ring structures among themselves, because of the presence of domain boundaries, they cannot identify all the cyclic structures. Therefore, the non-overlapping domain decomposition scheme misses cyclic structures 1, 4–8 and 11 as seen from Fig. 2.1(a).

The overlapping domain decomposition scheme implemented in MAFIA-MD (shown in Fig. 2.1(b)) remedies this. In Fig. 2.1(b), the whole computational domain is divided into five overlapping regions, i.e. A, B, C, D and the intersecting region between A, B, C, and D (referred to as subdomain ABCD). In this way, even though an individual domain misses some part of the ring structures (e.g., subdomain A misses cyclic structures 2, 6 and 9), other domains capture all of them (e.g., cyclic structures 2, 6 and 9 are captured by subdomains D, ABCD and B, respectively). It should be noted here, to remove any possible confusion, that the current version of the code is purely serial in nature. However, due to the independent nature of individual subdomains, the code can be easily parallelized to



(c) Minimum amount of overlap for a 6-carbon ring

Figure 2.1: Different domain decomposition strategies (2.1(a) and 2.1(b)) and minimum amount of overlap required for a 6-carbon ring (2.1(c)).

further speed up the analysis.

For the overlapping domain decomposition to work properly, the amount of overlap should be minimized to reduce unnecessary computations. The minimum amount of overlap necessary to capture all cyclic structures is equal to the maximum spatial footprint of the cyclic structures. A typical 6-membered ring structure is shown in Fig. 2.1(c), where “ a ” is the bond distance between two carbon atoms. This structure can be considered as a regular hexagon with sides of length a . As shown in the figure, the maximum spatial footprint of the 6-membered ring is twice the bond distance between two carbon atoms ($2a$). This is essentially the circumdiameter of a regular 6-membered polygon. Similarly, for ensuring the detection of 5- and 7-membered rings, the minimum overlap should be at least 1.7 and 2.3 times the bond distance between two carbon atoms respectively. Accounting for the rectangular shape of the sub-domains, an overlap **span** of 8Å is found to be optimum for ensuring detection of all the cyclic structures of different sizes.

The overlapping domain decomposition scheme shown in Fig. 2.1(b), has a potential drawback for counting same cyclic structure multiple times as it can reside in multiple subdomains simultaneously. For example, in Fig. 2.1(b), cyclic structure 6 resides in both subdomains B and ABCD. Therefore, the book-keeping of these cyclic structure requires special handling. An array of global carbon indices (actual indices of carbon atoms in the XYZ file) is maintained for keeping track of unique cyclic structures from all the subdomains. By cross-referencing the global indices, the duplicate counting problem introduced by the overlapping domain decomposition scheme is resolved. Every time a ring is detected, the constituent carbon coordinates and their global indices are inserted into an array. This array of carbon indices from each subdomain is then concatenated into a global array which contains all the carbon indices from the identified cyclic structures. When all the subdomains are traversed, the duplicate entries are deleted from the global array retaining only the unique cyclic structures detected throughout the computational domain. The total number of 5-, 6-, and 7-membered rings is calculated using this global array.

For smaller (typically less than 700 carbon atoms, the actual number will depend on the interconnectedness of the atoms) population of atoms, the overlapped domain

decomposition may not be necessary since the computational complexity is small enough to perform the ring detection in the entire domain without decomposition. In view of this, as well as to create an in-built verification tool, a separate functionality is implemented in **MAFIA-MD** named “**Sanity Check**” which, instead of decomposing the domain in several overlapped subdomains, considers the entire population of atoms as a whole. The user can chose to use this **Sanity Check** feature to either perform analysis on small atom population or to verify the accuracy of the overlapping domain decomposition by comparing the results from **Sanity Check** and overlapping domain decomposition. It should be noted here that, for larger cluster of atoms, the **Sanity Check** feature will take inconveniently long time to finish.

Creation of Directed Graphs and Determining the Elementary Circuits

Each subdomain carries the coordinates of carbon atoms inside the regions. Each point is sorted based on their distances from the global origin. Once the points are all sorted, a distance matrix [76] is calculated. The carbon atoms bonded with each other falls within a finite distances from one another. This distance is called the bond length. Single bond, double bond, triple bond and aromatic bonds between carbon atoms have different stable bond distances. Note that due to the atomic vibrations in the MD simulations, the bond length will not be the same as the exact theoretical value but will lie within a narrow range around the theoretical bond length. Therefore, the upper and lower limits of bond length, instead of an exact value, need to be specified for the bond length. To capture the bonded carbon atoms, the element in the distance matrix is converted to “1” if the distance falls between the upper and lower limit of valid bond distances specified in the beginning of the analysis by user. If the distance falls outside this bound, the element in the matrix element is replaced with “0”. This simplifies the distance matrix into an adjacency matrix [77], where each element with a value “1” represent a bond between two carbon atoms.

From the adjacency matrix generated in the earlier step, a directed graph is created using the open-source **Python** package “**NetworkX**” [78]. In a directed graph, the edges have direction and can have self-loops (a series of subsequent edges can point back to an earlier

node). The main idea is to find the cyclic structures (i.e., ring structures) from the graph generated from the carbon atoms. For achieving this, the algorithm proposed by Johnson [79] is used for finding all elementary cycles from a directed graph. This algorithm is essentially a depth-first-search algorithm [80, 81], optimized to find only elementary circuits (one vertex appear only once) and traversing an edge at most twice. The algorithm identifies the sets of vertices for all the elementary circuits that forms a closed cycle or ring. For soot relevant applications, currently only 5-, 6-, and 7-membered rings are identified in **MAFIA-MD**. However, the framework for identifying larger rings and nested rings or supersets of rings (along with a check for planarity) is already implemented for any future development. Following the identification of elementary circuits, individual 5-, 6-, and 7-membered rings are indexed and counted for each subdomain and added globally to get the total number of rings containing 5-, 6-, and 7- carbons at a given timestep. The algorithm for ring detection is presented in Algorithm 1. It should be noted that, although all Huckel conditions for aromaticity are not checked in the ring detection segment of **MAFIA-MD**, the actual bond order is calculated in the “**Chemical Characterization**” segment (Sec. 2.2.2) of the code to distinguish the aromatics from the detected alicyclic structure.

2.2.2 Chemical characterization of atom clusters

The outputs from section 2.2.1 are used to interpret the chemical characteristics of the given soot cluster obtained from RMD simulation. For this purpose, a universal structural conversion algorithm developed by Kim and Kim [69] is implemented to solve for the atomic connectivity and bond order of the existing molecules. The general chemical rules and valency information are used for this purpose. Initially, each atomic pair is assigned a bond order based on their valencies. After that, the degree of bond saturation is assigned by trial and error until the whole system of atoms is solved. Once the bond order of the atomic connectivity/network is solved, the information is exported in terms of simplified molecular-input line-entry system (SMILES) [70] or into a molar file format (SDF) [71]. An external tool called `xyz2mol` [82] is modified to work with the current implementation of **MAFIA-MD**. `xyz2mol` uses “**RDKIT**” [83], an open source cheminformatics interface, to export

Algorithm 1: Algorithm for ring identification and quantification.

Result: C/H ratio, percentage of alicyclics/aromatics, ring statistics

```

1 # Input:
2 Define the input parameters: upper and lower bond distance, span/overlapping
  distance, etc.
3 Open the trajectory (XYZ) file exported from RMD
4 Read the total number of atoms and timestamps from the trajectory (XYZ) file
5 # Pre-Processing:
6 Divide the entire domain into a number of subdomains
7 while NO subdomains are left do
8   Create distance matrix using upper and lower bond limits;
9   Simplify the distance matrix into an adjacency matrix;
10  # Elementary Cycle Detection:
11  Convert the adjacency matrix into a simply connected directed graph;
12  Apply Johnson’s algorithm [79] to extract all the simple cycles ;
13  Separate the cycles containing 5-, 6- or 7- vertices;
14  Create an array of sets containing the coordinates of all vertices of individual cycles;
15  Count and index the carbon atoms present in the ring structure;
16  Go to the next subdomain
17 end
18 Merge all rings from all subdomains;
19 Remove duplicate rings using global indices and extract the unique rings;
20 Pass the array of identified rings to the next segment of the code for optional chemical
  characterization and fringe analysis;
21 Print out the results;
22 # Output:
    1. C/H Ratio
    2. Percentage of alicyclics/aromatics
    3. Statistics of different rings
    4. Trajectory of alicyclic/aromatics and aliphatic carbons
  
```

the figures of the existing chemical structures inside the domain for visualization. The full functionality of the external tool “xyz2mol” is also kept in MAFIA-MD for batch processing of multiple .XYZ files at the same time by using the “Chemistry Only” functionality (see Sec. 2.3.4). This is particularly useful for the extraction of bond information from .XYZ files without requiring the large bond information files generated by RMD simulation.

2.2.3 Identification and analysis of molecular fringes

High Resolution Transmission Electron Microscopy (HRTEM) is often used to observe and characterize soot particles from different sources [66, 84, 85]. In HRTEM

images, nano-structures of soot particles are observed and characterized using the characteristics of the fringes formed. Fringes are formed due to the optical interaction of internal nanostructures within soot clusters during microscopy. During HRTEM imaging, the fringes are usually characterized based on fringe length, fringe spacing, and fringe tortuosity. Figure 2.2 presents examples of the important characteristics of fringes obtained from HRTEM images of soot. The characteristics of these fringes convey important information about the reactivity and stability of soot clusters [66]. The short fringes (fringes 1, 2, and 4 of Fig. 2.2) have more free edges per unit length compared to the long fringes (fringe 3 of Fig. 2.2). The atoms in these free edges are reactive and therefore, soot clusters exhibiting shorter fringe lengths are more reactive. Similarly, the fringes with wider spacing (fringe 2 of Fig. 2.2) have more spaces between the molecules to diffuse oxygen molecules compared to the fringes with narrow spacing (fringes 1, 3, and 4 of Fig. 2.2). Therefore, the soot clusters producing narrowly spaced fringes are less prone to oxidation, hence less reactive. The tortuosity of fringes is a representation of the amount of curvature observed in optical fringes. Molecules with higher curvature produce fringes with high tortuosity (fringe 4 of Fig. 2.2) during HRTEM imaging. Due to the higher curvature, the bond strain is higher in these molecules and they break easily which results in higher reactivity.

The current version of **MAFIA-MD** implements a scheme to calculate the fringe spacing of a cluster of carbon molecules. The calculation for fringe spacing is based on the orientation of different cyclic structure in the domain. Two structure is assumed to form a fringe if the structures are in close proximity ($3\text{\AA} - 6\text{\AA}$) [86] and parallel to each other. With respect to the parallelity constraints, a deviation up to an angle Φ is allowed. The value of Φ is hard-coded in the code as 10° . Algorithm 2 shows the algorithm implemented to calculate the fringe spacing histogram in **MAFIA-MD**. It is important to note that, for good statistics, a significant number of fringes must form in the soot cluster. Therefore, the analysis is only meaningful when the size of the cluster is large enough to contain multiple fringes.

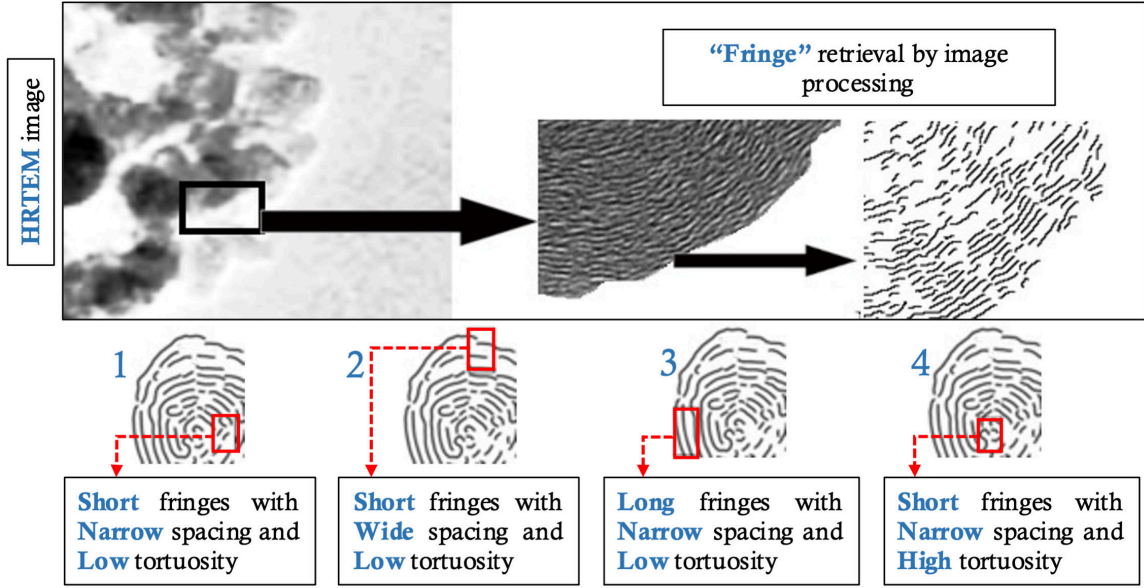


Figure 2.2: Examples and characterization of optical fringes obtained from a hypothetical HRTEM image of a soot particle.

Algorithm 2: Algorithm for calculating fringe spacing

Result: Fringe spacing histogram

```

1 # Input:
2 Coordinates of the carbon atoms inside the simulation domain;
3 List of all cyclic molecules in the domain from previous code segment;
4 # Pre-Processing:
5 Create a list of the coordinates of the centroids of each ring;
6 Create a list of all the vectors perpendicular to the existing rings through the centroids
  (surface vectors);
7 # Fringe spacing calculation:
8 fringeSpacing = [];
9 iter = 0 ;
10 while all the points in the centroid array is traversed do
11   if  $3\text{\AA} \leq \text{distance between two centroids} \leq 6\text{\AA}$  then
12     if  $(0 \leq \text{angle between the two surface vectors} \leq \Phi)$  or
13        $(180-\Phi \leq \text{angle between the two surface vectors} \leq 180)$  then
14       fringeSpacing[iter] = distance between two centroids ;
15       iter = iter +1 ;
16     end
17   end
18 end
19 Create histogram from fringeSpacing array with kernel density estimation (KDE) [87] ;
20 # Output:
21 Histogram of fringe spacing;
22 Kernel density estimation (KDE) [87];
23 Fringe spacing vs. Angle;

```

2.3 Workflow

2.3.1 Code structure

The core functionality of MAFIA-MD is implemented using Python 3. The class `FindRing()` is created for identification of rings and subsequent calculations as indicated in Algorithm 1. For chemical characterization discussed in Sec. 2.2.2, it uses an external tool `xyz2mol` [82] to extract the chemical information (SMILE strings [70] and SDF [71] files). The characterization and analysis of fringes (Sec. 2.2.3) requires the identification of rings to be performed beforehand and can be thought of as an extension of the ring identification functionality. The external tool `xyz2mol` requires the open-source cheminformatics interface “RDKit” [83], which in turn requires “anaconda” [88], an opensource package and environment management system for Python. The two segments of the code (i.e., ring identification and chemical characterization) are connected using a graphical user interface (GUI) using `tkinter` [89].

The main computational complexity of the ring identification portion of the code comes from the detection of simple cycles from the interconnected network of carbon atoms. This portion of the code takes almost 64% time to execute for the given set of trajectory files. This complexity can either increase or decrease based on the complexity of the network of molecules. Figure 2.3 depicts the schematics of the code structure and computational loads of relevant functions used in the code. The callgraph for MAFIA-MD is shown in the Fig. 2.4.

2.3.2 Program deployment

Any operating system containing `anaconda` and Python 3 can run MAFIA-MD. The present code is tested on popular operating systems like Windows 10, Ubuntu (16.04, 18.04 and 20.04), Fedora 34, CentOS 8, Debian 10, MacOS (Catalina and Big Sur). The deployment procedure is described below:

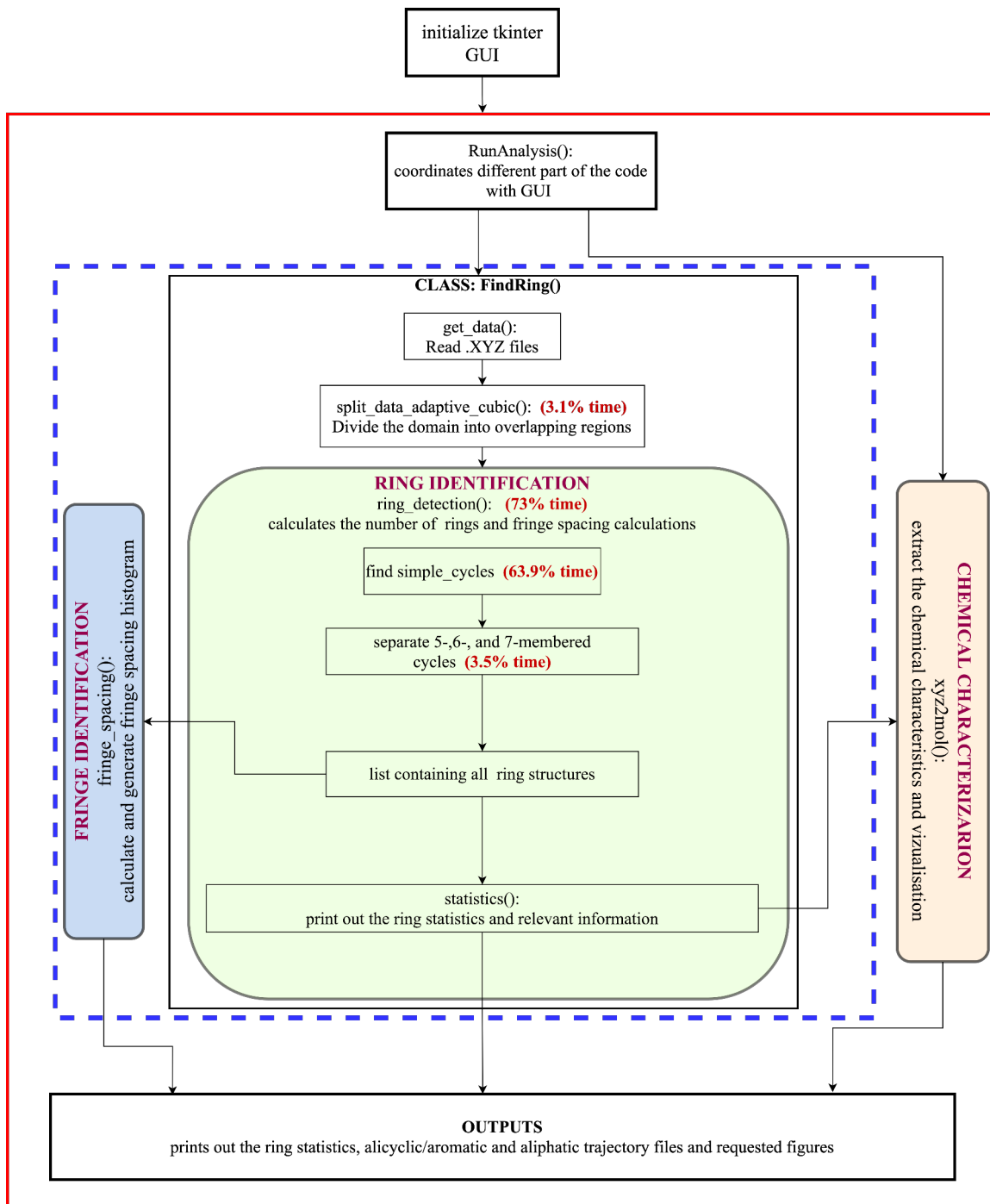


Figure 2.3: Code structure and relevant computational loads

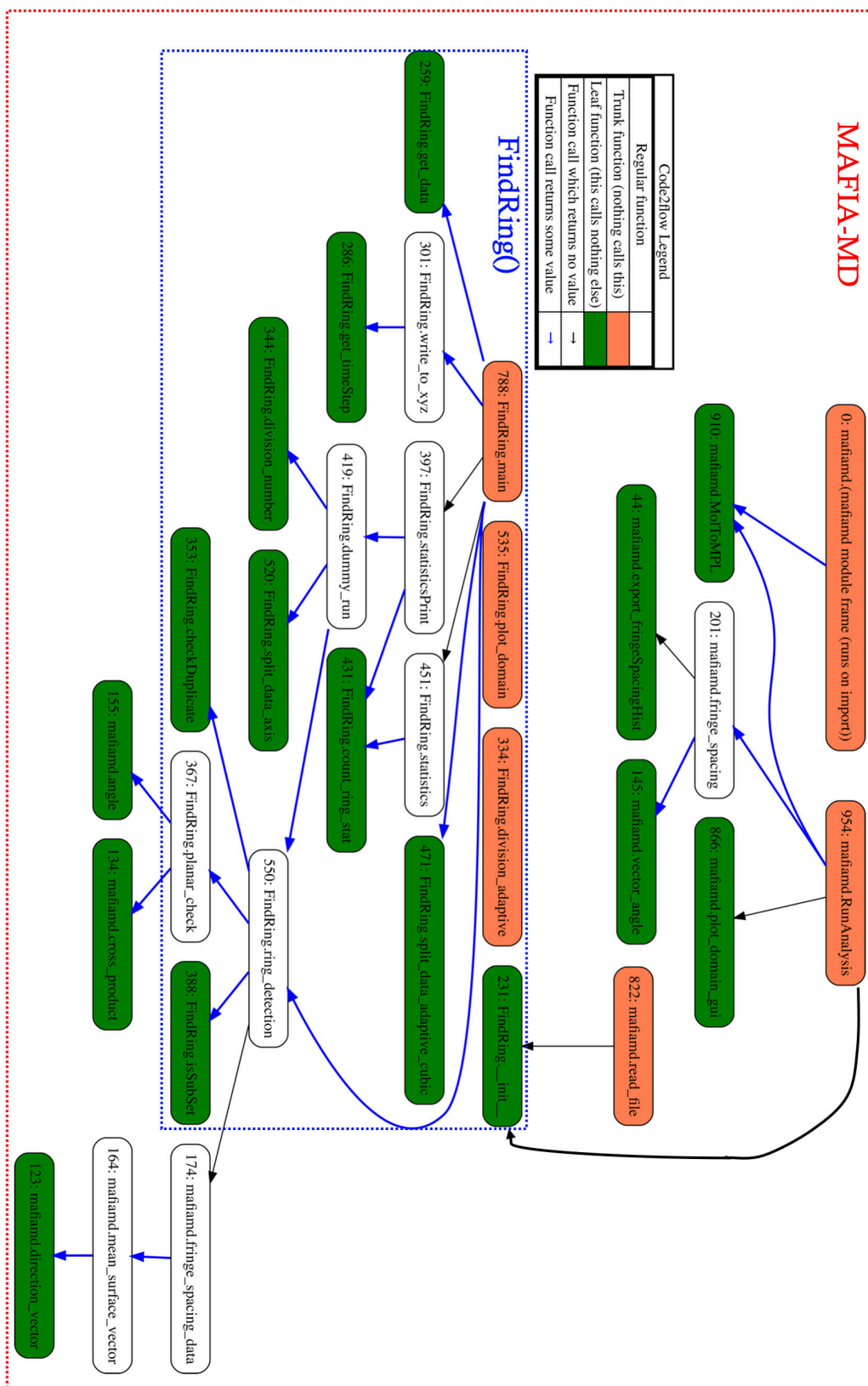


Figure 2.4: Callgraph of MAFIA-MD

1. Install `conda` or `miniconda` (installation instructions can be found in their respective websites [\[90, 91\]](#))
2. Download the `MAFIA-MD` repository. The installation files contain following files and directories
 - The parent directory:
 - `mafiamd.py`: The main Python code
 - `Makefile`: Makefile for Linux-like systems
 - `requirements.yml`: Build-requirements list
 - `LICENSE`: License file
 - `README.md`: Readme file
 - `callgraph.png`: The callgraph of the code
 - The `external_tool` directory: contains the `xyz2mol.py` package and its license
 - The `input` directory:
 - The `set1_validation_demo` directory: contains example files for validation (Sec. [2.3.5](#))
 - The `set2_fringe_analysis_demo` directory: contains example files for fringe analysis (Sec. [2.3.5](#))
 - The `output` directory: a blank demo directory for storing output
 - The `ancillary_script` directory: contains an ancillary script for splitting continous trajectories files in discrete individual timesteps and a demo example trajectory file
3. Go to the code's parent directory and create a virtual environment
 - Linux and MacOS: Use the Makefile


```
$ make
```
 - Windows: Use the `Anaconda Prompt` installed during the installation of `conda` or `miniconda` and execute


```
$ conda env create -f requirements.yml -p MAFIAMD
```

4. Activate the `conda` environment

```
$ conda activate ./MAFIAMD
```

5. Execute the code:

```
$ python mafiamd.py
```

This step (Step 5) will bring out the GUI (shown in Fig. 2.5) for specifying the analysis parameters discussed in Sec. 2.3.4. The execution time of the program can range from a few seconds to a few hour for each trajectory file depending on the size and complexity (i.e., interconnectivity of the carbon atoms) of the network. Once the execution is completed, close the program by using the **Quit** button on the GUI.

6. to deactivate the `conda` environment after execution

```
$ conda deactivate
```

2.3.3 Analysis modes

The user can perform the following types of analyses in **MAFIA-MD**:

1. Detection of cyclic structures and analysis: This is the main operational mode of **MAFIA-MD**. This mode will analyze the provided trajectory files and generate relevant information like ring statistics, C/H ratio, percentage of cyclic carbon atoms, etc. The result can be visualized in GUI by selecting the **Plot** option.
2. Sanity check: This mode is triggered by selecting the **Sanity Check** option. In this mode, the non-overlapping domain decomposition is performed *along with* the overlapping domain decomposition scheme from step 1.
3. Chemical characterization: This mode is triggered by selecting the **CHEM Calculation** option. This will perform step 1 and generate chemical information from the results in terms of SMILES or SDF files. The user can also visualize the existing alicyclic and aromatic rings by selecting the **Show Molecule** option.
4. Chemistry only: This mode is performed when only the chemical characterization of trajectory files are required and can be enabled by selecting the **Chemistry Only**

- option in the GUI. This will analyze the entire input trajectory and perform chemical characterization (step 3) without performing the ring detection analysis from step 1.
5. Fringe analysis: The user can use this functionality by selecting the **Fringe Spacing** option in GUI. This will perform step 1 and generate fringe spacing histogram for the input trajectory files along with kernel density estimation(KDE)[87].

2.3.4 Program input

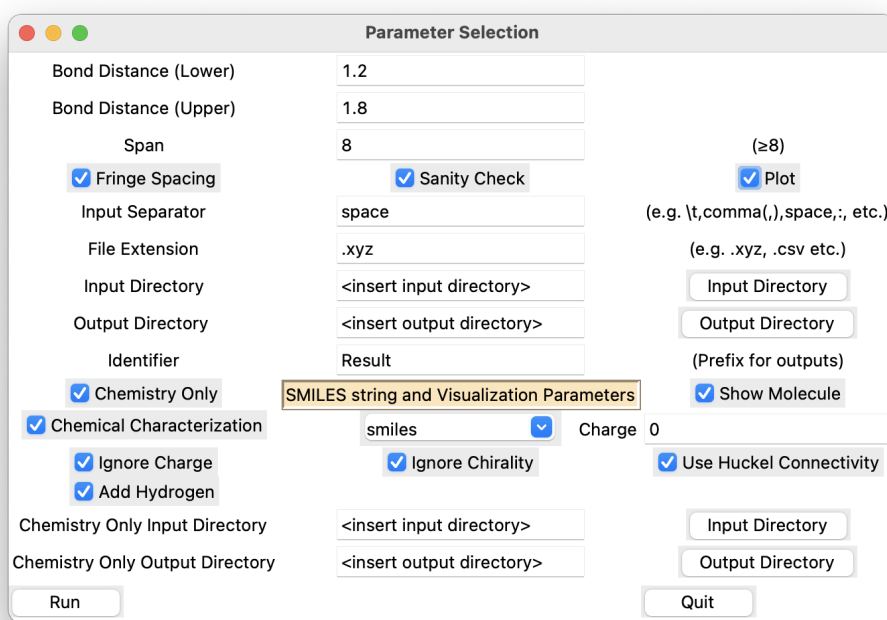


Figure 2.5: Program graphical user interface

- **Bond Distance (Lower); required:** minimum (lower limit) carbon-carbon bond distance (in Å). The default value is 1.2.
- **Bond Distance (Upper); required:** maximum (upper limit) carbon-carbon bond distance (in Å). The default value is 1.8.

- **Span; required:** twice the amount of overlap between two consecutive subdomains (in Å). A span of 8Å is found to be optimum for finding upto 7-membered rings. Therefore, the default value is 8.
- **Input Separator; required:** separator used between (X,Y,Z) co-ordinates in the input XYZ files. The default value is **space** (whitespace). The output trajectory files generated by the code is always tab separated.
- **File Extension; required:** extension of the input trajectory files. This option is kept for the cases when the input trajectory files are provided in comma separated format (.csv). The default value is **.xyz** (with the dot).
- **Input Directory; required:** directory containing the input trajectory files. The code will analyze *all* files kept in the specified input directory.
- **Output Directory; required:** directory where the output files (e.g., logs and plots) will be saved.
- **Identifier; required:** a text string prepended to each output file-name to differentiate between the separate instances of **MAFIA-MD** runs. The default value is **Result**.
- **Charge:** useful for **Chemistry Only** runs where a trajectory file contains a net charge.
- **Chemistry Only Input Directory:** directory containing the input trajectory files for **Chemistry Only** analysis.
- **Chemistry Only Output Directory:** directory where the output files will be saved for **Chemistry Only** analysis.
- **Analysis mode options:**
 - *Sanity Check:* when selected, the sanity check segment of the code will be executed along with the overlapping domain decomposition scheme.
 - *Plot:* when selected, the existing rings in the domain will be plotted in a separate window
 - *Fringe Spacing:* when selected, fringe spacing histogram is calculated and printed out.

- *Chemical Characterization*: when selected, the chemical characterization will be performed along with the ring analysis. Select “**smiles**” to get a SMILE string for the existing molecule or “**sdf**” to get individual SDF files representing the input trajectories. The following sub-options are available for chemical analysis.
 - *Show Molecule*: when selected, the ring structures found in the domain will be drawn on a separate window using RDKit.
 - *Ignore Charge*: when selected, the chemical calculation will not consider the effect of net charge during calculation
 - *Ignore Chirality*: when selected, the chemical calculation will not consider the effect of chirality during calculation. This is a functionality of the external tool xyz2mol.
 - *Use Huckel Connectivity*: when selected, the chemical calculation will use extended Huckel bond orders to locate bonds during calculation. Otherwise, van der Waals radii will be used. This is a functionality of the external tool xyz2mol.
 - *Add Hydrogen*: when selected, the output images will also contain hydrogen atoms. The default is to plot only carbon atoms.
- *Chemistry Only*: when selected, only the chemical analysis will be performed on the trajectory files considering both hydrogen and carbon atoms and it will not perform any ring identification. It is helpful for visualizing the general chemical structures inside the domain.

2.3.5 Numerical example

Example cases

Two sets of input trajectories, included with the code, are used to demonstrate the capabilities of MAFIA-MD. The first set contains three relatively small trajectories for the purpose of validation of the ring detection method. These three test trajectories are

`fabricated.xyz`, `fabricated2.xyz`, and `real_MD.xyz` and are kept in one single directory. These three input trajectory files are `whitespace` separated. Two of these are fabricated (`fabricated.xyz` and `fabricated2.xyz`) and the other one (`real_MD.xyz`) is selected from an actual molecular dynamics study performed by [57] [57] where an earlier version of MAFIA-MD [92] was used. The actual molecular dynamics trajectory file `real_MD.xyz` has 760 atoms. The fabricated trajectories are prepared using openbabel [93]. The details of these trajectories are

1. `fabricated.xyz`: 58 carbon atoms and 96 hydrogen atoms
 - Designed ring count:
 - 5-membered: 4, 6-membered: 8, 7-membered: 1
2. `fabricated2.xyz`: 110 carbon atoms and 190 hydrogen atoms
 - Designed ring count:
 - 5-membered: 5, 6-membered: 7, 7-membered: 2
3. `real_MD.xyz`: 559 carbon atoms and 201 hydrogen atoms
 - Manually counted ring count:
 - 5-membered: 14, 6-membered: 20, 7-membered: 6

In a second directory, the second set of test case is kept. This set consists of a single large trajectory file named `1769atoms.xyz` (1256 carbon atoms and 513 hydrogen atoms), which is also selected from [57]. This file is `tab` separated. Due to the size of this trajectory, it was not possible to manually count the rings in this trajectory for validation. This trajectory is used to show the capabilities of fringe analysis.

Validation tests

In the first set of tests, the directory containing the first set of trajectories (`fabricated.xyz`, `fabricated2.xyz`, and `real_MD.xyz`) was selected as input directory and MAFIA-MD performed analysis of all three trajectories simultaneously. The parameter selection and output for this set of trajectory files are shown in Fig. 2.6 and the output text file containing the result is shown in Listing 2.1. The output file contains all the requested

information and it correctly determines the exact number of rings present in the supplied trajectories, thereby validating the code.

Listing 2.1: Output file

```
-----
1:      fabricated.xyz
-----
```

```
Starting code for Adaptive Run
-----
```

```
-----
C/H Ratio:      0.6041666666666666
Total Alicyclic/Aromatic Carbon:      52
Total Aliphatic Carbon Number:      6
Total Existing Rings      {7: 1, 5: 4, 6: 8}
Percentage of Alicyclic/Aromatic components      : 0.896551724137931
Percentage of Aliphatic components      : 0.10344827586206895
-----
```

```
1:      Sanity Check for:      fabricated.xyz
-----
```

```
-----
Starting code for Axis:      0
-----
```

```
-----
C/H Ratio:      0.6041666666666666
Total Alicyclic/Aromatic Carbon:      52
Total Aliphatic Carbon Number:      6
Total Existing Rings      {5: 4, 6: 8, 7: 1}
Percentage of Alicyclic/Aromatic components      : 0.896551724137931
Percentage of Aliphatic components      : 0.10344827586206895
-----
```

```
1:      smiles_fabricated      : START
-----
```

```
[C]1=C=C=C=C=C2[C]=C=C=C12.[C]1=c2c([c]c3[c]c4c(c5[c]c#cc2c35)=C=C=C=4)C#C1.[C]1C#CC2=c3c(c4[c]c#cc5c4c4c(c#c[c]c34)=C=C=5)=C=C=C12
-----
```

1: smiles_fabricated : END

2: fabricated2.xyz

Starting code for Adaptive Run

C/H Ratio: 0.5789473684210527
 Total Alicyclic/Aromatic Carbon: 69
 Total Aliphatic Carbon Number: 41
 Total Existing Rings {5: 5, 6: 7, 7: 2}
 Percentage of Alicyclic/Aromatic components : 0.6272727272727273
 Percentage of Aliphatic components : 0.3727272727272727

2: Sanity Check for: fabricated2.xyz

Starting code for Axis: 0

C/H Ratio: 0.5789473684210527
 Total Alicyclic/Aromatic Carbon: 69
 Total Aliphatic Carbon Number: 41
 Total Existing Rings {6: 7, 5: 5, 7: 2}
 Percentage of Alicyclic/Aromatic components : 0.6272727272727273
 Percentage of Aliphatic components : 0.3727272727272727

2: smiles_fabricated2 : START

C1=C=C=c2c#cc#cc2=C=1.C1=C=C=c2c3[c]c#cc-3[c]c#cc2=C=1.[C]1=C2C(=C=C=c3c2[c]c(=C2C#CC#C2)c2c3=C=C=C=C=2)C#CC#C1.[C]1=C=C=C=C1C1=C=C=[C]C#C1.[C]1C#CC#C1.[C]1C#CC#C1

2: smiles_fabricated2 : END

```
-----
-----
3:      real_MD.xyz
-----
```

```
Starting code for Adaptive Run
-----
-----
```

```
C/H Ratio:      2.7810945273631837
Total Alicyclic/Aromatic Carbon:      158
Total Aliphatic Carbon Number:      401
Total Existing Rings      {5: 14, 6: 20, 7: 6}
Percentage of Alicyclic/Aromatic components      : 0.2826475849731664
Percentage of Aliphatic components      : 0.7173524150268336
-----
```

```
3:      Sanity Check for:      real_MD.xyz
-----
-----
```

```
Starting code for Axis:      0
-----
-----
```

```
C/H Ratio:      2.7810945273631837
Total Alicyclic/Aromatic Carbon:      158
Total Aliphatic Carbon Number:      401
Total Existing Rings      {6: 20, 5: 14, 7: 6}
Percentage of Alicyclic/Aromatic components      : 0.2826475849731664
Percentage of Aliphatic components      : 0.7173524150268336
-----
```

```
3:      smiles_real_MD : START
-----
```

```
C1#CC(=C2C#Cc3c#cc#cc32)C#C1.C1=C=c2[c]c3c#cc4c#cc5[c]c6c#cc=1c6c2c3c45.[C]1=C=C=C2C#CC3=[C][C@@
]34C3=C=C=c5c#cc6c7c5c3c(c3[c]c5[c]c#cc8[c]c9c#c[c]c%10c(c(c37)-c(c85)c9%10)=C=C=6)C1=C24.[C
]1=C=C=C2C#C[C]=C=C12.[C]1=C=C=c2[c]c3[c]c4c#cc#cc4[c]c4c5c6c(c#cc#cc6c(c21)c34)=C=C=C=5.[C
]1C#CC#C1.[C]1C#CC#C1.[C]1C#CC#C1.[C]1C#CC#C1.[C]1C#CC#C1.c1[c]c2[c]c-2c#1.c1c#cc#cc#1.c1c#
cc#cc#1
-----
```

```
3:      smiles_real_MD      : END
```

Demonstration of fringe analysis

The second test is done on the input directory containing only `1769atoms.xyz`. The fringe analysis capability is demonstrated in this test. The GUI for fringe spacing calculation mode and the output from MAFIA-MD is presented in Fig. 2.7(a). The output contains the fringe histogram and a list of fringe spacing vs. angle between the surface normal of two rings (to show the parallelity of the planes). The fringe spacing histogram generated by MAFIA-MD is shown in Fig. 2.7(b). The figure for the fringe spacing histogram with kernel density function (KDE) and a complete output log is stored in the user specified output directory along with the graphical output shown in Fig. 2.7. The output log for fringe spacing calculation is shown in Listing 2.2.

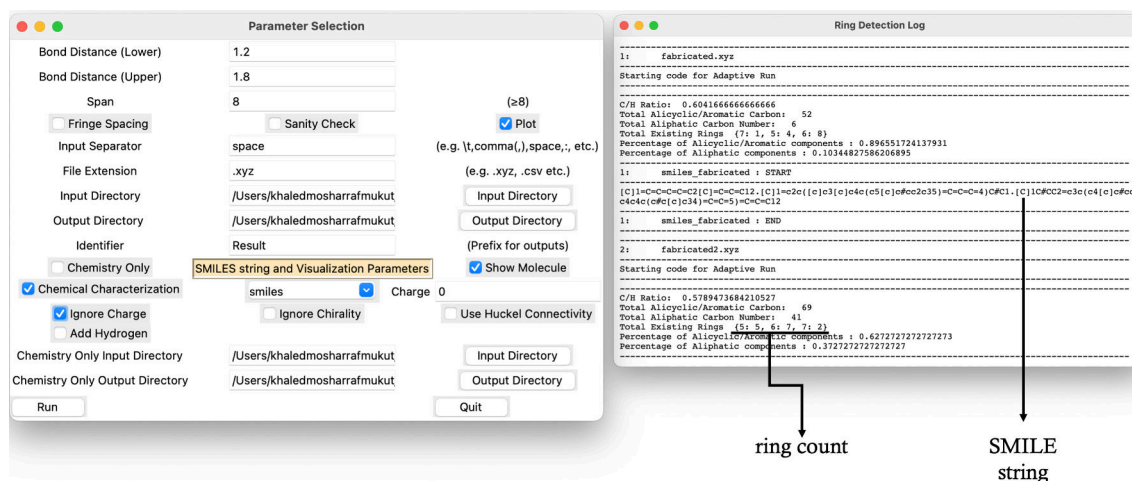
Listing 2.2: Fringe spacing calculation mode output log

```
1:      1769atoms.xyz
```

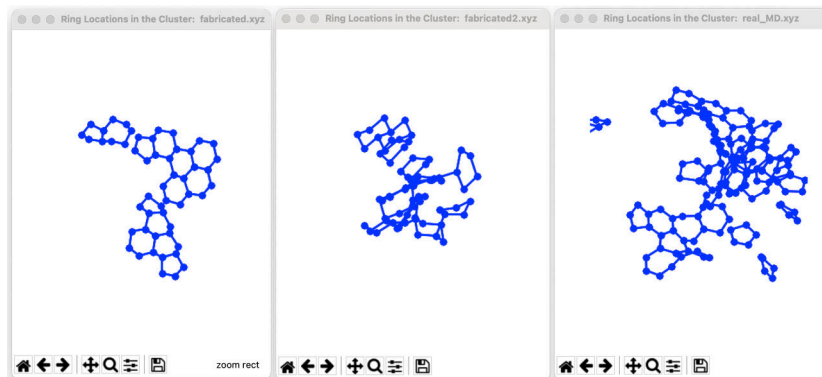
```
Starting code for Adaptive Run
```

```
C/H Ratio:      2.448343079922027
Total Alicyclic/Aromatic Carbon:      747
Total Aliphatic Carbon Number:      509
Total Existing Rings      {6: 103, 5: 42, 7: 54}
Percentage of Alicyclic/Aromatic components      : 0.5947452229299363
Percentage of Aliphatic components      : 0.40525477707006374
```

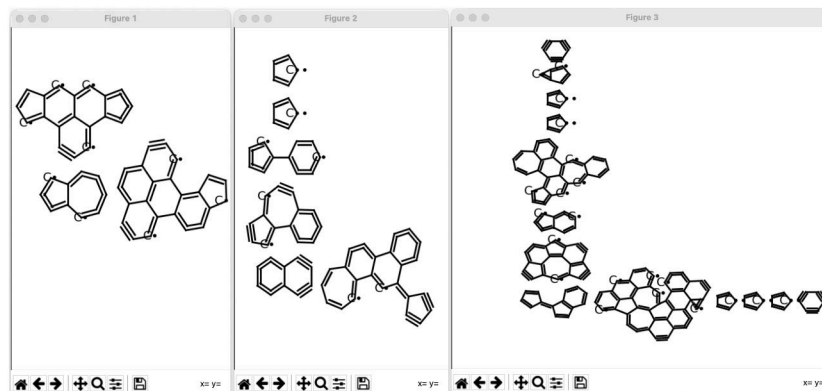
```
1:      Fringe Spacing_1769atoms      : START
```



(a) Parameter section and output window



(b) Plot of alicyclic and aromatic carbon atoms for fabricated.xyz, fabricated2.xyz and real_MD.xyz obtained from the GUI



(c) Molecules extracted from fabricated.xyz, fabricated2.xyz and real_MD.xyz respectively from left to right

Figure 2.6: Execution and output of MAFIA-MD

Fringe spacing: Fringe Spacing vs Angle

5.950012676836169	3.9242557263703115
5.290388731629016	0.0
4.629775482026949	5.429560409174904
4.419991063508438	9.246261423942366
5.612903574453927	1.6364418180798508
5.163402653281143	8.67062142732157
5.008754961693465	8.111938792537524
3.968375987304109	9.651232302073426
4.623928548881482	8.606666094895274
4.58422686097776	5.729706297680085
5.23611387847295	7.048151243502113
4.206558651539378	5.075347713313028
4.714066484238856	4.613551102989272
4.102726666254199	7.199702485266698
4.185954445820429	8.694128830794705
5.288683680451116	8.082443661468403
5.5246824885358565	9.121954527071836
3.9753826457241774	5.869899451671865
4.574647618386823	8.868308926350613
5.368046285546225	6.273100098502783
5.228879733893831	4.206006987785507
4.792224005753124	6.96472208665477
5.173526348729457	175.53685851260337
4.857647329007471	176.0514868131242
4.035298892998304	173.28483345021596
4.3219796323389685	170.746229983982
5.409863373609561	9.259381236595647
4.175217901573789	180.0
3.988760975726621	178.47288283324156
5.333240479100813	171.310888466995
5.126083238875523	8.463148316690887
3.979080838899854	8.951229787825891
5.055383120720362	5.3254081727125575
4.3724567234374465	5.943245793999107
4.783263574854577	5.033070521661905
5.869554203719395	5.342639370384733
5.239332298840521	9.795728302807293
3.4251790290727895	9.152348900213486

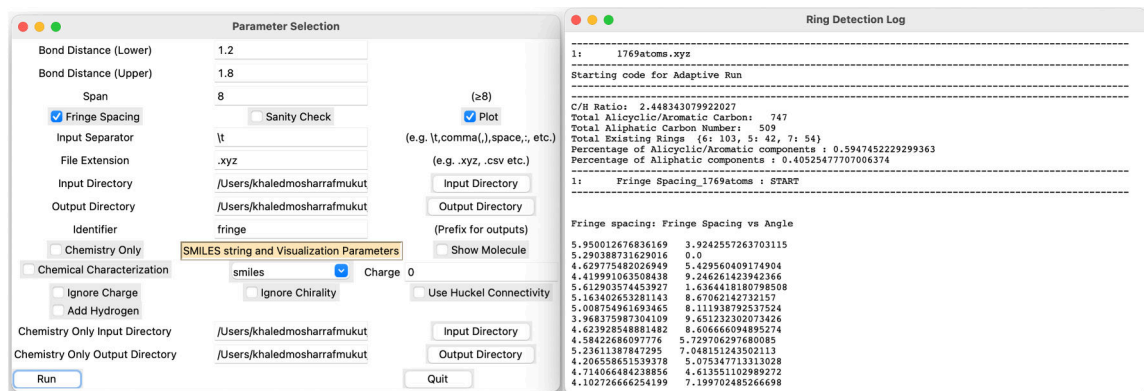
Fringe Spacing Histogram

Bins: [3. 3.25 3.5 3.75 4. 4.25 4.5 4.75 5. 5.25 5.5 5.75 6.]
 Hist: [0 1 0 4 5 3 5 3 8 5 2 2]

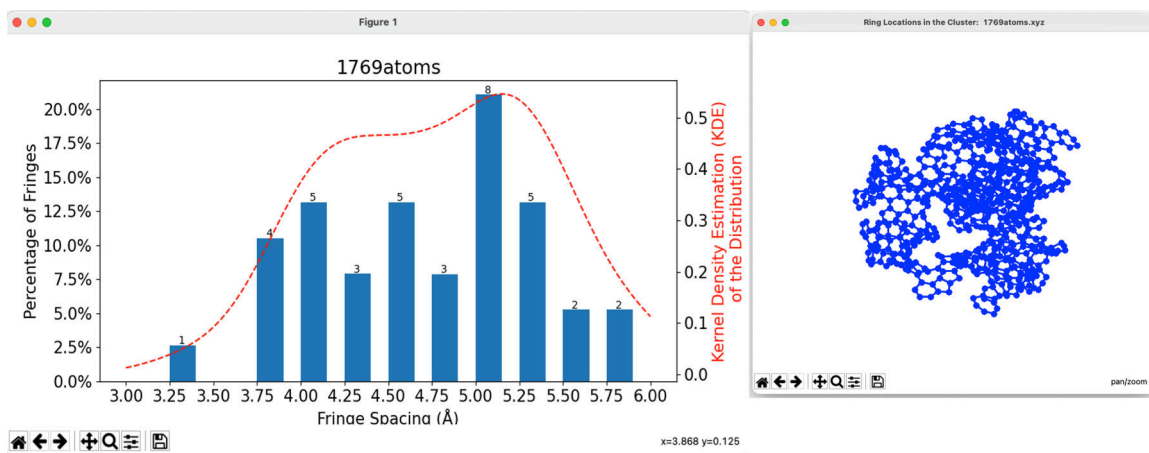
Fringe Spacing Histogram

End of Fringe spacing

1: Fringe Spacing_1769atoms : END



(a) Fringe spacing mode GUI (left) and output (right)



(b) Fringe spacing histogram (left) and rings inside domain (right)

Figure 2.7: Fringe spacing calculation mode outputs.

Table 2.2: Computational cost breakdown of MAFIA-MD

Trajectory Files		Computational Cost (s)		
Name	Composition	Ring Detection	Chemical Analysis	Fringe Spacing
fabricated.xyz	58 Carbon and 96 Hydrogen	0.2109	0.0201	NA
fabricated2.xyz	110 Carbon and 190 Hydrogen	0.2475	0.0313	NA
realmd.xyz	559 Carbon and 201 Hydrogen	1.6093	0.0944	NA
1769atoms.xyz	1256 Carbon and 513 Hydrogen	5.0159	1.5683	1.0928

Computational cost

The main computational cost of **MAFIA-MD** is in the ring detection part of the code as indicated in Fig. 2.3. The total computational cost of **MAFIA-MD** for the ring detection, chemical analysis and fringe spacing calculation for the example trajectory files when used on a system with intel 2.3GHz Core i5-8259U CPU is presented in Table 2.2.

2.4 Future works

The post-processing tool **MAFIA-MD** presented in this manuscript is capable of analyzing molecular dynamics trajectory files for hydrocarbon reactions. The current implementation is not parallelized. From Fig. 2.3, we see that the most computationally intensive (64%) part of the code is the part where it identifies the rings from the network of carbon atoms. The cycle-finding segment of the code is parallelizable and doing so will increase the functionality of **MAFIA-MD** in analyzing very large atomic clusters and reduce the runtime. The overlapping regions do not have any dependency on each other and therefore can take the advantage of many core systems like GPUs. This kind of optimization can enable real-time analysis of large scale reactive molecular dynamics simulation. The authors will continue working toward this goal in the future.

2.5 Summary

A useful post-processing utility called **MAFIA-MD** is presented in this manuscript for analyzing reactive molecular dynamics trajectory of hydrocarbon molecules. **MAFIA-MD** is written in **Python-3.7** and has been tested successfully in Windows, Linux, and MacOS. **MAFIA-MD** extracts the alicyclic and aromatic hydrocarbons by identifying cyclic structures

which are important in understanding complex physico-chemical phenomena in soot formation and growth. It also extracts the relevant soot related chemical information like C/H ratio and aliphatic to alicyclic/aromatic carbon ratio, etc. A methodology of calculating fringe spacing is implemented as well for diagnostic studies. **MAFIA-MD** have the ability to work with large number of simple trajectory files generated by reactive molecular dynamics simulation simultaneously. This capability of batch processing large number of files makes **MAFIA-MD** very useful for researchers working on large scale reactive molecular dynamics simulation.

Acknowledgements

The author would like thank Md. Hasanul Aziz for his guidance toward the development of **MAFIA-MD**.

CHAPTER 3

ACETYLENE PYROLYSIS CASE-I: BOUNDARY BETWEEN GAS PHASE SPECIES AND SOOT

This chapter is a reproduction of the article already presented at 12th US National Combustion Meeting as: "Mukut, K., Sharma, A., Goudeli, E., & Roy, S. (2021). A Closer Look into the Formation of Soot Particles: A Molecular Dynamics Study. In 12th US National Combustion Meeting, College Station, TX, USA, May 2021." [94]

3.1 Introduction

Soot or black carbon comprises tiny nano-particles and forms during incomplete combustion of hydrocarbon fuels [3]. Internal combustion engines, coal-burning power stations, wildfire are significant sources of soot emission. Soot particles can harm both public health and the environment. Exposure to the soot particles can cause critical health issues like DNA mutation, lung cancer [7], and cardio-vascular diseases [9]. Alongside the numerous adverse health effects, soot particles also play a detrimental role in the environment. It can heavily influence the radiative heat transfer in the atmosphere and contribute to global warming by absorbing sunlight and directly heating the atmosphere through radiative forcing. Soot can also increase the snow melting rate in the arctic ice by accumulating on top of it [10]. Therefore, it is essential to understand how these soot particles form, grow, age, and influence the atmosphere.

