SOOT

Markus Kraft
University of Cambridge
July 11-12, 2022

TSINGHUA-PRINCETON-COMBUSTION INSTITUTE

2022 SUMMER SCHOOL ON COMBUSTION

<table>
<tr><th colspan="6">Schedule</th></tr>
<tr><td>Beijing Time</td><td>July 11 (Mon.)</td><td>July 12 (Tue.)</td><td>July 13 (Wed.)</td><td>July 14 (Thu.)</td><td>July 15 (Fri.)</td></tr>
<tr><td>08:00 ~ 11:00</td><td></td><td></td><td>Mechanism Reduction and Stiff Chemistry Solvers
Tianfeng Lu
VMN: 52667557219</td><td></td><td>Mechanism Reduction and Stiff Chemistry Solvers
Tianfeng Lu
VMN: 52667557219</td></tr>
<tr><td>*10:00 ~ 12:00</td><td></td><td>Virtual Poster Session
10:00~12:00
VMN: 388239275</td><td></td><td>Virtual Lab Tour
10:00~12:00
VMN: 231842246</td><td></td></tr>
<tr><td>14:00 ~ 17:00 Session I</td><td colspan="3">Fundamental of Flames
Suk Ho Chung
VMN: 42399313194</td><td colspan="2">Combustion in Microgravity and Microscale
Kaoru Maruta
VMN: 71656262918</td></tr>
<tr><td>14:00 ~ 17:00 Session II</td><td colspan="2">Soot
Markus Kraft
VMN: 39404905340</td><td></td><td colspan="2">Current Status of Ammonia Combustion
William Roberts
VMN: 80506726244</td></tr>
<tr><td>19:00 ~ 22:00 Session I</td><td colspan="5">Combustion Chemistry and Kinetic Mechanism Development
Tiziano Faravelli
VMN: 35989357660</td></tr>
<tr><td>19:00 ~ 22:00 Session II</td><td colspan="4">Combustion Fundamentals of Fire Safety
José Torero
VMN: 57002781862</td><td></td></tr>
</table>

Note:

[1]Session I and Session II are simultaneous courses.

[2]VMN: Voov Meeting Number

Guidelines for Virtual Participation

1. **General Guidelines**

- Tencent Meeting software（腾讯会议）is recommended for participants whose IP addresses locate within Mainland China; Voov Meeting (International version of Tencent Meeting) is recommended for other IP addresses. The installation package can be found in the following links:

  a) 腾讯会议

  https://meeting.tencent.com/download/

  b) Voov Meeting

  https://voovmeeting.com/download-center.html?from=1001

- All the activities listed in the schedule are "registrant ONLY" due to content copyright.
- To facilitate virtual communications, each participant shall connect using stable internet and the computer or portable device shall be equipped with video camera, speaker (or earphone) and microphone.

2. **Lectures**

- The lectures are also "registrant ONLY". Only the students who registered for the course can be granted access to the virtual lecture room.
- To enter the course, each registered participant shall open the software and join the conference using the corresponding Voov Meeting Number (VMN) provided in the schedule; only participants who show unique identification codes and real names as "**xxxxxx-Last Name, First Name**" will be granted access to the lecture room; the identification code will be provided through email.
- During the course, each student shall follow the recommendation from the lecturer regarding the timing and protocol to ask questions or to further communicate with the lecturer.
- For technical or communication issues, the students can contact the TA in the virtual lecture or through emails.
- During the course, the students in general will not be allowed to use following functions in the software: 1) share screen; 2) annotation; 3) record.

3. **Lab Tour**

- The event will be hosted by graduate students from Center for Combustion Energy, Tsinghua University and live streamed using provided Voov Meeting Number.
- During the activity, the participants will not be allowed to use following functions in the software: 1) share screen; 2) annotation; 3) record.
- Questions from the virtual participants can be raised using the chat room.

4. **Poster Session**

- The event will be hosted by the poster authors (one Voov Meeting room per poster) and live streamed using provided Voov Meeting Number.
- During the activity, the participants will not be allowed to use following functions in the software: 1) share screen; 2) annotation; 3) record.
- Questions from the virtual participants can be raised using the chat room or request access to audio and video communication.