Soot formation is a highly complex process due to the massive complexity of the combustion reaction network. The actual soot formation process remains a topic of debate to this date, and a proper nucleation pathway is still unknown. The transformation of the gas-phase molecules to the fractal-like soot aggregates can be divided into a few steps to understand better and model soot formation, i.e.

1. formation of soot precursors including (but possibly not limited to) polycyclic aromatic hydrocarbons (PAHs) from the gas phase molecules,
2. inception/nucleation of soot particles from the soot precursors,

3. surface growth by addition of gas-phase molecules,
4. coagulation and coalescence of smaller soot particles to form large aggregates, and
5. oxidation of soot particles back to the gas phase.

Nucleation is the least understood process that occurs during soot formation. Frenklach et al. [14] proposed a detailed soot inception model based on a chemical kinetic description of acetylene pyrolysis and oxidation, where the aromatic molecules form and grow up to a certain size, and then grow by replication to form larger soot particles. At a later study, Frenklach et al. [12] expanded the nucleation model by assuming the sticking of PAH molecules of a certain size to form dimers, which then subsequently grow upon collision with each other forming trimers, tetramers, and so on. This dimerization or stacking of the precursor PAH molecules is often adopted as the nucleation mechanism during soot formation. Pyrene [95] and coronene [96] are the most commonly adopted soot precursor PAHs for numerical modeling purposes [21].

However, the dimerization assumption is being challenged by the recent findings. In contemporary studies, resonance stabilized radicals (RSRs) play the most critical role during soot nucleation through low-barrier hydrogen-abstraction and hydrogen-ejection reactions [67]. The RSRs can readily react with acetylene or vinyl radicals to grow and initiate soot clusters through radical chain reactions. Reflecting upon their earlier works, [65] presented a new concept for soot nucleation where a rotationally-activated dimer is proposed to be formed upon collisions between an aromatic molecule and an aromatic radical. The stabilization of these dimers is achieved by a doubly-bonded bridge between the molecules and the formation reaction is governed by a five-membered ring located at the edge of each molecule. With these recent discoveries and proposition, the crucial next step is to determine the rate constants for the series of reactions. The complete reaction mechanisms can pave the way to more accurate predictive soot models, which can be used to model soot formation in flames and practical combustion devices.

Considering the physical length and timescale of soot formation, reactive molecular dynamics (RMD) simulations are gaining more and more popularity in determining the

missing link in soot nucleation [61, 60, 62, 63]. The reactive force field (ReaxFF) proposed and developed by [51] shows the capability of capturing the chemical evolution in high pressure and high-temperature environment. ReaxFF is developed based on the bond order of atoms to describe bond breakage and formation.

The present study utilizes this ReaxFF potential for RMD simulation to investigate soot nucleation during pyrolysis of acetylene molecules. Incipient soot particles of size up to 3.5 nm are observed to form by clustering reactions among 1000 acetylene molecules at 1500 K. Various structural and compositional aspects of soot particles are investigated to characterize the incipient soot particles, such as C/H ratio, the abundance of 5-,6- and 7-membered rings, aromatic to aliphatic carbon ratio, the radius of gyration, atomic fractal dimension, and packing density.

3.2 Numerical Methodology

One thousand acetylene molecules are placed randomly in a domain ($83 \text{ \AA} \times 68 \text{ \AA} \times 74 \text{ \AA}$) at 1500 K. The temperature is chosen because of its relevance in a favorable soot formation condition [97]. The ReaxFF potential for hydrocarbons formulated by [98] is used to calculate the bond breaking and formation due to molecular collision. During the molecular dynamics calculation using ReaxFF, the bond lengths are adjusted based on the local chemical environment changes and hence update every timestep [99]. These repeated adjustments help accurately describe the bond cleavage and bond formation and consequently capture the chemical reaction leading to radicals formations during nucleation. The periodic boundary condition is assumed in all three directions to establish a semi-infinite approximation. Velocity-Verlet algorithm [100] is employed with a timestep of 0.25 fs [101] using the Nose-Hoover thermostat [102]. The damping parameter for the thermostat is set at 10 fs. Constant number, volume, and temperature (NVT) ensemble strategy are used to run RMD simulation up to 3.0 ns using Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [103] MD code.

3.3 Chemical structure of soot clusters

The chemical information of the soot structure is analyzed using a Python post-processing script developed in house [92]. This post-processing tool extracts the number of aromatic rings (i.e., 5-, 6- and 7-membered rings) in the simulation domain and gives the total number of aromatic carbon atoms. Once the aliphatic and aromatic carbon atoms are separated the code also utilizes a universal structure-conversion method for organic molecules [69] to identify each molecule present in the domain. The identified individual molecules are expressed in terms of simplified molecular-input line-entry system (SMILES) [70] and analyzed using the “RDKit” open-source cheminformatics interface [83].

3.4 Results and discussion

As mentioned earlier, one thousand acetylene molecules are kept in a semi-infinite domain at 1500 K, and allowed to react using the ReaxFF potential using RMD simulation. The evolution of incipient soot cluster growth inside the domain is depicted in figure 3.1. The hydrogen and carbon atoms are annotated using red and black dots, respectively. From Fig. 3.1, three distinct stages of soot evolution are observed: (a) formation of linear aliphatic chains ($t=0.05$ ns), (b) cyclization of the linear aliphatic chains into first aromatic ring structures ($t=0.3$ ns), and (c) continuous growth of soot particle due to surface reaction ($t \geq 0.4$ ns). The linear aliphatic chain structures created during the initial stage of nucleation collide to form larger branched PAH structures by giving away H-atoms and gaining C-mass. After that, the continuous condensation of gas-phase species onto the aromatic structure leads to larger and larger soot particles. A similar soot formation mechanism is also observed during the formation of carbon black at 2400 – 2500 K as reported by [104].

The evolution of the carbon-to-hydrogen (C/H) ratio as a function of molecular weight of incipient soot is presented in Fig. 3.2(a). The soot cluster obtained from the current molecular dynamics simulation is plotted side by side with the clusters obtained from [105, 106]. The soot particles’ chemistry is also analyzed based on the peri-condensed, or cata-condensed [16] nature of hydrocarbons. Peri-condensed PAHs have some of the carbon atoms situated in more than two molecular constitutive rings. Cata-condensed

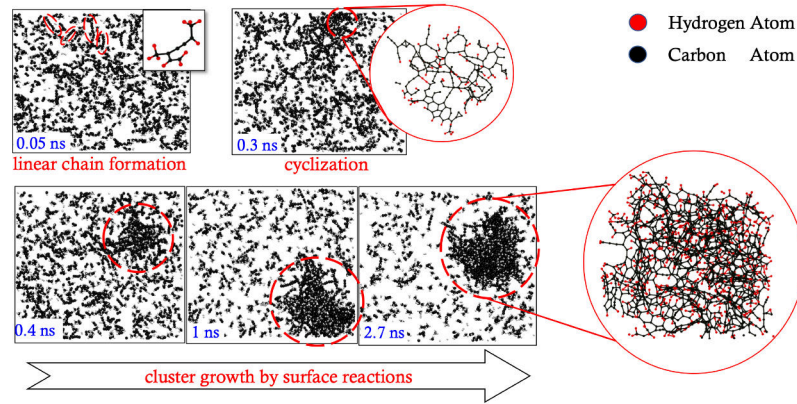


Figure 3.1: Evolution of incipient soot inside the domain from acetylene molecules

PAHs have all of their carbon atoms situated in one or a maximum of two constitutive molecular rings. The smaller soot particles (molecular weight > 1 kg/mol) are close to the peri-condensed boundary, and the larger ones are in between the peri-condensed and cata-condensed boundary, as observed from Fig. 3.2(a). This is consistent with contemporary studies and indicates that the incipient soot particles are composed of an intricate network of aromatic molecules containing 5-, 6-, and 7-member rings. With the growth of the soot particles, the aliphatic branches play an essential role in soot evolution. A similar conclusion can also be drawn from looking at the soot particle diameter shown in Fig. 3.2(b) where the smaller soot particles are observed to follow the peri-condensed PAH boundary.

Atomic fractal dimension (based on the orientation of carbon atoms in a soot cluster) of the incipient soot clusters can give an excellent insight into the soot clusters' shape. The distribution of the atomic fractal dimension for the soot clusters' molecular weight is presented in Fig. 3.3(a). The smaller soot particles exhibit smaller fractal dimensions ($D_f < 1.5$), indicating a chain-like structure. On the other hand, heavier soot clusters show larger fractal dimensions ($D_f \approx 2.8$), indicating sphere-like soot clusters. This observation is consistent with the soot particles obtained in SI engines [108]. The packing density is another vital characterization parameter to represent the compactness of a soot cluster. Packing density is the ratio of effective density of a soot cluster to bulk density

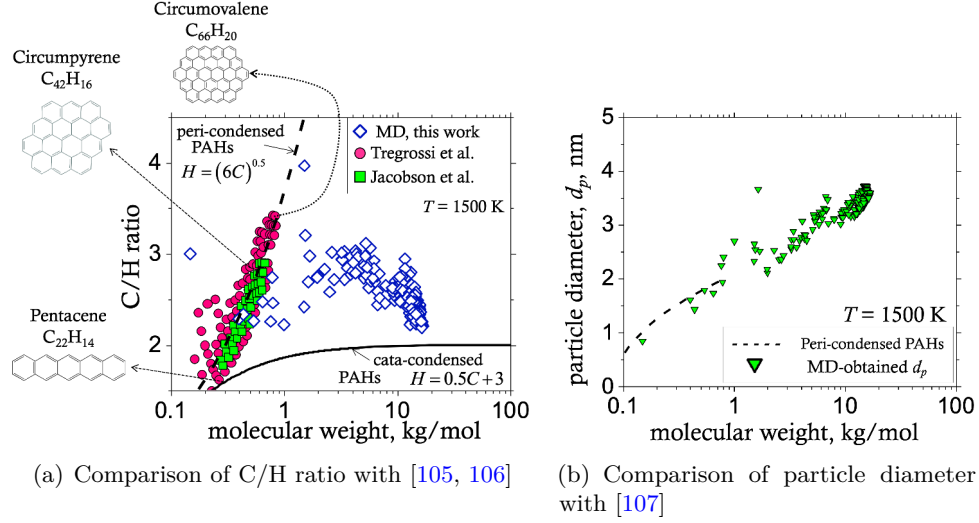


Figure 3.2: Comparison of (a) C/H ratio and (b) particle diameter (d_p) of the soot particles with results reported in literature.

(Eq.(3.1a)). The effective density (ρ_{eff}) and bulk density (ρ_{bulk}) are defined as shown in Eqs. (3.1b) and (3.1c) respectively [67].

$$\theta = \frac{\rho_{eff}}{\rho_{bulk}} \quad (3.1a)$$

$$\rho_{eff} = \frac{m_{particle}}{\frac{\pi d_p^3}{6}} \quad (3.1b)$$

$$\rho_{bulk} = (0.26088a^2c)^{-1} \left(\frac{w_C(C/H) + w_H}{C/H + 1} \right) \quad (3.1c)$$

Here,

$m_{particle}$ = particle mass

d_p = particle diameter assuming spherical particle

a = 2.46Å:graphite unit cell in the basal plane

c = 3.50Å:interlayer spacing of soot

w_C & w_H = molecular weights (g/mol) of carbon and hydrogen, respectively

The packing density of the soot clusters obtained in the present study is depicted in Fig. 3.3(b). The packing density of the smaller soot particles is also smaller and vice versa.

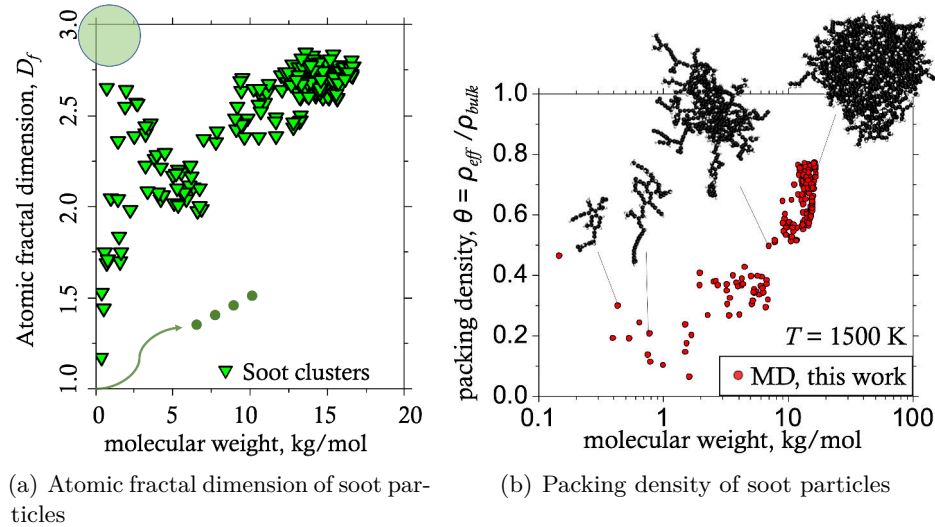


Figure 3.3: Distribution of (a) atomic fractal dimension and (b) packing density of the soot particles with molecular weight.

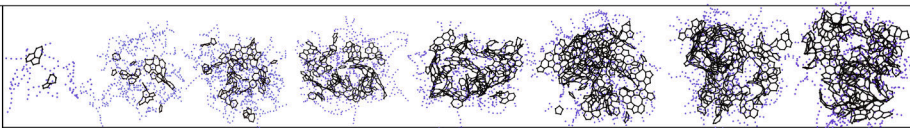
This indicates a transition of soot clusters from linear-like or porous-like structures ($D_f < 1.5$) to more compact ($D_f \approx 2.8$) clusters as they grow.

Table 3.1 presents some of the soot particles analyzed for the current study.

Different characterizing parameters such as aromatic-to-aliphatic carbon ratio, the radius of gyration, atomic fractal dimension, the total number of atoms, the fraction of 5-, 6- and 7-membered rings are also studied to see how soot particles evolve with time (0.25 ns to 2 ns). During the early stages of soot formation, the soot clusters comprise mostly aliphatic chain structures with fewer aromatic rings. The aromatic ring structures are located near the cluster's core while the aliphatic branches appear on the surface. This orientation is analogous with the proposed core-shell structure of soot particles [109]. With time, the soot particles' surface-to-volume ratio decreases, and the number of aromatics in the core increases. This results in a larger aromatic-to-aliphatic carbon ratio. The soot core composition can be characterized by fractions of 5-, 6-, and 7-membered rings. Initially, a steady rise in the 5- and 7-membered rings is observed. However, at later times, the 5- and 7-membered rings get converted into more stable 6-membered rings. This bolsters the importance of 5-membered rings during the intermediate stages of soot nucleation as championed by recent literature [64, 65]. In their new proposed soot nucleation proposition,

[65] emphasized the importance of such five-membered ring structures in the formation of bridge structures that stabilize the formation of rotationally-activated dimer due to the collision between different aromatic molecules and radicals.

Table 3.1: Characterization of soot particle evolution with time

	● Aliphatic Carbon Atoms				● Aromatic Carbon Atoms			
Primary structures								
Total atoms	178	972	1608	1450	1478	1754	1782	1800
C/H Ratio	2.423	1.809	1.5728	2.7958	2.596107	2.373	2.375	2.36
Radius of Gyration	8.3442	14.0543	16.3139	13.7136	13.5506	14.2946	14.135	14.2594
Fractal Dimension	2.1247	1.9507	2.1466	2.4643	2.6449	2.6342	2.7318	2.6188
5 membered rings	2	6	19	28	31	47	38	41
6 membered rings	1	8	29	53	81	76	76	85
7 membered rings	2	12	12	23	22	32	41	40
Time	0.25 ns	0.50 ns	0.75 ns	1.0 ns	1.25 ns	1.50 ns	1.75 ns	2.0 ns

The evolution of the chemical mixture inside the simulation domain is also analyzed to understand and isolate important molecules relevant for soot nucleation. The population distribution of the evolving molecular mixture is presented in Fig. 3.4 in terms of the number of constituting carbon (3.4(a)-3.4(d)) and molecular weight (MW) of existing species (3.4(e)-3.4(h)). During the early stages, many small molecules ($\#C \lesssim 20$ and $MW \lesssim 240$ kg/kmol) are observed, but no large molecule is present in the domain. This population of small molecules remain consistently dominant throughout the nucleation process. A few very large molecules beyond $\#C \approx 20$ are also seen at later times, but these molecules either get dissociated or get absorbed on the surface of the existing incipient soot cluster. Therefore, the molecular population can be distinctly differentiated in active and abundant small molecules and rare, transitional large molecules. This discrete jump in the relative abundance of molecules differentiated by number of constituting carbon atoms and

molecular weight signifies the importance of smaller hydrocarbon molecules and radicals ($\#C \lesssim 20$ and $MW \lesssim 240$ kg/kmol) as the building blocks of soot nucleation. Further analysis of these molecules is required to determine how they contribute to the soot nucleation mechanism and to what extent.

3.5 Conclusion and future works

The formation and evolution of soot nucleation during the pyrolysis of acetylene molecules were investigated using ReaxFF potential with RMD simulation. The MD-generated soot particles are characterized using several parameters like C/H ratio, aromatic-to-aliphatic carbon ratio, the radius of gyration, fractions of 5-,6- and 7-membered rings, packing density, particle diameter, atomic fractal dimension, and population distribution of constituting molecules. The following conclusions can be drawn based on the present investigation:

1. The soot particles comprise both aromatic and aliphatic carbons. The smaller particles mostly include peri-condensed PAHs.
2. The soot particles go through the transition from chain-line porous structure to spherical, more compact spherical form as they grow in size.
3. Mature soot resembles a core-shell-like structure. The core comprises an intricate network of aromatic molecules, while the shell contains mostly aliphatic branches.
4. Smaller hydrocarbon species ($\#C \lesssim 20$ and $MW \lesssim 240$) play a vital role during soot nucleation and act as the building block of larger soot particles.

Further investigation is required to identify and isolate the relevant molecules ($\#C \lesssim 20$ and $MW \lesssim 240$) and connect the missing dots to create a complete soot nucleation mechanism. Similar studies with the inclusion of oxygenating species will give further insight into the true nucleation pathway.

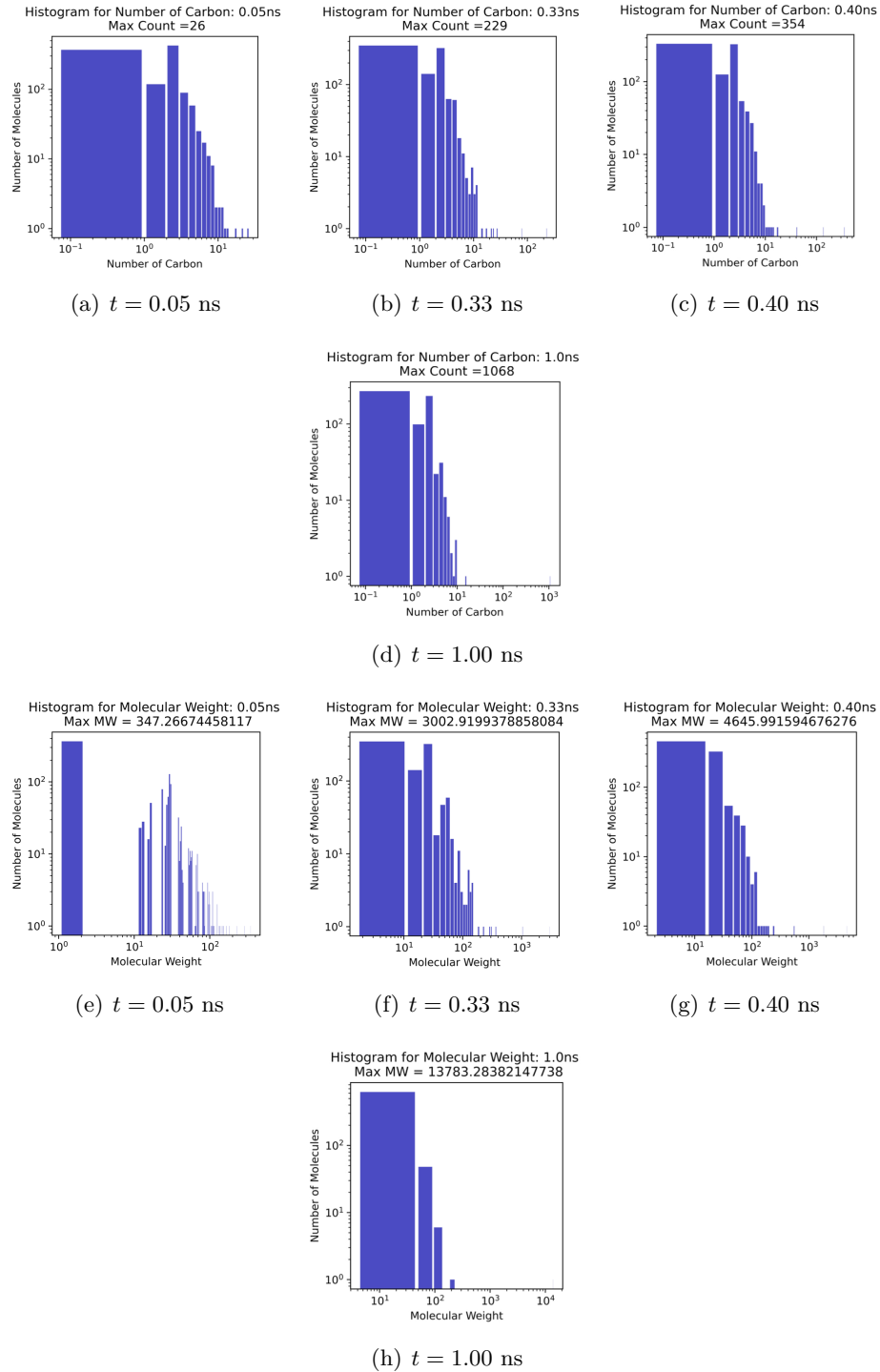


Figure 3.4: Population of hydrocarbon molecules present in the simulation domain at four different times based on (a-d) number of molecules and (e-h) molecular weight (kg/kmol).

CHAPTER 4

INTERNAL STRUCTURE OF INCIPIENT SOOT PARTICLES FROM
REACTIVE MOLECULAR DYNAMICS SIMULATIONS

This chapter is a reproduction of the article already published as: "Mukut, K. M., Ganguly, A., Goudeli, E., Kelesidis, G. A., & Roy, S. P. (2024). Internal Structure of Incipient Soot from Acetylene Pyrolysis Obtained via Molecular Dynamics Simulations. *J. Phys. Chem. A*, 128(26), 5175–5187. doi: 10.1021/acs.jpca.4c01548" [\[110\]](#)

4.1 Introduction

Soot is a harmful carbonaceous nanoparticle generated during combustion of hydrocarbon fuels. Soot, also known as black carbon, can cause serious health issues [\[7, 8\]](#) and acts as a major forcing factor in climate change [\[10, 6\]](#). The exact mechanism of the formation of soot particulates from gaseous precursors is still unknown due to the complex chemical nature of the hydrocarbon reaction network and time- and length-scale of the soot formation processes. According to the present understanding, soot formation occurs by a series of complex physicochemical events such as the formation of gas-phase soot precursors including, but not limited to, polycyclic aromatic hydrocarbons (PAHs), nucleation of incipient soot particles, growth and maturation of incipient soot particles due to surface reactions, aggregation by coagulation or coalescence, and decay of the particles by fragmentation and oxidation [\[111, 11, 21, 19, 112\]](#). The inception of soot particles is arguably the least understood phenomenon among these processes and the exact chemical reaction pathways of soot inception are not completely known yet. Researchers agree that soot formation starts with forming small gas-phase precursor molecules such as acetylene which leads to PAHs like benzene, pyrene, and coronene [\[113, 95, 96\]](#). The freshly formed PAHs then combine to form the solid or liquid-like incipient soot particles [\[3, 114, 115\]](#). These particles then start to grow by surface reactions and coalescence to form larger soot particles [\[3, 60, 116, 117, 118\]](#).

Due to the complexity and scale of incipient soot particles, their exact internal structures are not very well characterized yet. Recent studies have shown young soot particles tend to have a condensed core of ring-like structures surrounded by a shell of highly stacked large molecules while surrounded by a shell of less stacked smaller molecules [119]. As these incipient particles mature, their internal structures evolve, which in turn affects their physical and chemical properties.

There have been some recent breakthroughs in the experimental exploration of the internal structure of incipient soot. For example, Chang et al. [120] employed high-resolution transmission electron microscopy (HRTEM) and scanning electron microscopy (SEM) to investigate the structural evolution and fragmentation of coal-derived soot and carbon black particles under high-temperature air oxidation conditions. They also explored the onset of micropores and the internal graphitic microcrystals using X-ray diffraction (XRD) and Raman spectra. Morajkar et al. [121] utilized HRTEM, XRD, Raman spectroscopy, and inductively coupled plasma mass spectrometry (ICPMS) to examine the transmission of trace metals from biodiesel fuels to soot particles and the nanostructural irregularities of the soot. In their study, Gleason et al. [122] indicated that the formation of soot nuclei in an ethylene/nitrogen flame can be attributed exclusively to aromatic compounds comprising one or two rings. Carbone et al. [123] conducted a comprehensive investigation of soot inception in a laminar premixed ethylene flame and found that soot particles undergo an aging transformation from being nearly transparent in the visible spectrum to a more graphitic-like composition. Using low-fluence laser desorption ionization (LDI) in conjunction with HRTEM, Jacobson et al. [106], investigated the molecular composition of soot particles to determine the PAH concentration in soot particles. The capabilities of atomic force microscopy (AFM) were exploited by Barone et al. [124] to calculate particle size distribution functions under different sampling conditions. In their study, Schulz et al. [125] conducted an investigation into the initial phases of soot formation using AFM and observed the presence of multiple aromatic compounds, some of which displayed noticeable aliphatic side chains. Commodo et al. [126] investigated the initial phases of soot formation using X-ray, ultraviolet photoemission spectroscopy (UPS),

UV-visible, and Raman spectroscopy to show the coexistence of sp^3 carbon and a more advanced graphitic structure, which exhibits a slightly larger aromatic island, a reduced band gap, and an increased density of states. In another study, Commodo et al. [39] identified a noteworthy occurrence of aliphatic pentagonal rings in the early stages of soot formation, particularly in close proximity to the outer region of aromatic soot molecules, and it has been suggested that the elimination of hydrogen from these molecules can result in the creation of resonantly stabilized π -radicals [127]. This phenomenon has also been theorized by Johansson et al. [64], Gentile et al. [128], and Rundel et al. [42].

Even with such recent advancement in experimental findings, there is still a lot of unknowns about the internal structure of soot. The limitations of experimental methods can be compensated and complemented by first-principle modeling such as molecular dynamics. With the development of high-performance computational resources, reactive molecular dynamics (RMD) simulation has become more affordable for studying complex reactive networks. For soot-relevant RMD studies, the reactive force field (ReaxFF) potential developed by van Duin et al. [51] for carbon, hydrogen, and oxygen chemistry (CHO-parameters [129, 99]) is a popular choice. The ReaxFF potential can capture the physicochemical evolution of hydrocarbon systems in an extensive range of temperatures and pressures. It is based on the bond order between different atoms, which carry information related to bond breakage and formation. In recent years, RMD simulations have been used to investigate soot nucleation by pyrene dimerization [61], to shed light into the nucleation and growth of incipient soot from PAHs, such as naphthalene, pyrene, coronene, ovalene and circumcoronene [60], to explore the initial mechanism of soot nanoparticle formation [62] and to examine the effect of oxygenated additives on the reduction of diesel soot emissions [63].

Since RMD simulation provides detailed structural information at the atomic scale, it can be an excellent tool for analyzing the internal structure of incipient soot particles. For example, recently Pascazio et al. [130] looked into the internal structure and the mechanical properties of incipient soot particles using RMD simulation and quantified the amount of cross-linking in the core and shell region of developing and mature soot particles. Mature

soot primary particles exhibit a distinct core-shell structure with a disorderly condensed core of ring-like structures surrounded by a shell of chain-like structures [119, 131].

Process temperature plays an important role in the development and aging of soot particles. For example, in a recent study, Pathak et. al. [132] studied graphitization induced structural transformation of candle soot at different temperatures and found that increasing the temperature increases the rate of graphitization that leads to more spherical and mature soot and weakening of the correlation between graphitic nanostructure and surface functional groups (SFGs). SFGs have been found to be connected to the characteristics of soot aggregates, including the fractal dimension [133]. Since soot morphology, maturity and reactivity are expected to be influenced by temperature during acetylene pyrolysis [134], it is important to study the internal structure of incipient soot particles at different temperatures.

In the present study, a series of isothermal RMD simulations using the ReaxFF potential is conducted mimicking acetylene pyrolysis at different temperatures (1350, 1500, 1650, and 1800K). A variety of physicochemical features of these RMD-generated soot particles are then analyzed to shed light on different types of incipient soot particles and to characterize the internal structure of these particles obtained from RMD simulations.

4.2 Numerical Methodology

4.2.1 Simulation configurations

Following the methodology described in [57], 1000 acetylene molecules are randomly placed in a cubic domain ($75\text{\AA} \times 75\text{\AA} \times 75\text{\AA}$) at four different temperatures, i.e., 1350, 1500, 1650, and 1800 K. The temperatures are chosen to capture soot particles from various thermally activated systems. For statistical significance, at each temperature, simulations are performed for at least five times with different initial configurations. In total, 24 RMD simulations were performed for four different temperatures. The RMD simulations are performed using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [135] software. ReaxFF potential for hydrocarbons [51, 98] is used to capture

the chemical changes (bond breakage and formation) due to reactive molecular collisions during acetylene pyrolysis. The bond length between individual atoms is calculated at each timestep (0.25 fs) based on the changes in the chemical environment to describe bond cleavage and formation accurately [99]. This helps the model capture the chemical reactions leading to radical formation during soot nucleation. Periodic boundary conditions are assumed in all three dimensions. The coordinates of each atom are calculated and updated using the velocity-Verlet algorithm [100] in conjunction with the Nose-Hoover thermostat [102]. A constant number, volume, and temperature (NVT) ensemble strategy is used to run each simulation up to 10 ns. The simulation results are probed every 0.05 ns and the clusters of hydrocarbons that resembles primary soot particles are isolated, tabulated, and analyzed. Each of these extracted clusters has at least 20 carbon atoms and at least one 5-, 6-, or 7-membered ring structure following an earlier study by Mukut et al.[94]. Features such as surface area and volume of primary particles are calculated using MSMS software developed by Sanner [136] and other physicochemical characteristics are analyzed mostly using MAFIA-MD [22]. The open visualization tool (OVITO) [137] is used for visualization of the molecular clusters.

4.2.2 Workflow

The workflow in this study can be summarized as:

1. Conduct RMD simulations at different temperatures with various initial configurations. and extract incipient soot particles from the trajectory results.
2. Calculate chemical and morphological characteristics such as the number of atoms, C/H ratio, the radius of gyration, atomic fractal dimension, density, surface area, and volume.
3. Classify the soot particles based on *all* the calculated features using machine learning techniques such as k-means clustering [138] and t-distributed stochastic neighbor embedding (t-SNE) [139].
4. Investigate the internal distribution of several relevant features such as distribution of cyclic/non-cyclic molecules, C/H ratio, density etc. and find identifiable patterns in

the distribution.

4.2.3 Extraction of physicochemical properties

From the RMD simulations, we extract the coordinates of each atom present in the simulation box at regular time intervals via the trajectory file. Each individual timestep is investigated separately by analyzing the atom coordinates within the entire simulation domain, which contains both large molecular clusters and small molecules. The large soot-like molecular clusters are identified as the ones that have more than 20 carbon atoms and have at least one 5-, 6-, or 7-member ring [94]. It is noted that in our case, the smallest such cluster was found to have 65 carbon atoms. These clusters are isolated using the cluster analysis tool from the OVITO Python module [137] implemented in a unified Python script developed inhouse. Then the isolated clusters are analyzed individually to calculate their physical, morphological and chemical attributes. Some attributes are obtained trivially from the trajectory files, e.g., number of atoms (N), carbon to hydrogen ratio ($\Theta_{c/H}$), mass (M_p), and molar mass (M). Some other attributes like the radius of gyration (R_g), atomic fractal dimension (D_f), and density (ρ) are extracted by simple algebraic and geometric analysis or by using empirical correlations proposed in the literature and listed in A.2. The volume and the surface area of incipient particles are calculated using MSMS software [136] using a probe radius of 1.5 Å.

The identification and analysis of 5- /6- /7-member ring structures are done using MAFIA-MD [22]. MAFIA-MD can analyze RMD trajectory files to identify cyclic/ring structures in an atomic cluster. Not all cyclic structures identified are necessarily aromatic. As discussed in [22], it is difficult to exactly confirm which cyclic structures are aromatic as the information about aromaticity requires some approximations regarding the bond order of aromatic bonds and establishment of planarity. To remove any confusion, therefore, we used the terms “ring” or “cyclic” in this work instead of aromatic when discussing these internal structures in the soot clusters. The numbers of 5-, 6-, and 7-member rings are denoted as N_5 , N_6 , N_7 , respectively, and the total number of rings is denoted as N_{\odot} . Similarly, the number of carbons in rings is denoted as N_{\odot} and the number of non-cyclic

carbons in a particle is denoted as N_{C} .

A sample of two particles and a list of their respective properties are provided in A.3. This entire set of features is used in the classification of particles as discussed in Sec. 4.3.2. It must be noted here that while all the above-mentioned properties were evaluated for each particle, this article only focuses on the internal structure of the particles, which are characterized as discussed in Sec. 4.2.4. Therefore, beyond their use in the classification of particles, the detailed analysis of physical and morphological features such as volume, surface area, radius of gyration, and atomic fractal dimension are not the focus of this work.

4.2.4 Characterization of internal structure

We analyzed the internal structure of soot particles via the radial distribution of carbon atoms, C/H ratio, and density inside the particle. In order to compare different-sized particles on the same scale, we first normalized the radius of particles by scaling each particle by its radius of gyration (R_g). Then each particle is divided in equal number of radial bins. Each radial bin creates a spherical shell or strip as shown by the shaded yellow region of interest in Fig. 4.1. We calculate various internal features in these spherical strips and present them as a function of the normalized radial distance from the center of mass of each spherical strip ($r^* = r/R_g$).

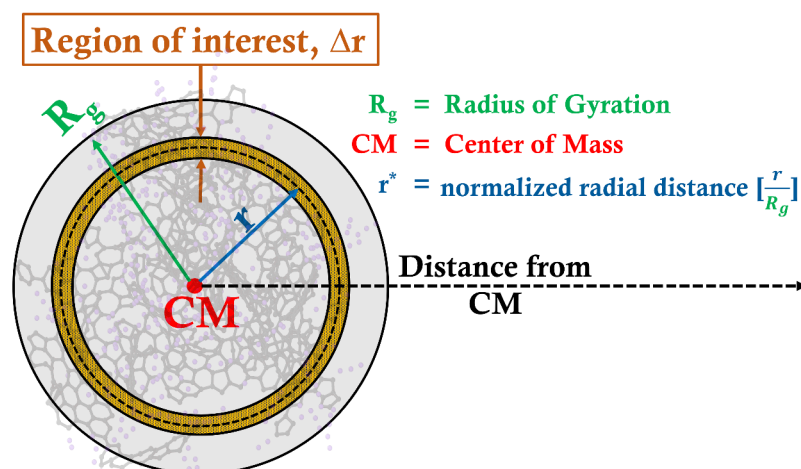


Figure 4.1: Schematic representation of the calculation of radial distribution of internal features in the soot particle.

For example, let's consider a strip of width Δr , whose midplane is distance r away from the center of mass. The number of cyclic and non-cyclic carbon atoms in this strip (i.e., within a radial distance of $r \pm \Delta r/2$) are counted and represented as a function of the normalized radius of the midplane ($r^* = r/R_g$), as $n_{\odot}(r^*)$ and $n_{\mathfrak{C}}(r^*)$, respectively. The radial distribution of cyclic and non-cyclic carbon per unit area (indicated by $''$) at a normalized distance r^* are then, respectively:

$$N_{\odot}''(r^*) = \frac{n_{\odot}(r^*)}{4\pi(r^* \times R_g)^2}; \quad N_{\mathfrak{C}}''(r^*) = \frac{n_{\mathfrak{C}}(r^*)}{4\pi(r^* \times R_g)^2} \quad (4.1)$$

In a similar manner, the radial distribution of C/H ratio is also analyzed for each particle. The C/H ratio of the entire particle is calculated as $\Theta_{C/H} = N_C/N_H$, where N_C and N_H are number of carbon and hydrogen atoms in the entire particle, respectively. This C/H ratio is termed as the *particle* C/H ratio ($\Theta_{C/H}$) to differentiate from the *local* C/H ratio ($\theta_{C/H}$), which is calculated using the number of carbon and hydrogen atoms in the spherical strips as shown in Fig. 4.1. The local C/H ratio ($\theta_{C/H}$) is determined by calculating the number of carbon ($n_C(r^*)$) and hydrogen ($n_H(r^*)$) atoms in a spherical strip with the midplane at a normalized distance r^* from the center of mass ($\theta_{C/H}(r^*) = n_C(r^*)/n_H(r^*)$). Finally, the local C/H ratio is normalized by the corresponding particle C/H ratio

$$\theta_{C/H}^*(r^*) = \frac{\theta_{C/H}(r^*)}{\Theta_{C/H}} \quad (4.2)$$

Similarly, the radial distribution of local density is also evaluated by dividing the simulated density of the thin spherical strip using Eqn. A.3 for the strip (referred as *local* density, $\varrho(r^*)$) by the simulated density of the particle (ρ_s) as

$$\rho_s^*(r^*) = \frac{\varrho(r^*)}{\rho_s} \quad (4.3)$$

4.3 Results and Discussion

4.3.1 Formation of incipient soot particles in RMD

During the RMD simulations, the system of atoms goes through different chemical and physical interactions resulting in the formation of larger atomic clusters due to the

pyrolysis of acetylene. The evolution of one of these atomic clusters is depicted in Fig. 4.2. Carbon and hydrogen atoms are represented using black and red dots, respectively. First, the acetylene molecules combine to form small linear chains (Fig. 4.2B: linearization) and then transform into cyclic structures (Fig. 4.2C: cyclization). After cyclization, the small clusters start growing due to both bond formation at the surface and internal reorganization. These larger atomic clusters resemble incipient soot particles (Fig. 4.2D–F). It is important to note that, the collisions are stochastic in nature and the time required for an event, i.e. linearization, cyclization, surface growth. etc., varies based on the initial configurations, and therefore are omitted from the figure for generality. A similar formation mechanism is also reported by Zhang et al. [104] for carbon-black simulations and Sharma et al [57] for acetylene pyrolysis simulations.

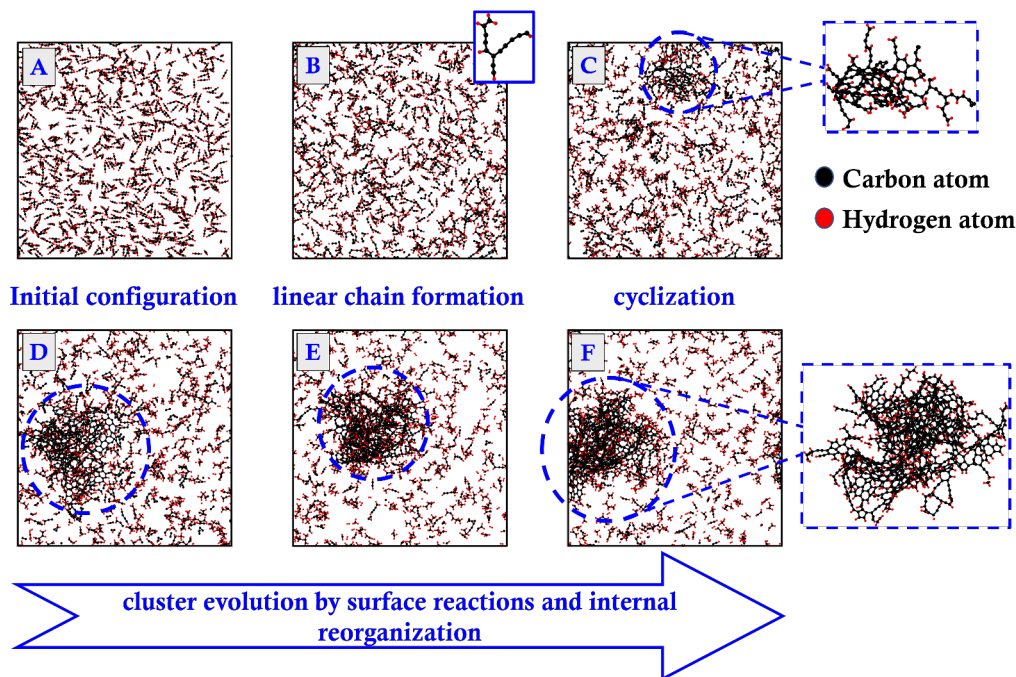


Figure 4.2: A general representation of steps during the formation and evolution of incipient soot cluster during acetylene pyrolysis (from a simulation performed at 1650 K).

The incipient soot clusters are extracted from the RMD simulations at different timesteps to capture the growth. Each simulation is run at least five times with a velocity

field initialized randomly in each case to generate soot particles with different evolutionary histories. In total, 3324 individual soot clusters are isolated from the RMD simulations (number of carbon atoms ranging from 65 to 1503). The C/H ratios ($\Theta_{C/H}$) of these particles are compared to the theoretical limits for PAHs in Fig. 4.3. Based on the compactness, the PAHs can be classified into two categories: (a) peri-condensed PAHs, where the carbon atoms in the aromatic structures can be shared by more than two aromatic rings and (b) cata-condensed PAHs, where the carbon atoms in the aromatic structure can be shared by at most two aromatic rings. Siegmann and Sattler [16] proposed a relationship between the number of carbon and hydrogen atoms for both peri-condensed and cata-condensed PAHs. Fig. 4.3 presents the C/H ratio and molar mass of the soot clusters from different temperatures and compares it with the peri-condensed and cata-condensed PAH zones derived from [16]. As observed from Fig. 4.3, the soot clusters fall between the peri-condensed and cata-condensed boundaries, indicating an intricate network of different types of aromatic and aliphatic structures in incipient particles. Using atomic force microscopy Commodo et. al. [39] showed that smaller aromatic clusters (number of carbon atoms ranging from 6 to 55, lower than what studied in this work) in the early stage of soot formation in a slightly sooting premixed ethylene flame tend to be close the peri-condensed line. However, in this work, we find that large clusters lie closer to cata-condensed limit than peri-condensed limit, potentially indicating significant presence of non-aromatic (i.e., aliphatic and alicyclic) structures.

4.3.2 Classification of incipient soot particles

We tagged, extracted, or calculated the physicochemical features (such as Temperature, number of carbon atoms, number of hydrogen atoms, and molar mass) for each particle. The complete set of features used in this study is listed in A.3.1, and two sample particles with the entire feature set are shown in A.3.2.

Our initial observation of trends of various internal and physicochemical features revealed a wide variation, indicating that these particles can be classified into multiple groups based on their features. The classification of the incipient soot particles is depicted

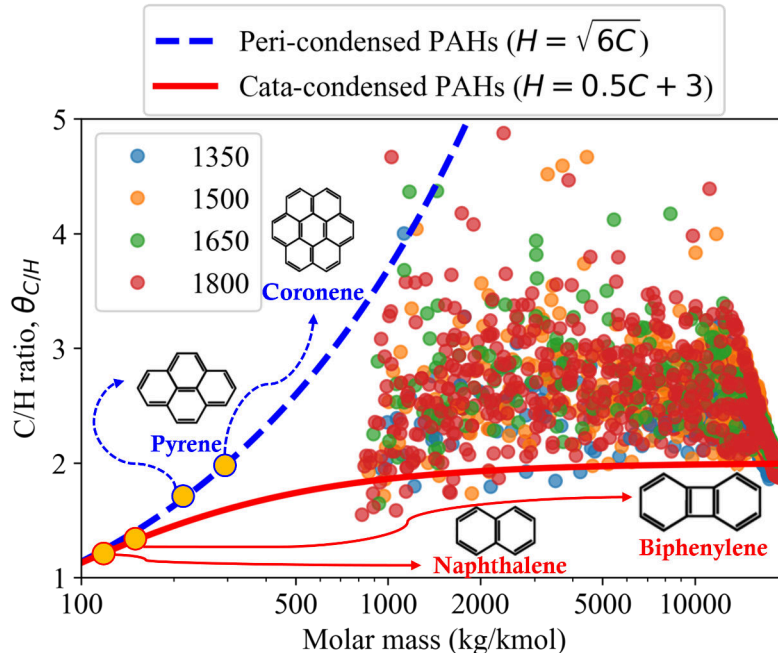
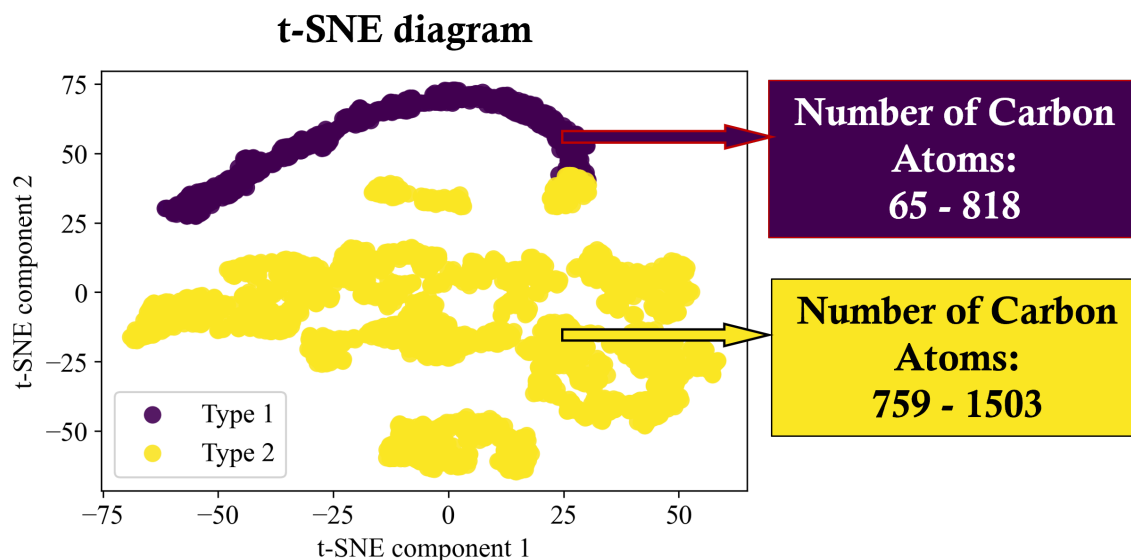


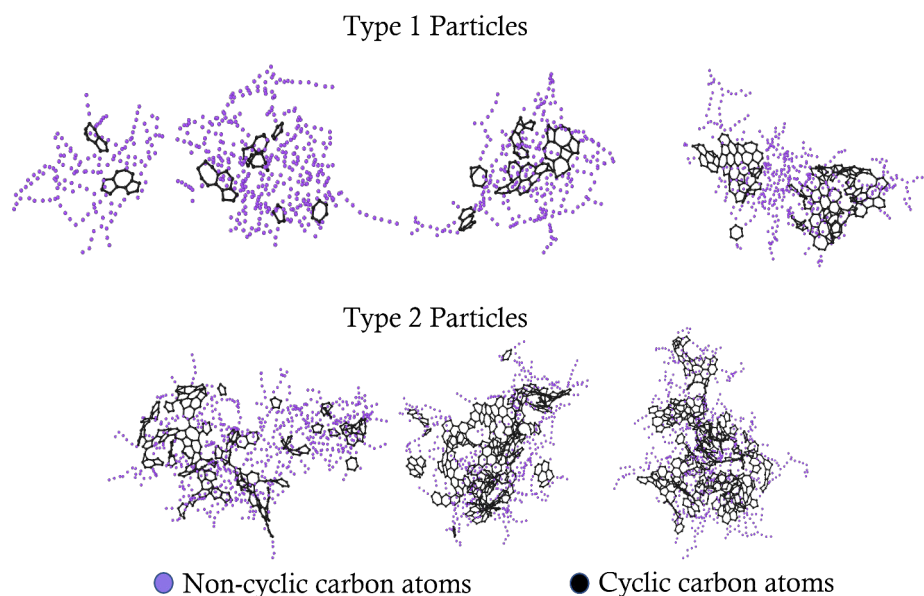
Figure 4.3: C/H ratio vs. molar mass of soot clusters at different temperatures.

in Fig. 4.4. We attempted two unsupervised machine-learning techniques to extract unique classifications that may exist in the incipient particle sample space. The first method is the k-means clustering algorithm [138], which is used to label particles of different classes based on all extracted features of the particles. Then we used the t-dispersion stochastic neighbor embedding (t-SNE)[139] plot to display the particle properties on a 2D map. In the t-SNE diagram, similar clusters (i.e., potentially belonging to the same class) are expected to be close to each other. Although the number of classes was not known a priori, trial and error with k-means clustering revealed good results with two classes. For identification purposes, these two classes are referred to as “type 1” and “type 2” particles, respectively. The resulting t-SNE diagram is shown in Figure 4.4(a).

Looking closely at the two classes, we see that the particles exhibiting similarity fall into a nearly continuous size range. For example, in the first class, (**type 1**) the incipient particles have a lower total number of carbon atoms (65 – 818) whereas in the second class (**type 2**) the particle have a higher number of total atoms (759 – 1503). This essentially points to the fact that the characteristics of the incipient particles change after a certain



(a) A t-SNE diagram generated using all the 3324 incipient soot clusters with two different k-means clusters.



(b) Some example particles from each class obtained from RMD simulations. The non-cyclic carbon atom structures are shown in purple dots and the cyclic structures are shown in black. Hydrogen atoms are omitted from the visualization for clarity.

Figure 4.4: Classification of incipient soot particles

level of growth: smaller particles (type 1) show different features and trends than larger particles (type 2). It should be noted here that the number of total carbon atoms is not a unique marker of the threshold between type 1 and type 2 – as indicated by a small overlap in the number of carbon atoms range between the types – but acts as a very good surrogate for threshold identifier. In total, we have obtained 670 type 1 and 2654 type 2 incipient particles from a total of 3324 particles. Fig. 4.4(b) depicts some example particles from the analyzed sample space. Here, the non-cyclic carbon atom structures are shown in purple dots and the cyclic structures are shown in black, while hydrogen atoms are omitted from the visualization for clarity.

4.3.3 Comparison with experimental data

The particles obtained in the current study show very good match with experimentally observed properties of incipient soot particles as reported in the literature. For example, the mean density of the particles in our study is calculated to be $1.53 \pm 0.08 \text{ g/cm}^3$, which is an excellent match with the empirical soot density of by Johansson et al [67], who reported the value to be 1.51 g/cm^3 . More detailed comparison of average particle properties are presented in [140].

4.3.4 Internal structure of incipient soot particles

The radial distributions of various quantities relevant to the analysis of the internal structure of the incipient particles are presented as box and whiskers plots in subsequent figures. In this visualization, every box includes the data points within the second and third quartiles, the horizontal line inside the box indicates the median value and whiskers represent the range of the data. The statistics were found to be insensitive to the process temperature. Hence aggregate data for all temperatures is presented here.

The radial distribution of cyclic and non-cyclic carbon atoms per unit area for type 1 (Figs. 4.5(a)-4.5(b)) and type 2 (Figs. 4.5(c)-4.5(d)) incipient particles as a function of the normalized radial distance from the center of mass is shown in Fig. 4.5. The blue vertical line depicts the location where the radial distance becomes equal to the radius of gyration

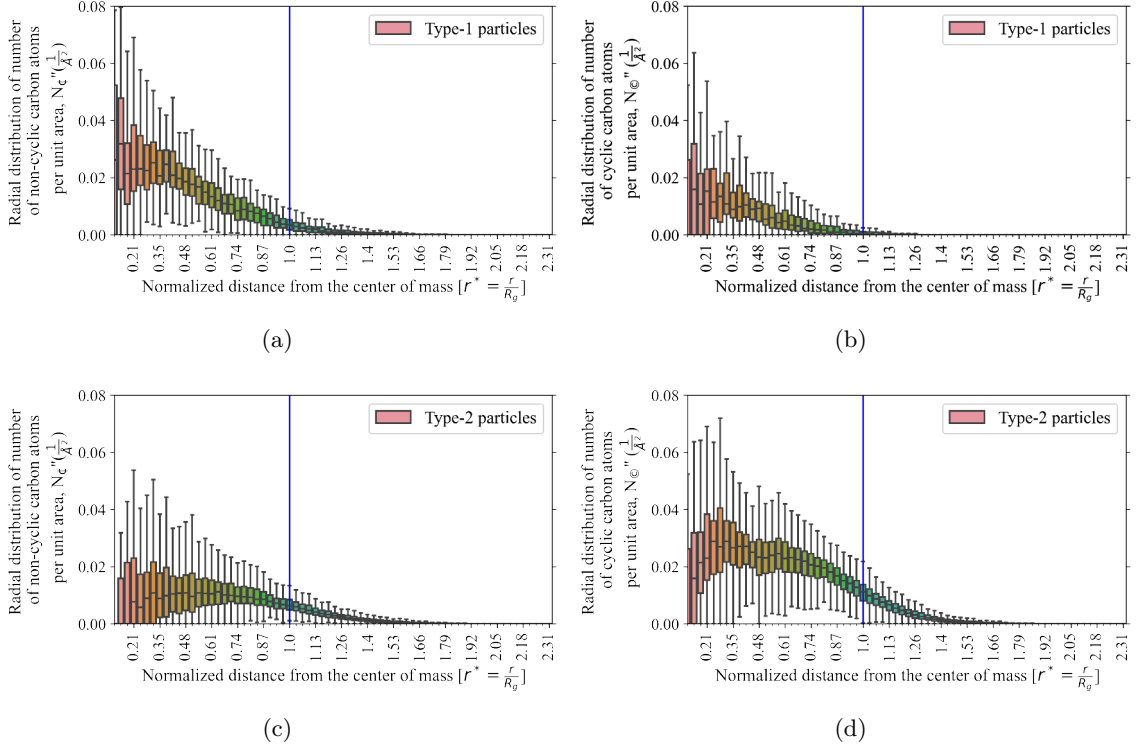


Figure 4.5: Radial distribution of (a,c) non-cyclic and (b,d) cyclic carbon atoms ($N_c''(r^*)$ and $N_c''(r^*)$) in (a,b) type 1 and (c,d) type 2 soot particles as a function of the normalized radial distance ($r^* = \frac{r}{R_g}$) from the center of mass. The blue vertical line is at a radial distance equivalent to R_g .

(R_g) of individual particles. For type 1 particles, an abundance of non-cyclic carbon atoms is observed in the central (i.e., less than 50% of R_g) region. The number of non-cyclic carbon atoms is almost twice the number of carbon atoms in cyclic structures in type 1 particles. Almost all the carbon atoms reside near the central region of type 1 particles and the number of carbon atoms quickly drops to zero as we go away from the center of mass. This indicates that in type 1 particles, i.e., at the very early stages of soot formation, the number of non-cyclic structures is significantly higher than the number of cyclic structures.

The type 2 particles, on the other hand, show a different trend where a very concentrated region of cyclic carbon atoms is observed in the central region of the particles. The concentration of cyclic carbon atoms slowly decreases as the distance from the center of mass increases. The number of non-cyclic carbon atoms increases from a very low value in

the central region, then reaches a steady value near the radius of gyration and then gradually drops to zero as the distance increases beyond the radius of gyration. This indicates that the non-cyclic carbon atoms are more likely to be present in the outer region (what can be presumably considered near the particle surface) of the type 2 incipient particles while the central region is dominated by the cyclic carbon atoms.

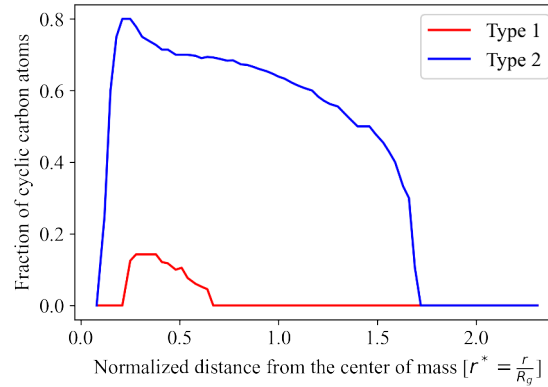


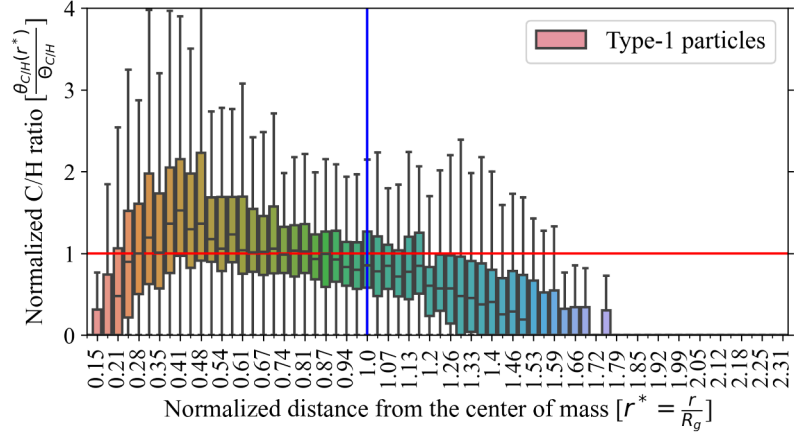
Figure 4.6: Radial distribution of the median fraction of cyclic carbon.

If the radial distribution of the median value of the fraction of cyclic carbon atoms with respect to total carbon atoms is analyzed (shown in Fig. 4.6), the dominance of cyclic carbon atoms in type 2 particles becomes very clear. Figure 4.6 shows a prominent central region where more than half of the carbon atoms belong to a ring structure. These evidences suggest that the internal chemical structure of incipient soot particles changes as the particles transition from type 1 to type 2. This indicates the development of core-shell structures as the incipient soot grows and matures. The presence of such core-shell structure has been theorized in the literature. For example, Michelsen et. al. [141] used a fractal core-shell model to explain the changes in the structure of soot aggregates and primary particles at different heights of a laminar co-flow ethylene-air flame. Kholghy et. al. [109] proposed a surface shell formation model to predict maturity of soot primary particles. More directly, recently Botero et. al. [119] studied the internal structure of soot particles using high-resolution transmission electron microscopy (HRTEM) to identify the PAH

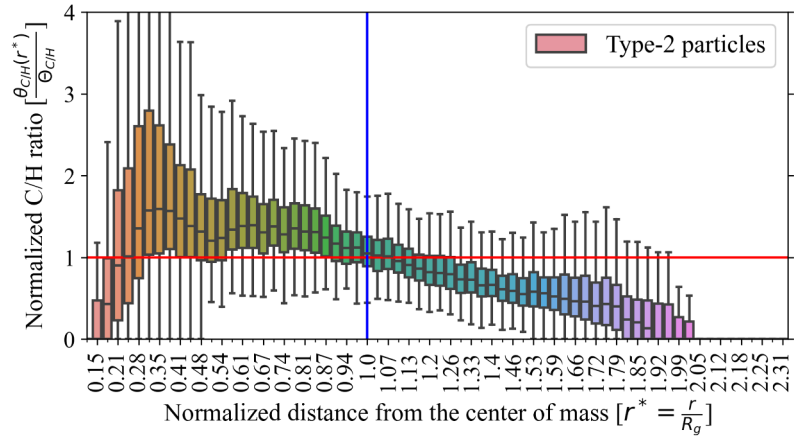
structures in the core and shell regions, and suggested the presence of a stabilized core region indicating nano-structural mobility. Pascazio et. al. [130] utilized RMD simulation and identified different levels of crosslinking in core and shell in hypothetical soot particles. Kelesidis et al. [142] also investigated oxidation dynamics of carbonaceous nanoparticles having various core-shell structures using lattice Monte Carlo simulations. The clear difference in the radial distribution of cyclic and non-cyclic carbon atoms between type 1 and type 2 particles supports these findings.

The changes in chemical properties inside the incipient soot particles can also be observed in the radial distribution of carbon to hydrogen ratio. The normalized local C/H ratio for type 1 (Fig. 4.7(a)) and type 2 (Fig. 4.7(b)) particles are plotted as a function of the normalized radial distance from the center of mass in Fig. 4.7 using box and whisker plots. The blue vertical line depicts the location where the radial distance becomes equal to the radius of gyration (R_g) of individual particles. The red horizontal line indicates where the local C/H ratio is equal to the particle C/H ratio.

The normalized local C/H ratio increases up to a certain distance from the center of mass and then starts to drop. For type 1 particles, the increase in the local C/H ratio takes longer distance from the center of mass, and the median value gradually reaches a peak value slightly higher than the particle C/H ratio (about 1.3 times). After that, the local C/H ratio starts to drop and the median value reaches a value close to the particle C/H ratio at around 65% of the radius of gyration. The value stays close to the particle C/H ratio up to the radius of gyration, and then slowly drops to zero. This indicates a very small or no dense core region in type 1 particles. The demarcation between the core and shell regions is not clear in type 1 particles because of the absence of a pronounced core as the local C/H ratio remains close to the particle C/H ratio. For type 2 particles, however, the local C/H ratio increases rapidly in the central region and the median value reaches a peak value of about 1.7 times the particle C/H ratio. After that, the local C/H ratio starts to drop and reaches a value equal to the global C/H ratio at the radius of gyration. Unlike type 1 particles, the region where the median of the normalized local C/H ratio is close to unity is very narrow in type 2 particles. This indicates that the dense core region of type 2



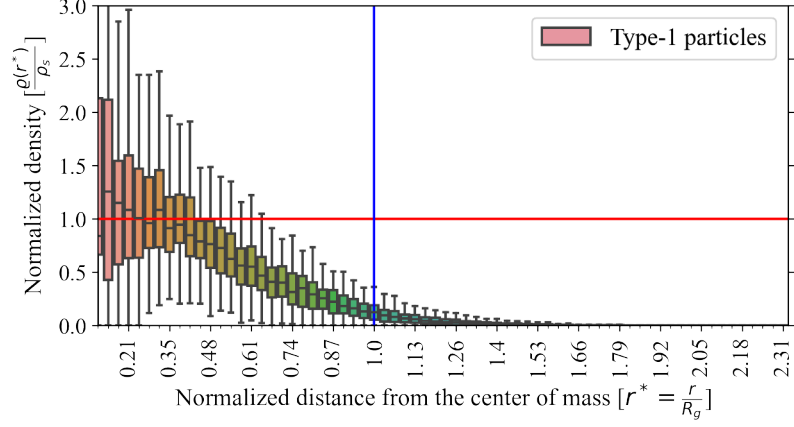
(a)



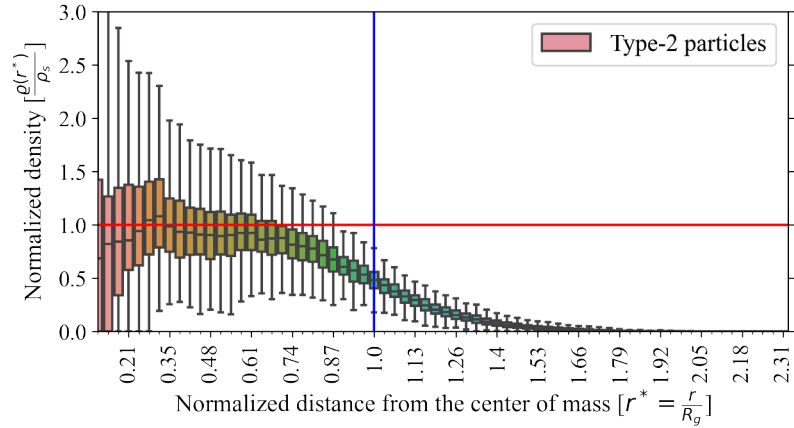
(b)

Figure 4.7: Radial distribution of normalized C/H ratio ($\theta_{c/H}^*(r^*) = \frac{\theta_{c/H}(r^*)}{\Theta_{c/H}}$) in incipient particles as a function of normalized radial distance ($r^* = \frac{r}{R_g}$) from the center of mass. The blue vertical line is at a radial distance equivalent to R_g . The $\frac{C}{H}$ ratio of the strip and the particle are the same along the red horizontal line.

particles is well-developed and the separation between the core and shell regions is more pronounced than in type 1 particles.



(a)



(b)

Figure 4.8: Radial distribution of normalized local density ($\rho^*(r^*) = \frac{\varrho(r^*)}{\rho_s}$) inside incipient particles as a function of normalized radial distance ($r^* = \frac{r}{R_g}$) from the center of mass. The blue vertical line is at a radial distance equivalent to R_g . The density of the strip and the particle are the same along the red horizontal line.

The radial distribution of the normalized density inside type 1 (Fig. 4.8(a)) and type 2 (Fig. 4.8(b)) incipient soot particles is presented in Fig. 4.8 as a function of the normalized radial distance from the center of mass of soot particles (r^*) as box and whisker plots. For type 1 particles, Fig. 4.8 shows a very small dense central region that extends up

to about 40% of the radius of gyration. The local density of type 1 particles is maximum near the center of mass, and it drops gradually as the distance from the center of mass increases. For type 2 particles, on the other hand, the dense core region is larger and extends up to about 50-60% of the radius of gyration. The local density of type 2 particles stays close to the particle density up to around 60% of the radius of gyration, and then it quickly drops.

4.3.5 The boundary between the core and shell

The results discussed so far indicate the presence of a core and shell structure in incipient soot particles. This is further examined by looking at the median values of normalized local density (ρ_s^*), normalized C/H ratio ($\theta_{c/H}^*$), and radial distribution of carbon atoms (N_{\odot}'' and N_{\oplus}'') in Fig. 4.9 for both type 1 (in red) and type 2 (in blue) incipient soot particles. No significant trend is noticed from the radial distribution of non-cyclic carbon per unit area (Fig. 4.9(d)). However, the normalized density (Fig. 4.9(a)), C/H ratio (Fig. 4.9(b)) and the radial distribution of cyclic carbon atoms per unit area (Fig. 4.9(c)) demonstrate unique common trends in type 2 particles, which is not observed in type 1 particles. All of the quantities

1. reach a maximum value near the center of mass and decrease gradually to a local minimum at a distance of about 50% of the radius of gyration,
2. show the presence of a plateau region between 50% and 60% of the radius of gyration,
3. and drop monotonously after approximately 60% of the radius of gyration.

The first region or the central region can be identified as the *core* of the soot as it is near the center of mass, it is denser and contains more rings than other regions (as seen in Fig. 4.5(b) and 4.5(d)). The local minima mark the beginning of the boundary between the *core* and the *shell*. The narrow plateau region can be thought of as the boundary region between *core* and *shell* regions. And finally, the gradual descent of these quantities indicates the *shell* region. In the type 1 particles, we can only see the gradual descent stage, indicating that the core-shell demarcation is not yet developed, i.e., there is no developed core. The reason for such a trend in density, C/H ratio and ring structures can be explained

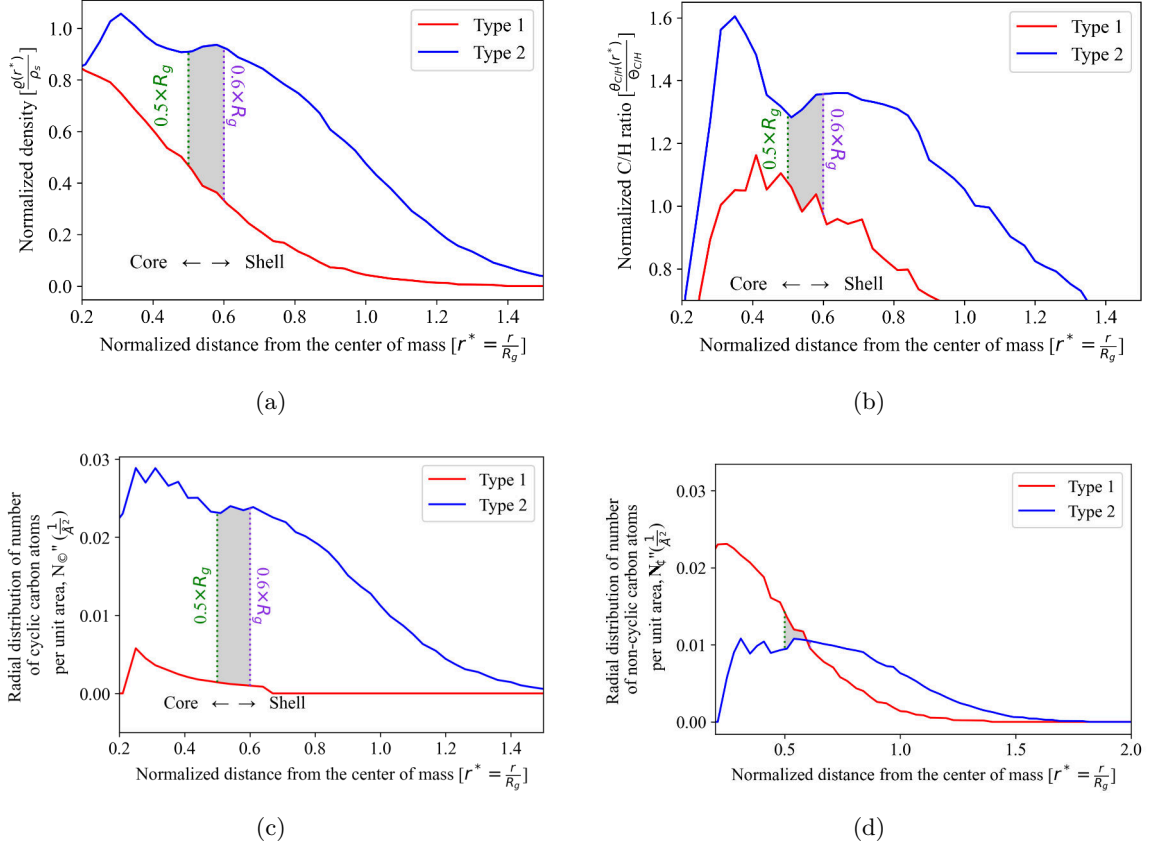


Figure 4.9: Identification of core and shell based on the radial distribution of normalized medians (a) density, (b) C/H ratio and radial distribution of the number of (c) cyclic and (d) non-cyclic carbon atoms per unit area in incipient particles as a function of normalized radial distance ($r^* = \frac{r}{R_g}$) from the center of mass.

by the nature of the stacking of cyclic molecules (disordered and ordered), as shown in Fig. 4.10. As observed in the schematic in Fig. 4.10, the core region comprises an interconnected cross-linked network of cyclic molecules while the shell region contains sheet-like organization of cyclic molecules. Such structural differences in core and shell are also supported by results presented by Pascazio et al. [130]. The core size of $0.5 - 0.60R_g$ obtained here is also consistent with the core size of $0.5 - 0.75R_g$ suggested by lattice Monte Carlo simulations [142].

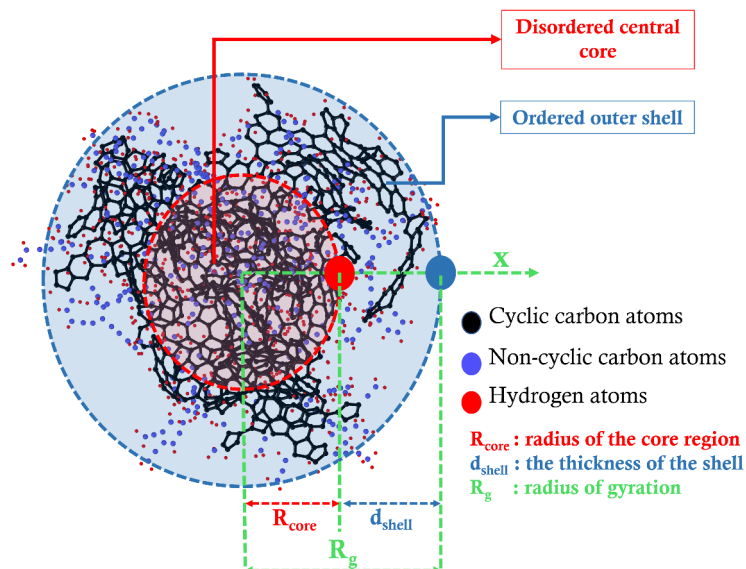


Figure 4.10: Schematic of core and shell structure of incipient particles.

4.4 Conclusion

A series of reactive molecular dynamics (RMD) simulations were performed to study the evolution of incipient soot particles during acetylene pyrolysis at four different temperatures. A total of 3324 incipient soot particles were obtained at different stages of evolution from these simulations. The mass, volume, surface area, radius of gyration, density, C/H ratio, and the number of cyclic structures were calculated for each particle. The internal structures of RMD-derived soot and their classification were investigated. Using unsupervised machine-learning techniques, the incipient soot particles are classified into two types – type 1 and type 2 – based on their morphological and chemical features. This classification was found to be very well predicted by the size of the particles with smaller particles in the type 1 class and larger particles in type 2. The internal structures of type 1 and type 2 particles show clearly distinct trends and features, indicating the two types correspond to early and late stages of incipient soot. The internal structure of the incipient particles from this investigation shows no direct or obvious sensitivity to temperature for temperature ranging from 1350 to 1800 K.

Other conclusions drawn from the study are

1. Incipient soot particles comprise both cyclic and non-cyclic structures.
2. The core of type 1 particles is dominated by non-cyclic structures while the core of type 2 particles is dominated by cyclic or ring structures.
3. The internal distribution of ring and non-ring structures indicates the presence of a dense core region and a less dense shell region in type 2 incipient particles.
4. The core comprises an interconnected cross-linked network of cyclic molecules while the shell region contains a sheet-like organization of cyclic molecules.
5. The core of type 2 particles extend up to approximately $0.5 - 0.6R_g$, followed by the shell region beyond $0.6R_g$.
6. The core and core-shell demarcation is not developed in type 1 particles.

Acknowledgments

The research benefited from computational resources provided through the NCMAS, supported by the Australian Government, The University of Melbourne's Research Computing Services and the Petascale Campus Initiative. K.M.M. and S.P.R. acknowledge funding support from the National Science Foundation as some of this material is based upon work supported by the National Science Foundation under Grant No. 2144290.

CHAPTER 5

**PHYSICOCHEMICAL EVOLUTION OF INCIPIENT SOOT PARTICLES IN
ACETYLENE PYROLYSIS: A REACTIVE MOLECULAR DYNAMICS
STUDY**

This chapter is a reproduction of the article already published as: "Mukut, K. M., Ganguly, A., Goudeli, E., Kelesidis, G. A., & Roy, S. P. (2024). Physical, chemical and morphological evolution of incipient soot obtained from molecular dynamics simulation of acetylene pyrolysis. *Fuel*, 373, 132197. doi: 10.1016/j.fuel.2024.132197" [140]

5.1 Introduction

Soot, also known as black carbon, is a particulate matter that is an unwanted byproduct of incomplete combustion of hydrocarbon fuels [3]. It greatly influences the radiative energy balance of the atmosphere and is a major forcing factor behind climate change [10, 6]. It also impacts public health and welfare and is one of the leading causes of mortality worldwide [143]. Exposure to soot or black carbon can lead to serious health issues such as cancer [7] and cardiovascular diseases [8]. Due to the ubiquitous presence of combustion events – both natural and anthropogenic – it is essential to understand the formation and evolution of soot particles in order to regulate and control their detrimental effects [144].

The precise mechanism underlying the formation of soot particulates from gaseous species remains uncertain, primarily due to the intricate chemical composition of the hydrocarbon system, the multiphysics interplay, and the multiscale nature of the soot formation process. The soot formation and growth involve a series of complex physical and chemical processes [111, 11]. These include the generation of gaseous precursor molecules such as polycyclic aromatic hydrocarbons (PAHs)[113, 95, 96], the inception of the incipient soot particles through the physical and chemical interactions among these precursor molecules [3, 21], the aging of these particles via surface growth due to dehydrogenation and carbon-addition from the gas phase and due to aggregation via coagulation and

coalescence [3, 67, 19], and the fragmentation and oxidation of soot particles [3, 21, 19, 145].

Due to the complex physicochemical interactions, soot undergoes significant alterations in its internal structure and physical properties during its evolution. These modifications occur as the particles progress from the initial stage of incipient soot to the intermediate stage of young soot, ultimately culminating in the final stage of mature soot [3]. Initially, clusters of gaseous PAHs come together to form the first incipient soot particles. During the later stages of formation, the incipient soot undergoes surface chemical reactions and surface condensation of PAHs, resulting in the transformation of incipient soot into young soot. The properties of young soot can be described as having a liquid-like structure and a relatively low level of graphitization [3, 146, 147]. Mature soot exhibits a highly organized graphitic structure that possesses a minimal level of curvature [146, 147].

The C/H ratio of soot particles increases as the particles get matured. For example, the C/H ratio of freshly formed soot particles is less than 2.0 [125] while the young soot particles can attain a C/H ratio ranging from 2.0 to 4.5 [148]. Mature soot particles usually have a C/H ratio of 4.5 or higher [148]. The average C/H ratio of soot particles from a premixed ethylene flame was also reported by Schulz et al. [125] to be around 2.33. The density of the soot particles also changes as the particles go through maturation. The empirical density of young and mature soot particle is measured to be 1.5 g/cm³ [149] and 1.7-1.9 g/cm³ [150], respectively. The values reported or assumed in multiple other studies [115, 151, 152, 67] are comparable to this.

The internal structure of soot consists of both aromatic and aliphatic carbon atoms. The presence of 5- and 6-member ring structures have been identified as an important feature of the soot core [67, 153, 65, 154], which influences mechanical properties of soot [130]. The physicochemical properties of mature soot particle that come out of a combustion system is strongly dependent on the internal structure of particles, which in turn depend on the physicochemical evolution process that the particle has gone through since its formation. [3].

In recent years, there have been remarkable experimental advances into understanding the internal structure of soot. Almost all of these recent experimental

findings show the importance of ring structures in various stages of soot. For example, Schulz et al. [125] observed multiple aromatic compounds with aliphatic side chains; Gleason et al. [122] showed that soot nuclei can form from aromatic compounds with only one or two rings; Cheng et al. [120] observed structural changes such as onset of micropores and graphitic microcrystals of carbon black and soot particles during their evolution; Carbone et al. [123] found that the optical properties of soot change as soot graphitizes during its evolution; Jacobson et al. [106] explored the molecular structure of soot and concentration of PAH in soot; Commodo et al. provided insights of the importance of sp^3 carbon and advanced graphitic structure [126] and aliphatic pentagonal rings [39] in the early stages of soot formation. When analyzing the radiation scattering by soot particles using refractive index, discrete element modeling with discrete dipole approximation simulations showed that the mass absorption coefficient of soot increases by up to 75% with increasing residence duration in premixed fires of low equivalency ratio [155]. Increased soot maturity is the cause of this, which is in great agreement with evidence obtained from laser induced incandescence in ethylene flames [156] and premixed methane [157].

However, despite these recent findings, our knowledge about the evolution of internal and physicochemical properties of soot is still incomplete. Since the exact physical and chemical pathways behind the formation and evolution of soot particles from incipient to mature are still not fully known, there is considerable uncertainty in the estimation of the physical, chemical, and morphological properties of soot at various stages. This poses difficulty down the line in engineering-scale modeling, where the goal is often to model soot emissions from combustion systems at a scale relevant to real-world devices.

Engineering-scale reactive CFD models are computational models that operate at the continuum scale and are designed to capture the continuum-scale dynamics of reactive flow. These models are capable of replicating the combustion behavior of real-world systems, including laboratory-scale flames, internal combustion engines, and fires, and can provide estimations of both local and global characteristics of the reactive flow field. Because of the inherent complexity involved in combustion modeling and the vast difference in scales, it is impractical for engineering-scale models to keep track of the evolution of the internal

atomic structures of the particle. Therefore, engineering-scale models of soot formation often depend on approximations. Notable numerical soot models such as the semi-empirical two-equation model [28], method of moments [12, 158, 159, 160, 161], stochastic soot model [108], discrete sectional model [162, 36, 163, 164, 165], etc. all contain several approximations and/or hypotheses to simplify the inception and growth of soot. These approximations, which include, but are not limited to, inception due to dimerization of soot precursor [12, 158], the spherical shape of particles [12, 28], constant density [108], etc. exist partly due to our incomplete knowledge of underlying soot-related processes and partly due to the complexity of combustion modeling.

In this work, we attempt to mitigate some of these shortcomings of engineering-scale soot models by providing an improved estimation – by way of reactive molecular dynamics simulation – of various soot properties as soot evolves in combustion systems. We extend our previous work on acetylene pyrolysis using reactive molecular dynamics (RMD) at four different process temperatures [110] to investigate how physical, chemical, and morphological properties of incipient soot evolve and how are they correlated with one another.

Reactive molecular dynamics allows for modeling the chemical evolution in a reacting system by tracking bond breakage and bond formation among atoms. This provides an unprecedented view of physicochemical transformations that take place during soot inception [61, 60, 62, 63]. One of the most common tools for modeling reactive hydrocarbon systems in RMD is the reactive force field (ReaxFF) [51] for carbon, hydrogen, and oxygen chemistry [129, 99]. RMD has been used to explore pyrene dimerization [61], which is often treated as a key step in soot nucleation. RMD has also been utilized to look at other mechanism of soot nanoparticle formation [62], the inception of soot from different PAHs such as naphthalene, pyrene, coronene, ovalene and circumcoronene [60]. Recently, the effect of oxygenated additives on diesel soot was also explored via RMD [63]. RMD also allows for a detailed atomic exploration of the internal structure of soot [130, 110]. For example, the amount of cross-linking in the core and shell structure of developing and mature soot particles were explored in [130]. In our previous study, we also identified the

presence of a denser core with the existence of ring structures at the center of larger incipient soot particles [110].

This study presents findings from a set of isothermal RMD simulations of acetylene pyrolysis at 1350, 1500, 1650, and 1800 K using the ReaxFF potential. We study four different process temperature because some researchers have indicated the influence of temperature on soot microstructure, surface reactivity, and degree of graphitization[132, 133]. The RMD simulations at the four temperatures yield numerous incipient soot particles at different stages of growth. The physicochemical characteristics of the soot particles are subsequently examined to explore the statistical measures that can be used to describe the development of the initial soot particle. The main objective of this article is to provide insights into the physicochemical and morphological characteristics of early-stage soot particles acquired through RMD simulations and to explore possible correlations among various properties which can potentially be used to improve engineering-scale soot models.

5.2 Numerical Methodology

5.2.1 Simulation configurations

The RMD approach taken in this study has been reported in [57], and the exact simulation configurations for this study have been previously described in [110]. Therefore, we only provide a brief summary here. 1000 acetylene molecules are randomly distributed within a $75\text{\AA} \times 75\text{\AA} \times 75\text{\AA}$ cubic domain at four different temperatures: 1350, 1500, 1650, and 1800 K. This temperature range is consistent with those used in laminar flames [166] or flow reactors [167]. Each configuration was simulated multiple times with different randomization. The simulations were done in Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) software [135] using the ReaxFF potential [51, 98] with a timestep of 0.25 fs following Mao et al. [60]. The simulations are conducted using the NVT (constant number, volume, and temperature) ensemble using the velocity-Verlet algorithm [100] along with the Nose-Hoover thermostat [102]. The simulation results are

examined at intervals of 0.05 ns, and large hydrocarbon clusters are identified, organized, and analyzed using a variety of tools including MSMS [136], MAFIA-MD [22], and OVITO [137].

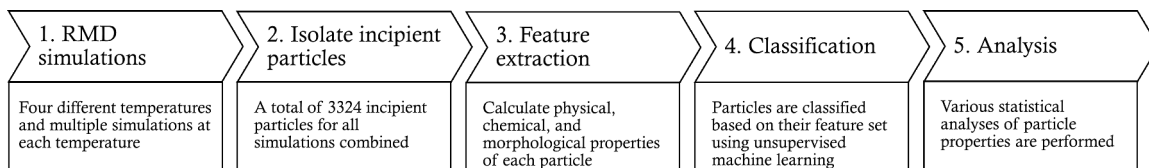


Figure 5.1: Overview of the workflow employed in this study. The blocks numbered (1) to (4) have been discussed in more detail in our previous work [110]. This work focuses on the block (5).

The overall workflow for this work is depicted in Fig. 5.1. A total of 24 RMD simulations are performed at four different temperatures. From each simulation, the large molecular clusters are identified as potential incipient soot at every 0.05 ns. These clusters are molecules having more than 20 carbon atoms and at least one 5-, 6-, or 7-member ring structure (indicating the possible presence of an aromatic component). These criteria are based on previous findings [94]. Incidentally, in our simulations, the smallest molecular cluster matching these criteria was found to have 65 carbon atoms. A total of 3324 soot particles are identified at various stages of their development from the 24 simulations. Then a variety of physical (e.g., density and mass), chemical (e.g., number of rings), and morphological (e.g., volume and surface area) features of each particle are quantified. These features are described in Sec. 5.2.2. Following this feature extraction, the particles are classified using two unsupervised machine-learning techniques: k-means clustering [138] and t-distributed stochastic neighbor embedding (t-SNE) [139]. As discussed in our previous work [110], this classification results in two distinct classes of incipient particles, referred to as Type 1 and Type 2 incipient particles. Finally, various statistics including the correlation among different properties are evaluated for the entire set of particles. The correlation between the physicochemical properties are estimated using Kendall’s Tau correlation [168]. Kendall’s Tau is a non-parametric measure of the relationship between columns of ranked

data. It estimates a correlation coefficient (τ) with a value between -1 and $+1$, where $+1$ means perfect positive correlation, -1 means perfect negative correlation, and 0 means no correlation between variables. It also provides a p-value that indicates the statistical significance of the correlation.

5.2.2 Extraction of physicochemical properties

Once the soot particles are identified and isolated using the cluster analysis tool from the OVITO python module [137], the physicochemical properties of each identified soot particle are evaluated using in-house Python script and other tools such as MSMS [136] and MAFIA-MD [22]. Key features such as number of atoms (N), carbon to hydrogen ratio ($\Theta_{c/H}$), and particle mass (M_p) can be obtained directly from the output log of RMD simulation (a trajectory file). The solvent-excluded volume (V) and solvent-excluded surface area (A) are calculated using MSMS [136] using a probe radius of 1.5\AA . Some other features are calculated by simple algebraic, or geometric analysis or by using correlations proposed in the literature as discussed below.

The radius of gyration (R_g) is calculated geometrically using the atomic coordinates extracted from the trajectory files using Eqn. 5.1:

$$R_g = \sqrt{\frac{\sum_{i=1}^N m_{p,i} r_i^2}{\sum_{i=1}^N m_{p,i}}}, \quad (5.1)$$

where r_i is the distance of the i^{th} atom from the center of mass of the molecular cluster, $m_{p,i}$ is the mass of i^{th} atom, and N is the total number of atoms in the cluster. In engineering-scale soot models, it is often difficult to capture the radius of gyration of incipient particles due to the lack of information on particle morphology and size distribution. However, the mass and volume of the particles are comparatively easier to obtain. Using these, the volume-equivalent radius (R_{eq}) can be calculated directly from particle volume from Eqn. 5.2:

$$R_{eq} = \left(\frac{3V}{4\pi} \right)^{1/3}. \quad (5.2)$$

The atomic fractal dimension (D_f) is calculated using the sandbox method [169, 170] using Eqn. 5.3.

$$D_f = \frac{\log M_p(r)}{\log r}, \quad (5.3)$$

where $M_p(r)$ is the mass of atoms in the cluster as a function of radial distance from center of mass (r). Atomic fractal dimension (D_f) is, therefore, the slope of the log-log plot of $M_p(r)$ vs. r . As discussed in [57], here the fractal dimension is calculated using the existing atoms in an incipient particle. Hence, the term “atomic” fractal dimension is used to remove any confusion with the fractal dimension of aggregates which is often used to calculate the level of geometric self-similarity in soot aggregates [171, 172, 173]. An atomic fractal dimension of 1 represents a linear structure while a value of 3 indicates a perfectly spherical shape [172].

The simulated primary particle density (shortened as simulated density, or ρ_s) of the incipient soot particles is obtained from the mass and volume of the incipient particle (Eqn. 5.4):

$$\rho_s = \frac{M_p}{V} \quad (5.4)$$

We also calculate a widely used metric – empirical density (ρ_e) also referred to as bulk density in the contemporary literatures, which can convey additional information about the maturity of the incipient particles. Empirical density of an incipient particle changes with the level of maturity of the particle and is calculated using Eqn. 5.5 [67, 174]:

$$\rho_e = (0.260884a^2c)^{-1} \left(\frac{w_C \Theta_{C/H} + w_H}{\Theta_{C/H} + 1} \right), \quad (5.5)$$

where w_C is the weight of carbon atoms (12.011 g/mol), w_H is the weight of hydrogen atoms (1.008 g/mol), a is the length of the graphite unit cell in the basal plane (2.46 Å), c is the interlayer spacing in Angstroms (3.50 Å for soot), and $\Theta_{C/H}$ represents the carbon to hydrogen ratio of the cluster.

Finally, MAFIA-MD [22] code is used to identify 5- /6- /7-membered ring structures in a molecular cluster. As discussed in detail in [22], MAFIA-MD does not strictly check for Huckel’s rules of aromaticity, and hence, to remove any confusion, we will use the terms

“ring” or “cyclic” in this manuscript instead of aromatic when discussing these internal structures in the soot clusters. The number of 5- /6- /7-membered rings are denoted as N_5 , N_6 , N_7 , respectively, and the total number of rings is denoted as N_{O} . Similarly, number of carbons in rings are denoted as N_{C} and the number of non-cyclic carbons in a particle is denoted as N_{C} .

5.3 Results and Discussion

5.3.1 Inception of soot and classification of incipient soot particles

The general sequence of events as seen from the RMD simulations leading up to the inception events have been discussed in detail in our earlier works [94, 110]. Initially, the acetylene molecules combine to form small linear chains. Subsequently, these linear chains undergo cyclization, transforming into cyclic structures. Following the process of cyclization, the small clusters start to grow through surface reactions that involve bond formation, as well as internal reorganization leading to incipient soot particles [94, 110]. A similar process has been also been reported by Zhang et al. [104] and Sharma et al. [57].

In the present study, being consistent with the current understanding of soot particles, an incipient soot particle is identified as a molecular cluster with more than 20 carbon atoms and with at least one ring structure (5-, 6-, or 7-member rings) [94]. The first such particle in our simulation was found to have 65 carbon atoms. A total of 3324 incipient soot particles, with carbon atoms ranging from 65 to 1503, were extracted at various stages of evolution from all the simulations at four temperatures. These particles showed significant variation in their physicochemical features and were therefore classified into two classes based on all the extracted features using unsupervised machine learning techniques (k-means clustering and t-SNE diagram) as discussed in [110]. For easier identification, the two classes of particles are denoted as “Type 1” and “Type 2” particles (see also Fig. 5.2). Type 1 particles exhibit a lower number of carbon atoms (65 – 818), whereas, in Type 2 particles, the number of carbon atoms is higher (759 – 1503). This observation highlights the fact that the physicochemical properties of the incipient particles

undergo a transition once a certain level of growth is reached [110]. It is important to mention that the classification is based on *all physicochemical features* and not based only on the number of carbon atoms. The number of carbon atoms, conveniently, serves as a good approximate indicator for the boundary between type 1 and type 2 particles. Some small type 2 particles may have a lower number of carbon atoms than large type 1 particles (there is a slight overlap in the range of the number of atoms between the two types). In our simulation, we identified a total of 670 type 1 and 2654 type 2 incipient particles. Figure 5.2 shows one sample Type 1 and one sample Type 2 particle. In this representation, the non-cyclic carbon atom structures are denoted by blue dots, while the cyclic structures are denoted by black dots.

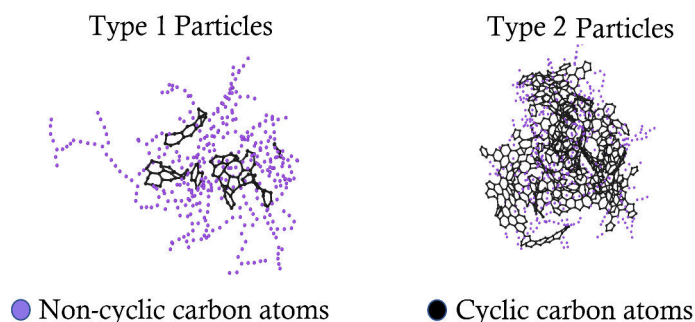


Figure 5.2: One example particle from each class obtained from RMD simulations.

It can be observed in Fig. 5.2 that there are discernible differences in the internal structure between type 1 and type 2 particles. Type 1 particles possess numerous small aromatic islands embedded within a plentiful network of aliphatic compounds. As the particles transition from type 1 to type 2, the smaller islands amalgamate into larger aromatic islands situated in the central region of the particles (growth centers [3]), resulting in a decrease in the proportion of aliphatics. In this manner, the liquid-like or young soot particles that are rich in aliphatic compounds (i.e., type 1 particles) transform and develop into partially graphitized soot particles (i.e., type 2 particles) with higher fraction of aromatic components. During this transformation, the soot particles acquire well-defined growth centers. Therefore, it is expected that the intercorrelations among various

physicochemical properties will change as the particle transitions from type 1 to type 2.

5.3.2 Physicochemical features of incipient particles

The physicochemical features of soot particles investigated include mass, number of atoms, atomic fractal dimension, volume, surface area, density, particle radius, and statistics of cyclic structures. The physical properties (for example, mass, volume, density, etc.) are often the features that are used in engineering-scale soot models for reactive CFD as well as in discrete element modeling [175] and dipole approximation [176, 155]. A better understanding of these features and their inter-correlation will help improve these engineering-scale soot models. The chemical properties such as the number of atoms, C/H ratio and statistics of cyclic/non-cyclic structures are important metrics by which the evolution and maturity of soot particles can be tracked. The evolution of the soot C/H ratio is essential to estimate the evolution of its optical properties [177] and assist its detection by laser diagnostics [177, 178].

In this work, we focused on the analysis for a set of select physical (mass, radius, surface area, and volume) and chemical (statistics of ring structures) features as the particle grows. As mentioned in Sec. 5.3.1, as the particle grows, it transitions from type 1 to type 2. Since the demarcation of these types can be very closely tracked by the number of carbon atoms in the particle, which correlates very well with the mass and size of the particle, we track the growth of a particle via either the total number of carbon atoms or the particle mass or the particle radius. The Kendall’s Tau [168] statistical test is performed to determine the level of correlation between the different properties. The correlation coefficients (τ) of the Kendall’s Tau statistical test are reported in each figure. It was found that whenever a good correlation (i.e. $|\tau| \gg 0$) was found between two variables, the test yielded a very small p-value (almost zero), indicating a high statistical significance. Therefore, the p-value is not reported in the figures for brevity. We visualize the data using a joint plot of scatter plot and histogram/distribution using the python package `seaborn` [179].

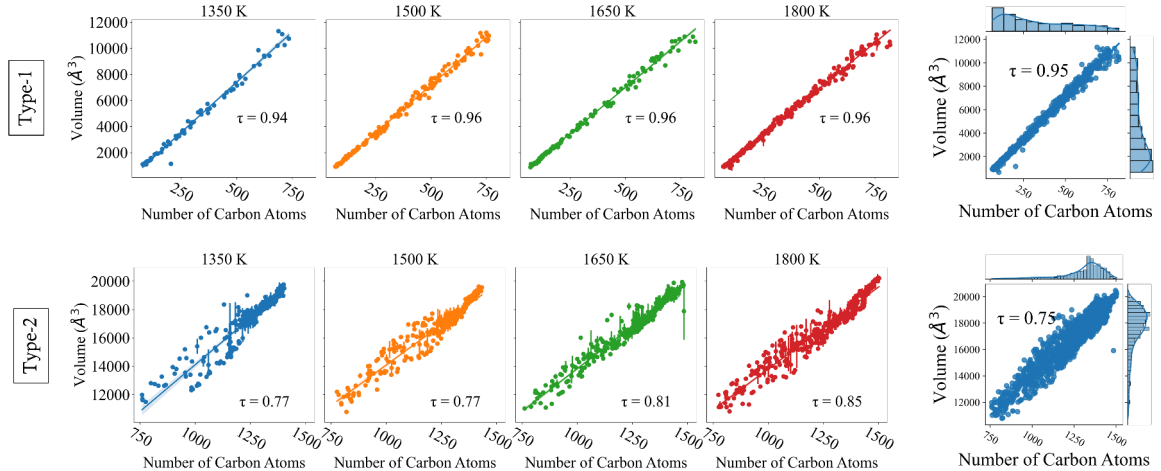


Figure 5.3: Evolution of particle volume with the number of carbon atoms at different temperatures. Top row: Type 1 particles, bottom row: Type 2 particles.

The evolution of particle volume and surface area is presented in Figs. 5.3 and 5.4 for both type 1 (top row) and type 2 particles (bottom row) as a function of the number of carbon atoms in the particles. The first four columns of plots correspond to simulation temperatures of 1350, 1500, 1650, and 1800 K, whereas the rightmost column combines all the temperatures together. A trendline is also shown on the scatter plots for each temperature. From Fig. 5.3, we see that both type 1 and type 2 particles have an excellent correlation between the particle volume and the number of carbon atoms in the particles across all temperatures. The correlation coefficient (τ) is very close to 1 for type 1 particles, representing an almost perfect correlation between the variables. For type 2 particles, the correlation coefficient (τ) is lower, but still close to 1. This indicates that the particle volume is well correlated with the number of carbon atoms in both type 1 and type 2 particles. Another important observation is that the particle volume and number of carbon atoms seem to be correlated in the same way for all simulation temperatures. The slopes of the lines that provide linear fit between the volume and the number of carbon atoms do not vary much with temperature (discussed further later in Table 5.1). This shows that the relationship between particle volume and number of carbon atoms is not affected by the temperature.

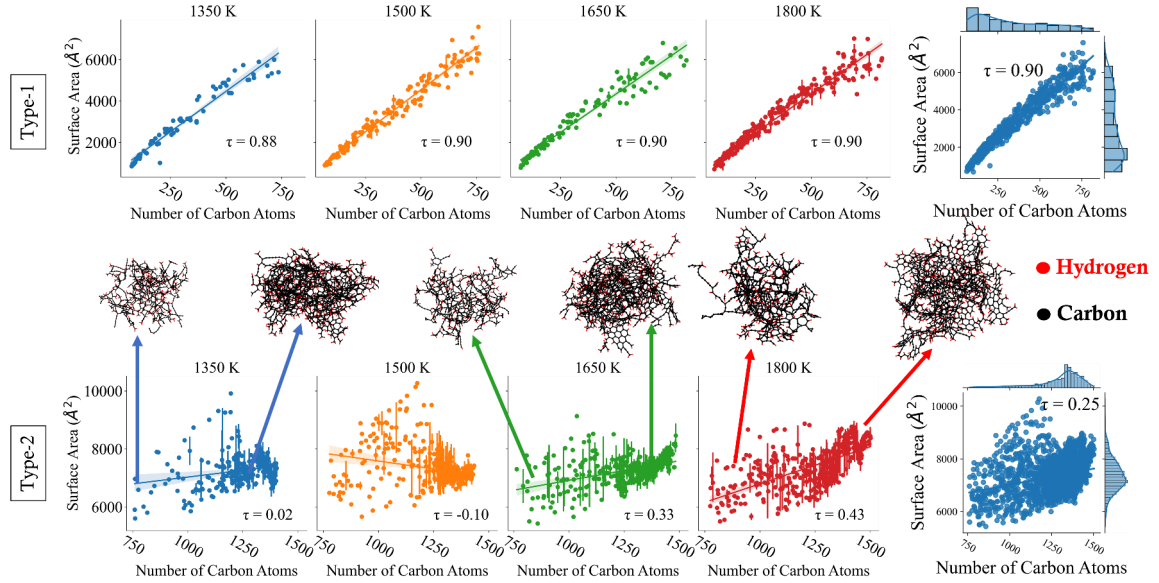


Figure 5.4: Evolution of particle surface area with the number of carbon atoms at different temperatures. Top row: Type 1 particles, bottom row: Type 2 particles.

For type 1 particles, a similar conclusion can also be drawn from Fig. 5.4 for the surface area: a very good correlation is observed with the number of carbon atoms. But for the type 2 particles, the surface area does not seem to be correlated with the number of carbon atoms. As the particle transitions from type 1 to type 2, a change from the initial linear growth to the extensive reorganization of the surface is observed. This is shown using some actual particles at the two extremes of the number of carbon atoms for 1350, 1650 and 1800 K in the insets of Fig. 5.4. As can be seen, while the topology of the surface shows significant difference between a small type 2 particle ($N_C \sim 800$) and a large type 2 particle ($N_C \sim 1400$), the actual difference in surface area is small. It shows that the soot surface goes through extensive rearrangement in type 2 particles, which makes the surface area hard to correlate with the number of carbon atoms. The rearrangement of the surface area occurs mostly because of graphitization of the surface as discussed in the internal structure of type 2 particles in [110]. The graphitization of the surface is a very complex process and is not well understood. However, it is known that the graphitization of the surface is a slow process, and it depends on the size of the primary particle itself [180]. We can also note that there is a lower number of ring structures in the similar-sized small type 2 particles

($N_C \sim 800 - 1000$) in 1800 K than in 1350 K. Temperature can accelerate surface graphitization [181] and, therefore, while comparing the small type 2 particles ($N_C \sim 800 - 1000$) we observe the formation of graphitized particles (indicated by a large presence of cyclic structures) earlier in 1800 K simulation than in 1350 K simulation. The fraction of cyclic carbon atoms is plotted against molar mass in Fig. 5.5. For type 1 particles, the fraction of cyclic carbon atoms is mostly insensitive to the particle mass. However, for type 2 particles, the fraction of cyclic carbon atoms increases rapidly with molar mass. This further supports the inference regarding the extensive reorganization of the surface in type 2 particles due to graphitization. This is consistent with results reported by Liu et al. [133] where it is shown that smaller soot particles take longer to graphitize compared to the larger ones. This lack of graphitization leads to a good correlation between surface area and the number of atoms in type 1 particles, which are primarily seen at the early stage of soot formation. Furthermore, the process temperature does not seem to affect the quality of correlation or lack thereof.