Teaching Assistants

- **Fundamentals of Flame (Prof. Suk Ho Chung)**

  TA1: Hengyi Zhou (周恒毅); zhouhy19@mails.tsinghua.edu.cn

  TA2: Xinyu Hu (胡馨予); hxy21@mails.tsinghua.edu.cn

- **Combustion Chemistry and Kinetic Mechanism Development (Prof. Tiziano Faravelli)**

  TA1: Shuqing Chen (陈舒晴); chen-sq19@mails.tsinghua.edu.cn

  TA2: Jingzan Shi (史京瓒); sjz21@mails.tsinghua.edu.cn

- **Current Status of Ammonia Combustion (Prof. William Roberts)**

  TA1: Yuzhe Wen (温禹哲); wyz20@mails.tsinghua.edu.cn

  TA2: Haodong Chen (陈皓东); chd20@mails.tsinghua.edu.cn

- **Soot (Prof. Markus Kraft)**

  TA1: Yuzhe Wen (温禹哲); wyz20@mails.tsinghua.edu.cn

  TA2: Haodong Chen (陈皓东); chd20@mails.tsinghua.edu.cn

- **Combustion Fundamentals of Fire Safety (Prof. José Torero)**

  TA1: Xuechun Gong (巩雪纯); gxc19@mails.tsinghua.edu.cn

  TA2: Weitian Wang (王巍添); wwt20@mails.tsinghua.edu.cn

- **Combustion in Microgravity and Microscale (Prof. Kaoru Maruta)**

  TA1: Hengyi Zhou (周恒毅); zhouhy19@mails.tsinghua.edu.cn

  TA2：Xinyu Hu (胡馨予); hxy21@mails.tsinghua.edu.cn

- **Mechanism Reduction and Stiff Chemistry Solvers (Prof. Tianfeng Lu)**

  TA1: Shuqing Chen (陈舒晴); chen-sq19@mails.tsinghua.edu.cn

  TA2: Jingzan Shi (史京瓒); sjz21@mails.tsinghua.edu.cn

Soot – Part 1

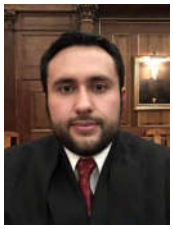
Markus Kraft

Computational Modelling Group Cambridge

Main Contributors:
Dr Jake Martin (Part 1)
Dr Angiras Menon (Part 2)
Dr Laura Pascazio (Part 3)
Dr Gustavo Leon (Part 4)

Jacob Martin

AngirasMenon



Gustavo Leon

Laura Pascazio

Part 1 Overview
Part 2 Quantum Chemistry
Part 3 Molecular Dynamics
Part 4 Kinetic Monte Carlo
Part 5 Stochastic Particle Methods
Part 6 Application – engine model

Part A: Illumination

1. Faraday and the chemical history of a candle
2. Important soot precursors

Part B: Pigment

1. Early art using soot/carbon blacks
2. Micro and nanostructure of soot/carbon black
3. Uses of carbon black as a material

Part C: Pollution

1. Smoke point and cleaner burning fuels
2. Electrical control of soot
3. Heidelberg Symposia 1990s
4. Impact on the climate and human health
5. Recent reviews and findings.

PART A: ILLUMINATION

"Direct evidence of early fire in archaeology remains rare, but from 1.5 Ma onward surprising numbers of sites preserve some evidence of burnt material. By the Middle Pleistocene, recognizable hearths demonstrate a social and economic focus on many sites.

The evidence of archaeological sites has to be evaluated against postulates of biological models such as the 'cooking hypothesis' or the 'social brain', and questions of social cooperation and the origins of language.

Although much remains to be worked out, it is plain that fire control has had a major impact in the course of human evolution."

Gowlett, John AJ. "The discovery of fire by humans: a long and convoluted process." Philosophical Transactions of the Royal Society B: Biological Sciences 371.1696 (2016): 20150164.