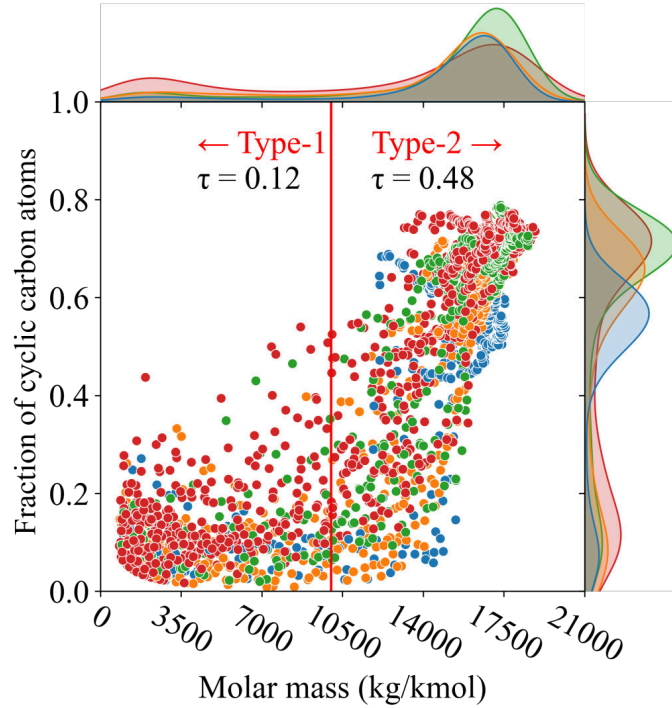


Figure 5.5: Evolution of fraction of carbon atoms in ring structures ($N_{@}/N_C$) with molar mass.

From Figs. 5.3 and 5.4, linear trends with the number of carbon atoms are observed for particle volume (type 1 and type 2) and surface area (type 1). A set of linear equations were fitted between particle volume and surface area with the number of carbon atoms for different temperatures and presented in Table 5.1. For particle volume, both type 1 and type 2 particles show a very good linear relationship with the number of carbon atoms (N_C) as indicated by high R^2 values across different temperatures and classes of particles. The R^2 , or the coefficient of determination ($R^2 \in [0, 1]$) is a statistical measure of how close the data are to the fitted regression line. As discussed previously, for surface area, a good correlation is only observed for type 1 particles and therefore, no correlation for surface area for type 2 particles is presented in Table 5.1. The slope of these linear correlations are very close to one another supporting the observation of the lack of influence of temperature on the quality of the correlation between these quantities. Therefore, we constructed a set of temperature-independent correlations for volume and surface area by aggregating data from all the temperature in Correlation Set 5.1.

Correlation Set 5.1: Equation of curves fitted to the temperature-aggregated data shown in Figs. 5.3 and 5.4. V is the particle volume in \AA^3 , A is the particle surface area in \AA^2 , and N_C is the total number of carbon atoms.

Type 1 particles:

$$V = 14.353N_C - 119.071, \quad R^2 = 0.992 \quad (5.6)$$

$$A = 7.733N_C + 547.914, \quad R^2 = 0.951 \quad (5.7)$$

Type 2 particles:

$$V = 11.825N_C + 2151.314, \quad R^2 = 0.897 \quad (5.8)$$

While we only show the results for particle volume and surface area, a similar conclusion about the effect of temperature can be drawn from the evolution of other morphological features, i.e. radius of gyration, volume equivalent radius, number/fraction of

Table 5.1: Equation of linear curves fitted to the data for different temperatures shown in Figs. 5.3 and 5.4. V is the particle volume in \AA^3 , A is the particle surface area in \AA^2 , and N_C is the total number of carbon atoms.

Temperature	Particle Volume (V)		Surface Area (A)	
	Type 1	Type 2	Type 1	Type 2
1350 K	$V=15.405 N_C - 297.596$ $R^2 = 0.987$	$V=13.172 N_C + 930.663$ $R^2 = 0.883$	$A=7.842 N_C + 551.652$ $R^2 = 0.944$	No meaningful correlation
1500 K	$V=15.405 N_C - 297.596$ $R^2 = 0.987$	$V=11.41 N_C + 2663.046$ $R^2 = 0.913$	$A=8.171 N_C + 464.281$ $R^2 = 0.960$	
1650 K	$V=14.207 N_C - 113.638$ $R^2 = 0.994$	$V=12.028 N_C + 1784.325$ $R^2 = 0.934$	$A=7.459 N_C + 614.673$ $R^2 = 0.944$	
1800 K	$V=14.014 N_C - 65.388$ $R^2 = 0.994$	$V=11.564 N_C + 2228.574$ $R^2 = 0.926$	$A=7.583 N_C + 560.486$ $R^2 = 0.953$	

cyclic structures, etc. as well . Discussion on these are omitted in this work for brevity.

Most engineering-scale soot models track the evolution of the soot population via the mass or size of soot particles. Therefore, the correlations of particle volume, surface area, and number of rings are shown with molar mass in Fig. 5.6 and with the radius of gyration in Fig. 5.7. The results are presented in the form of joint plots which show the distribution of the data points in the form of a scatter plot and the distribution of the individual properties in the form of a density plot along with the temperature markers. The vertical red line in each plot indicates the boundary between type 1 and type 2 particles and the correlation coefficients (τ) are reported separately for both types of particles on their respective zones. The transition of a particle from type 1 to type 2 occurs around a molar mass of 10050 kg/kmol or a radius of gyration of 13.8 Å, and the vertical red lines are drawn at those locations in these plots. The type 1 particles are on the left of this line and type 2 is on the right.

The volume, surface area, and radius of gyration of the type 1 particle are well correlated with the molar mass of the particles as seen in Fig. 5.6. For type 1 particles, volume and surface area show a linear correlation, while the radius of gyration shows a nonlinear correlation with the mass of the particles. For the total number of ring structures, however, the correlation is not as good as the other three properties. The nature of the correlations between variables remains unaffected by the temperature. The most important observation from Fig. 5.6 is the difference in the nature of the correlations between type 1 and type 2. In all cases, the data points become more clustered and the correlations become weaker except for the total number of ring structures in type 2 particles. A very sharp increase in the total number of ring structures is observed as the particles transition to type 2 and grow in mass. The total number of rings is also affected by temperature, even though temperature does not affect the other morphological features. A higher number of cyclic structures is observed with the increase in temperature as we can see from the temperature markers and density plots on the right side of the plot. A set of correlation equations by curve-fitting the data shown in Fig. 5.6 is presented in Corr. Set 5.2.

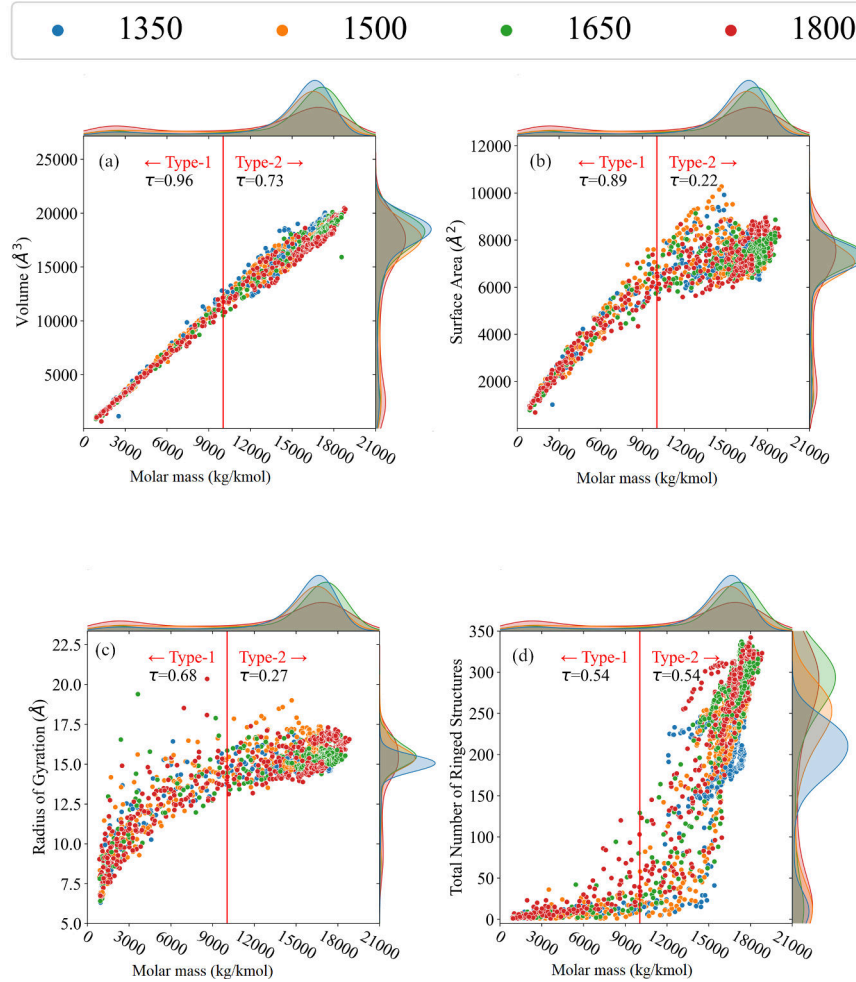


Figure 5.6: Relationship between molar mass and (a) volume, (b) surface area, (c) radius of gyration, and (d) total number of rings in incipient soot particles

Similar conclusions can be drawn using the radius of gyration as the independent variable as shown in Fig. 5.7. The volume and surface area, and mass all show a good correlation for type 1 particles. For the total number of ring structures, however, the correlation is not as good as the other three properties. The nature of the correlations between features remain unaffected by the temperature. Similar to Fig. 5.6, the type 2 particles show comparatively weaker correlations with the radius of gyration compared to type 1 particles. A set of correlation equations obtained by curve-fitting the data shown in Fig. 5.7 is presented in Corr. Set 5.3.

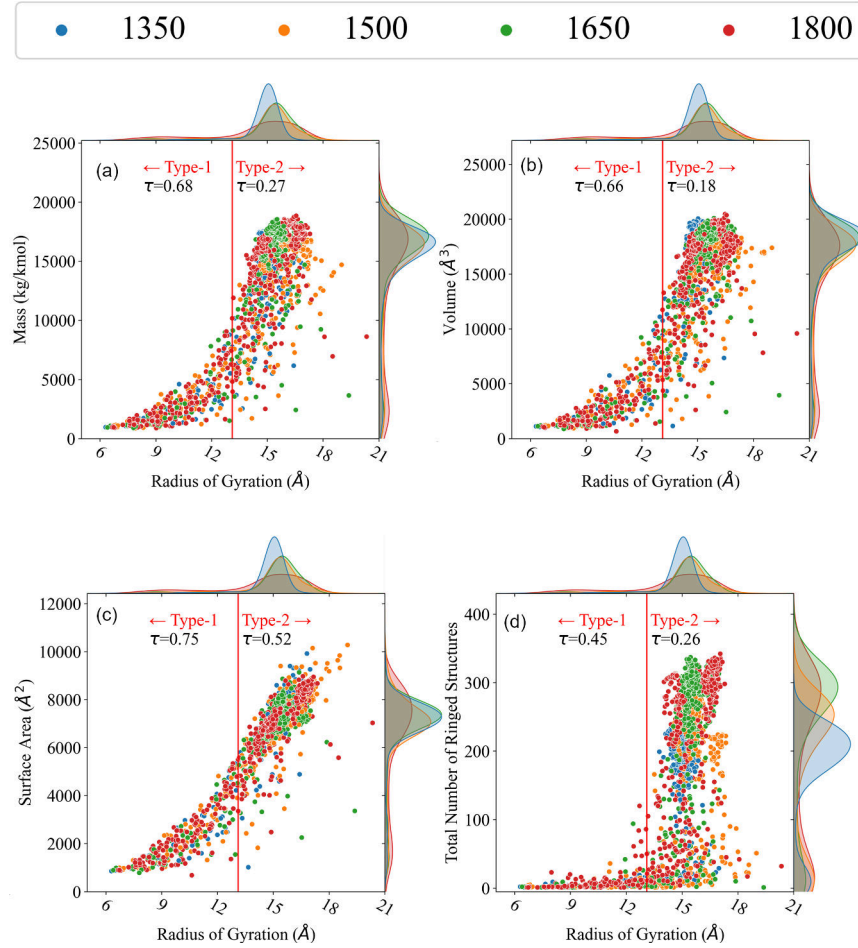


Figure 5.7: Relationship between radius of gyration and (a) mass, (b) volume, (c) surface area, and (d) total number of rings in incipient soot particles.

The nature of the physical and chemical properties of incipient soot changes as the particle grows from type 1 to type 2. Identification of this transition from type 1 and 2 particles is valuable in the context of engineering-scale soot models as it will allow for updating of property correlation in the soot model without analyzing the internal structure. This may lead to a more comprehensive understanding of soot evolution, including the morphology of primary particles. This is further evident in the discussion of ring structures presented in Sec. 5.3.3.

Correlation Set 5.2: Equation of curves fitted to the data shown in Fig. 5.6. M is the molar mass of the particle in kg/kmol, V is the particle volume in \AA^3 , A is the particle surface area in \AA^2 , R_g is the radius of gyration in \AA , and N_{O} is the total number of ring structures in the particle.

Type 1 particles:

$$V = 1.160M - 124.597, \quad R^2 = 0.993, \quad (5.9)$$

$$A = 0.625M - 545.410, \quad R^2 = 0.952, \quad (5.10)$$

$$R_g = 2.947 \ln(M) - 12.487, \quad R^2 = 0.717, \quad (5.11)$$

$$N_{\text{O}} = 0.004M - 5.964, \quad R^2 = 0.372. \quad (5.12)$$

Type 2 particles:

$$V = 0.935M - 2375.685, \quad R^2 = 0.905, \quad (5.13)$$

$$A = 0.098M - 5783.961, \quad R^2 = 0.090, \quad (5.14)$$

$$R_g = 1.933 \ln(M) - 3.293, \quad R^2 = 0.114, \quad (5.15)$$

$$N_{\text{O}} = 0.036M - 342.906, \quad R^2 = 0.704. \quad (5.16)$$

Correlation Set 5.3: Equation of curves fitted to the data shown in Fig. 5.6. V is the particle volume in \AA^3 , A is the particle surface area in \AA^2 , R_g is the radius of gyration in \AA , and N_{O} is the total number of ring structures in the particle.

Type 1 particles:

$$V = 995.677R_g - 6699.483, \quad R^2 = 0.647, \quad (5.17)$$

$$A = 596.486R_g - 3678.870, \quad R^2 = 0.767, \quad (5.18)$$

$$N_{\text{O}} = 2.945R_g - 22.272, \quad R^2 = 0.162. \quad (5.19)$$

Type 2 particles:

$$V = 807.096R_g + 4944.945, \quad R^2 = 0.106, \quad (5.20)$$

$$A = 608.281R_g + 2042.065, \quad R^2 = 0.541, \quad (5.21)$$

$$N_{\text{O}} = 25.862R_g - 169.784, \quad R^2 = 0.059. \quad (5.22)$$

Fig. 5.8 shows the comparison between the molar mass and the atomic fractal dimension (D_f) in the form of a joint scatter plot. As discussed in Eq. 5.3, the atomic fractal dimension is calculated using box counting method for individual incipient particles, not for an aggregate. The value of atomic fractal dimension (D_f) is an indication of the shape of the incipient primary particles. A value of D_f close to 3 indicates a spherical shape, while a value of D_f close to 1 indicates a linear shape. Contemporary engineering-scale soot models often assume spherical incipient particles [14, 12, 96].

For both type 1 and type 2 particles, we do not observe any specific correlation between molar mass and atomic fractal dimension as indicated by the low τ value in Fig. 5.8. As the molar mass increases, the atomic fractal dimension increases and approaches 3. This is because as the incipient particles grow, the shape of the particles becomes more spherical. The value of D_f captures the evolution of the shape of a single incipient particle before coalescence. In the inset of Fig. 5.8, some representative incipient particles are shown at various stages of their evolution.

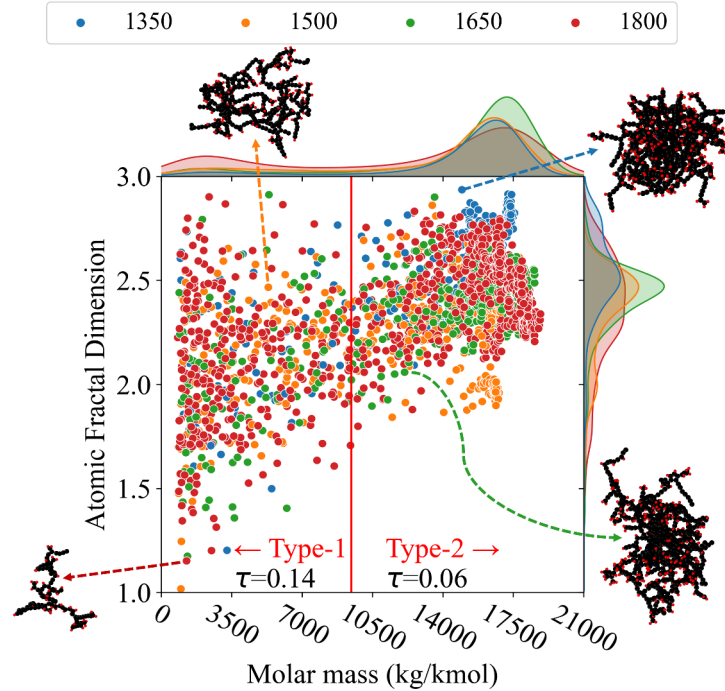


Figure 5.8: Evolution of atomic fractal dimension (D_f) with molar mass

The correlation between the volume equivalent radius and the radius of gyration is shown in Fig. 5.9(a). The transition from type 1 to type 2 particles occurs around a volume equivalent radius of 13.8 Å corresponding to the red vertical line in Fig. 5.9(a). An overall good correlation is observed between the two quantities for type 1 particles. Like the other morphological features, the type 2 particles show a comparatively weaker correlation compared to the type 1 particles. Since the particles tend to become spherical as they evolve (Fig. 5.8), the volume equivalent radius can be a good approximation for the radius of gyration. It should be noted here that although the particles become more spherical, they are not solid spheres, therefore relationships between the volume-equivalent radius and radius of gyration do not follow the classical $\sqrt{3/5}$ scaling (black solid line). Rather they closely follow the $R_g = R_{eq}$ scaling (blue dashed line) as shown in Fig. 5.9(a).

This becomes more evident if we look at the evolution of the ratio of radius of gyration and volume-equivalent radius (R_g/R_{eq}) as the particles grow in size. Figure 5.9(b) and 5.9(c) depict R_g/R_{eq} vs. the number of carbon atoms (N_c) for type 1 and type 2

particles respectively. For small type 1 particles, the ratio R_g/R_{eq} shows a wide range of values ranging from 1.0 to 2.1. But within the increase in particle size, i.e., increase in the number of carbon atoms (N_c) in the particle, the spread of the ratio R_g/R_{eq} becomes narrower and approaches unity as seen from Fig. 5.9(b). This is because the particles become more spherical and compact as they grow in size. For type 2 particles, the ratio R_g/R_{eq} steadily hovers near 1.0 with very little spread in the data as observed from Fig. 5.9(c). This is because the type 2 particles are more graphitized and have a more compact structure compared to type 1 particles.

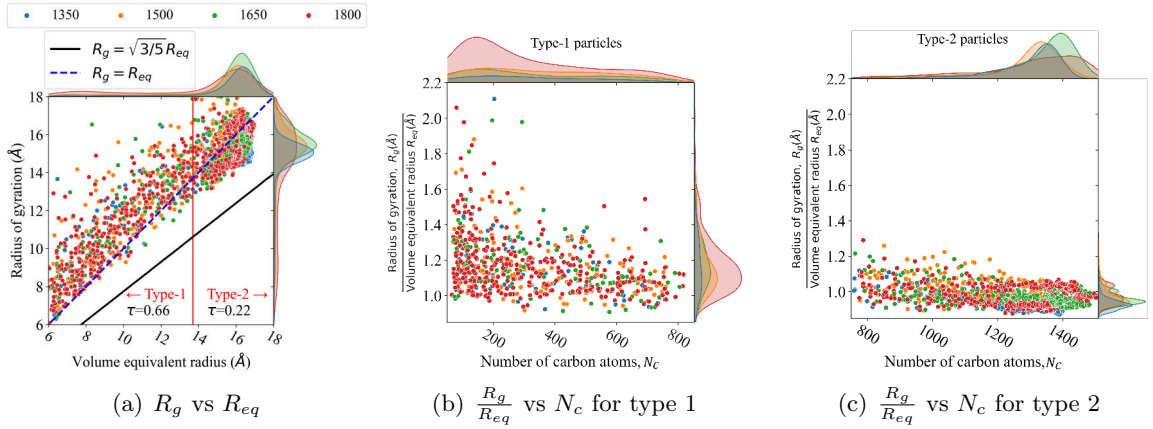


Figure 5.9: Evolution of radius of gyration (R_g) and volume equivalent radius (R_{eq}): Comparison between volume equivalent radius and radius of gyration for all particles (a); ratio of radius of gyration and volume equivalent radius vs number of carbon atoms for type 1 particles (b) and for type 2 particles (c).

5.3.3 Evolution of ring structures in incipient soot particles

The ring structures in soot particle play an important role in their evolution and impact their physical and chemical properties [182]. As seen in Figs. 5.6 and 5.7, the number of rings in incipient soot particles evolve differently from the morphological features. Therefore, we looked into the evolution of 5-/6-/7-membered rings more closely in this section. Fig. 5.10 shows the evolution of the total number of 5-/ 6-/ 7-membered rings (N_5 , N_6 , N_7) in the top row and their fraction with respect to the total number of rings (N_5/N_{O} ,

N_6/N_O , N_7/N_O) in the bottom row as a function of the molar mass of the incipient particles. The red vertical line indicates the threshold between type 1 and type 2 particles. The type 1 particles are on the left of the line and type 2 particles are on the right. The Kendall's correlation coefficient (τ) between the ring structures with particle mass are reported for each type on their respective zones.

The total number of 5-/6-/7-membered rings in type 1 particles increases only slightly with molar mass and no distinct difference is observed between different ring structures. However, in the type 2 regime, a rapid increase in the total number of rings is observed with mass. This increase is most prominent for the 6-membered rings. The 5- and 7-membered rings also increase with mass but at a much slower rate. A higher number of 6-membered rings is observed with increasing temperature in the type 2 regime as seen from the temperature markers and density plots on the right side of top middle plot of Fig 5.10(a).

Figure 5.10(b) shows the fraction of 5-/6-/7-membered rings in individual incipient particles and provides insights into how different rings interact and change during the evolution of incipient soot particles. In the type 1 regime of Fig. 5.10, the fraction of 5-/ 6-/ 7-membered rings does not show any order or correlation. Throughout the simulation domain, 5-, 6-, and 7-membered rings form independently due to the chemical interactions between small aliphatics without the influence of other cyclic structures. Once the particles grow enough to transition to type 2, the fractions of 5-/ 6-/ 7-membered rings start to change in an orderly manner. The fractions of 5- and 7-membered rings decrease with mass while the fraction of 6-membered rings increases. This is because the 6-membered rings are the most stable ring structures and as the particle grows, the 5- and 7-membered rings are more likely to break and form 6-membered rings¹. The fraction of rings is also affected by temperature. A higher fraction of 6-membered rings is observed with increasing

¹This shift in ring fractions is driven by the superior thermodynamic stability of 6-membered aromatic rings. While the absolute number of all ring types increases during particle growth (Fig. 5.10(a)), the chemical pathways preferentially lead to the formation of more stable 6-membered rings at a higher rate. This energetic preference acts as a thermodynamic sink, causing the proportion of 6-membered rings to increase at the expense of the less-stable 5- and 7-membered ring fractions. This is discussed in detail at Appendix E

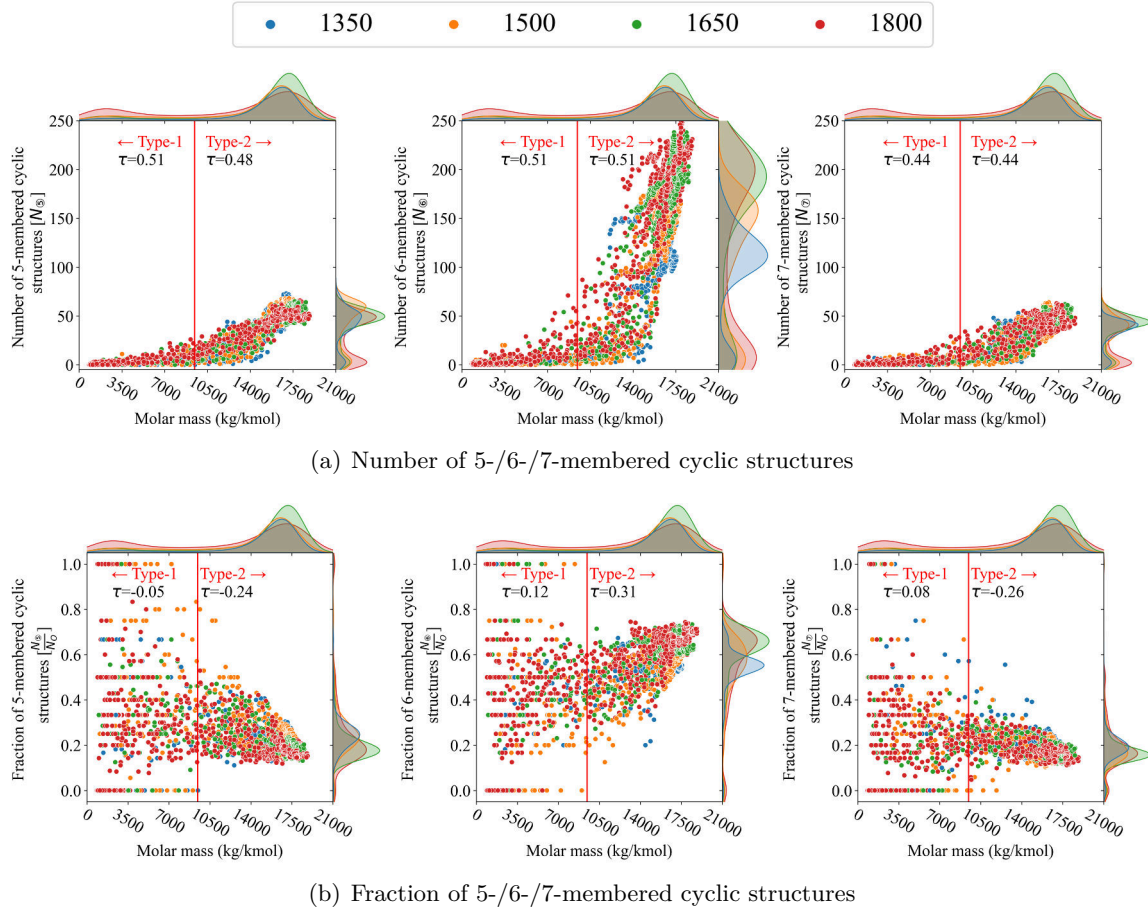


Figure 5.10: Evolution of cyclic structures in incipient soot particles.

temperature in the type 2 regime as seen from the temperature markers and density plots on the right side of the center plot of Fig. 5.10(b). The fraction of 5- and 7-membered cyclic structures, on the other hand, decreases slightly with increasing temperature.

5.3.4 Summary statistics

The summary of the statistics is presented in Table 5.2 in terms of mean, standard error to the mean (SEM) and standard deviation (SD) of the extracted features. These values can be useful in determining model parameters and model constants for engineering-scale soot model. The statistics are also separated by the type of the particles. It can be seen that the variation in physical features such as simulated density, radius of gyration, volume, and surface area is much smaller in type 2 particles than in type 1 particles. This indicates

that in type 1 regime, particles go through a strong physical and morphological evolution, but in type 2 regime, they go through a chemical restructuring. Due to these changes, from type 1 to type 2, the atomic fractal dimension increases from 2.1 to 2.5 (more spherical). The proportion of 6-membered rings increases significantly while the fractions of 5-/7-membered rings decrease as they restructured into the more stable 6-membered rings.

The calculated statistics provide a great validation for the RMD when compared with experimental data reported in the literature. Schulz et. al [125] reported that the average C/H ratio of soot particles from a premixed ethylene flame was 2.33 ± 0.16 , which matches very well with the mean C/H ratio of 2.36 with a standard deviation of 0.37 for the simulated particles in our study. It should be noted that, we do see a reduction in the C/H ratio as the particles transition from type 1 to type 2 particles. This can be explained by the internal structure of the incipient particles where the type 2 particles have a larger shell region consisting of mostly non-cyclic aliphatic chains compared to type 1 particles [110]. The average simulated density of the soot particle in our study is in the range of 1.50 (type 1) – 1.53 g/cm³ (type 2). This is very similar to values reported or assumed in several other studies [115, 151, 152]. Furthermore, Johansson et al. [67] and [149] found an empirical density of about 1.5 g/cm³ for incipient soot particles in their experiments. For PAHs with C/H ratio of 1.0-2.4, Minutolo et. al. [183] found the average density of soot to be 1.5 gm/cm³ using relationships from [184].

Finally, a summary correlation matrix between each pair of features for both type 1 and type 2 incipient particles are presented in Fig. 5.11. The correlation matrix is a square matrix with the number of rows and columns equal to the number of properties. The diagonal elements of the matrix are always 1, and the off-diagonal elements are the correlation coefficients between the corresponding properties. The correlation coefficients are calculated using Kendall’s Tau test. The correlation coefficients are presented in the form of a heat map in Fig. 5.11, where the color intensity is proportional to the value of the correlation coefficient. Both type 1 and type 2 correlations are presented in the square correlation matrix. The lower triangular region represents type 1 particles and the upper triangular region represents type 2 particles. For simplicity, the correlation matrix is

Table 5.2: Mean, Standard Error of the Mean (SEM), and Standard Deviation (SD) of the physicochemical features of soot clusters.

Property	Unit	Type 1			Type 2		
		Mean	± SEM	SD	Mean	± SEM	SD
Simulated Density	g/cm^3	1.502	± 0.005	0.139	1.533	± 0.001	0.060
Radius of gyration	Å	11.349	± 0.094	2.430	15.472	± 0.014	0.741
Volume equivalent radius	Å	9.793	± 0.090	2.318	16.063	± 0.012	0.597
Ratio of radius of gyration over volume equivalent radius	–	1.176	± 0.007	0.171	0.964	± 0.001	0.05
Molar mass	kg/kmol	4073.8	± 99.9	2584.0	16104.5	± 36.2	1866.3
Volume	Å^3	4600.3	± 116.3	3008.3	17432.3	± 35.6	1834.6
Surface area	Å^2	3090.6	± 64.0	1654.9	7369.2	± 11.9	612.2
Atomic fractal dimension	–	2.093	± 0.013	0.340	2.460	± 0.003	0.169
Fraction of 5-membered rings	–	0.374	± 0.011	0.294	0.218	± 0.001	0.063
Fraction of 6-membered rings	–	0.399	± 0.010	0.258	0.605	± 0.002	0.084
Fraction of 7-membered rings	–	0.227	± 0.009	0.221	0.177	± 0.001	0.041

presented only for a subset of features explored.

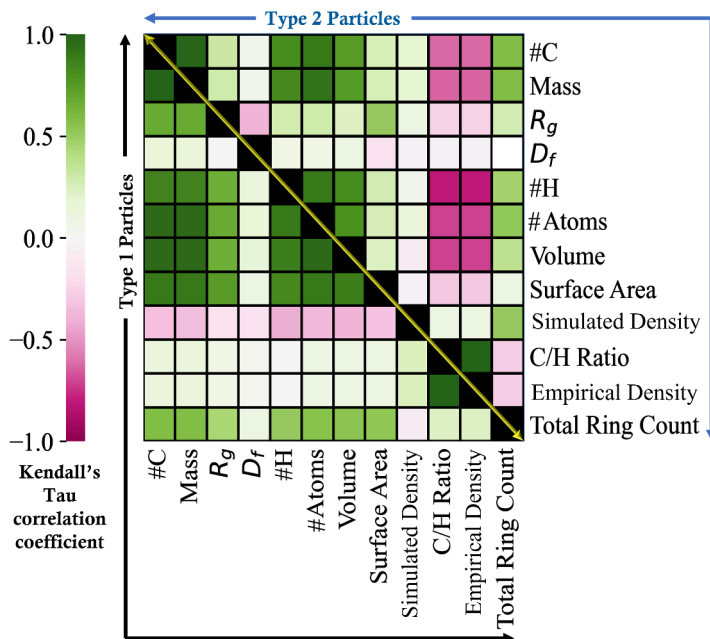


Figure 5.11: Kendall's rank correlation coefficient between different physicochemical features of type 1 (lower-left triangular matrix) and type 2 (upper-right triangular matrix) incipient soot particles.

If the nature of correlations were to be the same between the types of particles, the lower and upper triangular portions would be symmetric about the diagonal. However, as seen in Fig 5.11, the quality and trends of correlations change significantly between the types of particles. For example, the mass and empirical density is negatively correlated in type 2 particles whereas they are almost uncorrelated in type 1 particles. A correlation matrix like this can be useful in creating a reduced-order soot model in the future.

5.4 Conclusion

A series of molecular dynamics simulations was performed to study the evolution of incipient soot particles during acetylene pyrolysis at four different temperatures. The evolution of the incipient soot particles was studied via their physicochemical characteristics such as mass, volume, surface area, radius of gyration, density, and the number of

5-/6-/7-membered rings. Using unsupervised machine learning techniques, the incipient soot particles are classified into two types – type 1 and type 2 – based on their morphological and chemical features. This classification was very well predicted by the number of carbon atoms in the particles, indicating the two types corresponding to early and late stages of incipient soot. The RMD-derived incipient soot particle density and C/H ratio shows excellent agreement with experimental data.

The following conclusions were drawn from the study.

1. Morphological features of incipient particles are often well correlated with each other and with the size of the particle (indicated by the molar mass, number of atoms, or radius of gyration). However, the quality and nature of their correlations are different between different types of particles.
2. Type 1 particles usually show stronger correlations between different morphological features and size than type 2 particles. Morphological features in type 2 particles show much smaller variation than type 1 particles.
3. Morphological features of incipient particles, e.g., volume, surface area, radius of gyration etc. are not affected by temperature. However, the chemical characteristics, e.g., number of cyclic structure and 5-,6-,7-membered rings of incipient particles are affected by temperature.
4. At the early stage of incipient soot (type 1), 5-, 6-, and 7-membered rings are formed independently from one another. However, at the later stage (type 2), the 5- and 7-membered rings are more likely to break and form 6-membered rings.
5. The incipient particles evolve into a more spherical shape as they transition from type 1 to type 2.
6. Overall, type 1 particles show a more prominent morphological evolution (i.e. volume, surface area and radius of gyration) than type 2 particles with increasing particle mass. On the other hand, type 2 particles go through a more prominent chemical evolution (i.e. evolution of 5-,6-, and 7-membered cyclic structures) than type 1 particles.

The estimation of various quantitative properties and their correlation in incipient soot reported in this work can improve engineering-scale models for soot inception and growth.

Acknowledgments

The research benefited from computational resources provided through the NCMAS, supported by the Australian Government, The University of Melbourne's Research Computing Services and the Petascale Campus Initiative. K.M.M. and S.P.R. acknowledge funding support from the National Science Foundation as some of this material is based upon work supported by the National Science Foundation under Grant No. 2144290.

CHAPTER 6

ELUCIDATING PORE AND SURFACE FEATURES OF SOOT NANOPARTICLES USING MOLECULAR DYNAMICS SIMULATIONS

This chapter is a reproduction of currently under-review manuscript titled “Elucidating Pore and Surface Features of Soot Nanoparticles using Molecular Dynamics Simulations” by K. M. Mukut, E. Goudeli, G. A. Kelesidis, and S. P. Roy.

6.1 Introduction

Soot or black carbon is a harmful byproduct of incomplete combustion of hydrocarbon fuels [3]. It also impacts the radiative energy balance of the atmosphere and is a major forcing factor behind climate change [10, 6]. It impacts public health and welfare and is one of the leading causes of mortality worldwide [185].

The properties of the soot internal and external surface are important in defining how soot interacts with its surroundings. The interaction of soot with gases, pollutants, and water is influenced by its surface properties including chemical and physical properties, hydrophilicity, and charge. For instance, the type of the soot and the environment that surrounds it, both have an effect on the adsorption behavior of the soot [186, 187, 188].

The irregular nature of soot surface and shape is often measured and represented by a suitable definition of fractal dimension. Surface fractal dimension (D_s) can be thought of as a measure of the roughness of a surface [189]. The volumetric fractal dimension (D_v) on the other hand is a measure of compactness of a soot particle. Soot particles with high surface fractal dimension (D_s) have more surface area available for adsorption and chemical reaction leading to a higher reactivity in the atmosphere [190]. The morphology of soot aggregates are often characterized by a statistical mass-fractal relationship among the primary particles, which leads to an aggregate fractal dimension (D_f) [191]. The soot aggregates with lower aggregate fractal dimension (D_f) are prone to faster aggregation and coagulation [191]. Depending on the aggregate fractal dimension (D_f), soot can affect the radiative balance of the atmosphere and cloud formation differently [192, 173]. Higher aggregate fractal dimension (D_f) in soot aggregates increases the light scattering while

lower aggregate fractal dimension (D_f) increases absorption [192].

The distribution of interior cavities or pores is also crucial for understanding soot particles' interaction with the atmosphere. Cavities affect soot particle chemical reactivity, structural stability, and environmental impact. Porosity is a measure of empty space within a particle and defined as the ratio of empty space within a particle to the total volume of the particle. The empty space within a soot particle's volume aids in adsorbing atmospheric gases like SO_2 and NO_2 affecting the aging of soot particles which influences atmospheric chemistry and pollutant transport [193, 194]. Soot particles with higher porosity and irregular sphericity lead to higher reactivity which can influence cloud formation and precipitation by acting as cloud condensation nuclei [187]. Moreover, when inhaled, soot particles with greater porosity are more likely to enter the lungs deeply and cause harm to the body [195].

Depending on the size, pores are classified into micropores (pores ≤ 2 nm) and mesopores (pores between 2 and 50 nm) [196, 197]. Micropores are smaller, but they provide very large surface area for gas adsorption and interaction with atmospheric pollutants and water vapor [198, 199]. Mesopores on the other hand, promote diffusion of larger molecules through particles and increase gas transport into the micropores resulting in an increase in overall reactivity [199, 200, 201].

Based on the accessibility to the external surface, pores or cavities can again be classified into three groups: closed or isolated cavity that has no opening to the external surface; open or pocket cavity that has one opening to the external surface, and through or tunnel cavity which has two or more openings to the external surface [196, 197]. Figure 6.1 depicts a simplified diagram of a particle with different types of cavities.

The scale of soot particles can vary from a few nanometers [202] to a few micrometers [203] depending on fuel, environment and combustion conditions [204, 205]. Because of the small scale and dependency on the combustion conditions, it is difficult to employ experimental techniques to study different stages of soot formation and evolution of surface and internal properties [206, 207]. However, due to the recent advances in computational techniques, reactive molecular dynamics (RMD) simulations are becoming

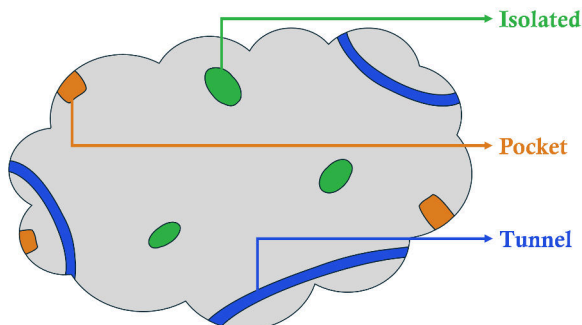


Figure 6.1: Schematic representation of different types of cavities inside an irregularly shaped particle.

increasingly popular tools for detailed atomic-level exploration of the internal structure of soot particles [130, 140, 110].

In this work, we investigate the surface and pore features of incipient soot primary particles using RMD simulations. We aim to provide insights into the surface and pore features of these particles and explore the correlation among these features. This study is a continuation of our previous works, where we investigated the internal structure of incipient soot particles [110] as well as the physical, chemical, and morphological evolution of these particles [140] using RMD simulations of acetylene pyrolysis. We believe that the results from the current study will provide a better understanding of the surface and pore properties of incipient soot particles and can be used to improve engineering-scale soot models. To the best of the authors' knowledge, this is the first study to investigate the surface and pore properties in such atomistic details and the methodology used here can be used as a guideline for future studies on soot particles.

6.2 Methodology

The overall workflow of the present study is depicted in Fig. 6.2. Blocks numbered from (1) to (3) in the top box are discussed in details in our previous works [110, 140]. The tasks done in the present study are shown in blocks numbered from (4) to (7) within the bottom box.

6.2.1 Simulation of incipient soot particles

The Reactive Molecular Dynamics (RMD) approach (Block 1) used in this work follows the methodology presented in [57], using the specific simulation settings previously reported in detail [110, 140]. Hence, only a brief synopsis of the methodology is presented here. One thousand acetylene molecules are randomly added to a cubic domain measuring $75 \text{ \AA} \times 75 \text{ \AA} \times 75 \text{ \AA}$ at four different temperatures: 1350, 1500, 1650, and 1800 K. This temperature range is similar to those commonly observed in flow reactors [167] or laminar flames [166]. Every configuration was simulated several times using various random initializations. The simulations were carried out using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) software [135], utilizing the ReaxFF potential [51, 98] at a timestep of 0.25 fs. The velocity-Verlet algorithm [100] and the Nose-Hoover thermostat [102] were used in the simulations, which were run under the NVT ensemble (constant number of particles, volume, and temperature).

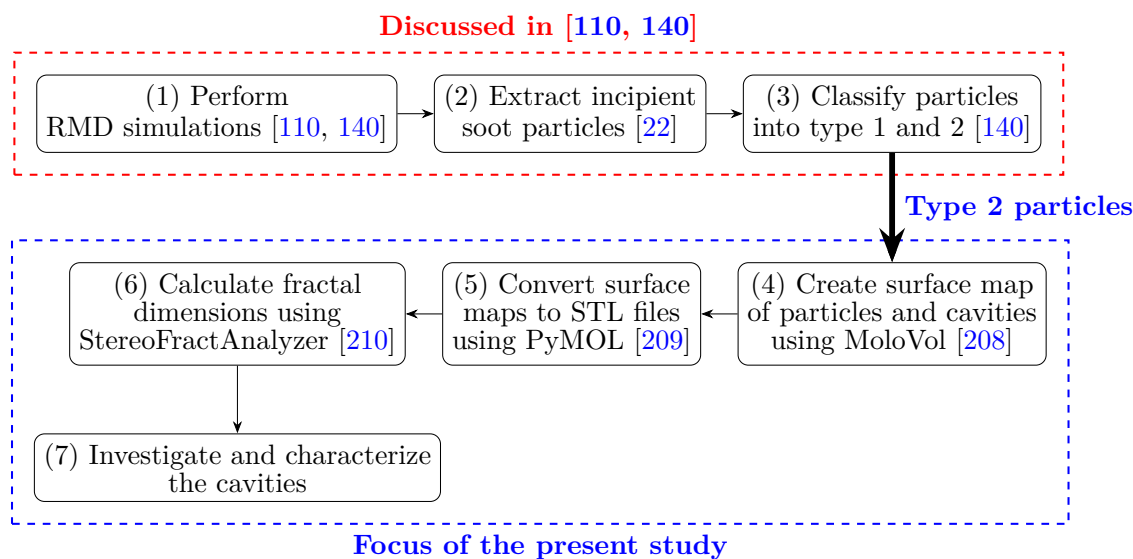


Figure 6.2: Overview of the workflow utilized in this work.

From each simulation, large molecular clusters were identified as incipient soot particles (Block 2). The incipient particles obtained from these simulations were validated

against experimental observation in our previous works [110, 140]. These particles were then classified using t-SNE [139] and k-means clustering [138], resulting in two distinct types: Type 1 and Type 2 (Block 3) [110]. Type 1 particles resemble the very early stage of soot formation where the particles are still in the process of growth by mass accumulation and surface reactions. Type 2 particles are in an advanced stage with clearly defined surface and internal structures. In this work, we only looked at the Type 2 particles because the internal structure of Type 1 particles are not yet well-developed.

6.2.2 Extraction of surface and pore information

From the RMD simulations, 2654 Type 2 incipient particles at four distinct temperatures were extracted. We used MoloVol [208] to extract the surface map of these particles' exterior surfaces and interior cavities (block (4) of Fig. 6.2). A single probe method is used to capture the surface area and surface map of the external surface while the two-probe method is used to capture both the external surface and internal cavities together. The small probe radius is set equal to the van der Waals radius of nitrogen (1.66 Å) [211] and the large probe radius is set to an arbitrary large value of 5.0 Å as discussed in [208]. The accuracy of the external surface area calculation by MoloVol was further verified with MSMS [136].

Surface maps generated using MoloVol contained details of external and internal surfaces and cavities. The surface map is then converted to a 3D surface mesh (STL format) using PyMOL [209] (block (5) of Fig. 6.2). This conversion is necessary to calculate the surface and volume fractal dimensions in the later steps. Two STL files are generated for each particle: one for the external surface and one for the cavities. An inhouse tool called StereoFractAnalyzer [210], which calculates the fractal dimension of the surface and volume of the particles and cavities using the box-counting method [212, 172, 213, 214], is used to calculate the surface and volume fractal dimensions from the STL files (block (6) of Fig. 6.2).

Additional analysis is done to explore and characterize the pores or cavities inside each particle (block (7) of Fig. 6.2). Fig. 6.3 shows the rendering of one example particle,

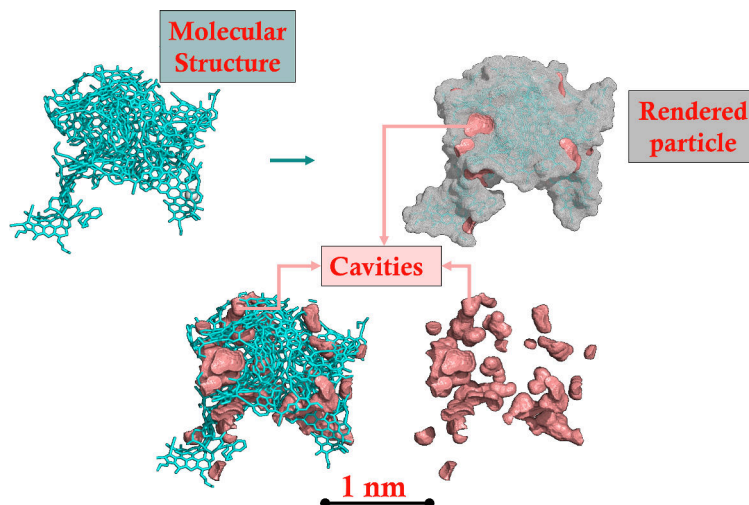


Figure 6.3: One example particle obtained from RMD simulations along with examples of surface renderings and identified cavities.

where the external surface and cavities are shown together. The cavities are colored according to their classification: isolated cavities (green), pockets (brown), and tunnels (blue). To the best of the authors' knowledge, this is the first study to offer such detailed information regarding incipient soot particle surfaces and cavities.

In this work, the words *pore*, *void*, and *cavity* are used interchangeably to denote an empty space inside a particle.

6.3 Results and Discussion

We start the results section with the bulk morphological properties of the particles, followed by a detailed analysis of the distribution of cavities and their properties. Finally, we explore the correlations between surface and volume features which can potentially be useful in engineering-scale models. For the purpose of summarizing the results, we have used histograms with probability density in conjunction with box and whisker plots to show the distribution of the data. The box represents the interquartile range (IQR) of the data, while the whiskers represent the range of the data. The line inside the box represents the median of the data. The statistics are reported in the form of **mean \pm standard error of mean (SEM)** [215] for each quantity. Standard deviation (SD) is also reported in the

summary table (Table 6.1) for completeness.

For all the results presented in this work, the authors didn't observe any significant dependency on the temperature. Therefore, the results are presented for all the temperature cases combined. However, the temperature dependency is discussed in the text where necessary and reported in the Supplementary Material.

6.3.1 Bulk morphological properties of incipient particles

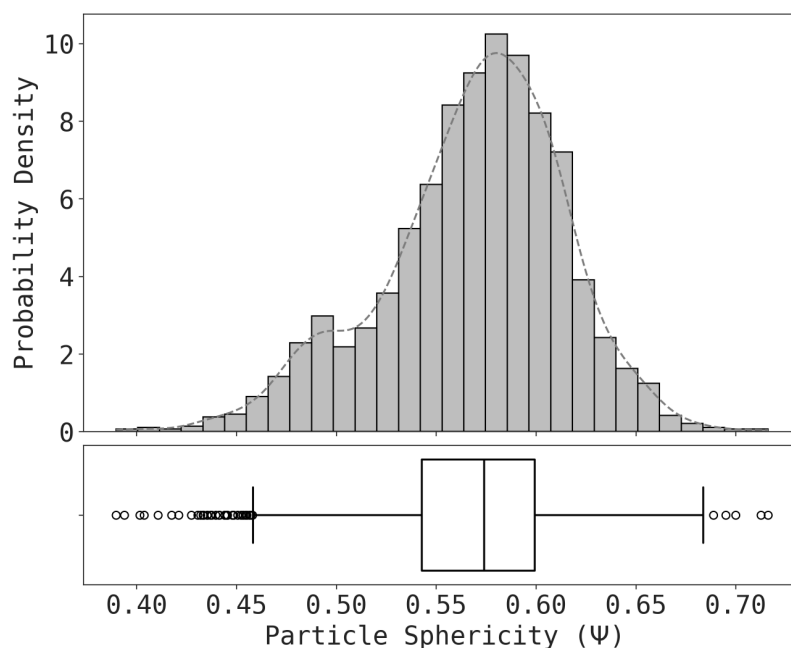


Figure 6.4: Distribution of sphericity (Ψ) of incipient particles.

Usually the soot primary particles are assumed to be perfectly spherical in engineering scale models [12, 28]. However, as observed in atomic force microscopy (AFM) [125] and high resolution transmission electron microscopy (HRTEM) [216, 217], the incipient soot particles are not perfectly spherical. The shape of the particle directly influences its reactivity particularly in oxidation processes [218, 176]. How closely the shape of a particle resembles a perfect sphere can be measured by sphericity (Ψ). Sphericity(Ψ) can also act as an indicator of the stage of maturity of soot particle [219]. In this work,

Eqn. B.1 is used to calculate the sphericity of the particles.

The sphericity of the incipient particles is found to be 0.57 ± 0.0008 . In contrast, the sphericity of a perfectly spherical non-porous particle is unity. The low value indicates a significantly high surface area available compared to the total volume. The average sphericity of 0.57 derived here by RMD is in reasonable agreement with the 0.63 ± 0.08 measured for larger young soot particles from premixed ethylene flames [220]. It must be noted here that presence of large number of micropores leads to a large surface area [194]. Additionally, since the surface area is calculated on an atomic level using an atomic probe, even small roughness on the surface can add to the overall surface area. Hence, even if the particle “looks” spherical, its sphericity can be significantly lower than unity. The sphericity distribution of the particles is shown in Fig. 6.4 using a histogram and a box-plot. The probability density of sphericity is also presented in Fig. 6.4 using the black-dashed line. The sphericity remains almost constant at all the temperatures (see Fig. C.1 in Supplementary Materials).

In practice, it is often difficult to measure the surface area of 3D microstructures. Therefore, a 2D equivalent quantity called circularity (σ) is also used in literature [221]. The definition of circularity (σ) is given in Eqn. B.2. In this study, because of asymmetry and irregularity in 3D shapes of particles, an average circularity ($\bar{\sigma}$) of each incipient particle is calculated by taking the projection of the 3D particle on to 10 evenly spaced planes in spherical coordinates as shown in Eqn. B.3. Figure 6.5 reports the average circularity of the incipient particles at four different temperatures including the standard deviation as error bars for each particle. Violin plots for the average circularity ($\bar{\sigma}$) of the particles at different temperatures are also presented in Fig. 6.5. The violin plots show that distribution is not affected by temperature (also see Fig. C.2 in the Supplementary Materials). From Fig. 6.5, the average circularity of the analyzed particle is found to be 0.66 ± 0.001 .