Egyptians invented the rush light. Rush dipped in animal fat

Romans developed tallow candle from animal fat dipped candles.

Middle ages saw beeswax cleaner burning

Whale oil lamps became important late 18th century.

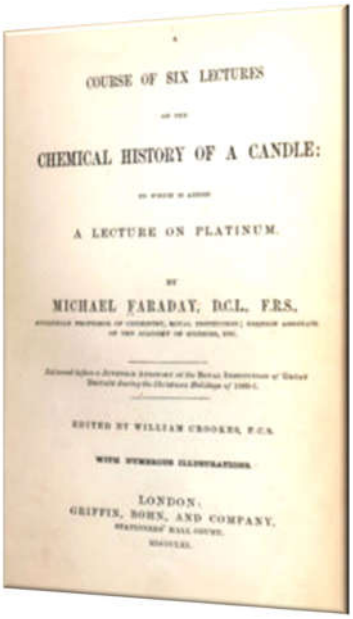
Michel Eugene Chevreul extracted Stearic acid in 1820's

Paraffin wax from petroleum in 1850s replaced other waxes.

Vegetable oil lamps were also used in India, China and Europe.

Focus of early candles were reducing sooting.

Chemical History of a Candle – 1861



COURSE OF SIX LECTURES

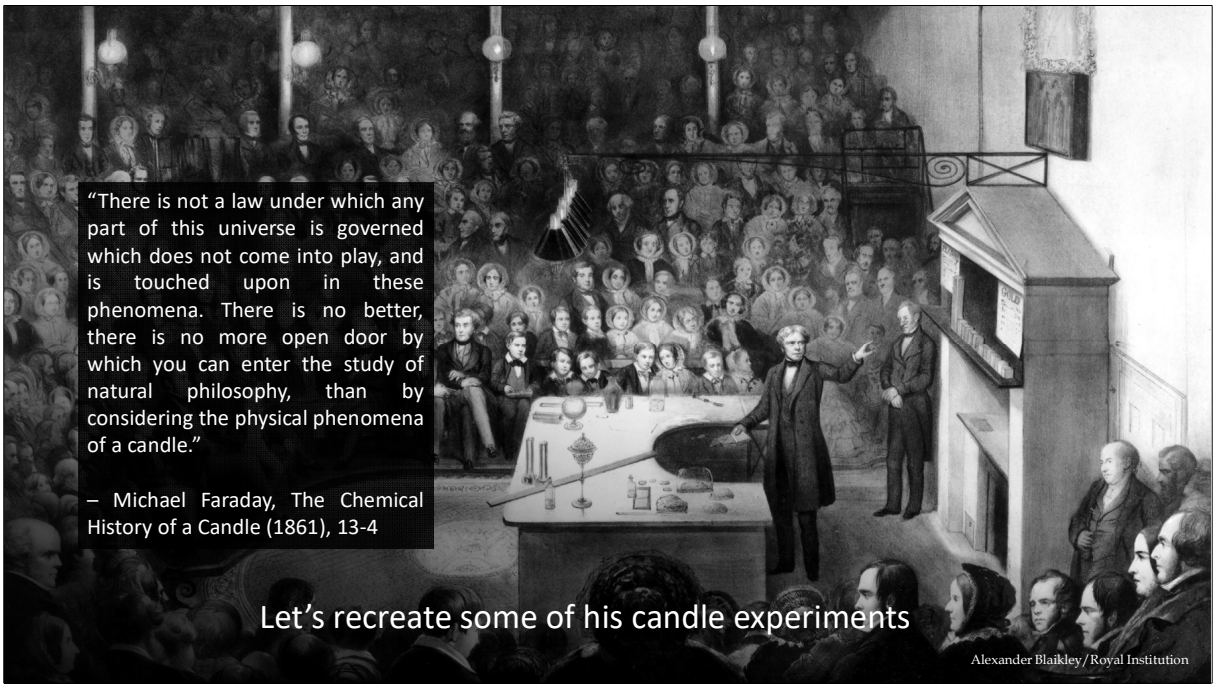
CHEMICAL HISTORY OF A CANDLE:

A LECTURE ON PLATINUM.