The porosity (Φ) of the particles is calculated as the ratio of the total pore volume (V_p) inside a particle to the bulk volume (V_B) of the particle (Eqn. B.4). Figure 6.6(a) depicts the distribution of porosity of the incipient particles using histogram with probability density and box plot. The observed porosity for the incipient particles in this

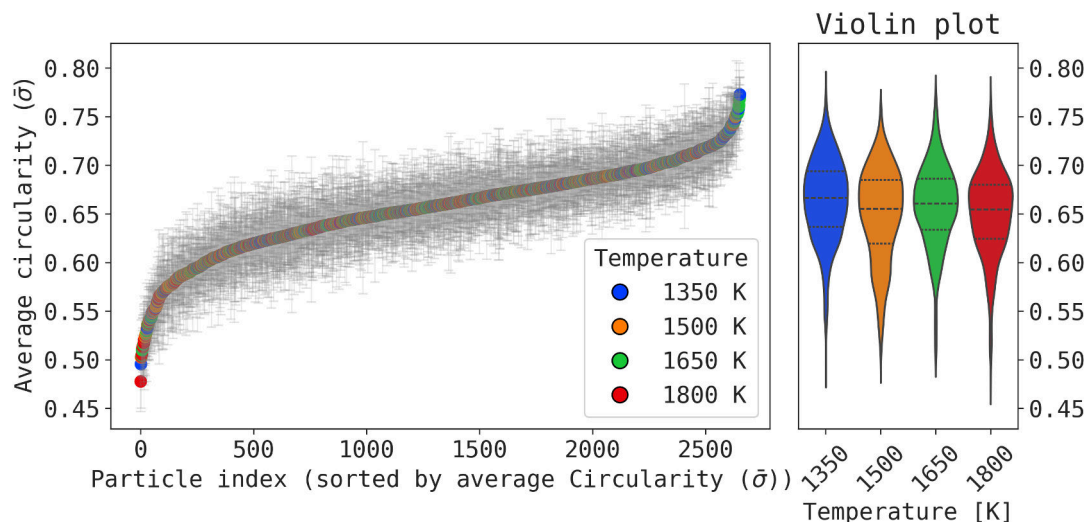
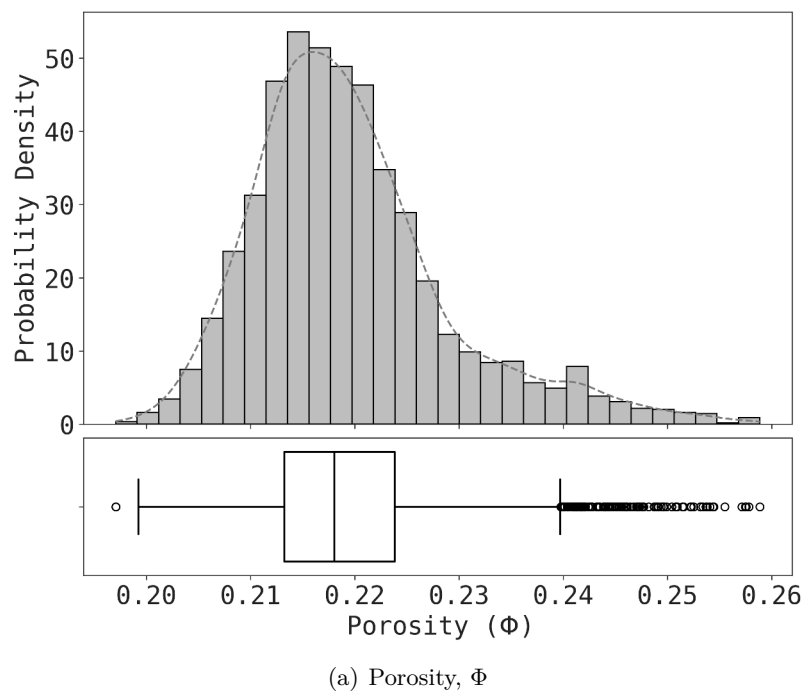
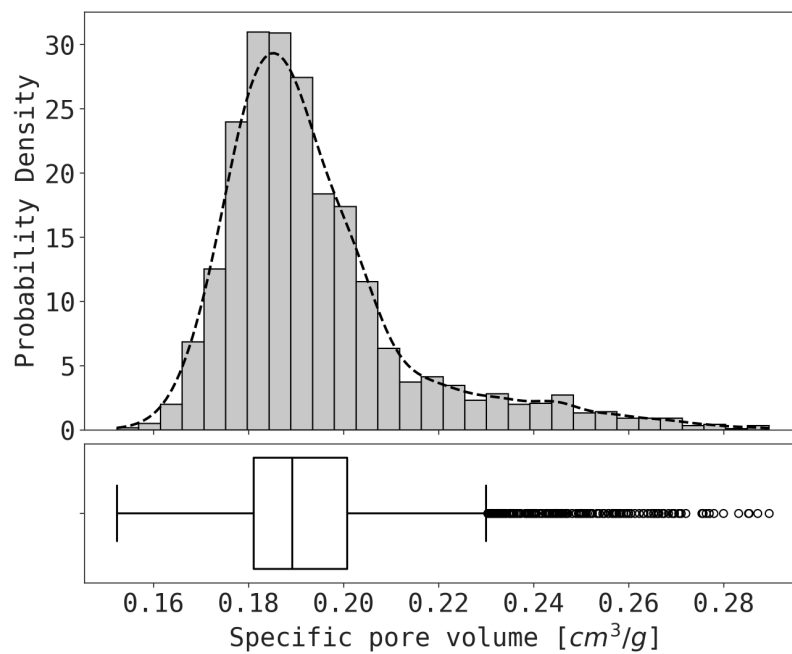


Figure 6.5: Average circularity ($\bar{\sigma}$) of incipient particles at four different temperatures.

study is found to be 0.22 ± 0.0002 , which is in good agreement with the measured soot porosity of 0.26 [222]. Temperature was found to have a very small effect on porosity (Fig. C.3 in Supplementary Materials).

Another way to look at porosity is to calculate the specific pore volume of the particles, which is a measure of available pore volume per unit mass of a particle. The specific pore volume obtained in this study is reported in Fig. 6.6(b). The average specific pore volume of the particles explored in this study is found to be $0.19 \pm 0.0004 \text{ cm}^3/\text{g}$. The reported values of specific pore volume in the literature vary widely depending on the sampling procedure, levels of maturity, and nature of pores. For example, in mature soot particles, Rockne et al. [200] reported a value of 0.004–0.08 cm^3/g for mesopores and 0.0009–0.013 cm^3/g for micropores. Whereas, Tripathi et al. [194] reported that particles with higher number of micropores can achieve specific pore volume as high as 0.7–2.3 cm^3/g . In our case, the particles are not fully mature and contain significant numbers of micropores, leading to a high value of porosity and specific pore volume. No significant temperature dependence was observed in specific pore volume (see Fig. C.4 in Supplementary Materials).

(a) Porosity, Φ 

(b) Specific pore volume

Figure 6.6: Distribution (a) porosity, Φ and (b) specific pore volume [cm^3/g] of incipient particles.

With the increase in porosity, the particles offer more accessible surface area. Specific surface area (SSA) is a measure of surface area available per unit mass. The specific surface area distribution of the incipient soot particles obtained in the current study is presented in Fig. 6.7. Contribution of isolated cavities is ignored in this calculation since this area is not accessible from the outside of the incipient particle. The average specific surface area of the particles is found to be $2652.36 \pm 7.39 \text{ m}^2/\text{g}$. This value is significantly higher than the specific surface area reported in contemporary literatures. This is potentially due to several reasons. The values reported in the literature, are mostly engine-out soot particles, which are mature, larger, and have gone through some interaction with the environment. For example, Rockne et al. [200] reported the specific surface area of soot particles for mature soot particles from different combustion sources to be in the range of 1 to $85 \text{ m}^2/\text{g}$. Ouf et al. [223] showed that, the specific surface area increases as the particle size decreases. Therefore, incipient particles explored in our study, which are very small and in the very early stage of formation, are expected to have higher specific surface area. Also, the presence of numerous micropores in the particles, which is the case in the particles in this work, can lead to higher surface area, which in turn lead to higher specific surface area [194]. Additionally, high values of SSA has been reported in high surface area carbon materials such as activated carbons derived from biomass (SSAs of up to $3386 \text{ m}^2/\text{g}$ [224]), metal-organic-framework(MOF)-derived carbons (up to $2872 \text{ m}^2/\text{g}$ [225]), and hypercrosslinked polymer-derived carbons (up to $4300 \text{ m}^2/\text{g}$ [226]). These examples illustrate that materials with high porosity can achieve exceptionally high SSAs, supporting the elevated SSA observed in our incipient soot particles.

The theoretical specific surface area ($SSA_{\text{Theoretical}}$) for the external surface of a spherical particle can be calculated using the volume-equivalent diameter (d_V) of the particle and its bulk density (ρ_b) as shown in Eqn. B.5. The average theoretical specific surface area ($\overline{SSA}_{\text{Theoretical}}$) of the particles is found to be $1493.77 \text{ m}^2/\text{g}$ and is shown in Fig. 6.7 as a red dashed vertical line. Equation B.5 describes the theoretical minimum external specific surface area of a perfect smooth sphere with no surface cavities. Notably, the measured specific surface areas for our incipient soot particles exceed this theoretical

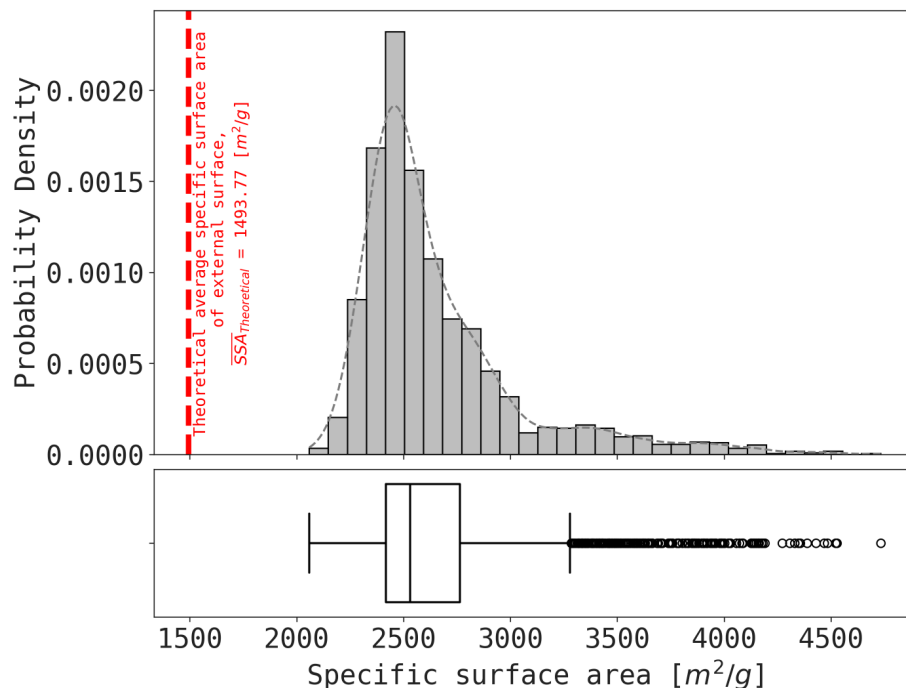


Figure 6.7: Distribution of specific surface area [m^2/g] of incipient particles.

minimum by almost a factor of two. This highlights the significant contribution of porosity and complex surface morphology to the total accessible surface area. Additionally, Eqn. B.5 shows that SSA is inversely proportional to the volume-equivalent diameter (d_V), meaning that smaller particles inherently possess higher external surface area per unit mass. Since our incipient particles are in the very early stage of formation and small (diameters < 5 nm), their baseline external surface area (A_{Ext}) is already high. However, the fact that our measured SSAs surpass the theoretical value clearly illustrates that besides direct surface effects, porosity further elevates the overall accessible surface area. No specific trend is observed with temperature (Supplementary Materials, Fig. C.5) for specific surface area.

To further elucidate the influence of internal porosity on the overall accessible surface area, we decompose the total SSA into contributions from distinct cavity types as well as the exterior surface area (ESA), which represents the surface area calculated without any cavity contributions (note the difference between *external* and *exterior* surface area: exterior surface does not include any contributions from cavities, whereas external surface includes the exposed surface of tunnels and pockets). The SSA values presented in Fig. 6.7

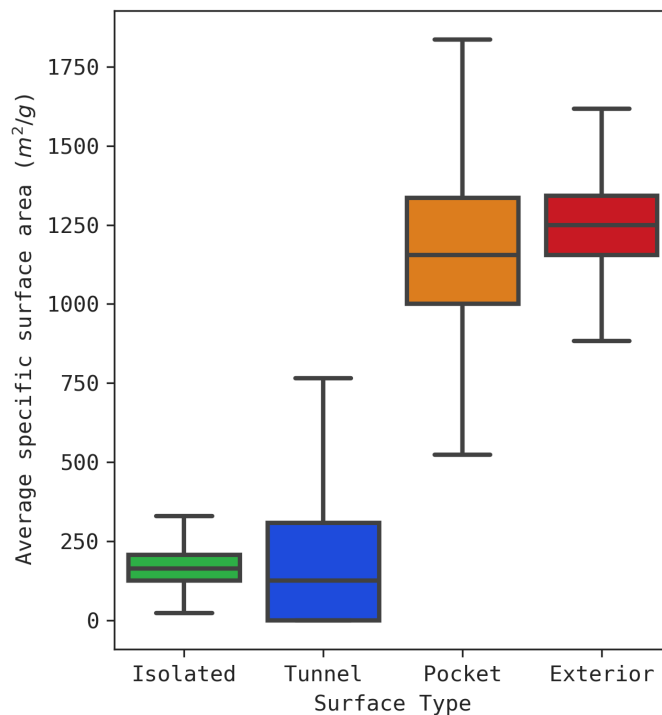


Figure 6.8: Contribution of different cavities to specific surface area [m^2/g] of incipient particles.

include contributions from tunnel and pocket cavities. Figure 6.8 separately illustrates the SSA for each cavity type (tunnel, pocket, isolated) alongside the ESA. The comparison between the ESA and the cavity contributions clearly demonstrates that the enhanced SSA observed in our incipient soot particles is largely due to the presence of pockets. The average specific surface areas are: $172.36 \pm 1.34 \text{ m}^2/\text{g}$ for isolated cavities, $212.12 \pm 5.43 \text{ m}^2/\text{g}$ for tunnel cavities, $1190.94 \pm 5.71 \text{ m}^2/\text{g}$ for pocket cavities, and $1249.29 \pm 2.84 \text{ m}^2/\text{g}$ for the exterior surface. This decomposition not only emphasizes the significant role played by tunnel and pocket cavities in increasing the accessible surface area, but also provides insight into the interplay between particle size and internal morphology. No specific trend is observed with temperature (Supplementary Materials, Fig. C.6).

6.3.2 Bulk morphological properties of pores

As discussed earlier, depending on the access to the surface of incipient particles, the pores are classified into three groups: tunnels, pockets, and isolated cavities. These classifications provide insight into the way different cavities interact with the environment. Following Eqn. B.1, the sphericity of each cavity can also be calculated. To compare the statistics of all the particles together, we calculated the average pore sphericity ($\bar{\psi}$) of different classes of cavities for each particle using Eqn. B.6.

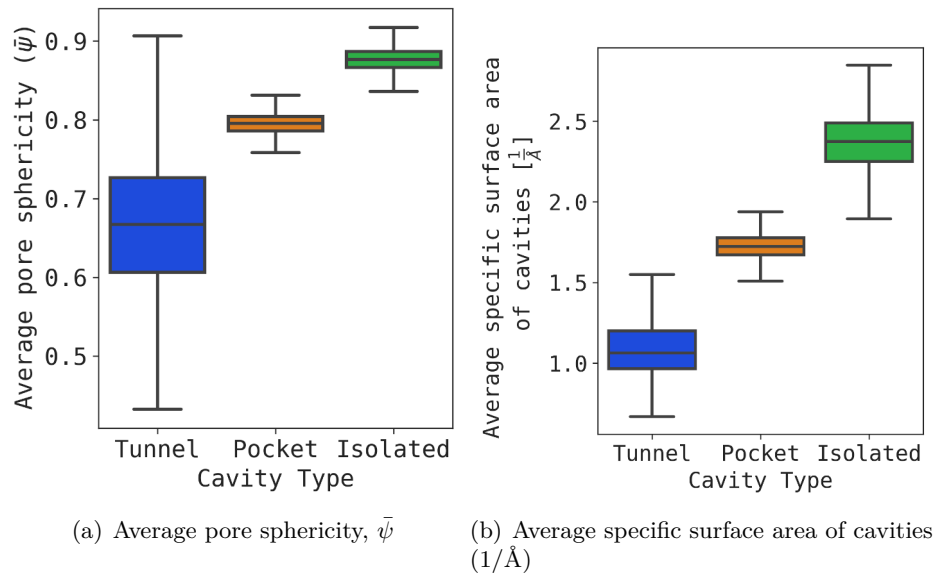


Figure 6.9: (a) Average pore sphericity, $\bar{\psi}$ and (b) average specific surface area of cavities [$1/\text{\AA}$] in incipient particles.

Figure 6.9(a) shows clear distinction among the average pore sphericity of different cavities. The average pore sphericity of tunnels, pockets, and isolated cavities are found to be 0.67 ± 0.002 , 0.80 ± 0.0002 , and 0.88 ± 0.00003 , respectively. This indicates that, the isolated cavities are more spherical than the pockets and tunnels, which are more elongated due to their openings to the external surface.

Since cavities do not have any mass, the surface area of different cavities are quantified using a volume-based specific surface area which is defined as the amount of

cavity surface area per unit volume of the cavity. Thus, the unit for this volume-based surface area is \AA^{-1} . Figure 6.9(b) depicts the average volume-based specific surface area of different cavities. Unlike Fig. 6.8, where SSA is computed by aggregating the contributions of all cavities across all particles, here the volume-based SSA is first averaged over all cavities within each individual particle and then further averaged across all particles. This procedure preserves information regarding the typical size or footprint of each cavity type within incipient soot particles. As observed with pore sphericity, different cavities shows clear distinction. The average specific surface area of tunnels, pockets, and isolated cavities are found to be 1.1 ± 0.005 , 1.72 ± 0.002 , and $2.38 \pm 0.03 \text{\AA}^{-1}$, respectively. Since isolated cavities are more spherical than the other two types, they also have the highest specific surface area than pockets and tunnels. Both average pore sphericity ($\bar{\psi}$) and specific surface area of the cavities are found to be independent of temperature (See Supplementary Materials Fig. C.7).

6.3.3 Fractal characteristics of incipient particles and cavities

The fractal dimension is an important metric for characterizing the complexity and self-similarity of a structure. In the context of soot particles, usually the fractal dimension is reported for the soot aggregates, which are formed by the agglomeration of primary particles. This fractal dimension, also known as aggregate fractal dimension (D_f) is calculated using a statistical mass-fractal relationships [192, 227]. In this work, we are not investigating this definition of fractal dimension of an aggregate, rather we are focusing on the fractal dimensions of an individual primary particle. While the fractal dimension of soot aggregates has been extensively studied, the fractal dimensions of primary particles is less commonly reported. Understanding the fractal nature of individual primary particle morphology provides valuable insights into the early stages of soot formation, which can inform more accurate models of soot reactivity and growth. In this work, we analyzed surface and volume fractal dimension (D_S and D_V) of incipient primary particles.

The surface fractal dimension (D_S) of the incipient particle describes how the particle's surface scales as its size increases. A smooth surface would have a surface fractal

dimension value of 2. A higher surface fractal dimension indicates a rougher, more irregular surface, which increases the available surface area for chemical interactions. The surface fractal dimension of the particles is found to be 2.22 ± 0.001 , which is in excellent agreement with the $D_S = 2.25 \pm 0.09$ measured for carbon black [228, 229] and indicating a highly irregular and rough surface of the incipient particles. The distribution of the surface fractal dimension of the incipient particles is shown in Fig. 6.10(a). The surface fractal dimension of the particles is found to be independent of temperature (Supplementary Materials, Fig. C.8).

The volume fractal dimension (D_V) can be thought of as the ability to self-similarly fill the space by a solid [230]. With a higher value of volume fractal dimension (D_V), the particle is more compact and less porous. The distribution of volume fractal dimension of the incipient particles is shown in Fig. 6.10(b). The volume fractal dimension of the particles is found to be 2.53 ± 0.0006 in the present study. This indicates that the incipient particles are highly porous and have a complex internal structure. The volume fractal dimension of the particles is found to be independent of temperature (Supplementary Materials, Fig. C.9). It should be noted here that previously Mukut et al. [110, 140] reported an atomic fractal dimension of these incipient primary particles, which was calculated as a mass-based fractal dimension (i.e., ability of a particle to self-similarly contribute to particle mass) for primary particle.

The fractal characteristics of the cavities provide insights into the complexity of the internal pore networks. These networks influence the ability of the particles to adsorb gases and engage in chemical reactions. A lower surface fractal dimension for the cavities (D_{SC}) suggests a smoother internal surface compared to the external particle surface, which may affect the overall reactivity and interaction with the environment. The surface fractal dimension (D_{SC}) and volume fractal dimension (D_{VC}) of different cavities are calculated using the same method as for the particles. As shown in Fig. 6.11(a), the surface fractal dimension of the cavities is found to be 2.17 ± 0.006 . In contrast to the particle surface, the cavity surface is slightly more regular and less rough. This is also reflected in the higher average cavity sphericity ($\bar{\psi}$) value shown in Fig. 6.9(a).

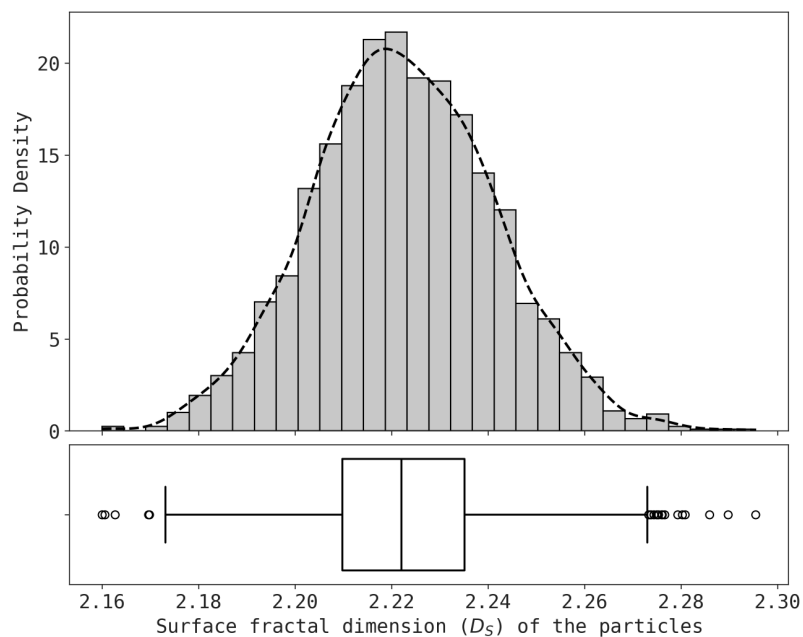
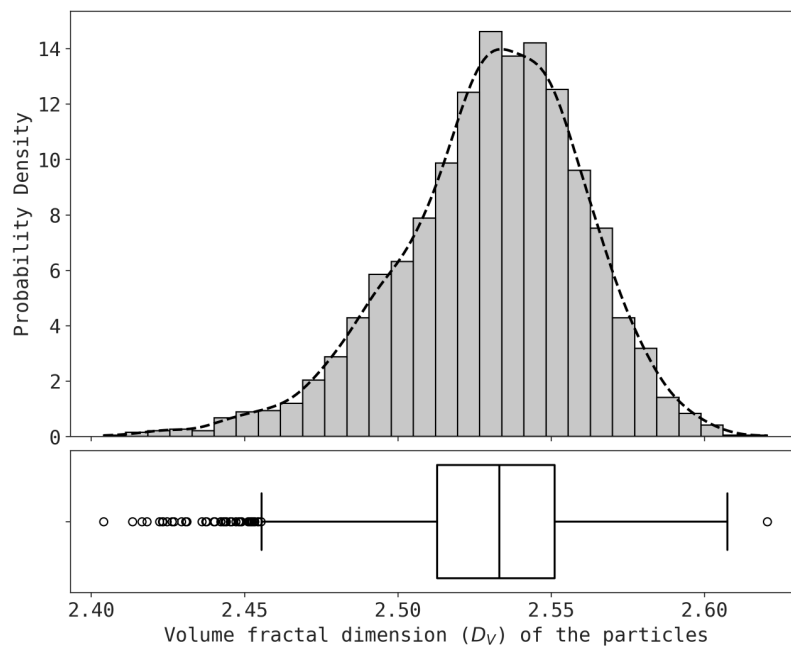
(a) Surface fractal dimension (D_S) of the particles(b) Volume fractal dimension (D_V) of the particles

Figure 6.10: Distribution of (a) surface fractal dimension (D_S), and (b) volume fractal dimension (D_V) of incipient particles.

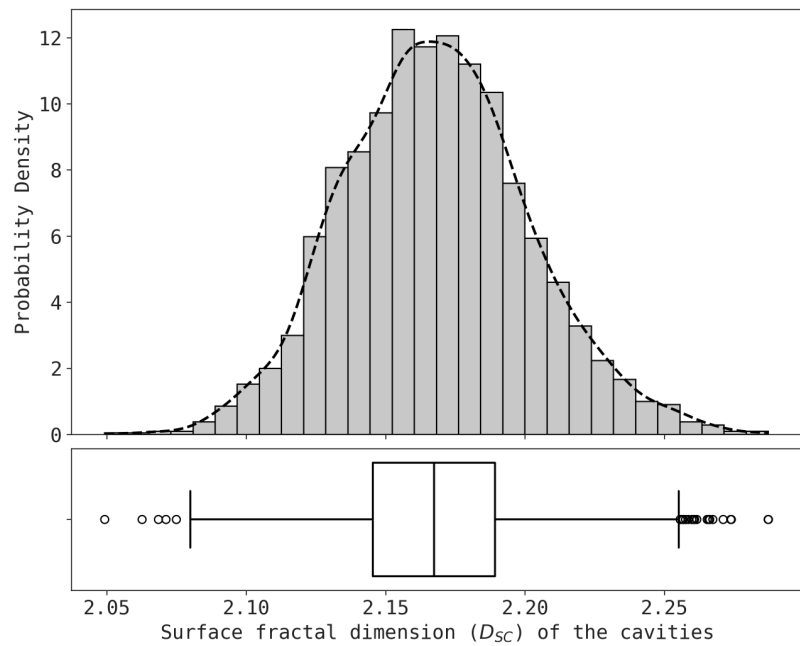
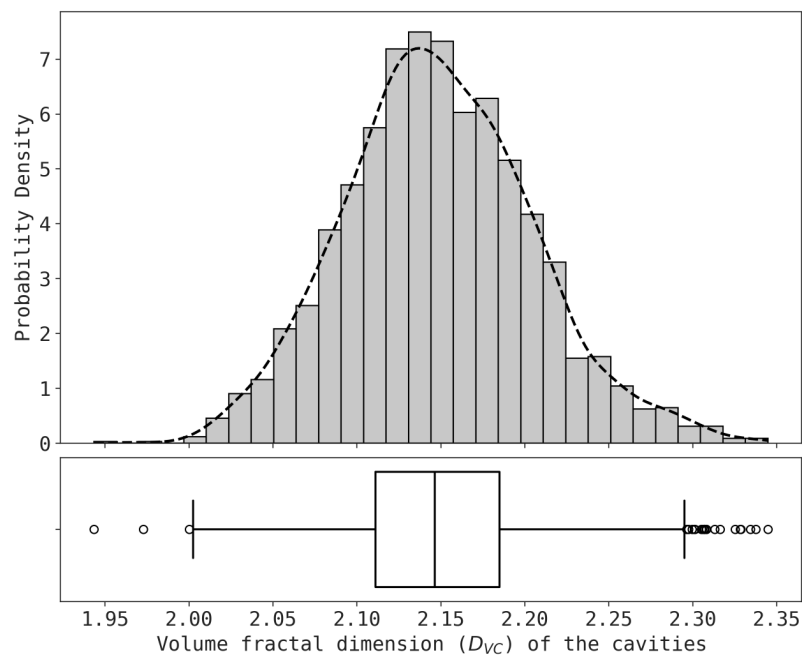
(a) Surface fractal dimension (D_{SC}) of the cavities(b) Volume fractal dimension (D_{VC}) of the cavities

Figure 6.11: Distribution of (a) surface fractal dimension (D_{SC}), and (b) volume fractal dimension (D_{VC}) of the cavities within incipient particles.

The volume fractal dimension of cavities (D_{VC}) is commonly referred to as the pore fractal dimension. This is a measure of complexity of internal pore network inside soot particles which quantify how the empty space within a primary particle is self-similarly distributed within the particle [231]. It has been reported in the literature that the pores of soot and carbon blacks form a fractal-like network with pore fractal dimension of between 2–2.5 as measured using small angle X-ray scattering (SAXS) [232], Ar, CO₂, and N₂ adsorption studies [233, 234], and numerical studies of soot oxidation [142]. The volume fractal dimension (D_{VC}) of the cavities in this work is found to be 2.15 ± 0.001 , which is in excellent agreement with the reported values. The distribution of volume fractal dimension of cavities (D_{VC}) observed in the present study is shown in Fig. 6.11(b). Both the surface and volume fractal dimensions of the cavities are found to be independent of temperature (Supplementary Materials, Figs. C.10 and C.11).

6.3.4 Pore size distribution inside incipient particles

Understanding the distribution of pore sizes within soot particles can provide insights into their formation mechanisms and behavior in different combustion environments. The cavity size in this work is represented by the volume-equivalent pore diameter (d_p), which represents the diameter of a sphere with an equivalent volume to the cavity. The distribution of individual cavity sizes (d_p) in the entire population is shown in Fig. 6.12. The distribution is log-normal, which indicates that although there are a few larger cavities, most cavities are small: which is typical for porous structures. Additionally, it can be seen that all cavities identified in this study are micropores (≤ 2 nm), with the majority measuring less than 1 nm. This is expected since all the incipient particles investigated here are in the very early stage of formation and have a diameter smaller than 5 nm. This essentially makes micropores the primary focus of this analysis. As has been discussed earlier, such large presence of micropores increases the surface area leading to high specific surface area, high porosity, and low sphericity values observed in the incipient particles, which is typical of soot particles in the early stages of formation [194].

This distribution of cavity size (i.e., d_p) can also be analyzed by calculating the average values for cavity sizes ($\overline{d_p}$) inside each particle. The tunnel cavities have the largest average size (8.92 ± 0.05 Å), followed by pockets (4.85 ± 0.006 Å) and isolated (3.08 ± 0.005 Å) cavities. The size distribution of average cavity sizes ($\overline{d_p}$) (shown in Supplementary Materials, Fig. C.12) is a Gaussian distribution, as is expected, since averaging tends to smooth out variability.

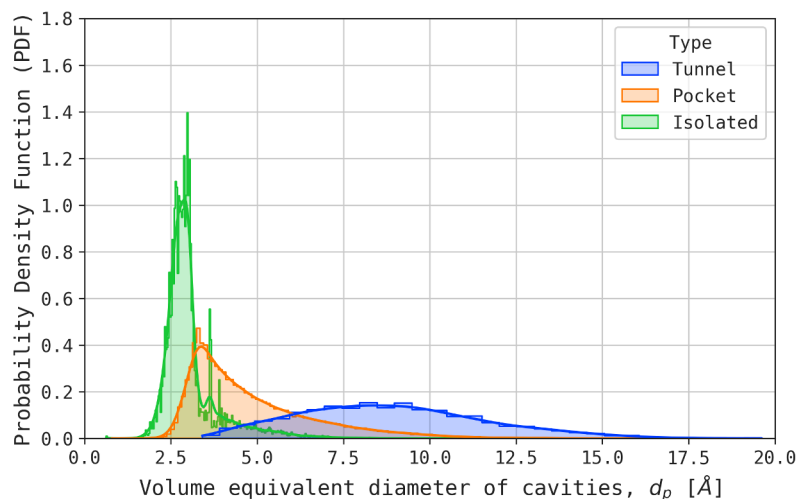


Figure 6.12: Cavity size (d_p) distribution of all cavities across all the particles.

To get a more complete understanding of the pore network of the particles, the average fraction of void volume occupied by cavities of different sizes within incipient particles is calculated and shown in Fig. 6.13. Figure 6.13 shows that the pocket cavities occupy the largest fraction of void volume, followed by isolated cavities and tunnels. Although tunnels are larger on average (Supplementary Materials, Fig. C.12), their lower abundance in the incipient particles results in them occupying a smaller fraction of the overall void volume. In contrast, the higher abundance of pocket cavities allows them to occupy a larger portion of the void space, despite their smaller size. The pockets and tunnels present in the bulk particles are critical for the oxidation of soot and carbon black at low temperatures (e.g. 550 °C) [234]. At these conditions, O₂ can diffuse through these

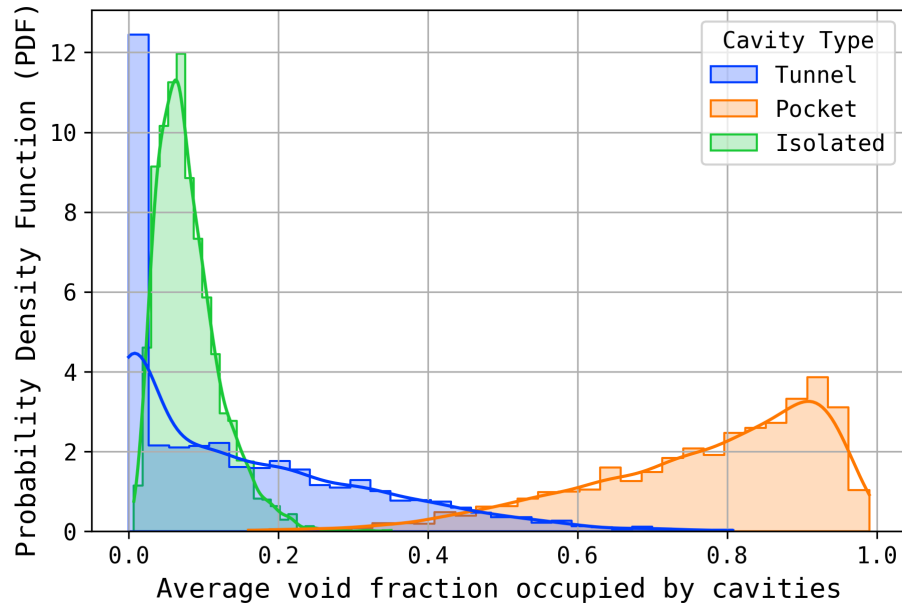


Figure 6.13: Distribution of average fraction of void volume occupied by cavities of different sizes within incipient particles.

pockets and tunnels and internally oxidise soot and carbon black nanoparticles, resulting in hollow spheres [234, 142].

In the porous material literature, the pore size distribution (PSD) is often reported in terms of cumulative void volume (per unit mass) as a function of pore size [235, 236]. Cumulative void volume ($V_C(d_p)$) refers to the total volume of empty spaces accumulated within a particle up to specific cavity sizes (d_p) and is calculated using Eqn. B.7.

Each particle will produce a different pore size distribution. To present the data in a compact form, we have calculated the average PSD from all the particles by dividing the diameter range into discrete size bins and calculating the cumulative volume of cavities in each bin for all the particles. The average PSD of different cavities is shown in Fig. 6.14. The shaded region in Fig. 6.14 represents the standard error of the mean (SEM). The average PSD of isolated cavities is shifted towards smaller pore sizes, while that of pocket and tunnel cavities is shifted towards larger pore sizes. For the sake of completeness, the un-averaged PSD of the entire population of cavities is also provided in Fig. C.13.

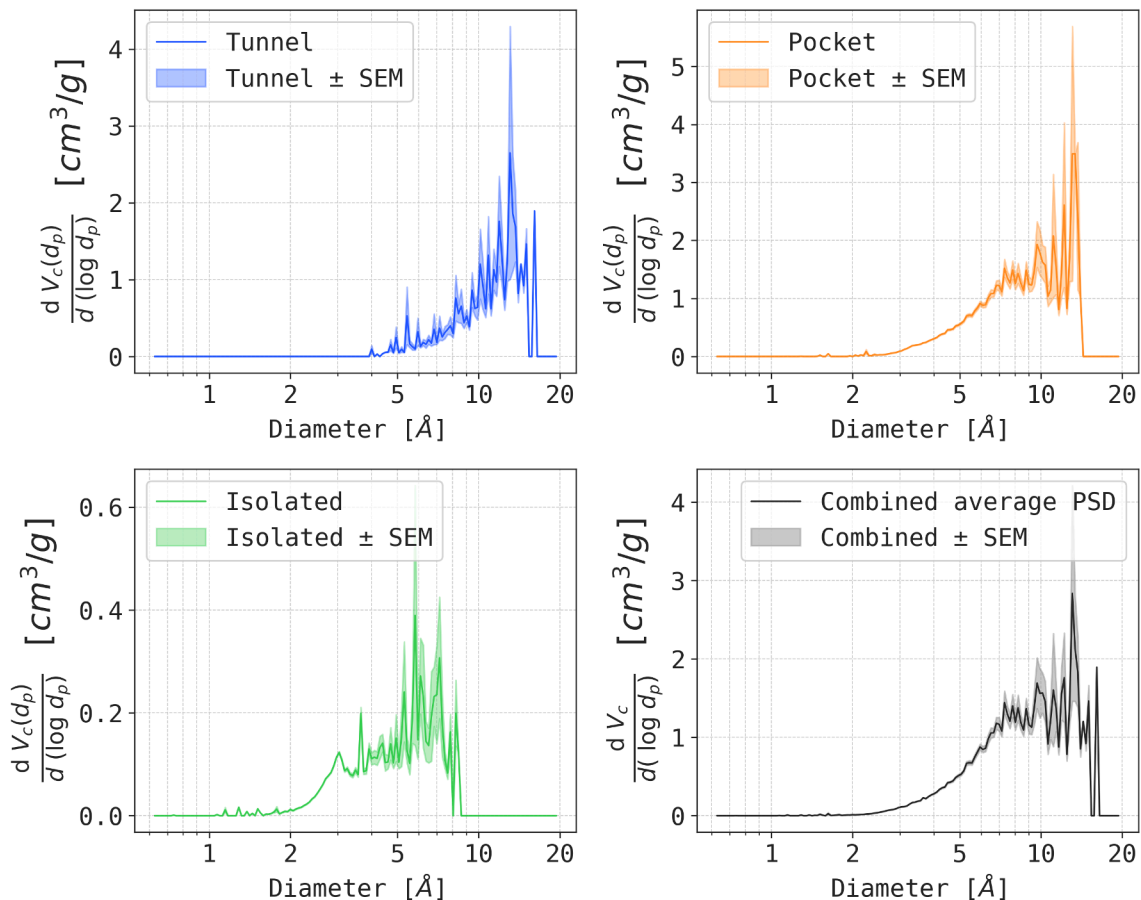


Figure 6.14: Average pore size distribution (PSD) of different cavities within incipient soot primary particles.

6.3.5 Correlations for pore area and pore volume

Engineering scale soot models usually carry information about the volume of soot particles [28, 14], which does not, on its own, include any information on the pores. If some correlations can be drawn between the volume and surface area of pores and the particle volume, some critical information on the pore network within a particle can be extracted, which can improve the modeling of chemical reactivities of soot particles. A good positive correlation ($R^2 \approx 0.70$) between total pore volume (V_p) and total pore surface area (A_p) was found (see Supplementary Materials, Fig. C.14). No significant temperature dependence is observed in this correlation.

Similarly, the particle's material volume (V) and total pore volume (V_p) (Supplementary Material, Fig. C.15), shows a positive and generally linear correlation. However, the correlation shows a weaker correlation coefficient ($R^2 \approx 0.60$) than the correlation between pore volume (V_p) and pore surface area (A_p). This suggests that while particle size is a significant factor in determining pore volume (V_p), other factors may also play a role. A slight change in correlation is also observed with process temperature, as shown in the Supplementary Materials (Fig. C.15).

Finally, individual pore volume (V_{pi}) and pore surface area (A_{pi}) shows a strong positive and linear correlation, with high correlation coefficients ($R^2 \approx 0.98$, depicted in Supplementary Materials, Fig. C.16), indicating that larger cavities have proportionally greater surface areas. This correlation is independent of cavity types, highlighting that cavity size is a dominant factor in determining pore surface area, regardless of the type of cavity. The equations for the linear correlations (for all temperature combined) are given in Correlation Set 6.1. An exponential and a quadratic fit were also explored, but no significant improvement of correlation coefficient were observed. These correlations provide valuable insights into how particle morphology and pore structure evolve, with implications for modeling the surface area and reactivity of soot particles in different environmental and combustion scenarios.

Correlation Set 6.1: Temperature-independent equations of curves fitted to the pore volume, pore surface area, and particle volume (data shown in Supplementary Materials, Figs. C.14, C.15, and C.16). Here, V is the particle’s material volume [\AA^3], V_p is the total pore volume within an incipient particle [\AA^3], A_p is the total pore surface area [\AA^2], V_{pi} is the volume of an individual pore [\AA^3], and A_{pi} is the surface area of an individual pore [\AA^2].

$$A_p = 1.39V_p - 335.53, \quad R^2 = 0.70, \quad (6.1)$$

$$V_p = 0.16V + 277.52, \quad R^2 = 0.57, \quad (6.2)$$

$$A_{pi} = 0.89V_{pi} + 30.20, \quad R^2 = 0.98 \quad (6.3)$$

6.3.6 Summary statistics

The summary statistics is presented in Table 6.1, which shows the mean, standard error of the mean (SEM) and standard deviation (SD) of the surface and pore features of incipient particles. The table provides a comprehensive overview of the key properties of the particles, including sphericity, porosity, specific pore volume, specific surface area, average pore sphericity, average specific surface area of different cavities, surface fractal dimension, volume fractal dimension, and average volume equivalent diameter of different cavities. These statistics can be used to develop more accurate models of soot formation and reactivity, with implications for environmental and combustion studies.

6.4 Conclusion

We present a novel strategy to extract detailed information of the external surface and internal pores from the atomic structures of the incipient soot particles obtained from reactive molecular dynamics (RMD) simulations. By integrating RMD with atomistic-scale pore and surface analysis, this work offers a transferable framework for characterizing nanostructured carbon materials ranging from soot and carbon black to biochars and

Table 6.1: Mean, Standard Error of the Mean (SEM), and Standard Deviation (SD) of the surface and pore features of incipient particles.

Property	Unit	Statistics		
		Mean	± SEM	SD
Particle sphericity, Ψ	–	0.57	± 0.0009	0.046
Particle average circularity, $\bar{\sigma}$	–	0.655	± 0.001	0.044
Porosity, Φ	–	0.22	± 0.0002	0.01
Specific pore volume	cm ³ /g	0.194	± 0.0004	0.02
Specific surface area of particle	m ² /g	2652.36	± 7.388	380.62
Average pore sphericity for tunnels cavities, ψ_{Tunnel}	–	0.665	± 0.002	0.088
Average pore sphericity for pockets cavities, ψ_{Pocket}	–	0.795	± 0.0002	0.014
Average pore sphericity for isolated cavities $\psi_{I\,isolated}$	–	0.877	± 0.0003	0.016
Average specific surface area of tunnel cavities (mass-based)	m ² /g	212.12	± 5.43	279.63
Average specific surface area of pocket cavities (mass-based)		1190.94	± 5.71	294.17
Average specific surface area of isolated cavities (mass-based)		172.36	± 1.34	69.07
Average specific surface area of tunnel cavities (volume-based)	$\frac{1}{\text{\AA}}$	1.096	± 0.005	0.196
Average specific surface area of pocket cavities (volume-based)		1.723	± 0.002	0.084
Average specific surface area of isolated cavities (volume-based)		2.374	± 0.004	0.193
Surface fractal dimension of particle, D_S	–	2.223	± 0.0003	0.019
Volume fractal dimension of particle, D_V	–	2.531	± 0.001	0.03
Surface fractal dimension of cavity, D_{SC}	–	2.168	± 0.001	0.033
Volume (Pore) fractal dimension of cavity, D_{VC}	–	2.149	± 0.001	0.056
Average volume equivalent diameter of tunnels cavities	Å	8.918	± 0.055	2.359
Average volume equivalent diameter of pockets cavities		4.854	± 0.006	0.329
Average volume equivalent diameter of isolated cavities		3.078	± 0.005	0.262

porous carbons used in catalysis and adsorption. The detailed quantification of fractal dimensions, porosity, and surface area provides valuable descriptors for interfacial behavior, with implications for modeling gas-solid interactions, pollutant adsorption, and transport phenomena in complex carbonaceous systems.

In this study, we conducted an in-depth investigation into the surface and pore characteristics of incipient soot particles obtained from previously validated RMD simulations of acetylene pyrolysis at 1350, 1500, 1650, and 1800 K [110, 140]. Features explored in this work include surface fractal dimension, volume fractal dimension, porosity, pore size distribution, and characterization of different types of cavity within incipient particles. In all analysis presented here, no significant impact of process temperature was found among the temperatures studied. This minimal temperature dependence observed suggests a stable early-stage structural configuration across different combustion conditions.

The particles were found to have a rough surface and a complex shape as indicated by a surface fractal dimension of approximately 2.22 and an average circularity of 0.66. The particles were found to be significantly porous with a porosity of 0.22. All pores in the particle were found to be micropores (i.e., of size less than 2 nm). Such large presence of micropores potentially increased the available surface area, leading to a high value of specific surface area of approximately 2652.36 m²/g and a low value of sphericity (around 0.57).

Cavities within the soot particles were categorized into tunnels, pockets, and isolated cavities based on their connectivity to the external surface. Isolated cavities were notably more spherical, whereas tunnels and pockets exhibited more elongated geometries. The presence and distribution of these cavities enhance the particles' surface area and influence gas diffusion and adsorption processes. The pore size distribution showed a log-normal distribution indicating the presence of a large number of small pores alongside a few larger ones. Pockets, despite their smaller size, occupied a larger fraction of the total void volume due to their higher abundance, while tunnels – though larger – were less prevalent. A positive correlation was identified between volume and surface area of pores. The volume of pore was also found to be correlated with particle's material volume. A set of correlations were proposed to connect these quantities.

The insights gained from this study enhance our understanding of the morphological evolution of soot particles during their incipient stages. The high porosity and specific surface area observed suggest that early-stage soot particles possess significant potential for gas adsorption and chemical reactivity, which are critical for modeling soot behavior in combustion and atmospheric environments. Additionally, the proposed correlations between pore volume and pore surface area and particle volume offer practical way to incorporate detailed morphological characteristics into larger-scale models, thereby improving their predictive accuracy.

6.5 Acknowledgments

K.M.M. and S.P.R. wish to thank Mr. Anindya Ganguly (University of Melbourne) for providing the RMD trajectory files from which the particles were extracted for analysis. The research benefited from computational resources provided through the NCMAS, supported by the Australian Government, The University of Melbourne's Research Computing Services and the Petascale Campus Initiative. K.M.M. and S.P.R. acknowledge funding support from the National Science Foundation as some of this material is based upon work supported by the National Science Foundation under Grant No. 2144290.

CHAPTER 7

SUMMARY AND FUTURE WORKS

7.1 Summary

This dissertation has presented a comprehensive exploration of soot formation mechanisms and morphological evolution from a molecular modeling perspective, addressing critical gaps in current knowledge. Employing state-of-the-art reactive molecular dynamics (RMD) simulations, this research has provided significant advancements in understanding the intricate chemical and physical transformations that occur during soot inception and subsequent early growth.

Key contributions of this work include:

1. **Development and Application of MAFIA-MD:** A sophisticated computational analysis tool (MAFIA-MD) was developed to accurately identify and quantify detailed molecular structures and morphological features from RMD simulations. This tool provides valuable insights into the chemical composition, internal structure, and surface characteristics of soot particles, significantly enhancing our capability to interpret simulation results in relation to practical soot formation processes.
2. **Gas-to-Particle Boundary Characterization:** Detailed molecular-level analysis established clear markers distinguishing gas-phase species from particulate soot both in terms of mass and number of carbons.
3. **Internal Structure Evolution of Incipient Soot Particles:** A comprehensive characterization of internal morphological transitions during soot nucleation and growth was carried out. Critical insights into particle density, structural complexity, and the internal chemical variations were obtained, revealing a new understanding of soot particle formation dynamics.
4. **Physicochemical Evolution of Incipient Soot Particles:** Investigating various physicochemical properties of incipient soot particles which provided valuable data regarding particle growth dynamics, ring structures, fractal dimensions, and particle

morphology. This analysis provides a foundation for more physically accurate soot models by establishing clear quantitative relationships between different physicochemical attributes.

5. **Pore and Surface Feature Analysis:** The study extensively characterized soot particle surface morphology and internal pore structures, revealing critical diagnostics relationships between particle morphology and surface properties. The quantification of surface areas, pore volumes, fractal characteristics, and pore size distributions significantly contributes to a deeper understanding of soot particle reactivity and maturation. The methodology developed for pore and surface feature analysis has a use beyond soot, potentially applicable to a wide range of porous materials in various fields, including catalysis, energy storage, and environmental science.

Through these contributions, this dissertation has clearly addressed existing limitations and provided robust molecular-level understanding essential for refining some of the soot formation models and enhancing strategies for controlling soot emissions and future engineering scale models.

7.2 Future Works

This dissertation provides significant insights into the formation, evolution, and morphological characteristics of soot particles obtained from acetylene (C_2H_2) pyrolysis using a series of reactive molecular dynamics simulations (RMD). However, several avenues for future exploration still remain, which can further enhance our understanding of soot formation at both molecular and macro scales. The following segment will discuss the possible future works on top of the methodology discussed in this dissertation.

7.2.1 Expansion of Reactive Molecular Dynamics (RMD) Simulations

The present study solely focuses on the pyrolysis of acetylene (C_2H_2) using RMD simulation. Future work should extend this to other hydrocarbon fuels, such as: ethylene (C_2H_4), n-heptane (C_7H_{16}) and n-dodecane ($C_{12}H_{26}$). Having this knowledge would allow for a more comprehensive comprehension of the generation of soot across a variety of fuels

and situations. Research should also be conducted on oxygenated fuels such as biodiesel in order to determine the extent to which these fuels contribute to the reduction of soot emissions. Previous studies have shown that oxygenated additives have the capability to alter the formation of soot by reducing the quantities of polycyclic aromatic hydrocarbon (PAH) precursors and delaying the nucleation process [237]. Also, the oxygenated fuel blends emit smaller and compact soot particles [238]. It is possible that the examination of these compounds could be useful in discovering alternative fuels that produce fewer emissions for the environment.

7.2.2 Temperature and Pressure Effects on Soot Evolution

The current work investigated the generation of soot at specific fixed temperatures, but it is recommended that future research investigate a wider range of temperatures and pressures, particularly under dynamic conditions that are observed in real combustion systems such as engines. Temperature has an impact on the physical characteristics of the particles, which in turn affects the graphitization process and the reactivity of soot particles. Research on pressure effects can also shed light on the factors that drive soot formation under high pressure conditions, such as those seen in rocket engines and gas turbines [239, 240, 241, 242].

7.2.3 Modeling Soot Aggregate and Coagulation

This study primarily focused on the early stages of soot formation, where incipient particles form through nucleation and surface growth processes. However, to gain a more comprehensive understanding of soot formation, future research should expand into the later stages of soot formation, i.e. aggregation and coagulation, which significantly alter soot particles' physical and chemical properties. As smaller particles cluster into larger, more complex aggregates, their optical properties, surface roughness, and internal pore structures change, influencing their interactions with light, pollutants, and atmospheric processes. Putting together reactive molecular dynamics (RMD) and mesoscale models, like the discrete element method (DEM), would help us understand how these processes of

aggregation and coagulation work [176]. This integration could help model the complete evolution of soot, from nucleation to mature aggregates, and assess how environmental factors like temperature, pressure, and turbulence influence these stages. By bridging atomistic-level details with macroscopic behavior, this approach would improve the predictive accuracy of engineering-scale combustion models and atmospheric simulations, leading to better predictions of soot's environmental and health impacts.

7.2.4 Further Investigation into Reaction Pathways

Reactive molecular dynamics (RMD) simulations provide detailed atomic-level data on chemical changes during soot formation, including bond breaking, radical formation, and surface reactions. However, extracting dominant reaction pathways from this vast data is challenging. Advanced techniques like reaction network analysis or machine learning can simplify this by identifying key molecular transformations and recurring patterns, particularly in the role of polycyclic aromatic hydrocarbons (PAHs) and resonance-stabilized radicals (RSRs), which are critical in soot nucleation. Further analysis of how temperature and pressure affect reaction pathways could refine soot formation models for various combustion environments. Validating these findings with experimental data, such as laser-induced incandescence or mass spectrometry, can enhance the accuracy of soot models, leading to better predictions of soot emissions and their environmental impacts.

7.2.5 Integration RMD Data into Engineering-Scale Model

Integrating knowledge from reactive molecular dynamics (RMD) simulations into engineering-scale soot models is a crucial future path. RMD offers comprehensive atomic-level data, including soot particle characteristics and chemical pathways, which may be converted into parameters for sectional or method of moments (MOM) models. For instance, the precision of soot growth forecasts can be improved by using RMD-derived data on particle size, surface area, and chemical processes (such as PAH generation and HACA mechanisms). Multi-scale modeling will be essential for bridging the gap between RMD and larger combustion simulations, possibly by coupling RMD with mesoscopic

models like the discrete element method (DEM) to simulate aggregation and feed into computational fluid dynamics (CFD) models. This method can enhance CFD's depiction of soot oxidation rates, fractal dimension, and particle morphology [243].

Machine learning (ML) may also help integrate large datasets from RMD into engineering models by identifying significant trends and generating reduced-order models for use in CFD. Lastly, it will be crucial to verify these integrated models against experimental data, like mass spectrometry or laser-induced incandescence, in order to guarantee accuracy and enhance the entire approach.

By bridging the gap between macroscopic behavior and molecular-scale insights, this integration will enhance our ability to control and reduce soot emissions. Additionally, it will improve the ability of soot models to anticipate various combustion situations.

BIBLIOGRAPHY

- [1] Anthi Liati, Benjamin T. Brem, Lukas Durdina, Melanie Vögtli, Yadira Arroyo Rojas Dasilva, Panayotis Dimopoulos Eggenschwiler, and Jing Wang. Electron Microscopic Study of Soot Particulate Matter Emissions from Aircraft Turbine Engines. *Environ. Sci. Technol.*, 48(18):10975–10983, September 2014. ISSN 0013-936X. doi: 10.1021/es501809b.
- [2] Rui Tang, Jing Shang, Xinghua Qiu, Jicheng Gong, Tao Xue, and Tong Zhu. Origin, Structural Characteristics, and Health Effects of Atmospheric Soot Particles: A Review. *Curr. Pollution Rep.*, 10(3):532–547, September 2024. ISSN 2198-6592. doi: 10.1007/s40726-024-00307-9.
- [3] Hope A. Michelsen, Meredith B. Colket, Per-Erik Bengtsson, Andrea D’Anna, Pascale Desgroux, Brian S. Haynes, J. Houston Miller, Graham J. Nathan, Heinz Pitsch, and Hai Wang. A Review of Terminology Used to Describe Soot Formation and Evolution under Combustion and Pyrolytic Conditions. *ACS Nano*, 14(10):12470–12490, October 2020. ISSN 1936-0851. doi: 10.1021/acsnano.0c06226.
- [4] Thomas R. Barfknecht. Toxicology of soot. *Prog. Energy Combust. Sci.*, 9(3):199–237, January 1983. ISSN 0360-1285. doi: 10.1016/0360-1285(83)90002-3.
- [5] Hamid Omidvarborna, Ashok Kumar, and Dong-Shik Kim. Recent studies on soot modeling for diesel combustion. *Renewable and Sustainable Energy Reviews*, 48: 635–647, Aug 2015. ISSN 1364-0321. doi: 10.1016/j.rser.2015.04.019.
- [6] T. C. Bond, S. J. Doherty, D. W. Fahey, P. M. Forster, T. Berntsen, B. J. DeAngelo, M. G. Flanner, S. Ghan, B. Kärcher, D. Koch, S. Kinne, Y. Kondo, P. K. Quinn, M. C. Sarofim, M. G. Schultz, M. Schulz, C. Venkataraman, H. Zhang, S. Zhang, N. Bellouin, S. K. Guttikunda, P. K. Hopke, M. Z. Jacobson, J. W. Kaiser, Z. Klimont, U. Lohmann, J. P. Schwarz, D. Shindell, T. Storelvmo, S. G. Warren, and C. S. Zender. Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research: Atmospheres*, 118(11):5380–5552, June 2013. ISSN 2169-897X. doi: 10.1002/jgrd.50171.
- [7] Nicklas Raun Jacobsen, Giulio Pojana, Paul White, Peter Møller, Corey Alexander Cohn, Karen Smith Korsholm, Ulla Vogel, Antonio Marcomini, Steffen Loft, and Håkan Wallin. Genotoxicity, cytotoxicity, and reactive oxygen species induced by single-walled carbon nanotubes and C60 fullerenes in the FE1-Muta™Mouse lung epithelial cells. *Environ. Mol. Mutagen.*, 49(6):476–487, July 2008. ISSN 0893-6692. doi: 10.1002/em.20406.
- [8] Michael Jerrett, Richard T. Burnett, Bernardo S. Beckerman, Michelle C. Turner, Daniel Krewski, George Thurston, Randall V. Martin, Aaron van Donkelaar, Edward Hughes, Yuanli Shi, Susan M. Gapstur, Michael J. Thun, and C. Arden Pope Iii. Spatial Analysis of Air Pollution and Mortality in California. *Am. J. Respir. Crit. Care Med.*, August 2013. URL

<https://www.atsjournals.org/doi/10.1164/rccm.201303-0609OC>.

- [9] Michael Jerrett, Richard T. Burnett, Bernardo S. Beckerman, Michelle C. Turner, Daniel Krewski, George Thurston, Randall V. Martin, Aaron van Donkelaar, Edward Hughes, Yuanli Shi, Susan M. Gapstur, Michael J. Thun, and C. Arden Pope Rd. Spatial analysis of air pollution and mortality in California. *American Journal of Respiratory and Critical Care Medicine*, 188(5):593–599, Sep 2013. ISSN 1535-4970. doi: 10.1164/rccm.201303-0609OC.
- [10] James Hansen and Larissa Nazarenko. Soot climate forcing via snow and ice albedos. *Proc. Natl. Acad. Sci. U.S.A.*, 101(2):423–428, January 2004. ISSN 0027-8424. doi: 10.1073/pnas.2237157100.
- [11] Jörg Appel, Henning Bockhorn, and Michael Frenklach. Kinetic modeling of soot formation with detailed chemistry and physics: laminar premixed flames of C2 hydrocarbons. *Combust. Flame*, 121(1):122–136, April 2000. ISSN 0010-2180. doi: 10.1016/S0010-2180(99)00135-2.
- [12] Michael Frenklach. Reaction mechanism of soot formation in flames. *Phys. Chem. Chem. Phys.*, 4(11):2028–2037, May 2002. ISSN 1463-9076. doi: 10.1039/B110045A.
- [13] Michael Frenklach and Stephen J. Harris. Aerosol dynamics modeling using the method of moments. *Journal of Colloid and Interface Science*, 118(1):252–261, July 1987. doi: 10.1016/0021-9797(87)90454-1. URL [https://doi.org/10.1016/0021-9797\(87\)90454-1](https://doi.org/10.1016/0021-9797(87)90454-1).
- [14] Michael Frenklach and Hai Wang. Detailed modeling of soot particle nucleation and growth. *Symp. Combust.*, 23(1):1559–1566, January 1991. ISSN 0082-0784. doi: 10.1016/S0082-0784(06)80426-1.
- [15] H Richter and J.B Howard. Formation of polycyclic aromatic hydrocarbons and their growth to soot—a review of chemical reaction pathways. *Progress in Energy and Combustion Science*, 26(4-6):565–608, August 2000. doi: 10.1016/S0360-1285(00)00009-5. URL [https://doi.org/10.1016/S0360-1285\(00\)00009-5](https://doi.org/10.1016/S0360-1285(00)00009-5).
- [16] K. Siegmann and K. Sattler. Formation mechanism for polycyclic aromatic hydrocarbons in methane flames. *J. Chem. Phys.*, 112(2):698–709, January 2000. ISSN 0021-9606. doi: 10.1063/1.480648.
- [17] Dale R. Tree and Kenth I. Svensson. Soot processes in compression ignition engines. *Progress in Energy and Combustion Science*, 33(3):272–309, June 2007. doi: 10.1016/j.pecs.2006.03.002. URL <https://doi.org/10.1016/j.pecs.2006.03.002>.
- [18] Ian M Kennedy. Models of soot formation and oxidation. *Progress in Energy and Combustion Science*, 23(2):95–132, January 1997. doi: 10.1016/S0360-1285(97)00007-5. URL [https://doi.org/10.1016/S0360-1285\(97\)00007-5](https://doi.org/10.1016/S0360-1285(97)00007-5).

- [19] Stelios Rigopoulos. Modelling of Soot Aerosol Dynamics in Turbulent Flow. *Flow, Turbulence and Combustion*, 103(3):565–604, 2019. ISSN 15731987. doi: 10.1007/s10494-019-00054-8.
- [20] Murray J. Thomson. Modeling soot formation in flames and reactors: Recent progress and current challenges. *Proc. Combust. Inst.*, 39(1):805–823, January 2023. ISSN 1540-7489. doi: 10.1016/j.proci.2022.07.263.
- [21] Hai Wang. Formation of nascent soot and other condensed-phase materials in flames. *Proc. Combust. Inst.*, 33(1):41–67, January 2011. ISSN 1540-7489. doi: 10.1016/j.proci.2010.09.009.
- [22] Khaled Mosharraf Mukut, Somesh Roy, and Eirini Goudeli. Molecular arrangement and fringe identification and analysis from molecular dynamics (MAFIA-MD): A tool for analyzing the molecular structures formed during reactive molecular dynamics simulation of hydrocarbons. *Comput. Phys. Commun.*, 276:108325, July 2022. ISSN 0010-4655. doi: 10.1016/j.cpc.2022.108325.
- [23] R. J. Gill and D. B. Olson. Estimation of soot thresholds for fuel mixtures. *Combustion Science and Technology*, 40(5-6):307–315, September 1984. doi: 10.1080/00102208408923814. URL <https://doi.org/10.1080/00102208408923814>.
- [24] D.M Probert I. M Khan, G. Greeves. Prediction of soot and nitric oxide concentrations in diesel engine exhaust. *Air Pollution Control in Transport Engines, Institute of Mechanical Engineers, London*, Paper C142:205–217, 1971.
- [25] Xiaowei Li, Peiling Ke, He Zheng, and Aiyang Wang. Structural properties and growth evolution of diamond-like carbon films with different incident energies: A molecular dynamics study. *Applied Surface Science*, Complete(273):670–675, 2013. ISSN 0169-4332. doi: 10.1016/j.apsusc.2013.02.108.
- [26] Scott A. Skeen, Julien Manin, Lyle M. Pickett, Emre Cenker, Gilles Bruneaux, Katsufumi Kondo, Tets Aizawa, Fredrik Westlye, Kristine Dalen, Anders Ivarsson, Tiemin Xuan, Jose M Garcia-Oliver, Yuanjiang Pei, Sibendu Som, Wang Hu, Rolf D. Reitz, Tommaso Lucchini, Gianluca D'Errico, Daniele Farrace, Sushant S. Pandurangi, Yuri M. Wright, Muhammad Aqib Chishty, Michele Bolla, and Evatt Hawkes. A progress review on soot experiments and modeling in the engine combustion network (ECN). *SAE International Journal of Engines*, 9(2):883–898, April 2016. doi: 10.4271/2016-01-0734. URL <https://doi.org/10.4271/2016-01-0734>.
- [27] Ian M. Kennedy, Wolfgang Kollmann, and J.-Y. Chen. A model for soot formation in a laminar diffusion flame. *Combustion and Flame*, 81(1):73–85, July 1990. doi: 10.1016/0010-2180(90)90071-x. URL [https://doi.org/10.1016/0010-2180\(90\)90071-x](https://doi.org/10.1016/0010-2180(90)90071-x).
- [28] K. M. Leung, R. P. Lindstedt, and W. P. Jones. A simplified reaction mechanism for soot formation in nonpremixed flames. *Combust. Flame*, 87(3):289–305, December 1991. ISSN 0010-2180. doi: 10.1016/0010-2180(91)90114-Q.

- [29] P.A. Tesner, E.I. Tsygankova, L.P. Guilazetdinov, V.P. Zuyev, and G.V. Loshakova. The formation of soot from aromatic hydrocarbons in diffusion flames of hydrocarbon-hydrogen mixtures. *Combustion and Flame*, 17(3):279–285, December 1971. doi: 10.1016/s0010-2180(71)80049-4. URL [https://doi.org/10.1016/s0010-2180\(71\)80049-4](https://doi.org/10.1016/s0010-2180(71)80049-4).
- [30] Fred Gelbard, Yoram Tambour, and John H Seinfeld. Sectional representations for simulating aerosol dynamics. *Journal of Colloid and Interface Science*, 76(2):541–556, August 1980. doi: 10.1016/0021-9797(80)90394-x. URL [https://doi.org/10.1016/0021-9797\(80\)90394-x](https://doi.org/10.1016/0021-9797(80)90394-x).
- [31] Michael Frenklach. Method of moments with interpolative closure. *Chemical Engineering Science*, 57(12):2229–2239, 2002. ISSN 0009-2509.
- [32] Robert McGraw. Description of aerosol dynamics by the quadrature method of moments. *Aerosol Science and Technology*, 27(2):255–265, January 1997. doi: 10.1080/02786829708965471. URL <https://doi.org/10.1080/02786829708965471>.
- [33] Michael Balthasar and Michael Frenklach. Monte-carlo simulation of soot particle coagulation and aggregation: the effect of a realistic size distribution. *Proceedings of the Combustion Institute*, 30(1):1467–1475, January 2005. doi: 10.1016/j.proci.2004.07.035. URL <https://doi.org/10.1016/j.proci.2004.07.035>.
- [34] Andrei Kazakov and Michael Frenklach. Dynamic modeling of soot particle coagulation and aggregation: Implementation with the method of moments and application to high-pressure laminar premixed flames. *Combustion and Flame*, 114(3-4):484–501, August 1998. doi: 10.1016/s0010-2180(97)00322-2. URL [https://doi.org/10.1016/s0010-2180\(97\)00322-2](https://doi.org/10.1016/s0010-2180(97)00322-2).
- [35] D. Mitrakos, E. Hinis, and C. Housiadas. Sectional Modeling of Aerosol Dynamics in Multi-Dimensional Flows. *Aerosol Sci. Technol.*, November 2007. URL <https://www.tandfonline.com/doi/full/10.1080/02786820701697804>.
- [36] Zhijie Huo, Matthew J. Cleary, Mariano Sirignano, and Assaad R. Masri. A sectional soot formation kinetics scheme with a new model for coagulation efficiency. *Combust. Flame*, 230:111444, August 2021. ISSN 0010-2180. doi: 10.1016/j.combustflame.2021.111444.
- [37] M Frenklach and H Wang. Detailed Mechanism and Modeling of Soot Particle Formation. In H Bockhorn, editor, *Soot Formation in Combustion: Mechanisms and Models*. Springer-Verlag, New York, 1994. URL https://link.springer.com/chapter/10.1007/978-3-642-85167-4_10.
- [38] M.Y. Choi, A. Hamins, G.W. Mulholland, and T. Kashiwagi. Simultaneous optical measurement of soot volume fraction and temperature in premixed flames. *Combustion and Flame*, 99(1):174–186, October 1994. doi: 10.1016/0010-2180(94)90088-4. URL [https://doi.org/10.1016/0010-2180\(94\)90088-4](https://doi.org/10.1016/0010-2180(94)90088-4).

- [39] Mario Commодо, Katharina Kaiser, Gianluigi De Falco, Patrizia Minutolo, Fabian Schulz, Andrea D’Anna, and Leo Gross. On the early stages of soot formation: Molecular structure elucidation by high-resolution atomic force microscopy. *Combust. Flame*, 205:154–164, July 2019. ISSN 0010-2180. doi: 10.1016/j.combustflame.2019.03.042.
- [40] Nancy J. Brown, Kenneth L. Revzan, and Michael Frenklach. Detailed kinetic modeling of soot formation in ethylene/air mixtures reacting in a perfectly stirred reactor. *Symposium (International) on Combustion*, 27(1):1573–1580, January 1998. doi: 10.1016/s0082-0784(98)80566-3. URL [https://doi.org/10.1016/s0082-0784\(98\)80566-3](https://doi.org/10.1016/s0082-0784(98)80566-3).
- [41] Leonard-Alexander Lieske, Mario Commодо, Jacob W. Martin, Katharina Kaiser, Vasiliki Benekou, Patrizia Minutolo, Andrea D’Anna, and Leo Gross. Portraits of Soot Molecules Reveal Pathways to Large Aromatics, Five-/Seven-Membered Rings, and Inception through π -Radical Localization. *ACS Nano*, 2023, July 2023. ISSN 1936-0851. doi: 10.1021/acsnano.3c02194.
- [42] James A. Rundel, Charlotte M. Thomas, Paul E. Schrader, Kevin R. Wilson, K. Olof Johansson, Ray P. Bambha, and Hope A. Michelsen. Promotion of particle formation by resonance-stabilized radicals during hydrocarbon pyrolysis. *Combust. Flame*, 243: 111942, September 2022. ISSN 0010-2180. doi: 10.1016/j.combustflame.2021.111942.
- [43] Scott A. Hollingsworth and Ron O. Dror. Molecular dynamics simulation for all. *Neuron*, 99(6):1129, Sep 2018. doi: 10.1016/j.neuron.2018.08.011.
- [44] Xiaowei Li, Aiyang Wang, and Kwang-Ryeol Lee. Comparison of empirical potentials for calculating structural properties of amorphous carbon films by molecular dynamics simulation. *Computational Materials Science*, 151:246–254, Aug 2018. ISSN 0927-0256. doi: 10.1016/j.commatsci.2018.04.062.
- [45] Basile Curchod and Todd J. Martínez. Ab Initio Nonadiabatic Quantum Molecular Dynamics. *Chemical Reviews*, 118(7):3305–3336, Apr 2018. ISSN 0009-2665. doi: 10.1021/acs.chemrev.7b00423.
- [46] Paul Erhart and Karsten Albe. Analytical potential for atomistic simulations of silicon, carbon, and silicon carbide. *Physical Review B*, 71(3):035211, Jan 2005. ISSN 2469-9969. doi: 10.1103/PhysRevB.71.035211.
- [47] J. Tersoff. New empirical approach for the structure and energy of covalent systems. *Physical Review B*, 37(12):6991–7000, Apr 1988. ISSN 2469-9969. doi: 10.1103/PhysRevB.37.6991.
- [48] Donald W Brenner, Olga A Shenderova, Judith A Harrison, Steven J Stuart, Boris Ni, and Susan B Sinnott. A second-generation reactive empirical bond order (REBO) potential energy expression for hydrocarbons. *Journal of Physics: Condensed Matter*, 14(4):783–802, jan 2002. doi: 10.1088/0953-8984/14/4/312. URL <https://doi.org/10.1088/0953-8984/14/4/312>.

- [49] Donald W. Brenner. Erratum: Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films. *Physical Review B*, 46(3): 1948, Jul 1992. ISSN 2469-9969. doi: 10.1103/PhysRevB.46.1948.2.
- [50] Thomas C. O'Connor, Jan Andzelm, and Mark O. Robbins. AIREBO-M: A reactive model for hydrocarbons at extreme pressures. *Journal of Chemical Physics*, 142(2): 024903, Jan 2015. ISSN 0021-9606. doi: 10.1063/1.4905549.
- [51] Adri C. T. van Duin, Siddharth Dasgupta, Francois Lorant, and William A. Goddard. ReaxFF: A Reactive Force Field for Hydrocarbons. *J. Phys. Chem. A*, 105(41): 9396–9409, October 2001. ISSN 1089-5639. doi: 10.1021/jp004368u.
- [52] N. A. Marks. Thin film deposition of tetrahedral amorphous carbon: a molecular dynamics study. *Diamond & Related Materials*, 8(14):1223–1231, 2005. ISSN 0925-9635. doi: 10.1016/j.diamond.2004.10.047.
- [53] Difan Zhang, Michael R. Dutzer, Tao Liang, Alexandre F. Fonseca, Ying Wu, Krista S. Walton, David S. Sholl, Amir H. Farmahini, Suresh K. Bhatia, and Susan B. Sinnott. Computational investigation on CO₂ adsorption in titanium carbide-derived carbons with residual titanium. *Carbon*, 111:741–751, Jan 2017. ISSN 0008-6223. doi: 10.1016/j.carbon.2016.10.037.
- [54] V. S. Dozhdikov, A. Yu Basharin, and P. R. Levashov. Structure of amorphous carbon quenched from liquid in the pressure range 1–40 GPa: Molecular dynamic modeling. *Journal of Physics: Conference Series*, 946(1):012086, Jan 2018. ISSN 1742-6596. doi: 10.1088/1742-6596/946/1/012086.
- [55] Tian-Bao Ma, Lin-Feng Wang, Yuan-Zhong Hu, Xin Li, and Hui Wang. A shear localization mechanism for lubricity of amorphous carbon materials - Scientific Reports. *Scientific Reports*, 4(3662):1–6, Jan 2014. ISSN 2045-2322. doi: 10.1038/srep03662.
- [56] Minwoong Joe, Myoung-Woon Moon, Jungsoo Oh, Kyu-Hwan Lee, and Kwang-Ryeol Lee. Molecular dynamics simulation study of the growth of a rough amorphous carbon film by the grazing incidence of energetic carbon atoms. *Carbon*, 50(2): 404–410, Feb 2012. ISSN 0008-6223. doi: 10.1016/j.carbon.2011.08.053.
- [57] Akaash Sharma, Khaled Mosharraf Mukut, Somesh P. Roy, and Eirini Goudeli. The coalescence of incipient soot clusters. *Carbon*, 180:215–225, August 2021. ISSN 0008-6223. doi: 10.1016/j.carbon.2021.04.065.
- [58] Amar M. Kamat, Adri C. T. van Duin, and Alexei Yakovlev. Molecular Dynamics Simulations of Laser-Induced Incandescence of Soot Using an Extended ReaxFF Reactive Force Field. *Journal of Physical Chemistry A*, 114(48):12561–12572, Dec 2010. ISSN 1089-5639. doi: 10.1021/jp1080302.
- [59] Sharmin Shabnam. ReaxFF reactive force field investigations in combustion and nascent soot formation, Feb 2020. URL <https://etda.libraries.psu.edu/catalog/17410sbs5576>. [Online; accessed 21.

Aug. 2021].

- [60] Qian Mao, Adri C. T. van Duin, and K. H. Luo. Formation of incipient soot particles from polycyclic aromatic hydrocarbons: A ReaxFF molecular dynamics study. *Carbon*, 121:380–388, September 2017. ISSN 0008-6223. doi: 10.1016/j.carbon.2017.06.009.
- [61] Charles A. Schuetz and Michael Frenklach. Nucleation of soot: Molecular dynamics simulations of pyrene dimerization. *Proc. Combust. Inst.*, 29(2):2307–2314, January 2002. ISSN 1540-7489. doi: 10.1016/S1540-7489(02)80281-4.
- [62] Song Han, Xiaoxia Li, Fengguang Nie, Mo Zheng, Xiaolong Liu, and Li Guo. Revealing the Initial Chemistry of Soot Nanoparticle Formation by ReaxFF Molecular Dynamics Simulations. *Energy Fuels*, 31(8):8434–8444, August 2017. ISSN 0887-0624. doi: 10.1021/acs.energyfuels.7b01194.
- [63] Cheng Chen and Xi Jiang. Molecular dynamics simulation of soot formation during diesel combustion with oxygenated fuel addition. *Phys. Chem. Chem. Phys.*, 22(36):20829–20836, September 2020. ISSN 1463-9076. doi: 10.1039/D0CP01917H.
- [64] K. O. Johansson, M. P. Head-Gordon, P. E. Schrader, K. R. Wilson, and H. A. Michelsen. Resonance-stabilized hydrocarbon-radical chain reactions may explain soot inception and growth. *Science*, 361(6406):997–1000, September 2018. ISSN 0036-8075. doi: 10.1126/science.aat3417.
- [65] Michael Frenklach and Alexander M. Mebel. On the mechanism of soot nucleation. *Physical Chemistry Chemical Physics*, 22(9):5314–5331, 2020.
- [66] Joonsik Hwang, Felix Sebastian Hirner, Choongsik Bae, Chetankumar Patel, Tarun Gupta, and Avinash Kumar Agarwal. HRTEM evaluation of primary soot particles originated in a small-bore biofuel compression-ignition engine. *Applied Thermal Engineering*, 159:113899, Aug 2019. ISSN 1359-4311. doi: 10.1016/j.applthermaleng.2019.113899.
- [67] K. O. Johansson, F. El Gabaly, P. E. Schrader, M. F. Campbell, and H. A. Michelsen. Evolution of maturity levels of the particle surface and bulk during soot growth and oxidation in a flame. *Aerosol Sci. Technol.*, 51(12):1333–1344, December 2017. ISSN 0278-6826. doi: 10.1080/02786826.2017.1355047.
- [68] The MDAnalysis Development Team. XYZ trajectory — MDAnalysis User Guide (release 2.0.0), May 2021. URL <https://userguide.mdanalysis.org/stable/formats/reference/xyz.html>. [Online; accessed 5. Jun. 2021].
- [69] Yeonjoon Kim and Woo Youn Kim. Universal Structure Conversion Method for Organic Molecules: From Atomic Connectivity to Three-Dimensional Geometry. *Bulletin of the Korean Chemical Society*, 36(7):1769–1777, Jul 2015. ISSN 1229-5949. doi: 10.1002/bkcs.10334.

- [70] David Weininger. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of Chemical Information and Computer Sciences*, 28(1):31–36, Feb 1988. ISSN 0095-2338. doi: 10.1021/ci00057a005.
- [71] Arthur Dalby, James G. Nourse, W. Douglas Hounshell, Ann K. I. Gushurst, David L. Grier, Burton A. Leland, and John Laufer. Description of several chemical structure file formats used by computer programs developed at Molecular Design Limited. *Journal of Chemical Information and Computer Sciences*, 32(3):244–255, May 1992. ISSN 0095-2338. doi: 10.1021/ci00007a012.
- [72] Sébastien Le Roux and Philippe Jund. Ring statistics analysis of topological networks: New approach and application to amorphous ges2 and sio2 systems. *Computational Materials Science*, 49(1):70–83, 2010. ISSN 0927-0256. doi: <https://doi.org/10.1016/j.commatsci.2010.04.023>. URL <https://www.sciencedirect.com/science/article/pii/S0927025610002363>.
- [73] Martin Brehm and Barbara Kirchner. TRAVIS - a free analyzer and visualizer for Monte Carlo and molecular dynamics trajectories. *Journal of Chemical Information and Modeling*, 51(8):2007–2023, 2011. doi: 10.1021/ci200217w.
- [74] M. Brehm, M. Thomas, S. Gehrke, and B. Kirchner. TRAVIS – a free analyzer for trajectories from molecular simulation. *The Journal of Chemical Physics*, 152(16): 164105, 2020. ISSN 0021-9606. doi: 10.1063/5.0005078. URL <http://aip.scitation.org/doi/10.1063/5.0005078>.
- [75] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms, Third Edition*. The MIT Press, Cambridge, MA, USA, Jul 2009. ISBN 9780262033848. URL <https://mitpress.mit.edu/books/introduction-algorithms-third-edition>.
- [76] Hugo Flávio. 1.2) Creating a distances matrix, Jan 2021. URL https://cran.r-project.org/web/packages/actel/vignettes/a-2_distances_matrix.html. [Online; accessed 6. Jun. 2021].
- [77] György Turán. On the succinct representation of graphs. *Discrete Applied Mathematics*, 8(3):289–294, Jul 1984. ISSN 0166-218X. doi: 10.1016/0166-218X(84)90126-4.
- [78] The NetworkX Developers. NetworkX Documentation, Apr 2021. URL <https://networkx.org/documentation/stable/>. [Online; accessed 6. Jun. 2021].
- [79] Donald B. Johnson. Finding All the Elementary Circuits of a Directed Graph. *SIAM Journal on Computing*, Jul 2006. URL <https://epubs.siam.org/doi/10.1137/0204007>.
- [80] Dexter Kozen. *The Design and Analysis of Algorithms*. Springer-Verlag, New York, NY, USA, 1992. ISBN 978-0-387-97687-7. doi: 10.1007/978-1-4612-4400-4.

- [81] R. Tarjan. Depth-First Search and Linear Graph Algorithms. *undefined*, 1972. URL <https://www.semanticscholar.org/paper/Depth-First-Search-and-Linear-Graph-Algorithms-Tarjan/385742fffcf113656f0d3cf6c06ef95cb8439dc6>.
- [82] jensengroup. xyz2mol, Jun 2021. URL <https://github.com/jensengroup/xyz2mol>. [Online; accessed 23. Jun. 2021].
- [83] Greg Landrum. RDKit, Oct 2019. URL <http://www.rdkit.org>. [Online; accessed 23. Jun. 2021].
- [84] Randy L. Vander Wal, Aleksey Yezerets, Neal W. Currier, Do Heui Kim, and Chong Min Wang. HRTEM Study of diesel soot collected from diesel particulate filters. *Carbon*, 45(1):70–77, Jan 2007. ISSN 0008-6223. doi: 10.1016/j.carbon.2006.08.005.
- [85] Chethan K. Gaddam, Chung-Hsuan Huang, and Randy L. Vander Wal. Quantification of nano-scale carbon structure by HRTEM and lattice fringe analysis. *Pattern Recognition Letters*, 76:90–97, Jun 2016. ISSN 0167-8655. doi: 10.1016/j.patrec.2015.08.028.
- [86] Randy L. Vander Wal. Soot Nanostructure: Definition, Quantification and Implications. *SAE Transactions*, 114:429–436, 2005. ISSN 0096-736X. URL <http://www.jstor.org/stable/44720970>.
- [87] Yen-Chi Chen. A tutorial on kernel density estimation and recent advances, 2017.
- [88] Anaconda software distribution, 2020. URL <https://docs.anaconda.com/>.
- [89] Fredrik Lundh. An introduction to tkinter. URL: *www.pythonware.com/library/tkinter/introduction/index.htm*, 1999.
- [90] Installation guide for conda, Jun 2021. URL <https://docs.conda.io/projects/conda/en/latest/user-guide/install/index.html>. [Online; accessed 5. Jul. 2021].
- [91] Miniconda documentation, Jun 2021. URL <https://docs.conda.io/en/latest/miniconda.html>. [Online; accessed 5. Jul. 2021].
- [92] K. M Mukut and S. Roy. kmmukut/RingDetection: RingDetection. <https://doi.org/10.5281/zenodo.4283067>, 2020. v1.01.
- [93] Open Babel development team. Open babel, Jun 2021. URL http://openbabel.org/wiki/Main_Page. [Online; accessed 5. Jul. 2021].
- [94] K. M. Mukut, A. Sharma, E. Goudeli, and S. P. Roy. A closer look into the formation of soot particles: A molecular dynamics study. In *12th US National Combustion Meeting, College Station, TX, USA*, May 2021.

- [95] M. Balthasar and M. Kraft. A stochastic approach to calculate the particle size distribution function of soot particles in laminar premixed flames. *Combust. Flame*, 133(3):289–298, May 2003. ISSN 0010-2180. doi: 10.1016/S0010-2180(03)00003-8.
- [96] Buyu Wang, Sebastian Mosbach, Sebastian Schmutzhard, Shijin Shuai, Yaqing Huang, and Markus Kraft. Modelling soot formation from wall films in a gasoline direct injection engine using a detailed population balance model. *Appl. Energy*, 163: 154–166, February 2016. ISSN 0306-2619. doi: 10.1016/j.apenergy.2015.11.011.
- [97] Somesh Prasad Roy. *Aerosol-dynamics-based soot modeling of flames*. PhD thesis, Pennsylvania State University, Feb 2014. URL <https://etda.libraries.psu.edu/catalog/20625>. [Online; accessed 1. Apr. 2021].
- [98] Fidel Castro-Marciano, Amar M. Kamat, Michael F. Russo, Adri C. T. van Duin, and Jonathan P. Mathews. Combustion of an Illinois No. 6 coal char simulated using an atomistic char representation and the ReaxFF reactive force field. *Combust. Flame*, 159(3):1272–1285, March 2012. ISSN 0010-2180. doi: 10.1016/j.combustflame.2011.10.022.
- [99] Kimberly Chenoweth, Adri C. T. van Duin, and William A. Goddard. ReaxFF Reactive Force Field for Molecular Dynamics Simulations of Hydrocarbon Oxidation. *J. Phys. Chem. A*, 112(5):1040–1053, February 2008. ISSN 1089-5639. doi: 10.1021/jp709896w.
- [100] William C. Swope, Hans C. Andersen, Peter H. Berens, and Kent R. Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.*, 76(1):637–649, January 1982. ISSN 0021-9606. doi: 10.1063/1.442716.
- [101] Kimberly Chenoweth, Adri C. T. van Duin, Siddharth Dasgupta, and William A. Goddard III. Initiation Mechanisms and Kinetics of Pyrolysis and Combustion of JP-10 Hydrocarbon Jet Fuel. *Journal of Physical Chemistry A*, 113(9):1740–1746, Mar 2009. ISSN 1089-5639. doi: 10.1021/jp8081479.
- [102] D. J. Evans and B. L. Holian. The Nose–Hoover thermostat. *J. Chem. Phys.*, 83(8): 4069–4074, October 1985. ISSN 0021-9606. doi: 10.1063/1.449071.
- [103] Steve Plimpton. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *Journal of Computational Physics*, 117(1):1–19, Mar 1995. ISSN 0021-9991. doi: 10.1006/jcph.1995.1039.
- [104] Chaoyang Zhang, Chi Zhang, Yu Ma, and Xianggui Xue. Imaging the C black formation by acetylene pyrolysis with molecular reactive force field simulations. *Phys. Chem. Chem. Phys.*, 17(17):11469–11480, April 2015. ISSN 1463-9076. doi: 10.1039/C5CP00926J.
- [105] A. Tregrossi, A. Ciajolo, and R. Barbella. The combustion of benzene in rich premixed flames at atmospheric pressure. *Combustion and Flame*, 117(3):553–561,

- May 1999. ISSN 0010-2180. doi: 10.1016/S0010-2180(98)00157-6.
- [106] Rachelle S. Jacobson, Andrew R. Korte, Akos Vertes, and J. Houston Miller. The Molecular Composition of Soot. *Angew. Chem. Int. Ed.*, 59(11):4484–4490, March 2020. ISSN 1433-7851. doi: 10.1002/anie.201914115.
 - [107] Abhijeet Raj, Markus Sander, Vinod Janardhanan, and Markus Kraft. A study on the coagulation of polycyclic aromatic hydrocarbon clusters to determine their collision efficiency. *Combust. Flame*, 157(3):523–534, March 2010. ISSN 0010-2180. doi: 10.1016/j.combustflame.2009.10.003.
 - [108] Sebastian Mosbach, Matthew S. Celnik, Abhijeet Raj, Markus Kraft, Hongzhi R. Zhang, Shuichi Kubo, and Kyoung-Oh Kim. Towards a detailed soot model for internal combustion engines. *Combust. Flame*, 156(6):1156–1165, June 2009. ISSN 0010-2180. doi: 10.1016/j.combustflame.2009.01.003.
 - [109] Mohammad Reza Kholghy, Armin Veshkini, and Murray John Thomson. The core-shell internal nanostructure of soot – A criterion to model soot maturity. *Carbon*, 100:508–536, April 2016. ISSN 0008-6223. doi: 10.1016/j.carbon.2016.01.022.
 - [110] Khaled Mosharraf Mukut, Anindya Ganguly, Eirini Goudeli, Georgios A. Kelesidis, and Somesh P. Roy. Internal Structure of Incipient Soot from Acetylene Pyrolysis Obtained via Molecular Dynamics Simulations. *J. Phys. Chem. A*, 128(26):5175–5187, July 2024. ISSN 1089-5639. doi: 10.1021/acs.jpca.4c01548.
 - [111] Cornelia Irimiea, Alessandro Faccinnetto, Xavier Mercier, Ismael-Kenneth Ortega, Nicolas Nuns, Eric Therssen, Pascale Desgroux, and Cristian Focsa. Unveiling trends in soot nucleation and growth: When secondary ion mass spectrometry meets statistical analysis. *Carbon*, 144:815–830, April 2019. ISSN 0008-6223. doi: 10.1016/j.carbon.2018.12.015.
 - [112] Darson D. Li, Cheng Wang, Qing N. Chan, and Guan H. Yeoh. Soot: A review of computational models at different length scales. *Experimental and Computational Multiphase Flow*, 5(1):1–14, mar 2023. ISSN 2661-8869. doi: 10.1007/s42757-021-0124-4. URL <https://link.springer.com/10.1007/s42757-021-0124-4>.
 - [113] Richard A. Dobbins and Haran Subramaniasivam. Soot Precursor Particles in Flames. In *Soot Formation in Combustion*, pages 290–301. Springer, Berlin, Germany, 1994. doi: 10.1007/978-3-642-85167-4_16.
 - [114] Randall L. Vander Wal. Soot precursor carbonization: Visualization using LIF and LII and comparison using bright and dark field TEM. *Combust. Flame*, 112(4): 607–616, March 1998. ISSN 0010-2180. doi: 10.1016/S0010-2180(97)00171-5.
 - [115] Bin Zhao, Kei Uchikawa, and Hai Wang. A comparative study of nanoparticles in premixed flames by scanning mobility particle sizer, small angle neutron scattering, and transmission electron microscopy. *Proc. Combust. Inst.*, 31(1):851–860, January 2007. ISSN 1540-7489. doi: 10.1016/j.proci.2006.08.064.

- [116] Menghui Chen, Wei Li, Houjun Zhang, Menghui Liu, Jinli Zhang, Xiangyuan Li, and You Han. Recent ReaxFF MD studies on pyrolysis and combustion mechanisms of aviation/aerospace fuels and energetic additives. *Energy Advances*, 2(1):54–72, 2023. doi: 10.1039/D2YA00285J.
- [117] Hassan Sabbah, Ludovic Biennier, Stephen J. Klippenstein, Ian R. Sims, and Bertrand R. Rowe. Exploring the Role of PAHs in the Formation of Soot: Pyrene Dimerization. *J. Phys. Chem. Lett.*, 1(19):2962–2967, October 2010. doi: 10.1021/jz101033t.
- [118] Nazly E. Sánchez, Alicia Callejas, Ángela Millera, Rafael Bilbao, and María U. Alzueta. Polycyclic Aromatic Hydrocarbon (PAH) and Soot Formation in the Pyrolysis of Acetylene and Ethylene: Effect of the Reaction Temperature. *Energy Fuels*, 26(8):4823–4829, August 2012. ISSN 0887-0624. doi: 10.1021/ef300749q.
- [119] Maria L. Botero, Yuan Sheng, Jethro Akroyd, Jacob Martin, Jochen A. H. Dreyer, Wenming Yang, and Markus Kraft. Internal structure of soot particles in a diffusion flame. *Carbon*, 141:635–642, January 2019. ISSN 0008-6223. doi: 10.1016/j.carbon.2018.09.063.
- [120] Qinghua Chang, Rui Gao, Ming Gao, Guangsuo Yu, and Fuchen Wang. The structural evolution and fragmentation of coal-derived soot and carbon black during high-temperature air oxidation. *Combust. Flame*, 216:111–125, June 2020. ISSN 0010-2180. doi: 10.1016/j.combustflame.2019.11.045.
- [121] Pranay P. Morajkar, Moataz K. Abdrabou, Abhijeet Raj, Mirella Elkadi, Sasi Stephen, and Mohamed Ibrahim Ali. Transmission of trace metals from fuels to soot particles: An ICP-MS and soot nanostructural disorder study using diesel and diesel/Karanja biodiesel blend. *Fuel*, 280:118631, November 2020. ISSN 0016-2361. doi: 10.1016/j.fuel.2020.118631.
- [122] Kevin Gleason, Francesco Carbone, Andrew J. Sumner, Brian D. Drollette, Desiree L. Plata, and Alessandro Gomez. Small aromatic hydrocarbons control the onset of soot nucleation. *Combust. Flame*, 223:398–406, January 2021. ISSN 0010-2180. doi: 10.1016/j.combustflame.2020.08.029.
- [123] Francesco Carbone, Kevin Gleason, and Alessandro Gomez. Probing gas-to-particle transition in a moderately sooting atmospheric pressure ethylene/air laminar premixed flame. Part I: gas phase and soot ensemble characterization. *Combust. Flame*, 181:315–328, July 2017. ISSN 0010-2180. doi: 10.1016/j.combustflame.2017.01.029.
- [124] A. C. Barone, A. D’Alessio, and A. D’Anna. Morphological characterization of the early process of soot formation by atomic force microscopy. *Combust. Flame*, 132(1): 181–187, January 2003. ISSN 0010-2180. doi: 10.1016/S0010-2180(02)00434-0.
- [125] Fabian Schulz, Mario Commодо, Katharina Kaiser, Gianluigi De Falco, Patrizia Minutolo, Gerhard Meyer, Andrea D’anna, and Leo Gross. Insights into incipient soot formation by atomic force microscopy. *Proc. Combust. Inst.*, 37(1):885–892,

January 2019. ISSN 1540-7489. doi: 10.1016/j.proci.2018.06.100.

- [126] Mario Commодо, Andrea D’Anna, Gianluigi De Falco, Rosanna Larciprete, and Patrizia Minutolo. Illuminating the earliest stages of the soot formation by photoemission and Raman spectroscopy. *Combust. Flame*, 181:188–197, July 2017. ISSN 0010-2180. doi: 10.1016/j.combustflame.2017.03.020.
- [127] Dorian S. N. Parker, Fangtong Zhang, Y. Seol Kim, Ralf. I. Kaiser, and Alexander M. Mebel. On the Formation of Resonantly Stabilized C₅H₃ Radicals—A Crossed Beam and Ab Initio Study of the Reaction of Ground State Carbon Atoms with Vinylacetylene. *J. Phys. Chem. A*, 115(5):593–601, February 2011. ISSN 1089-5639. doi: 10.1021/jp109800h.
- [128] Francesco Silvio Gentile, Francesca Picca, Gianluigi De Falco, Mario Commодо, Patrizia Minutolo, Mauro Causà, and Andrea D’Anna. Soot inception: A DFT study of σ and π dimerization of resonantly stabilized aromatic radicals. *Fuel*, 279:118491, November 2020. ISSN 0016-2361. doi: 10.1016/j.fuel.2020.118491.
- [129] Chowdhury Ashraf and Adri C. T. van Duin. Extension of the ReaxFF Combustion Force Field toward Syngas Combustion and Initial Oxidation Kinetics. *J. Phys. Chem. A*, 121(5):1051–1068, February 2017. ISSN 1520-5215. doi: 10.1021/acs.jpca.6b12429.
- [130] Laura Pascazio, Jacob W. Martin, Kimberly Bowal, Jethro Akroyd, and Markus Kraft. Exploring the internal structure of soot particles using nanoindentation: A reactive molecular dynamics study. *Combust. Flame*, 219:45–56, September 2020. ISSN 0010-2180. doi: 10.1016/j.combustflame.2020.04.029.
- [131] Tomoji Ishiguro, Yoshiki Takatori, and Kazuhiro Akihama. Microstructure of diesel soot particles probed by electron microscopy: First observation of inner core and outer shell. *Combust. Flame*, 108(1):231–234, January 1997. ISSN 0010-2180. doi: 10.1016/S0010-2180(96)00206-4.
- [132] Anil D. Pathak, Darshna Potphode, and Chandra S. Sharma. Graphitization induced structural transformation of candle soot carbon into carbon nano-onion as a functional anode for metal-ion batteries. *Mater. Adv.*, 3(8):3610–3619, April 2022. ISSN 2633-5409. doi: 10.1039/D2MA00042C.
- [133] Ye Liu, Gang Lv, Chenyang Fan, Na Li, and Xiaowei Wang. Surface functional groups and graphitization degree of soot in the sooting history of methane premixed flame. In *WCXTM 17: SAE World Congress Experience*. SAE International, mar 2017. doi: <https://doi.org/10.4271/2017-01-1003>. URL <https://doi.org/10.4271/2017-01-1003>.
- [134] M. Ruiz, R. Villoria, A. Millera, M. Alzueta, and R. Bilbao. Influence of the temperature on the properties of the soot formed from c₂h₂ pyrolysis. *Chemical Engineering Journal*, 127:1–9, 2007. doi: 10.1016/J.CEJ.2006.09.006.
- [135] Aidan P. Thompson, H. Metin Aktulga, Richard Berger, Dan S. Bolintineanu, W. Michael Brown, Paul S. Crozier, Pieter J. In ’t Veld, Axel Kohlmeyer, Stan G.

- Moore, Trung Dac Nguyen, Ray Shan, Mark J. Stevens, Julien Tranchida, Christian Trott, and Steven J. Plimpton. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput. Phys. Commun.*, 271:108171, February 2022. ISSN 0010-4655. doi: 10.1016/j.cpc.2021.108171.
- [136] Michel F. Sanner, Arthur J. Olson, and Jean-Claude Spehner. Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers*, 38(3):305–320, March 1996. ISSN 0006-3525. doi: 10.1002/(SICI)1097-0282(199603)38:3<305::AID-BIP4>3.0.CO;2-Y.
- [137] Alexander Stukowski. Visualization and analysis of atomistic simulation data with OVITO—the Open Visualization Tool. *Model. Simul. Mater. Sci. Eng.*, 18(1):015012, December 2009. ISSN 0965-0393. doi: 10.1088/0965-0393/18/1/015012.
- [138] S. Lloyd. Least squares quantization in PCM. *IEEE Trans. Inf. Theory*, 28(2): 129–137, March 1982. ISSN 1557-9654. doi: 10.1109/TIT.1982.1056489.
- [139] Laurens van der Maaten and Geoffrey Hinton. Visualizing Data using t-SNE. *Journal of Machine Learning Research*, 9(86):2579–2605, 2008. ISSN 1533-7928. URL <https://www.jmlr.org/papers/v9/vandermaaten08a.html>.
- [140] Khaled Mosharraf Mukut, Anindya Ganguly, Eirini Goudeli, Georgios A. Kelesidis, and Somesh P. Roy. Physical, chemical and morphological evolution of incipient soot obtained from molecular dynamics simulation of acetylene pyrolysis. *Fuel*, 373:132197, October 2024. ISSN 0016-2361. doi: 10.1016/j.fuel.2024.132197.
- [141] Hope A. Michelsen, Matthew F. Campbell, K. Olof Johansson, Ich C. Tran, Paul E. Schrader, Ray P. Bambha, Emre Cencer, Joshua A. Hammons, Chenhui Zhu, Eric Schaible, and Anthony van Buuren. Soot-particle core-shell and fractal structures from small-angle X-ray scattering measurements in a flame. *Carbon*, 196:440–456, August 2022. ISSN 0008-6223. doi: 10.1016/j.carbon.2022.05.009.
- [142] Georgios A. Kelesidis, Patrizia Crepaldi, and Sotiris E. Pratsinis. Oxidation dynamics of soot or carbon black accounting for its core-shell structure and pore network. *Carbon*, 219:118764, February 2024. ISSN 0008-6223. doi: 10.1016/j.carbon.2023.118764.
- [143] Stephen S. Lim, Theo Vos, Abraham D. Flaxman, Goodarz Danaei, Kenji Shibuya, Heather Adair-Rohani, Markus Amann, H. Ross Anderson, Kathryn G. Andrews, Martin Aryee, Charles Atkinson, Loraine J. Bacchus, Adil N. Bahalim, Kalpana Balakrishnan, John Balmes, Suzanne Barker-Collo, Amanda Baxter, Michelle L. Bell, Jed D. Blore, Fiona Blyth, Carissa Bonner, Guilherme Borges, Rupert Bourne, Michel Boussinesq, Michael Brauer, Peter Brooks, Nigel G. Bruce, Bert Brunekreef, Claire Bryan-Hancock, Chiara Bucello, Rachelle Buchbinder, Fiona Bull, Richard T. Burnett, Tim E. Byers, Bianca Calabria, Jonathan Carapetis, Emily Carnahan, Zoe Chafe, Fiona Charlson, Honglei Chen, Jian Shen Chen, Andrew Tai Ann Cheng, Jennifer Christine Child, Aaron Cohen, K. Ellicott Colson, Benjamin C. Cowie, Sarah

Darby, Susan Darling, Adrian Davis, Louisa Degenhardt, Frank Dentener, Don C. Des Jarlais, Karen Devries, Mukesh Dherani, Eric L. Ding, E. Ray Dorsey, Tim Driscoll, Karen Edmond, Suad Eltahir Ali, Rebecca E. Engell, Patricia J. Erwin, Saman Fahimi, Gail Falder, Farshad Farzadfar, Alize Ferrari, Mariel M. Finucane, Seth Flaxman, Francis Gerry R. Fowkes, Greg Freedman, Michael K. Freeman, Emmanuela Gakidou, Santu Ghosh, Edward Giovannucci, Gerhard Gmel, Kathryn Graham, Rebecca Grainger, Bridget Grant, David Gunnell, Hialy R. Gutierrez, Wayne Hall, Hans W. Hoek, Anthony Hogan, H. Dean Hosgood, Damian Hoy, Howard Hu, Bryan J. Hubbell, Sally J. Hutchings, Sydney E. Ibeanusi, Gemma L. Jacklyn, Rashmi Jasrasaria, Jost B. Jonas, Haidong Kan, John A. Kanis, Nicholas Kassebaum, Norito Kawakami, Young Ho Khang, Shahab Khatibzadeh, Jon Paul Khoo, Cindy Kok, Francine Laden, Ratilal Laloo, Qing Lan, Tim Lathlean, Janet L. Leasher, James Leigh, Yang Li, John Kent Lin, Steven E. Lipshultz, Stephanie London, Rafael Lozano, Yuan Lu, Joelle Mak, Reza Malekzadeh, Leslie Mallinger, Wagner Marcenes, Lyn March, Robin Marks, Randall Martin, Paul McGale, John McGrath, Sumi Mehta, George A. Mensah, Tony R. Merriman, Renata Micha, Catherine Michaud, Vinod Mishra, Khayriyyah Mohd Hanafiah, Ali A. Mokdad, Lidia Morawska, Dariush Mozaffarian, Tasha Murphy, Mohsen Naghavi, Bruce Neal, Paul K. Nelson, Joan Miquel Nolla, Rosana Norman, Casey Olives, Saad B. Omer, Jessica Orchard, Richard Osborne, Bart Ostro, Andrew Page, Kiran D. Pandey, Charles D.H. Parry, Erin Passmore, Jayadeep Patra, Neil Pearce, Pamela M. Pelizzari, Max Petzold, Michael R. Phillips, Dan Pope, C. Arden Pope, John Powles, Mayuree Rao, Homie Razavi, Eva A. Rehfuess, Jürgen T. Rehm, Beate Ritz, Frederick P. Rivara, Thomas Roberts, Carolyn Robinson, Jose A. Rodriguez-Portales, Isabelle Romieu, Robin Room, Lisa C. Rosenfeld, Ananya Roy, Lesley Rushton, Joshua A. Salomon, Uchechukwu Sampson, Lidia Sanchez-Riera, Ella Sanman, Amir Sapkota, Soraya Seedat, Peilin Shi, Kevin Shield, Rupak Shivakoti, Gitanjali M. Singh, David A. Sleet, Emma Smith, Kirk R. Smith, Nicolas J.C. Stapelberg, Kyle Steenland, Heidi Stöckl, Lars Jacob Stovner, Kurt Straif, Lahn Straney, George D. Thurston, Jimmy H. Tran, Rita Van Dingenen, Aaron Van Donkelaar, J. Lennert Veerman, Lakshmi Vijayakumar, Robert Weintraub, Myrna M. Weissman, Richard A. White, Harvey Whiteford, Steven T. Wiersma, James D. Wilkinson, Hywel C. Williams, Warwick Williams, Nicholas Wilson, Anthony D. Woolf, Paul Yip, Jan M. Zielinski, Alan D. Lopez, Christopher J.L. Murray, and Majid Ezzati. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, 380(9859):2224–2260, 2012. doi: 10.1016/S0140-6736(12)61766-8.

- [144] Cracking the code to soot formation: Scientists unlock mystery to help reduce hazardous emissions, February 2023. URL <https://www.sciencedaily.com/releases/2018/09/180906141629.htm>. [Online; accessed 8. Feb. 2023].
- [145] C. Russo, A. Tregrossi, and A. Ciajolo. Dehydrogenation and growth of soot in premixed flames. *Proc. Combust. Inst.*, 35(2):1803–1809, January 2015. ISSN 1540-7489. doi: 10.1016/j.proci.2014.05.024.

- [146] Alberto Baldelli, Una Trivanovic, Timothy A. Sipkens, and Steven N. Rogak. On determining soot maturity: A review of the role of microscopy- and spectroscopy-based techniques. *Chemosphere*, 252:126532, August 2020. ISSN 0045-6535. doi: 10.1016/j.chemosphere.2020.126532.
- [147] Yuxin Zhou, Mengda Wang, Qingyan He, and Xiaoqing You. Experimental investigation on the size-dependent maturity of soot particles in laminar premixed ethylene burner-stabilized stagnation flames. *Proc. Combust. Inst.*, 39(1):1147–1155, January 2023. ISSN 1540-7489. doi: 10.1016/j.proci.2022.09.030.
- [148] P. Minutolo, G. Gambi, and A. D’Alessio. The optical band gap model in the interpretation of the UV-visible absorption spectra of rich premixed flames. *Symp. Combust.*, 26(1):951–957, January 1996. ISSN 0082-0784. doi: 10.1016/S0082-0784(96)80307-9.
- [149] Joaquin Camacho, Changran Liu, Chen Gu, He Lin, Zhen Huang, Quanxi Tang, Xiaoqing You, Chiara Saggese, Yang Li, Heejung Jung, Lei Deng, Irenaeus Wlokas, and Hai Wang. Mobility size and mass of nascent soot particles in a benchmark premixed ethylene flame. *Combust. Flame*, 162(10):3810–3822, October 2015. ISSN 0010-2180. doi: 10.1016/j.combustflame.2015.07.018.
- [150] Kihong Park, David B. Kittelson, and Peter H. McMurry. Structural Properties of Diesel Exhaust Particles Measured by Transmission Electron Microscopy (TEM): Relationships to Particle Mass and Mobility. *Aerosol Sci. Technol.*, September 2004. ISSN 1066-2163. URL <https://www.tandfonline.com/doi/abs/10.1080/027868290505189>.
- [151] Armin Veshkini. *Understanding Soot Particle Growth Chemistry and Particle Sizing Using a Novel Soot Growth and Formation Model*. PhD thesis, University of Toronto, November 2015. URL <https://tspace.library.utoronto.ca/handle/1807/71398>.
- [152] Pascale Desgroux. Recent advances in soot nucleation understanding. In *European Combustion Meeting*, Lisboa, Portugal, Apr 2019. URL <https://hal.archives-ouvertes.fr/hal-03326326>.
- [153] Alexander S. Semnikhin, Anna S. Savchenkova, Ivan V. Chechet, Sergey G. Matveev, Michael Frenklach, and Alexander M. Mebel. Transformation of an Embedded Five-Membered Ring in Polycyclic Aromatic Hydrocarbons via the Hydrogen-Abstraction–Acetylene-Addition Mechanism: A Theoretical Study. *The Journal of Physical Chemistry A*, 125(16):3341–3354, 2021. ISSN 1089-5639. doi: 10.1021/acs.jpca.1c00900. URL <https://pubs.acs.org/doi/10.1021/acs.jpca.1c00900>.
- [154] Michael Frenklach, Alexander S. Semnikhin, and Alexander M. Mebel. On the Mechanism of Soot Nucleation. III. The Fate and Facility of the E-Bridge. *The Journal of Physical Chemistry A*, 125(31):6789–6795, aug 2021. doi: 10.1021/acs.jpca.1c04936. URL

<https://pubs.acs.org/doi/10.1021/acs.jpca.1c04936>.

- [155] Georgios A. Kelesidis and Sotiris E. Pratsinis. Soot light absorption and refractive index during agglomeration and surface growth. *Proc. Combust. Inst.*, 37(1): 1177–1184, January 2019. ISSN 1540-7489. doi: 10.1016/j.proci.2018.08.025.
- [156] Nils-Erik Olofsson, Johan Simonsson, Sandra Török, Henrik Bladh, and Per-Erik Bengtsson. Evolution of properties for aging soot in premixed flat flames studied by laser-induced incandescence and elastic light scattering. *Appl. Phys. B*, 119(4): 669–683, June 2015. ISSN 1432-0649. doi: 10.1007/s00340-015-6067-3.
- [157] Salma Bejaoui, Sébastien Batut, Eric Therssen, Nathalie Lamoureux, Pascale Desgroux, and Fengshan Liu. Measurements and modeling of laser-induced incandescence of soot at different heights in a flat premixed flame. *Appl. Phys. B*, 118(3):449–469, March 2015. ISSN 1432-0649. doi: 10.1007/s00340-015-6014-3.
- [158] S. P. Roy, P. G. Arias, V. R. Lecoustre, D. C. Haworth, H. G. Im, and A. Trouvé. Development of High Fidelity Soot Aerosol Dynamics Models using Method of Moments with Interpolative Closure. *Aerosol Sci. Technol.*, 48(4):379–391, April 2014. ISSN 0278-6826. doi: 10.1080/02786826.2013.878017.
- [159] Michael Balthasar and Michael Frenklach. Detailed kinetic modeling of soot aggregate formation in laminar premixed flames. *Combust. Flame*, 140(1):130–145, January 2005. ISSN 0010-2180. doi: 10.1016/j.combustflame.2004.11.004.
- [160] M E Mueller, G Blanquart, and H Pitsch. Hybrid Method of Moments for Modeling Soot Formation and Growth. *Combust. Flame*, 156:1143–1155, 2009. doi: 10.1016/j.combustflame.2009.01.025.
- [161] D L Marchisio and R O Fox. Solution of Population Balance Equations using the Direct Quadrature Method of Moments. *Aerosol Sci. Technol.*, 36:43–73, 2005.
- [162] Abhijit Kalbhor, Daniel Mira, and Jeroen van Oijen. A computationally efficient approach for soot modeling with discrete sectional method and FGM chemistry. *Combust. Flame*, 255:112868, September 2023. ISSN 0010-2180. doi: 10.1016/j.combustflame.2023.112868.
- [163] F Gelbard, Y Tambour, and J H Seinfeld. Sectional Representation for Simulating Aerosol Dynamics. *J. Colloid Interface Sci.*, 76:541–556, 1980.
- [164] J J Wu and R C Flagan. A Discrete Sectional Solution to the Aerosol Dynamic Equation. *J. Colloid Interface Sci.*, 123:339–352, 1988.
- [165] Somesh P Roy and Daniel C Haworth. A Systematic Comparison of Detailed Soot Models and Gas-phase Chemical Mechanisms in Laminar Premixed Flames. *Combustion Science and Technology*, 188(7):1021–1053, 2016. doi: 10.1080/00102202.2016.1145117.

- [166] Bin Zhao, Zhiwei Yang, Zhigang Li, Murray V. Johnston, and Hai Wang. Particle size distribution function of incipient soot in laminar premixed ethylene flames: effect of flame temperature. *Proc. Combust. Inst.*, 30(1):1441–1448, January 2005. ISSN 1540-7489. doi: 10.1016/j.proci.2004.08.104.
- [167] Junyu Mei, Yuxin Zhou, Xiaoqing You, and Chung K. Law. Formation of nascent soot during very fuel-rich oxidation of ethylene at low temperatures. *Combust. Flame*, 226:31–41, April 2021. ISSN 0010-2180. doi: 10.1016/j.combustflame.2020.11.031.
- [168] M. G. Kendall. A new measure of rank correlation. *Biometrika*, 30(1-2):81–93, June 1938. ISSN 0006-3444. doi: 10.1093/biomet/30.1-2.81.
- [169] James Theiler. Estimating fractal dimension. *J. Opt. Soc. Am. A, JOSAA*, 7(6):1055–1073, June 1990. ISSN 1520-8532. doi: 10.1364/JOSAA.7.001055.
- [170] S. R. Forrest and T. A. Witten, Jr. Long-range correlations in smoke-particle aggregates. *J. Phys. A: Math. Gen.*, 12(5):L109, May 1979. ISSN 0305-4470. doi: 10.1088/0305-4470/12/5/008.
- [171] Andres Suarez. Numerical simulation of multi-dimensional fractal soot aggregates. Master’s thesis, KTH, School of Engineering Sciences in Chemistry, Biotechnology and Health (CBH), 2018.
- [172] Rui Wang, Abhinandan Kumar Singh, Subash Reddy Kolan, and Evangelos Tsotsas. Fractal analysis of aggregates: Correlation between the 2D and 3D box-counting fractal dimension and power law fractal dimension. *Chaos, Solitons Fractals*, 160:112246, July 2022. ISSN 0960-0779. doi: 10.1016/j.chaos.2022.112246.
- [173] Yuanyuan Wang, Fengshan Liu, Cenlin He, Lei Bi, Tianhai Cheng, Zhili Wang, Hua Zhang, Xiaoye Zhang, Zongbo Shi, and Weijun Li. Fractal Dimensions and Mixing Structures of Soot Particles during Atmospheric Processing. *Environ. Sci. Technol. Lett.*, 4(11):487–493, November 2017. doi: 10.1021/acs.estlett.7b00418.
- [174] Peter F. DeCarlo, Jay G. Slowik, Douglas R. Worsnop, Paul Davidovits, and Jose L. Jimenez. Particle Morphology and Density Characterization by Combined Mobility and Aerodynamic Diameter Measurements. Part 1: Theory. *Aerosol Sci. Technol.*, 38(12):1185–1205, January 2004. ISSN 0278-6826. doi: 10.1080/027868290903907.
- [175] Georgios A. Kelesidis, Eirini Goudeli, and Sotiris E. Pratsinis. Flame synthesis of functional nanostructured materials and devices: Surface growth and aggregation. *Proc. Combust. Inst.*, 36(1):29–50, January 2017. ISSN 1540-7489. doi: 10.1016/j.proci.2016.08.078.
- [176] Georgios A. Kelesidis and Sotiris E. Pratsinis. Estimating the internal and surface oxidation of soot agglomerates. *Combust. Flame*, 209:493–499, November 2019. ISSN 0010-2180. doi: 10.1016/j.combustflame.2019.08.001.
- [177] Georgios A. Kelesidis, Christian A. Bruun, and Sotiris E. Pratsinis. The impact of organic carbon on soot light absorption. *Carbon*, 172:742–749, February 2021. ISSN

- 0008-6223. doi: 10.1016/j.carbon.2020.10.032.
- [178] Georgios A. Kelesidis and Sotiris E. Pratsinis. Santoro flame: The volume fraction of soot accounting for its morphology & composition. *Combust. Flame*, 240:112025, June 2022. ISSN 0010-2180. doi: 10.1016/j.combustflame.2022.112025.
 - [179] Michael Waskom, Olga Botvinnik, Drew O’Kane, Paul Hobson, Saulius Lukauskas, David C Gemperline, Tom Augspurger, Yaroslav Halchenko, John B. Cole, Jordi Warmenhoven, Julian de Ruiter, Stephan Hoyer, Jake Vanderplas, Santi Villalba, Gero Kunter, Eric Quintero, Pete Bachant, Marcel Martin, Kyle Meyer, Alistair Miles, Yoav Ram, Cameron Pye, Tal Yarkoni, Mike Lee Williams, Constantine Evans, Clark Fitzgerald, Brian, Chris Fonnesbeck, Antony Lee, and Adel Qalieh. mwaskom/seaborn: v0.8.0 (july 2017), July 2017. URL <https://doi.org/10.5281/zenodo.824567>.
 - [180] Barbara Apicella, P. Pré, J. N. Rouzaud, J. Abrahamson, R. L. Vander Wal, A. Ciajolo, A. Tregrossi, and C. Russo. Laser-induced structural modifications of differently aged soot investigated by HRTEM. *Combust. Flame*, 204:13–22, June 2019. ISSN 0010-2180. doi: 10.1016/j.combustflame.2019.02.026.
 - [181] Y Kobayashi, S Tanaka, and M Arai. Paks behavior and graphitization degree of soot in a hexane diffusion flame. *J Nanosci Nanotechnol*, 2:1–12, 2018.
 - [182] Maurin Salamanca, Maria L. Botero, Jacob W. Martin, Jochen A. H. Dreyer, Jethro Akroyd, and Markus Kraft. The impact of cyclic fuels on the formation and structure of soot. *Combust. Flame*, 219:1–12, September 2020. ISSN 0010-2180. doi: 10.1016/j.combustflame.2020.04.026.
 - [183] Patrizia Minutolo, Mario Commodo, and Andrea D’Anna. Optical properties of incipient soot. *Proc. Combust. Inst.*, October 2022. ISSN 1540-7489. doi: 10.1016/j.proci.2022.09.019.
 - [184] Hope A. Michelsen. Effects of maturity and temperature on soot density and specific heat. *Proc. Combust. Inst.*, 38(1):1197–1205, January 2021. ISSN 1540-7489. doi: 10.1016/j.proci.2020.06.383.
 - [185] Bert Brunekreef, Maciej Strak, Jie Chen, Zorana J. Andersen, Richard Atkinson, Mariska Bauwelink, Tom Bellander, Marie-Christine Boutron, Jorgen Brandt, Iain Carey, Giulia Cesaroni, Francesco Forastiere, Daniela Fecht, John Gulliver, Ole Hertel, Barbara Hoffmann, Kees de Hoogh, Danny Houthuijs, Ulla Hvidtfeldt, Nicole Janssen, Jeanette Jorgensen, Klea Katsouyanni, Matthias Ketzel, Jochem Klompmaker, Norun Hjertager Krog, Shuo Liu, Petter Ljungman, Amar Mehta, Gabriele Nagel, Bente Oftedal, Goran Pershagen, Annette Peters, Ole Raaschou-Nielsen, Matteo Renzi, Sophia Rodopoulou, Evi Samoli, Per Schwarze, Torben Sigsgaard, Massimo Stafoggia, Danielle Vienneau, Gudrun Weinmayr, Kathrin Wolf, and Gerard Hoek. Mortality and Morbidity Effects of Long-Term Exposure to Low-Level PM_{2.5}, BC, NO₂, and O₃: An Analysis of European Cohorts in the ELAPSE Project. *Research Report (Health Effects Institute)*, 2021 (Research Report

- (Health Effects Institute), 208):1–127, 2021. URL <https://pubmed.ncbi.nlm.nih.gov/36106702>. PMID: 36106702.
- [186] A. R. Chughtai, M. M. O. Atteya, J. Kim, B. K. Konowalchuk, and D. M. Smith. Adsorption and adsorbate interaction at soot particle surfaces. *Carbon*, 36(11):1573–1589, November 1998. ISSN 0008-6223. doi: 10.1016/S0008-6223(98)00116-X.
 - [187] Olga Popovicheva, Natalia M. Persiantseva, Natalia K. Shonija, Paul DeMott, Kirsten Koehler, Markus Petters, Sonia Kreidenweis, Victoria Tishkova, Benjamin Demirdjian, and Jean Suzanne. Water interaction with hydrophobic and hydrophilic soot particles. *Phys. Chem. Chem. Phys.*, 10(17):2332–2344, April 2008. ISSN 1463-9076. doi: 10.1039/B718944N.
 - [188] Maria Eugenia Monge, Barbara D’Anna, Linda Mazri, Anne Giroir-Fendler, Markus Ammann, D. J. Donaldson, and Christian George. Light changes the atmospheric reactivity of soot. *Proc. Natl. Acad. Sci. U.S.A.*, 107(15):6605–6609, April 2010. doi: 10.1073/pnas.0908341107.
 - [189] Peter Pfeifer. Fractal dimension as working tool for surface-roughness problems. *Applications of Surface Science*, 18(1):146–164, May 1984. ISSN 0378-5963. doi: 10.1016/0378-5963(84)90042-4.
 - [190] Zheng Li, Chonglin Song, Jinou Song, Gang Lv, Surong Dong, and Zhuang Zhao. Evolution of the nanostructure, fractal dimension and size of in-cylinder soot during diesel combustion process. *Combust. Flame*, 158(8):1624–1630, August 2011. ISSN 0010-2180. doi: 10.1016/j.combustflame.2010.12.006.
 - [191] Georgios Skillas, Stefan Künzel, Heinz Burtscher, Urs Baltensperger, and Konstantin Siegmann. High fractal-like dimension of diesel soot agglomerates. *J. Aerosol Sci.*, 29(4):411–419, April 1998. ISSN 0021-8502. doi: 10.1016/S0021-8502(97)00448-5.
 - [192] Li Liu, Michael I. Mishchenko, and W. Patrick Arnott. A study of radiative properties of fractal soot aggregates using the superposition T-matrix method. *J. Quant. Spectrosc. Radiat. Transfer*, 109(15):2656–2663, October 2008. ISSN 0022-4073. doi: 10.1016/j.jqsrt.2008.05.001.
 - [193] Bing Hu and Umit Koylu. Size and Morphology of Soot Particulates Sampled from a Turbulent Nonpremixed Acetylene Flame. *Aerosol Sci. Technol.*, October 2004. ISSN 1038-5861. URL <https://www.tandfonline.com/doi/abs/10.1080/027868290519111>.
 - [194] Pranav K. Tripathi, Mingxian Liu, Yunhui Zhao, Xiaomei Ma, Lihua Gan, Owen Noonan, and Chengzhong Yu. Enlargement of uniform micropores in hierarchically ordered micro–mesoporous carbon for high level decontamination of bisphenol A. *J. Mater. Chem. A*, 2(22):8534–8544, May 2014. ISSN 2050-7488. doi: 10.1039/C4TA00578C.
 - [195] Atar Singh Pipal, Himanshi Rohra, Rahul Tiwari, and Ajay Taneja. Particle size distribution, morphometric study and mixing structure of accumulation and ultrafine

- aerosols emitted from indoor activities in different socioeconomic micro-environment. *Atmos. Pollut. Res.*, 12(4):101–111, April 2021. ISSN 1309-1042. doi: 10.1016/j.apr.2021.02.015.
- [196] Borislav D. Zdravkov, Jiří J. Čermák, Martin Šefara, and Josef Janků. Pore classification in the characterization of porous materials: A perspective. *Cent. Eur. J. Chem.*, 5(2):385–395, June 2007. ISSN 1644-3624. doi: 10.2478/s11532-007-0017-9.
 - [197] J. Rouquerol, D. Avnir, C. W. Fairbridge, D. H. Everett, J. M. Haynes, N. Pernicone, J. D. F. Ramsay, K. S. W. Sing, and K. K. Unger. Recommendations for the characterization of porous solids (Technical Report). *Pure Appl. Chem.*, 66(8): 1739–1758, 1994. ISSN 0033-4545. doi: 10.1351/pac199466081739.
 - [198] Jerzy Choma, Jacek Jagiello, and Mietek Jaroniec. Assessing the contribution of micropores and mesopores from nitrogen adsorption on nanoporous carbons: Application to pore size analysis. *Carbon*, 183:150–157, October 2021. ISSN 0008-6223. doi: 10.1016/j.carbon.2021.07.020.
 - [199] Guochang Wang and Yiwen Ju. Organic shale micropore and mesopore structure characterization by ultra-low pressure N₂ physisorption: Experimental procedure and interpretation model. *J. Nat. Gas Sci. Eng.*, 27:452–465, November 2015. ISSN 1875-5100. doi: 10.1016/j.jngse.2015.08.003.
 - [200] Karl J. Rockne, Gary L. Taghon, and David S. Kosson. Pore structure of soot deposits from several combustion sources. *Chemosphere*, 41(8):1125–1135, October 2000. ISSN 0045-6535. doi: 10.1016/S0045-6535(00)00040-0.
 - [201] Tatsuhiko Miyata, Akira Endo, Takao Ohmori, Takaji Akiya, and Masaru Nakaiwa. Evaluation of pore size distribution in boundary region of micropore and mesopore using gas adsorption method. *J. Colloid Interface Sci.*, 262(1):116–125, June 2003. ISSN 0021-9797. doi: 10.1016/S0021-9797(02)00254-0.
 - [202] Urs Mathis, Martin Mohr, Ralf Kaegi, Andrea Bertola, and Konstantinos Boulouchos. Influence of Diesel Engine Combustion Parameters on Primary Soot Particle Diameter. *Environ. Sci. Technol.*, 39(6):1887–1892, March 2005. ISSN 0013-936X. doi: 10.1021/es049578p.
 - [203] Hao Jiang, Tie Li, Yifeng Wang, and Pengfei He. Morphology and nano-structure analysis of soot particles sampled from high pressure diesel jet flames under diesel-like conditions. *Meas. Sci. Technol.*, 29(4):045801, March 2018. ISSN 0957-0233. doi: 10.1088/1361-6501/aaa667.
 - [204] A. Mészáros. The number concentration and size distribution of the soot particles in the 0.02–0.5 μm radius range at sites of different pollution levels. *Sci. Total Environ.*, 36:283–288, July 1984. ISSN 0048-9697. doi: 10.1016/0048-9697(84)90278-X.
 - [205] N. P. Ivleva, U. McKeon, R. Niessner, and U. Pöschl. Raman Microspectroscopic Analysis of Size-Resolved Atmospheric Aerosol Particle Samples Collected with an ELPI: Soot, Humic-Like Substances, and Inorganic Compounds. *Aerosol Sci.*

- Technol.*, June 2007. URL
<https://www.tandfonline.com/doi/full/10.1080/02786820701376391>.
- [206] H. A. Michelsen. Probing soot formation, chemical and physical evolution, and oxidation: A review of in situ diagnostic techniques and needs. *Proc. Combust. Inst.*, 36(1):717–735, January 2017. ISSN 1540-7489. doi: 10.1016/j.proci.2016.08.027.
 - [207] T. M. Gruenberger, M. Moghiman, P. J. Bowen, and N. Syred. Dynamics of soot formation by turbulent combustion and thermal decomposition of natural gas. *Combust. Sci. Technol.*, May 2002. ISSN 1008-0547. URL
<https://www.tandfonline.com/doi/abs/10.1080/713713038>.
 - [208] J. B. Maglic and R. Lavendomme. MoloVol: an easy-to-use program for analyzing cavities, volumes and surface areas of chemical structures. *J. Appl. Crystallogr.*, 55(4): 1033–1044, August 2022. ISSN 1600-5767. doi: 10.1107/S1600576722004988.
 - [209] LLC Schrödinger and Warren DeLano. Pymol, Jun 2022. URL
<http://www.pymol.org/pymol>. [Online; accessed 5. Jun. 2022].
 - [210] StereoFractAnalyzer, October 2024. URL
<https://github.com/comp-comb/StereoFractAnalyzer>. [Online; accessed 3. Oct. 2024].
 - [211] Santiago Alvarez. A cartography of the van der Waals territories. *Dalton Trans.*, 42 (24):8617–8636, May 2013. ISSN 1477-9226. doi: 10.1039/C3DT50599E.
 - [212] Antonio Giorgilli, Danilo Casati, Luisella Sironi, and Luigi Galgani. An efficient procedure to compute fractal dimensions by box counting. *Phys. Lett. A*, 115(5): 202–206, April 1986. ISSN 0375-9601. doi: 10.1016/0375-9601(86)90465-2.
 - [213] Measuring the fractal dimensions of surfaces, January 2024. URL
<https://apps.dtic.mil/sti/tr/pdf/ADA129664.pdf>. [Online; accessed 5. Oct. 2024].
 - [214] Keith C. Clarke. Computation of the fractal dimension of topographic surfaces using the triangular prism surface area method. *Comput. Geosci.*, 12(5):713–722, January 1986. ISSN 0098-3004. doi: 10.1016/0098-3004(86)90047-6.
 - [215] Dong Kyu Lee, Junyong In, and Sangseok Lee. Standard deviation and standard error of the mean. *Korean Journal of Anesthesiology*, 68(3):220, June 2015. doi: 10.4097/kjae.2015.68.3.220.
 - [216] Paul D. Teini, Darshan M. A. Karwat, and Arvind Atreya. Observations of nascent soot: Molecular deposition and particle morphology. *Combust. Flame*, 158(10): 2045–2055, October 2011. ISSN 0010-2180. doi: 10.1016/j.combustflame.2011.03.005.
 - [217] Maria L. Botero, Dongping Chen, Silvia González-Calera, David Jefferson, and Markus Kraft. HRTEM evaluation of soot particles produced by the non-premixed combustion of liquid fuels. *Carbon*, 96:459–473, January 2016. ISSN 0008-6223. doi:

10.1016/j.carbon.2015.09.077.

- [218] Joaquin Camacho, Yujie Tao, and Hai Wang. Kinetics of nascent soot oxidation by molecular oxygen in a flow reactor. *Proc. Combust. Inst.*, 35(2):1887–1894, January 2015. ISSN 1540-7489. doi: 10.1016/j.proci.2014.05.095.
- [219] Angela Violi. Modeling of soot particle inception in aromatic and aliphatic premixed flames. *Combust. Flame*, 139(4):279–287, December 2004. ISSN 0010-2180. doi: 10.1016/j.combustflame.2004.08.013.
- [220] Marina Schenk, Sydnie Lieb, Henning Vieker, André Beyer, Armin Götzhäuser, Hai Wang, and Katharina Kohse-Höinghaus. Morphology of nascent soot in ethylene flames. *Proc. Combust. Inst.*, 35(2):1879–1886, January 2015. ISSN 1540-7489. doi: 10.1016/j.proci.2014.05.009.
- [221] John R. Grace and Arian Ebneyamini. Connecting particle sphericity and circularity. *Particuology*, 54:1–4, February 2021. ISSN 1674-2001. doi: 10.1016/j.partic.2020.09.006.
- [222] Qiang Song, Bailei He, Qiang Yao, Zhongwei Meng, and Changhe Chen. Influence of Diffusion on Thermogravimetric Analysis of Carbon Black Oxidation. *Energy Fuels*, 20(5):1895–1900, September 2006. ISSN 0887-0624. doi: 10.1021/ef0600659.
- [223] F.-X. Ouf, S. Bourrous, C. Vallières, J. Yon, and L. Lintis. Specific surface area of combustion emitted particles: Impact of primary particle diameter and organic content. *J. Aerosol Sci.*, 137:105436, November 2019. ISSN 0021-8502. doi: 10.1016/j.jaerosci.2019.105436.
- [224] Juntong Sun, Jin Niu, Mengyue Liu, Jing Ji, Meiling Dou, and Feng Wang. Biomass-derived nitrogen-doped porous carbons with tailored hierarchical porosity and high specific surface area for high energy and power density supercapacitors. *Appl. Surf. Sci.*, 427:807–813, January 2018. ISSN 0169-4332. doi: 10.1016/j.apsusc.2017.07.220.
- [225] Danmiao Kang, Qinglei Liu, Jiajun Gu, Yishi Su, Wang Zhang, and Di Zhang. “Egg-Box”-Assisted Fabrication of Porous Carbon with Small Mesopores for High-Rate Electric Double Layer Capacitors. *ACS Nano*, 9(11):11225–11233, November 2015. ISSN 1936-0851. doi: 10.1021/acs.nano.5b04821.
- [226] Jet-Sing M. Lee, Michael E. Briggs, Tom Hasell, and Andrew I. Cooper. Hyperporous Carbons from Hypercrosslinked Polymers. *Adv. Mater.*, 28(44):9804–9810, November 2016. ISSN 0935-9648. doi: 10.1002/adma.201603051.
- [227] C. M. Sorensen. Light Scattering by Fractal Aggregates: A Review. *Aerosol Sci. Technol.*, January 2001. ISSN 1038-4153. URL <https://www.tandfonline.com/doi/abs/10.1080/02786820117868>.
- [228] David Avnir, Dina Farin, and Peter Pfeifer. Molecular fractal surfaces. *Nature*, 308: 261–263, March 1984. ISSN 1476-4687. doi: 10.1038/308261a0.

- [229] David Avnir, Dina Farin, and Peter Pfeifer. Chemistry in noninteger dimensions between two and three. II. Fractal surfaces of adsorbents. *J. Chem. Phys.*, 79(7): 3566–3571, October 1983. ISSN 0021-9606. doi: 10.1063/1.446211.
- [230] Christian Ahl and Jürgen Niemeyer. The fractal dimension of the pore-volume inside soils. *Z. Pflanzenernähr. Bodenkd.*, 152(5):457–458, January 1989. ISSN 0044-3263. doi: 10.1002/jpln.19891520512.
- [231] H. P. Tang, J. Z. Wang, J. L. Zhu, Q. B. Ao, J. Y. Wang, B. J. Yang, and Y. N. Li. Fractal dimension of pore-structure of porous metal materials made by stainless steel powder. *Powder Technol.*, 217:383–387, February 2012. ISSN 0032-5910. doi: 10.1016/j.powtec.2011.10.053.
- [232] Angelo W. Kandas, I. Gokhan Senel, Yiannis Levendis, and Adel F. Sarofim. Soot surface area evolution during air oxidation as evaluated by small angle X-ray scattering and CO₂ adsorption. *Carbon*, 43(2):241–251, January 2005. ISSN 0008-6223. doi: 10.1016/j.carbon.2004.08.028.
- [233] Françoise Ehrburger-Dolle, Manfred Holz, and J. Lahaye. Use of N₂, Ar and CO₂ adsorption for the determination of microporosity and surface fractal dimension of carbon blacks and silicas. *Pure Appl. Chem.*, 65(10):2223–2230, January 1993. ISSN 1365-3075. doi: 10.1351/pac199365102223.
- [234] Georgios A. Kelesidis, Nicola Rossi, and Sotiris E. Pratsinis. Porosity and crystallinity dynamics of carbon black during internal and surface oxidation. *Carbon*, 197:334–340, September 2022. ISSN 0008-6223. doi: 10.1016/j.carbon.2022.06.020.
- [235] Viktória Hegedűs, Farkas Kerényi, Róbert Boda, Dóra Horváth, István Lázár, Enikő Tóth-Györi, Balázs Dezső, and Csaba Hegedus. β -Tricalcium phosphate-silica aerogel as an alternative bioactive ceramic for the potential use in dentistry. *Adv. Appl. Ceram.*, 119(5-6):364–371, July 2020. ISSN 1743-6753. doi: 10.1080/17436753.2019.1625567.
- [236] Hakan Demiral and İlknur Demiral. Preparation and characterization of carbon molecular sieves from chestnut shell by chemical vapor deposition. *Adv. Powder Technol.*, 29(12):3033–3039, December 2018. ISSN 0921-8831. doi: 10.1016/j.appt.2018.07.015.
- [237] Puneet Verma, Mohammad Jafari, S. M. Ashrafur Rahman, Edmund Pickering, Svetlana Stevanovic, Ashley Dowell, Richard Brown, and Zoran Ristovski. The impact of chemical composition of oxygenated fuels on morphology and nanostructure of soot particles. *Fuel*, 259:116167, January 2020. ISSN 0016-2361. doi: 10.1016/j.fuel.2019.116167.
- [238] Puneet Verma, Edmund Pickering, Mohammad Jafari, Yi Guo, Svetlana Stevanovic, Joseph F. S. Fernando, Dmitri Golberg, Peter Brooks, Richard Brown, and Zoran Ristovski. Influence of fuel-oxygen content on morphology and nanostructure of soot particles. *Combust. Flame*, 205:206–219, July 2019. ISSN 0010-2180. doi: 10.1016/j.combustflame.2019.04.009.

- [239] Décio S. Bento, Kevin A. Thomson, and Ömer L. Gülder. Soot formation and temperature field structure in laminar propane–air diffusion flames at elevated pressures. *Combust. Flame*, 145(4):765–778, June 2006. ISSN 0010-2180. doi: 10.1016/j.combustflame.2006.01.010.
- [240] Ahmet E. Karataş and Ömer L. Gülder. Soot formation in high pressure laminar diffusion flames. *Prog. Energy Combust. Sci.*, 38(6):818–845, December 2012. ISSN 0360-1285. doi: 10.1016/j.pecs.2012.04.003.
- [241] Dezhi Zhou and Suo Yang. Soot-based Global Pathway Analysis: Soot formation and evolution at elevated pressures in co-flow diffusion flames. *Combust. Flame*, 227: 255–270, May 2021. ISSN 0010-2180. doi: 10.1016/j.combustflame.2021.01.007.
- [242] Mario Commодо, Ahmet E. Karataş, Gianluigi De Falco, Patrizia Minutolo, Andrea D’Anna, and Ömer L. Gülder. On the effect of pressure on soot nanostructure: A Raman spectroscopy investigation. *Combust. Flame*, 219:13–19, September 2020. ISSN 0010-2180. doi: 10.1016/j.combustflame.2020.04.008.
- [243] J. Morán, A. Poux, F. Cepeda, F. Escudero, A. Fuentes, L. Gallen, E. Riber, B. Cuenot, and J. Yon. Multi-scale soot formation simulation providing detailed particle morphology in a laminar coflow diffusion flame. *Combust. Flame*, 256:112987, October 2023. ISSN 0010-2180. doi: 10.1016/j.combustflame.2023.112987.
- [244] Robert L. Murry, Douglas L. Strout, Gregory K. Odom, and Gustavo E. Scuseria. Role of sp³ carbon and 7-membered rings in fullerene annealing and fragmentation. *Nature*, 366:665–667, December 1993. ISSN 1476-4687. doi: 10.1038/366665a0.
- [245] Sumio Iijima, Toshinari Ichihashi, and Yoshinori Ando. Pentagons, heptagons and negative curvature in graphite microtubule growth. *Nature*, 356:776–778, April 1992. ISSN 1476-4687. doi: 10.1038/356776a0.
- [246] Jacob W. Martin, Maurin Salamanca, and Markus Kraft. Soot inception: Carbonaceous nanoparticle formation in flames. *Prog. Energy Combust. Sci.*, 88: 100956, January 2022. ISSN 0360-1285. doi: 10.1016/j.pecs.2021.100956.
- [247] Jie Ma, Dario Alfè, Angelos Michaelides, and Enge Wang. Stone-Wales defects in graphene and other planar sp²-bonded materials. *Phys. Rev. B*, 80(3):033407, July 2009. doi: 10.1103/PhysRevB.80.033407.
- [248] J. Kotakoski, J. C. Meyer, S. Kurasch, D. Santos-Cottin, U. Kaiser, and A. V. Krasheninnikov. Stone-Wales-type transformations in carbon nanostructures driven by electron irradiation. *Phys. Rev. B*, 83(24):245420, June 2011. doi: 10.1103/PhysRevB.83.245420.
- [249] Izaac Mitchell, Lu Qiu, Lowell D. Lamb, and Feng Ding. High Temperature Accelerated Stone–Wales Transformation and the Threshold Temperature of IPR-C60 Formation. *J. Phys. Chem. A*, 125(21):4548–4557, June 2021. ISSN 1089-5639. doi: 10.1021/acs.jpca.1c02151.

- [250] Chen-Li Zhang and Hui-Shen Shen. Self-healing in defective carbon nanotubes: a molecular dynamics study. *J. Phys.: Condens. Matter*, 19(38):386212, August 2007. ISSN 0953-8984. doi: 10.1088/0953-8984/19/38/386212.
- [251] Michael Frenklach and Alexander M. Mebel. On the Mechanism of Soot Nucleation. IV. Molecular Growth of the Flattened E-Bridge. *J. Phys. Chem. A*, 126(49): 9259–9267, December 2022. ISSN 1089-5639. doi: 10.1021/acs.jpca.2c06819.
- [252] Rui Wang, Jinjing Li, and Yanming Wang. Atomistic Mechanisms of Ring Formation during Catalyzed Carbon Nanotube Growth. *J. Phys. Chem. C*, 128(12):5112–5119, March 2024. ISSN 1932-7447. doi: 10.1021/acs.jpcc.3c05694.

APPENDIX A

SYMBOLS, NOMENCLATURE, AND DEFINITIONS USED IN CHAPTER 4

A.1 Symbols and nomenclature

a : A parameter in Eqn A.4

A : Surface area of a particle (\AA^2)

c : A parameter in Eqn A.4

D_f : Atomic fractal dimension of a particle

$m_{p,i}$: Mass of i^{th} particle (kg)

M : Molar mass of a particle (kg/kmol)

M_p : Mass of a particle (kg)

n_{\odot} : Number of cyclic carbon atoms at a specific location

n_{\P} : Number of non-cyclic carbon atoms at a specific location

N : Total number of atoms in the entire particle

N_{\odot} : Number of rings in the entire particle

N_5 : Number of 5-membered rings in the entire particle

N_6 : Number of 6-membered rings in the entire particle

N_7 : Number of 7-membered rings in the entire particle

N_C : Number of carbon atoms in the entire particle

N_H : Number of hydrogen atoms in the entire particle

N_{\odot} : Number of cyclic carbon atoms in the entire particle

N_{\P} : Number of non-cyclic carbon atoms in the entire particle

r : Local radius (\AA)

R_{eq} : Volume equivalent radius (\AA)

R_g : Radius of gyration (\AA)

ρ_s : Simulated density of a particle (kg/m^3)

ρ_e : Empirical density of a particle (kg/m^3)

ϱ : Local (actual) density (kg/m^3)

T : Temperature (K)

$\theta_{C/H}$: Local C/H ratio

$\Theta_{C/H}$: C/H ratio of the entire particle

V : Volume of a particle (\AA^3)

w_C : Mass of a carbon atom (kg)

w_H : Mass of a hydrogen atom (kg)

Superscripts:

*: Denotes normalized value

//: Denotes per unit area value

A.2 Expressions for physical properties of soot particles

The trajectory files obtained from RMD simulations contain coordinates of each atom with reference to a global reference frame. This coordinate information along with the mass of each atom is used to calculate the coordinate of the center of mass of each particle. The mass of a particle M_p is calculated by summing up the mass of the atoms in the cluster. The volume (V) is calculated using MSMS [136] with a pore size of 1.5 \AA . The volume equivalent radius of a particle with volume V is calculated via Eqn. A.1.

$$R_{eq} = \left(\frac{3V}{4\pi} \right)^{1/3} \quad (\text{A.1})$$

The radius of gyration (R_g) is calculated following the standard definition using Eqn. A.2.

$$R_g = \sqrt{\frac{\sum_{i=1}^N m_{p,i} r_i^2}{\sum_{i=1}^N m_{p,i}}}, \quad (\text{A.2})$$

where r_i is the distance of the i^{th} atom from the center of mass, $m_{p,i}$ is the mass of individual atoms, and N is the total number of atoms in the cluster.

The simulated density (ρ_s) is calculated using the particle mass (M_p) and volume (V) of the incipient particle using Eqn. A.3.

$$\rho_s = \frac{M_p}{V} \quad (\text{A.3})$$

Empirical (bulk) density [67, 174] of an incipient particle is calculated using Eqn. A.4.

$$\rho_e = (0.260884a^2c)^{-1} \left(\frac{w_C \Theta_{C/H} + w_H}{\Theta_{C/H} + 1} \right), \quad (\text{A.4})$$

where w_C and w_H are the molar masses of a carbon and hydrogen atoms, a is the length of the graphite unite cell in the basal plane, c is the interlayer spacing in Angstroms, and $\Theta_{C/H}$ represents the carbon to hydrogen ratio of the cluster. More details can be found in [67, 174].

The atomic fractal dimension (D_f), following the approach used in [57], is calculated using the sandbox method [169, 170] using Eqn. A.5.

$$D_f = \frac{\log M_p(r)}{\log r}, \quad (\text{A.5})$$

where $M_p(r)$ is the mass of atoms in the cluster as a function of radial distance from the center of mass. Please note that this “atomic” fractal dimension is for a single incipient particle and is different from the traditional fractal dimension used in aggregate characterization [57].

A.3 Physicochemical data used and analyzed in this study

A.3.1 Feature set

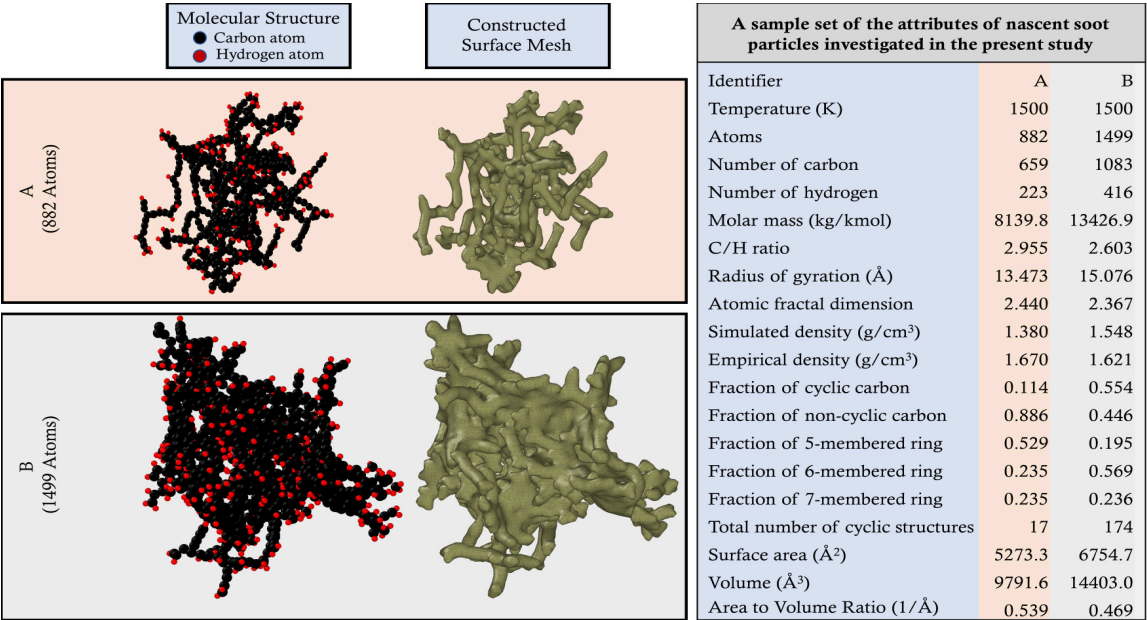
The feature set used in this study for each particle includes the following

1. Temperature (T)
2. Number of carbon atoms (N_C)
3. Number of hydrogen atoms (N_H)
4. Number of atoms (N)
5. Molar mass (M)
6. C/H ratio ($\Theta_{C/H}$)
7. Radius of gyration (R_g)
8. Atomic fractal dimension (D_f)
9. Simulated density (ρ_s)

10. Empirical density (ρ_e) (also referred to as the bulk density in literature)
11. Total number of cyclic structures (N_{O})
12. Fraction of cyclic carbon atoms (N_{O}/N_C)
13. Fraction of 5-member rings (N_5/N_{O})
14. Fraction of 6-member rings (N_6/N_{O})
15. Fraction of 7-member rings (N_7/N_{O})
16. Surface area (A)
17. Volume (V)
18. Area to volume ratio (A/V)

A.3.2 Sample data

Figure A.1 shows two sample soot clusters and their properties as examples. These clusters, labeled as A and B, were extracted from a simulation at 1500 K at two different times. The left side shows the molecular structure of the particle and the three-dimensional volumetric representation by constructing a surface mesh using OVITO [137] (this is what the incipient particles would actually look like). The physicochemical properties of these particles as analyzed in this work for classification via machine-learning are tabulated on the right side of the figure.



APPENDIX B

SYMBOLS AND DEFINITIONS USED IN CHAPTER 6

B.1 Mathematical Equations

Below is the collection of equations used throughout the manuscript, arranged in the order they are discussed in the main text.

Sphericity (Ψ , dimensionless) is calculated using

$$\Psi = \frac{\pi^{1/3}(6 \times V_B)^{2/3}}{A_{Ext}} \quad (\text{B.1})$$

where V_B is the bulk volume of the particle (\AA^3), and A_{Ext} is the external surface area (\AA^2).

Circularity (σ , dimensionless) and average circularity ($\bar{\sigma}$, dimensionless) are calculated as

$$\sigma = \frac{P_c}{P} \quad (\text{B.2})$$

$$\bar{\sigma} = \frac{1}{N} \sum_{i=1}^N \left(\frac{P_c}{P} \right)_i, \quad N = 10 \quad (\text{B.3})$$

where P_c is the perimeter of a circle of the same projected area, P is the actual perimeter of the projected area. Average circularity is calculated over 10 projections.

Porosity (Φ , dimensionless) is calculated using

$$\Phi = \frac{V_p}{V_B} \quad (\text{B.4})$$

where V_p is the total pore volume (\AA^3), and V_B is the bulk volume of the particle (\AA^3).

Theoretical specific surface area ($SSA_{Theoretical}$, m^2/g) is evaluated using

$$\begin{aligned} SSA_{Theoretical} &= \frac{\text{Surface area}}{\text{Bulk Volume} \times \text{Bulk Density}} \\ &= \frac{\pi d_V^2}{\frac{1}{6} \pi d_V^3 \times \rho_b} \\ &= \frac{6}{d_V \times \rho_b} \end{aligned} \quad (\text{B.5})$$

where d_V is the volume-equivalent diameter (\AA), and ρ_b is the bulk density (g/cm^3).

Average pore sphericity ($\bar{\psi}$, dimensionless) for different types of pores are calculated using

$$\bar{\psi} = \frac{1}{N_{particles}} \sum_i^{N_{particles}} \frac{1}{n_i} \sum_j^{n_i} \psi_j \quad (\text{B.6})$$

where $N_{particles}$ is the total number of particles, n_i is the number of cavities in particle i , and ψ_j is the sphericity of the j^{th} cavity in particle i .

Cumulative void volume per unit mass ($V_C(d_p)$, cm^3/g) is caculated using

$$V_C(d_p) = \sum_i^{N_{particle}} \frac{1}{M_i} \sum_j^{n_i} v_{c,j} \mathcal{H}(d_{p,j} - d_p) \quad (\text{B.7})$$

where M_i is mass of particle i , and \mathcal{H} is the Heaviside step function, and d_p represents the pore size.

B.2 List of Symbols and Definitions

Table B.1: List of symbols and their definitions used in the main text.

Symbol	Definition	Scope	Equation No.
Ψ	Sphericity of an incipient particle	Particle	B.1
ψ	Sphericity of a single pore (cavity)	Cavity	B.6
Φ	Porosity of an incipient particle	Particle	B.4
σ	Circularity of an incipient particle	Particle	B.2
D_S	Surface fractal dimension of an incipient particle	Particle	
D_V	Volume fractal dimension of an incipient particle	Particle	
D_f	Aggregate fractal dimension of a soot aggregate	N/A	
A_p	Total pore surface area (\AA^2) within a particle	Particle	6.1
V_p	Total pore volume (\AA^3) in a particle	Particle	6.1, B.4
A_{Ext}	External surface area (\AA^2) of a particle	Particle	B.1
A_{pi}	Surface area of the i^{th} pore (\AA^2)	Cavity	6.3
V_{pi}	Volume of the i^{th} pore (\AA^3)	Cavity	6.3
V	Material volume of an incipient particle (\AA^3)	Particle	6.2
V_B	Bulk volume of the particle (\AA^3) including void volume (i.e., $V + V_p$)	Particle	B.1, B.4
ρ_b	Bulk density (g/cm^3) of an incipient particle (calculated using bulk volume and mass of a particle)	Particle	B.5
d_V	Volume-equivalent diameter (\AA) of an incipient particle (based on bulk volume)	Particle	B.5
$SSA_{Theoretical}$	Theoretical specific surface area (m^2/g) of an incipient particle	Particle	B.5
D_{SC}	Surface fractal dimension of a cavity	Cavity	
D_{VC}	Volume fractal dimension of a cavity	Cavity	
ψ_{Tunnel}	Average pore sphericity for tunnel cavities in an ensemble of incipient particles	Ensemble	B.6
ψ_{Pocket}	Average pore sphericity for pocket cavities in an ensemble of incipient particles	Ensemble	B.6
$\psi_{Isolated}$	Average pore sphericity for isolated cavities in an ensemble of incipient particles	Ensemble	B.6
d_p	Volume-equivalent diameter of a cavity (\AA)	Cavity	B.7
$V_C(d_p)$	Cumulative void volume per unit mass (cm^3/g) within an incipient particle up to cavity size d_p	Particle	B.7
$\overline{d_p}$	Average volume-equivalent diameter of cavities within an incipient particle	Particle	

B.3 Definitions and Key Parameters

- **Types of Cavities:**

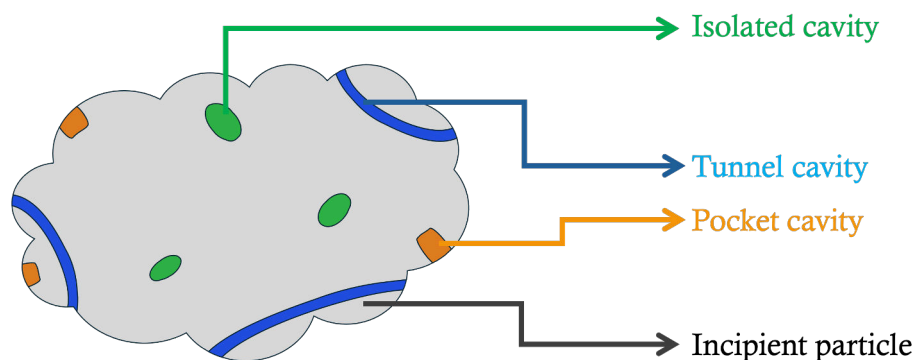


Figure B.1: Types of cavities in soot particles.

- **Tunnel Cavity:** A pore with two or more openings to the external surface, allowing gas flow and enhancing mass transport.
- **Pocket Cavity:** A pore with a single opening to the external surface, forming a semi-enclosed void that contributes to accessible surface area.
- **Isolated Cavity:** A completely enclosed void with no external openings, contributing only to bulk porosity.

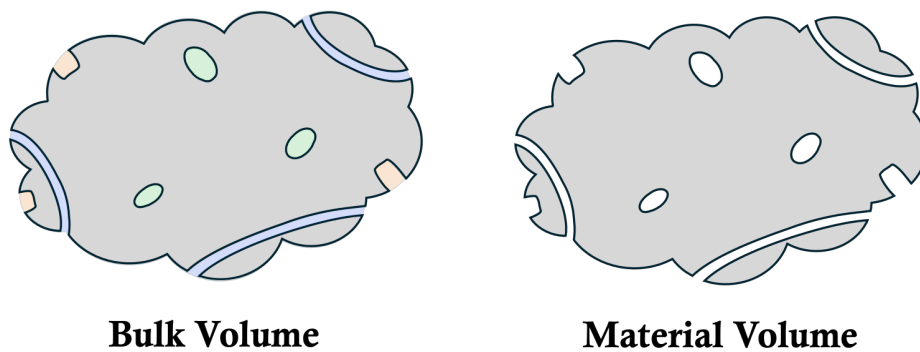


Figure B.2: Material and Bulk Volumes of a Soot Particle.

- **Material Volume (V):** The volume of a particle that is completely composed of solid material, independent of porosity contributions.
- **Bulk Volume (V_B):** The total volume occupied by a particle, considering both solid material and void spaces.
- **Total Pore Volume (V_p):** The cumulative internal void volume of all pores in a particle. This is the difference between the bulk volume and the material volume, i.e. $V_p = V_B - V$.

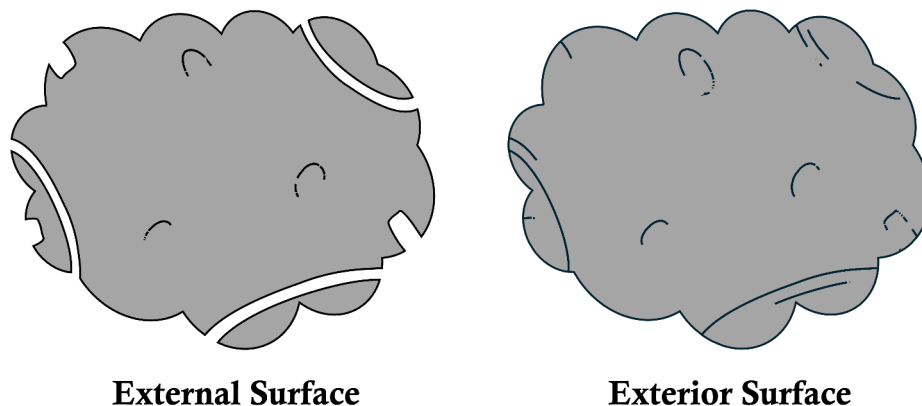


Figure B.3: External and Exterior Surface Areas of a Soot Particle.

- **External Surface Area (A_{Ext}):** The surface area of a particle that is directly exposed to the external environment.
- **Exterior Surface Area ($A_{Exterior}$):** The surface area of a particle that is directly exposed to the external environment but not part of any cavities.
- **Specific Surface Area (SSA):** The total surface area available per unit mass of a particle, significantly influenced by the presence of tunnel and pocket cavities.
- **Surface Fractal Dimension (D_S):** A measure of the complexity of the particle's surface, affecting its reactivity and adsorption capacity.
- **Volume Fractal Dimension (D_V):** Describes the self-similar distribution of mass inside a particle, influencing pore structure and internal morphology.
- **Aggregate Fractal Dimension (D_f):** Characterizes the structural arrangement of primary particles in a soot aggregate, following a statistical fractal relationship.

- **Porosity (Φ):** Defined as the ratio of the total pore volume (V_p) inside a particle to the bulk volume (V_B), representing the degree of void spaces.
- **Cumulative Void Volume (V_c):** The accumulated total volume of void spaces within a particle up to a given cavity size, often used in pore size distribution analysis.
- **Volume-Equivalent Diameter (d_p):** Represents the diameter of a sphere with the same volume as a given cavity, providing a measure of pore size.
- **Bulk Density (ρ_b):** The mass per unit volume (including any void) of a particle, relevant for evaluating particle compactness.
- **Pore Size Distribution (PSD):** A statistical distribution describing how pore sizes are spread within a particle, affecting gas diffusion and reactivity.
- **Sphericity (Ψ):** A dimensionless measure of how closely the shape of an incipient soot particle resembles a sphere. Lower values indicate more irregular shapes.
- **Pore Sphericity (ψ):** A dimensionless measure of the regularity of a pore's shape.
- **Circularity (σ):** A metric quantifying how close the projection of a soot particle is to a perfect circle, based on its 2D projection.
- **Heaviside Step Function (H):** A mathematical function used to define cavity distributions in pore size calculations.
- **Surface Area of an Individual Pore (A_{pi}):** The total exposed surface area of a single (i^{th}) pore.
- **Volume of an Individual Pore (V_{pi}):** The total enclosed volume of a single (i^{th}) pore.
- **Total Pore Surface Area (A_p):** The cumulative internal surface area of all pores within a particle.
- **Averaged Theoretical Specific Surface Area ($\overline{SSA}_{Theoretical}$):** The mean theoretical estimation of SSA over multiple particles.
- **Average Pore Size ($\overline{d_p}$):** The mean pore size across all cavities in a particle.

APPENDIX C

SUPPLEMENTARY DATA FOR CHAPTER 6

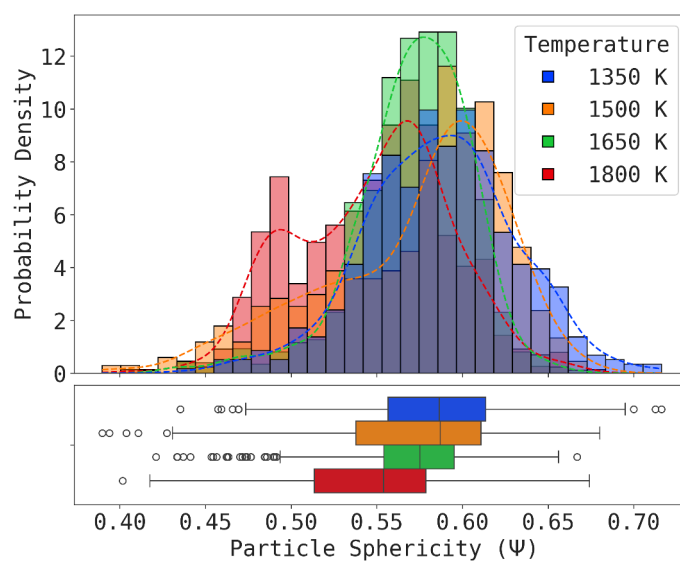


Figure C.1: Distribution of particle sphericity at different temperatures.

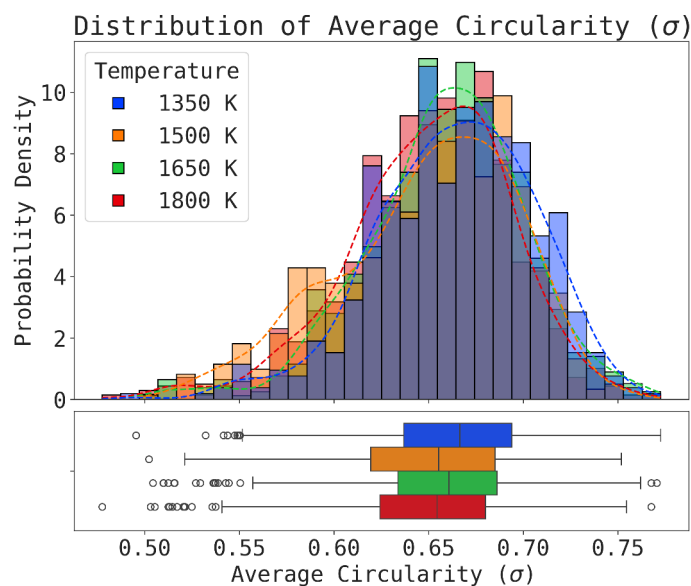


Figure C.2: Distribution of average circularity of particles at different temperatures.

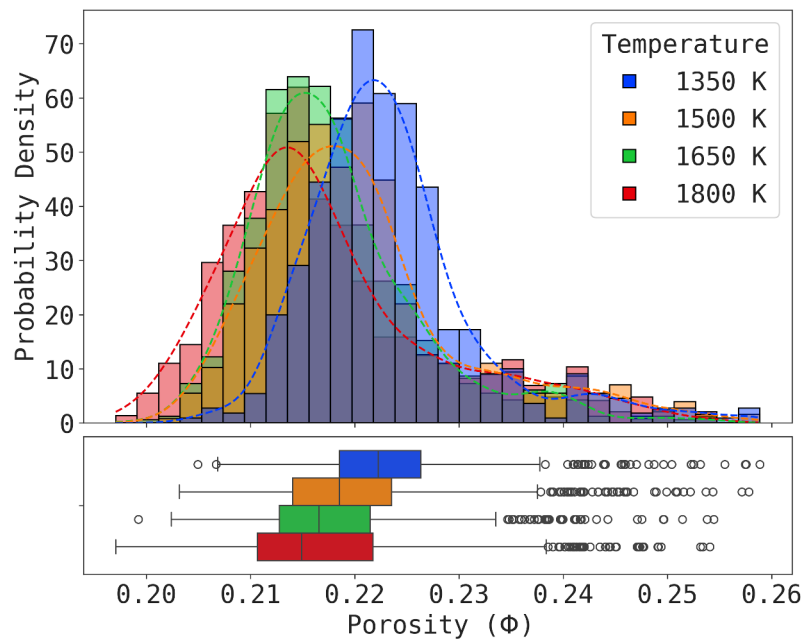


Figure C.3: Distribution of porosity, Φ of particles at different temperatures.

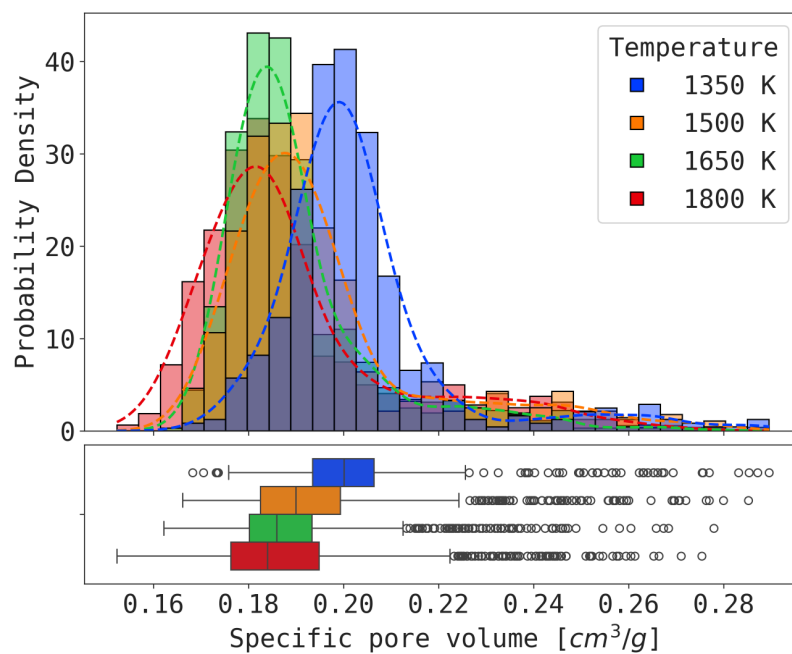


Figure C.4: Distribution of specific pore volume [cm^3/g] of particles at different temperatures.

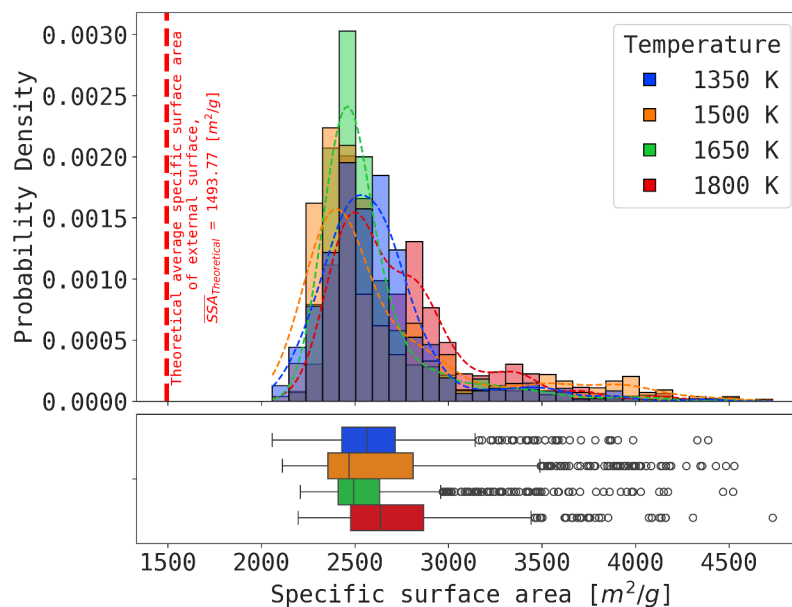


Figure C.5: Distribution of specific surface area of incipient particles at four different temperatures.

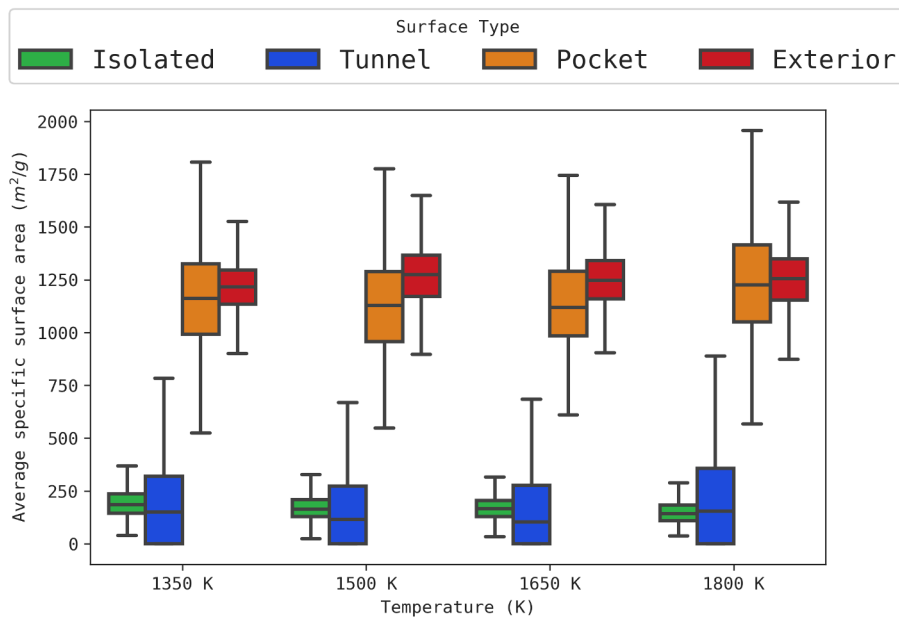


Figure C.6: Contribution of different cavities to specific surface area [m^2/g] of incipient particles.

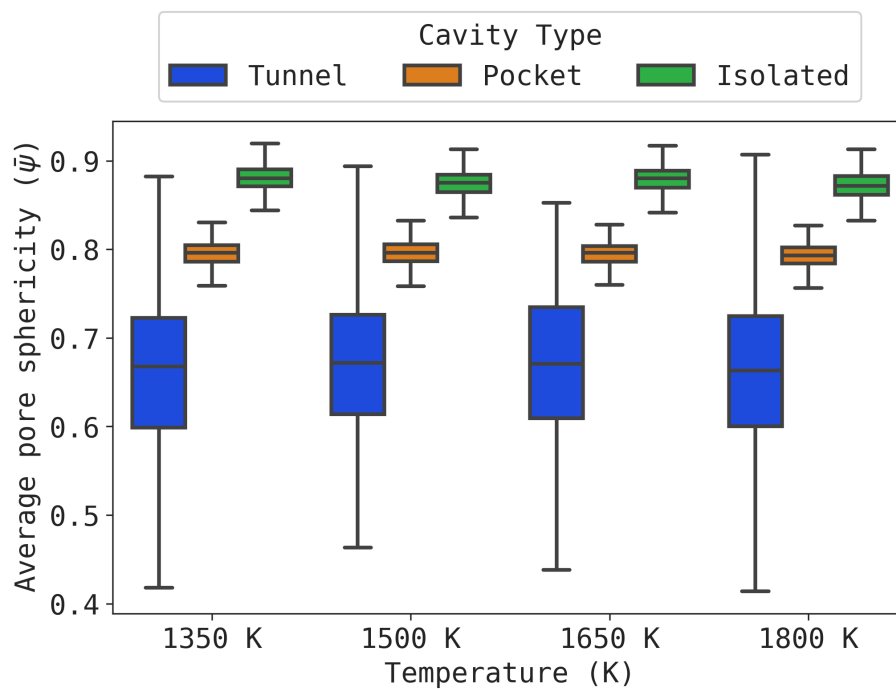
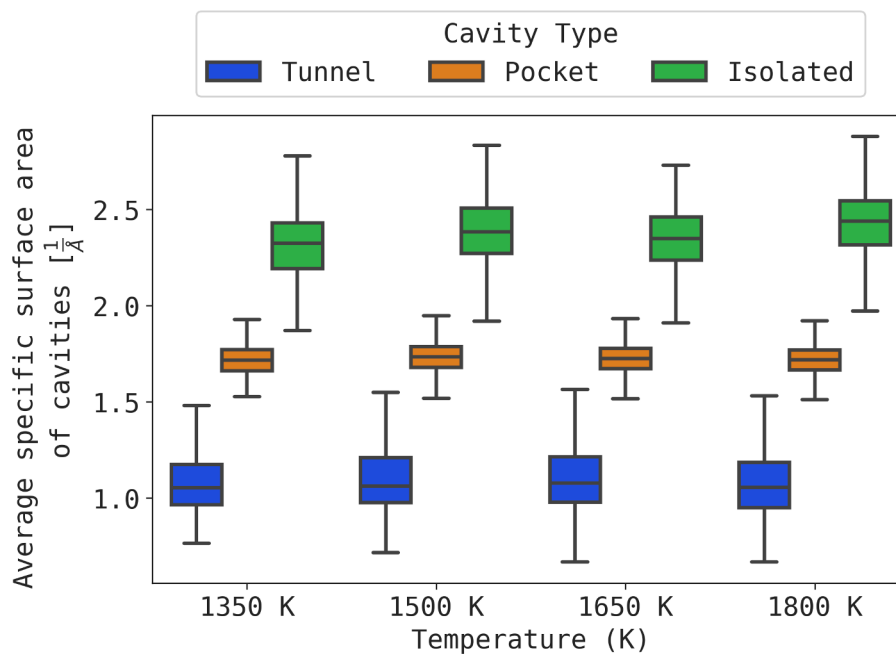
(a) Average pore sphericity, $\bar{\psi}$ (b) Average specific surface area of cavities ($1/\text{\AA}$)

Figure C.7: (a) Average pore sphericity, $\bar{\psi}$ and (b) average specific surface area of cavities [$1/\text{\AA}$] in incipient particles at different temperatures.

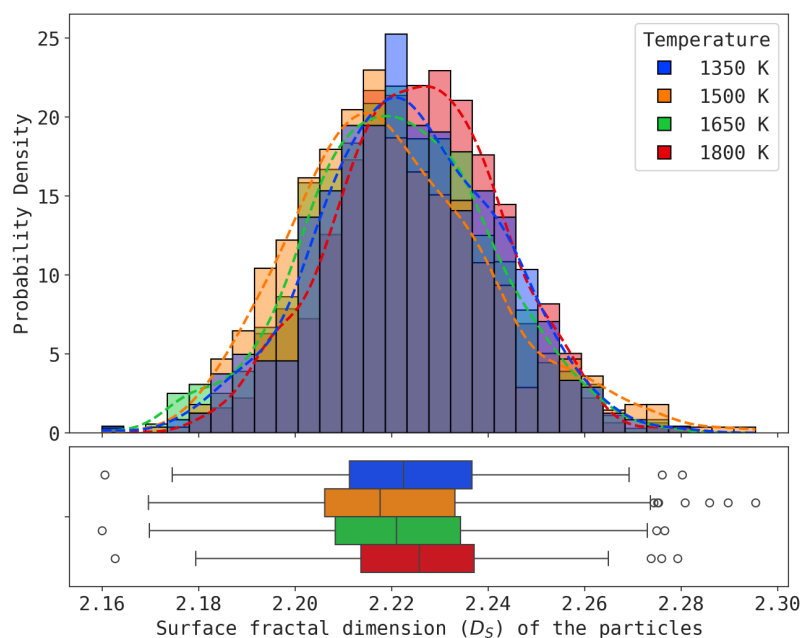


Figure C.8: Distribution of surface fractal dimension (D_S) of incipient particles at four different temperatures.

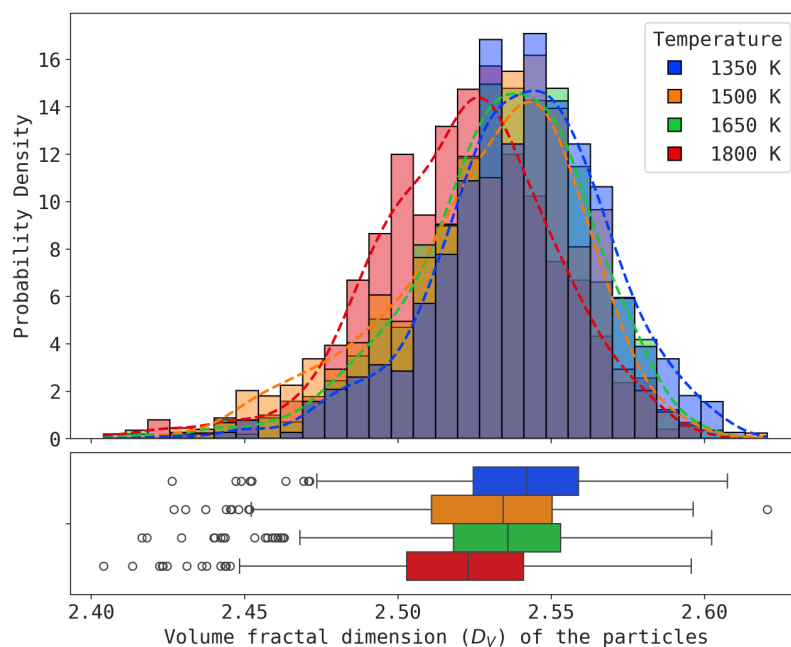


Figure C.9: Distribution of volume fractal dimension (D_V) of incipient particles at four different temperatures.

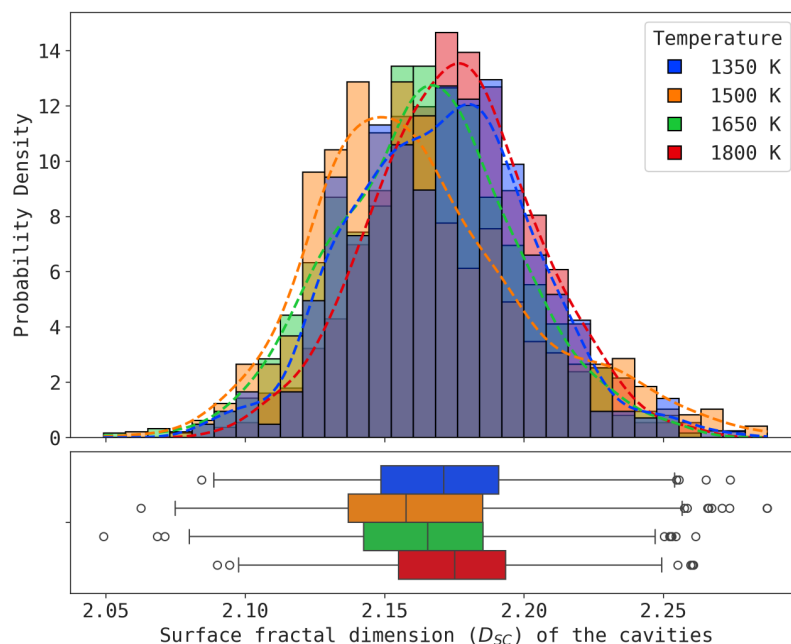


Figure C.10: Distribution of surface fractal dimension (D_{SC}) of cavities at four different temperatures.

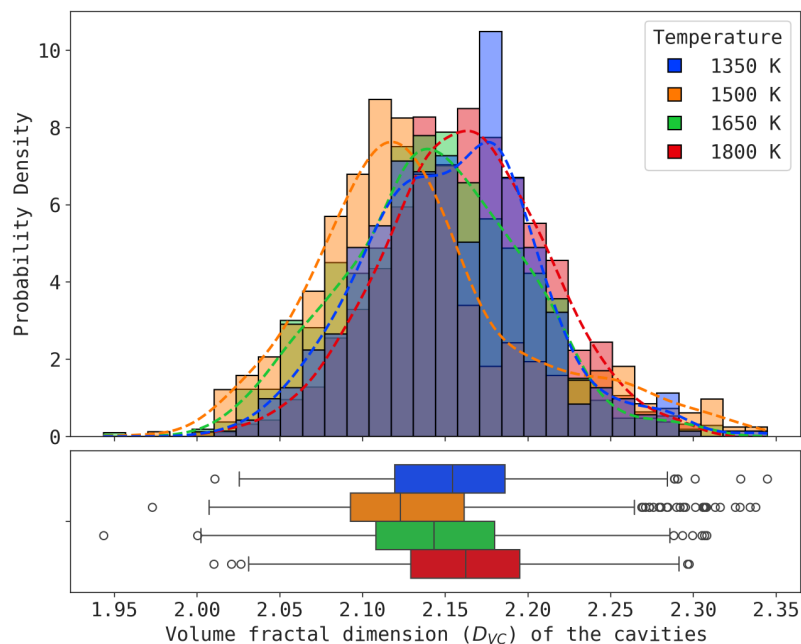


Figure C.11: Distribution of volume fractal dimension (D_{VC}) of cavities at four different temperatures.

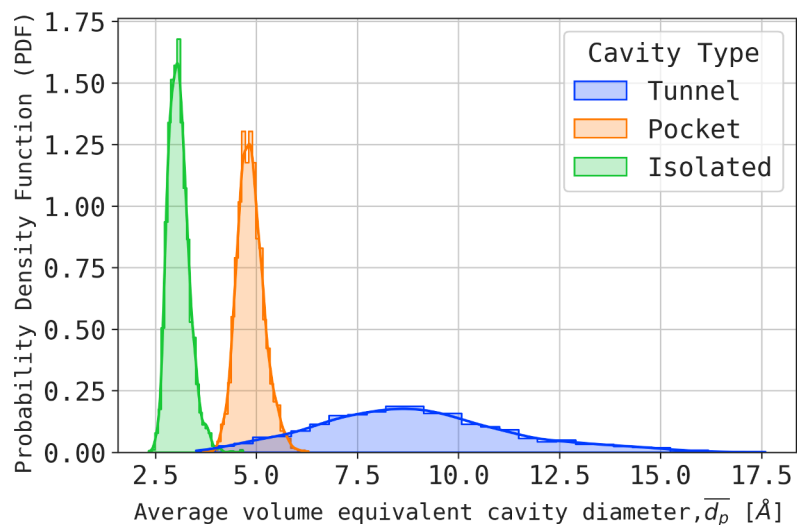


Figure C.12: Distribution of average cavity size (\bar{d}_p) distribution of different cavities within an incipient soot primary particle.

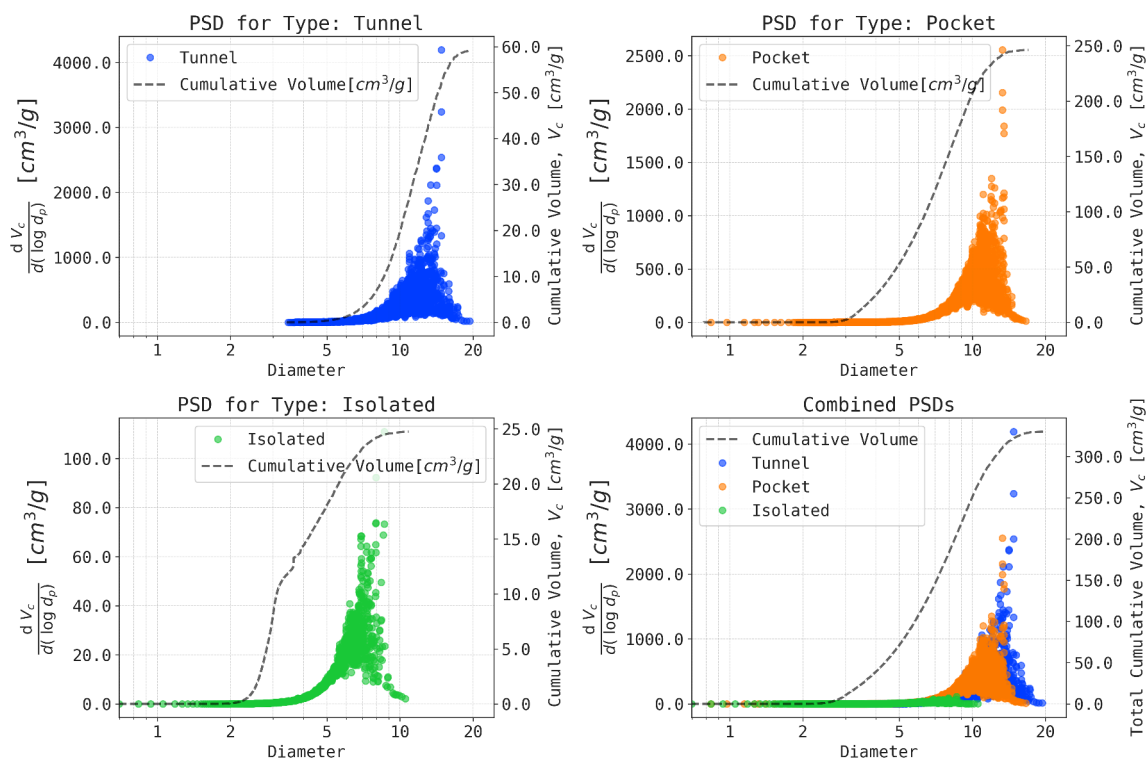


Figure C.13: Pore size distribution (PSD) of the entire population of cavities from all incipient soot primary particles.

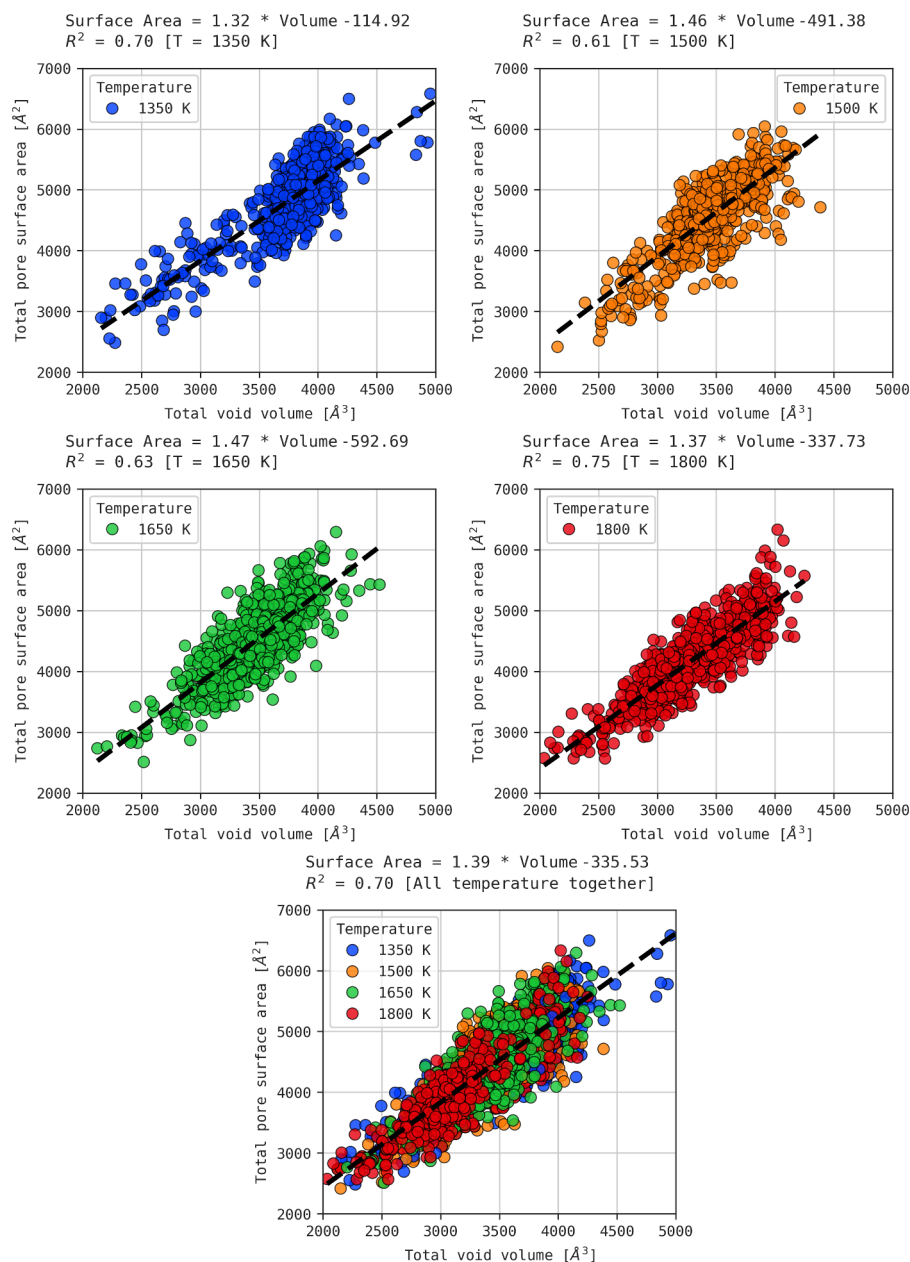


Figure C.14: Correlation between total pore volume and pore surface area within incipient particles at different temperatures. The equation for linear fit to the data also added to the plot. An exponential and a quadratic fit were also explored, but no significant improvement of correlation coefficient was observed.

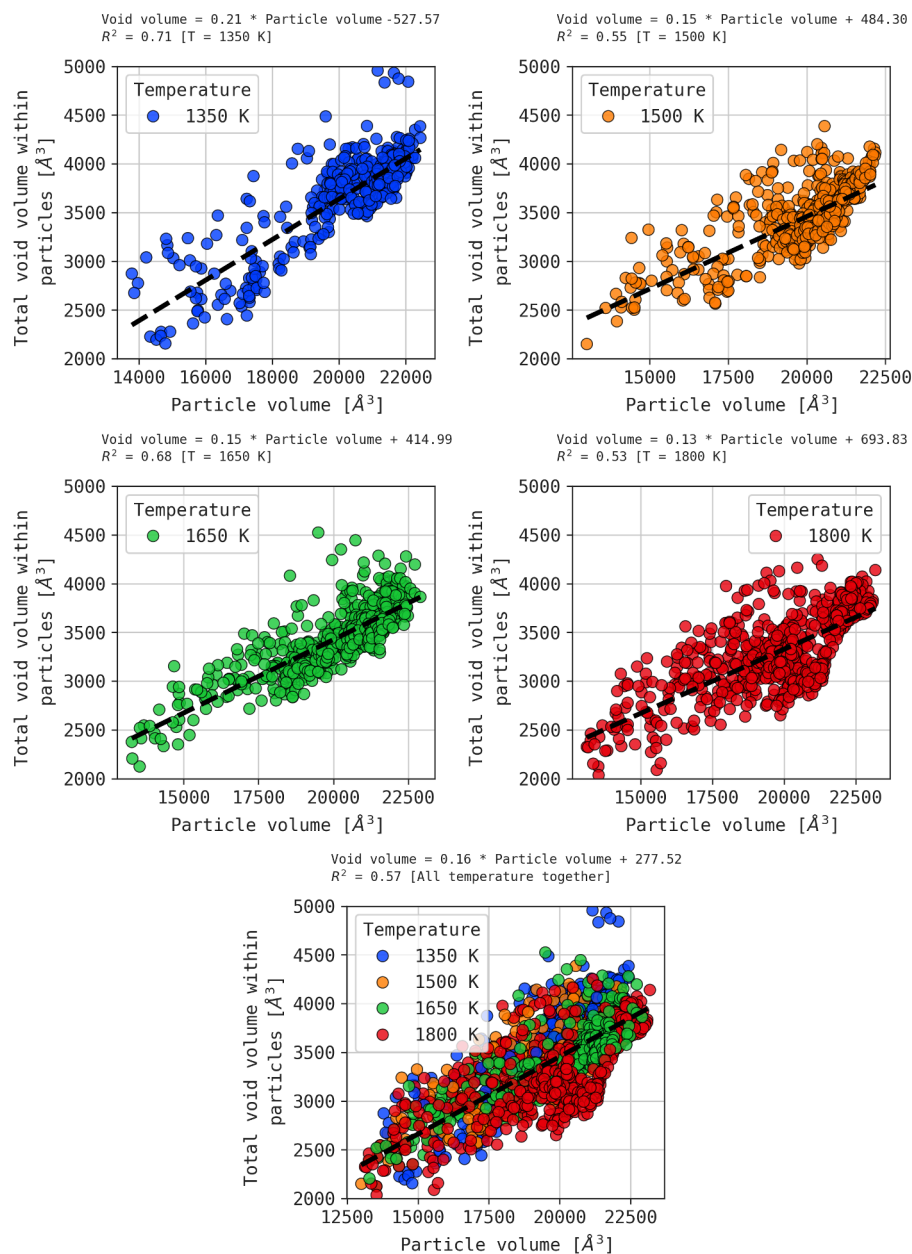


Figure C.15: Correlation between particle volume and pore volume within incipient particles at different temperatures. The equation for linear fit to the data also added to the plot. An exponential and a quadratic fit were also explored, but no significant improvement of correlation coefficient was observed.

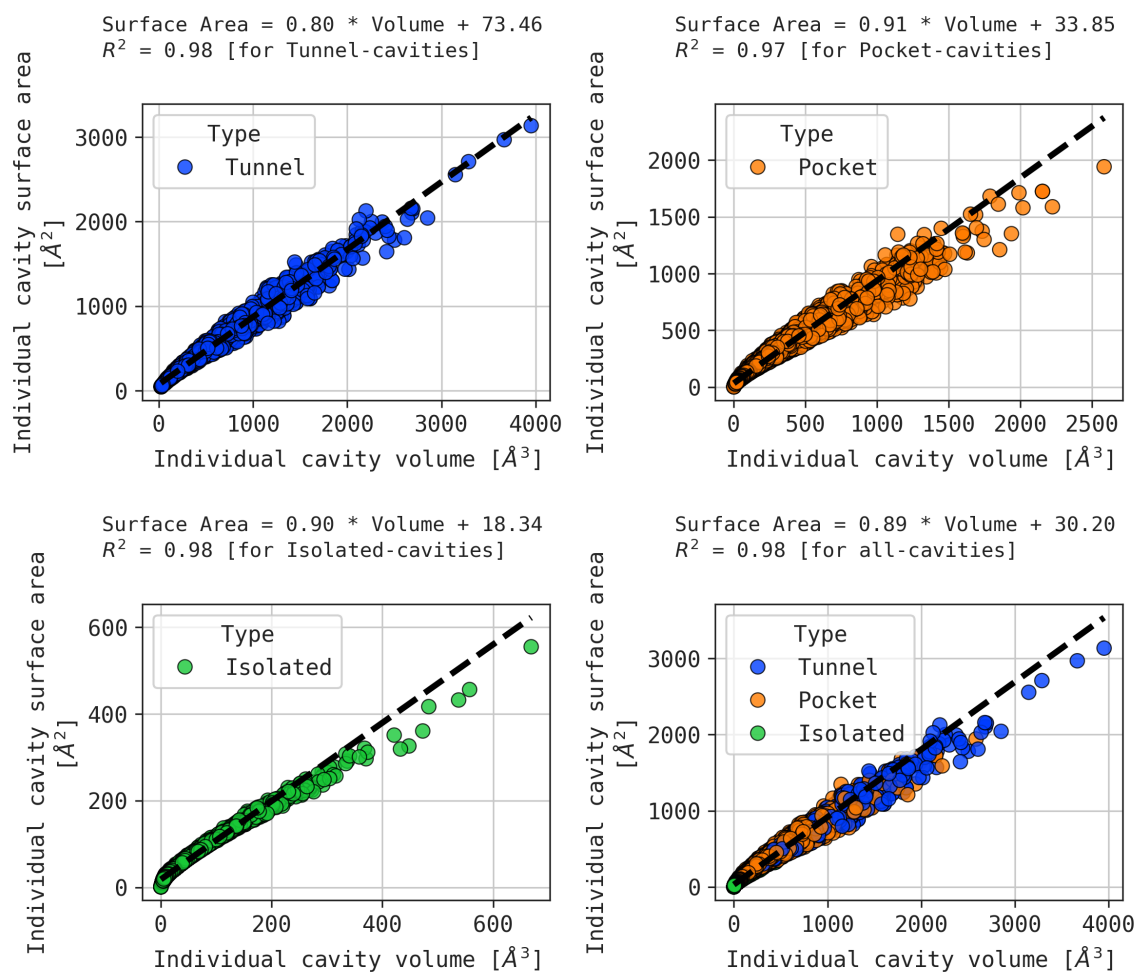


Figure C.16: Correlation between individual pore volume and pore surface area for different cavity types.

APPENDIX D

DEVELOPED SOFTWARE AND TOOLS

This chapter details the computational tools and software packages developed in support of the research presented in this dissertation. All tools are open-source and available for public use.

D.1 MAFIA-MD: Molecular Arrangements and Fringe Identification and Analysis from Molecular Dynamics

MAFIA-MD is a Python-based analysis tool developed to process molecular dynamics simulation outputs. Its primary function is to identify and quantify the number of aromatic ring structures within a system, which is critical for tracking the evolution of polycyclic aromatic hydrocarbons (PAHs) and soot precursors.

- **GitHub Repository:** <https://github.com/comp-comb/MAFIA-MD>

D.2 StereoFractAnalyzer: 2D and 3D Fractal Analyzer

StereoFractAnalyzer is a computational package for the detailed geometric and topological characterization of molecular structures. It is designed to perform fractal analysis, providing key insights into the three-dimensional complexity and morphology of molecular systems.

- **GitHub Repository:** <https://github.com/comp-comb/StereoFractAnalyzer>
- **PyPI Package:** <https://pypi.org/project/StereoFractAnalyzer/>

D.3 PyPack: A Molecular System Builder for LAMMPS

PyPack is a Python toolkit for building molecular simulation systems from SMILES strings or structure files. It enables collision-free packing into 3D periodic boxes with support for LAMMPS and XYZ outputs, with on-the-fly conformer generation for maximum diversity.

- **GitHub Repository:** <https://github.com/kmmukut/PyPack>

APPENDIX E

EVIDENCE FOR THE RESTRUCTURING OF NON-HEXAGONAL RINGS

This appendix provides supporting evidence from the literature for the interpretation of the ring statistics presented in Chapter 5. Specifically, it addresses the observation that the fraction of 6-membered rings increases with particle mass at the expense of 5- and 7-membered rings. The evidence presented here establishes a strong basis for explaining this trend as a result of a dynamic restructuring process, where less-stable non-hexagonal rings are converted into more stable hexagonal structures at high temperatures.

E.1 Thermodynamic Driving Force: Ring Stability

The fundamental reason for the prevalence of 6-membered rings is their superior thermodynamic stability. In sp^2 -hybridized carbon networks, hexagonal rings are the lowest-energy, strain-free configuration. The incorporation of non-hexagonal rings, such as pentagons and heptagons, introduces ring strain and raises the system’s energy. This principle is well-established and has been quantified in numerous studies. For example, Murry et al. [244] demonstrated through quantum-chemical calculations that annealing processes are energetically favorable pathways for fullerenes to rearrange their bonding structure, a process intimately linked with the presence of strained, non-hexagonal rings[244].

While 5- and 7-membered rings are energetically less favorable, they are critical intermediates in the formation of three-dimensional carbon structures from planar polycyclic aromatic hydrocarbons (PAHs). The inclusion of pentagons is essential for introducing positive curvature (allowing the flat sheets to form bowl-like structures), while heptagons can introduce negative curvature [245]. A comprehensive review by Martin et al. [246] highlights that the presence of pentagons creates “considerable strain”, establishing the energetic driving force for subsequent rearrangements into more stable forms.

E.2 The Stone-Wales Transformation

A primary mechanism for the interconversion of rings is the “Stone-Wales (SW) transformation”, which involves the rotation of a C-C bond to transform four adjacent hexagons into a pair of pentagons and a pair of heptagons (a 5-7-7-5 defect) [247, 248]. Critically, this process is reversible. While creating a 5-7 pair from hexagons is energetically costly, the reverse reaction—the conversion of a 5-7 pair back into two stable hexagons—is also possible and represents a pathway for “healing” the less stable defects [249, 250].

E.3 Collision and Conversion of 5- and 7-Membered Rings

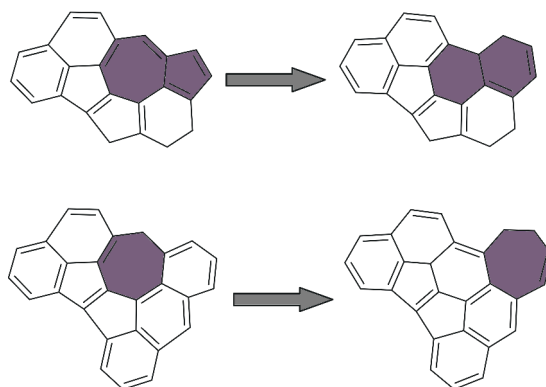


Figure E.1: Illustration of the “collision” of edge seven- and five-membered rings transforming into two edge six-membered rings (top), and the migration of a seven-membered ring (bottom). This demonstrates a direct pathway for the conversion of non-hexagonal rings into more stable hexagonal structures. (Adapted from Frenklach and Mebel [251].)

Beyond the intramolecular SW transformation, direct interaction between non-hexagonal rings provides another pathway for conversion. Frenklach and Mebel [251] performed kinetic Monte Carlo simulations that explicitly showed the “collision” of a migrating 7-membered ring with a 5-membered ring on a PAH edge, resulting in their mutual conversion into a pair of 6-membered rings (illustrated in Fig. E.1). This study provides direct mechanistic evidence that encounters between these less-stable rings can lead to their annihilation and the formation of the more stable hexagonal structure.

E.4 Evidence from Nanotube Growth Simulations

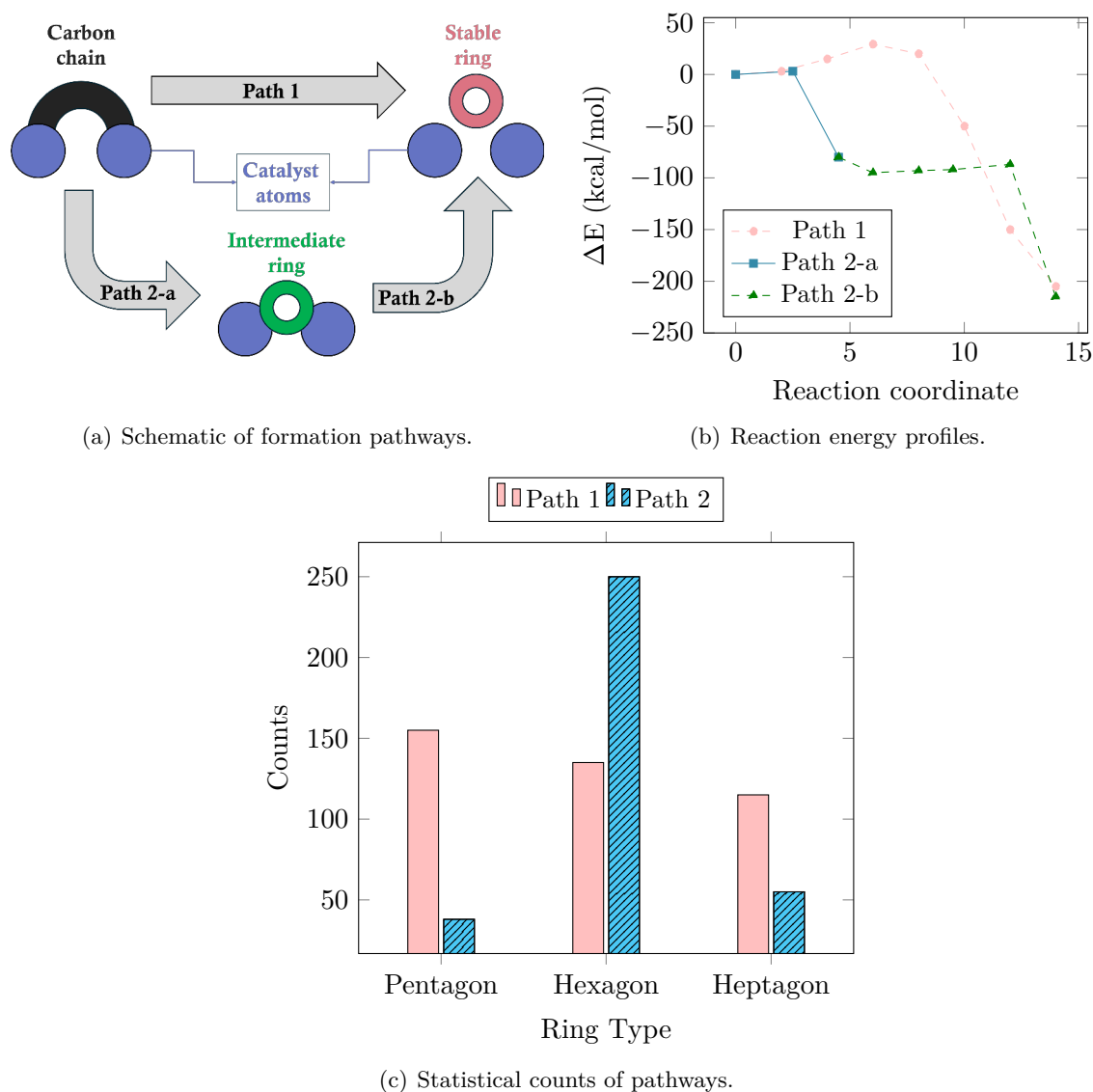


Figure E.2: Recreation of the key data illustrating the ring formation pathways during carbon growth. The figure shows (a) the schematics of the direct (Path 1) and indirect (Path 2) pathways, (b) the significantly lower energy barrier for Path 2, (c) the statistical dominance of Path 2 for hexagon formation. Data adapted from Wang et al. [252].

Further direct evidence comes from reactive molecular dynamics simulations of carbon nanotube (CNT) growth, a process with analogous carbon chemistry. A study by

Wang et al. [252] analyzed the specific reaction pathways for the formation of different ring types, providing clear, quantitative evidence of the conversion process.

Figure E.2(c) illustrates the statistical outcomes of two different formation pathways. Panel (c) is the most critical piece of evidence: it shows that while pentagons and heptagons are primarily formed directly from carbon chains (Path 1), hexagons are overwhelmingly formed via an indirect, two-step mechanism (Path 2). This directly indicates that a less-stable intermediate is being converted into the final, more stable hexagonal structure.

The energetic reason for this preference is also detailed in Fig. E.2. The energy profile in Fig. E.2(b) shows that the activation energy barrier for the direct formation of a hexagon (Path 1) is very high (29.33 kcal/mol). In contrast, the pathway that proceeds through a pentagon intermediate (Path 2) has much lower energy barriers (3.1 and 6.82 kcal/mol). This demonstrates that it is kinetically and energetically far more favorable for the system to first form a less-stable pentagon and then convert it into a hexagon, rather than forming a hexagon directly.

E.5 Summary

In summary, the literature provides strong, multi-faceted support for interpreting the increasing fraction of 6-membered rings as a dynamic restructuring process. This is based on a clear thermodynamic driving force (the inherent stability of hexagons) and, most importantly, direct simulation evidence that explicitly quantifies and energetically favors the conversion of less-stable 5- and 7-membered rings into 6-membered rings during high-temperature carbon growth.