MICHAEL FARADAY, D.C.L., F.R.S.

EDITED BY WILLIAM CROOKES, F.C.S.

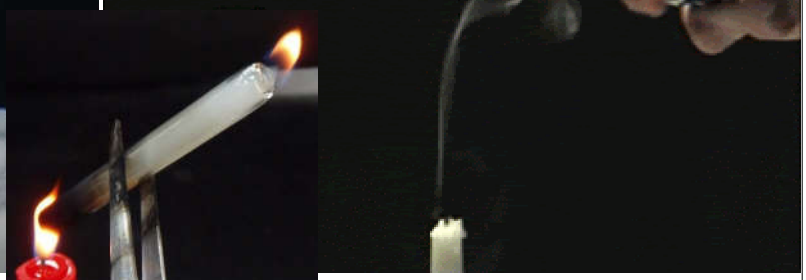
LONDON:
GRIFFIN, BOHN, AND COMPANY.



What burns in a candle?

1. Blow out a candle

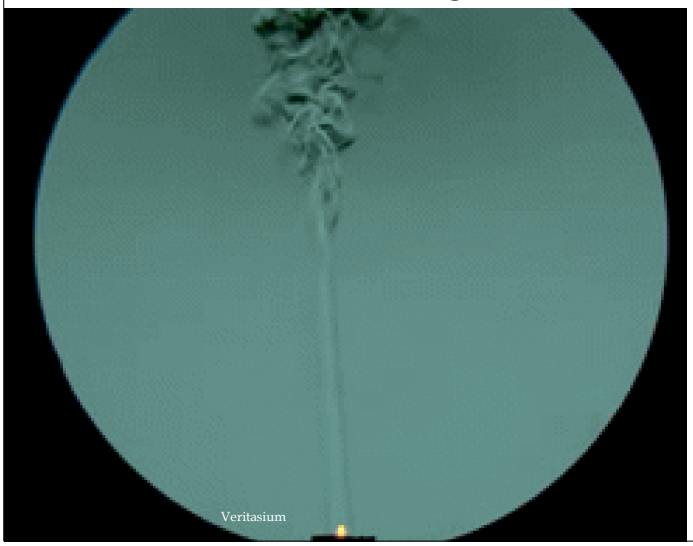
2. Reignite the wax trail



3. Glass tube can extract wax to split the candle in two

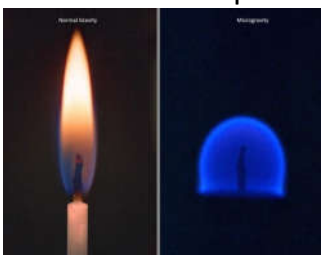
Why is the flame a tulip shape?

Schlieren video showing air flow



Hot air rises

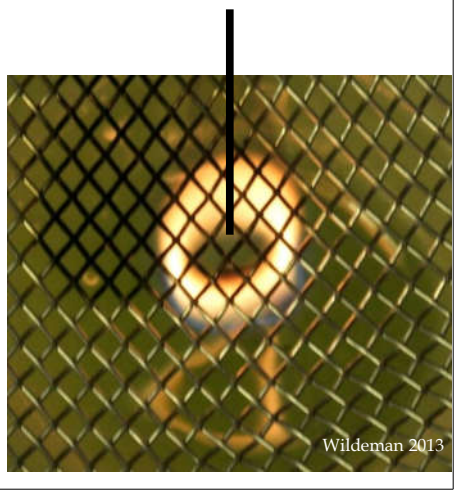
What about in space?



What is inside of a flame?

Wire mesh extinguishes flame

Flame is hollow with a dark zone in the middle



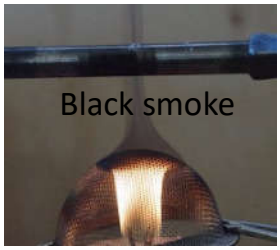
Safety lamp

Humphrey Davy was tasked with building a lamp that would be safe in coal mines with methane leaks.

Methane and candle

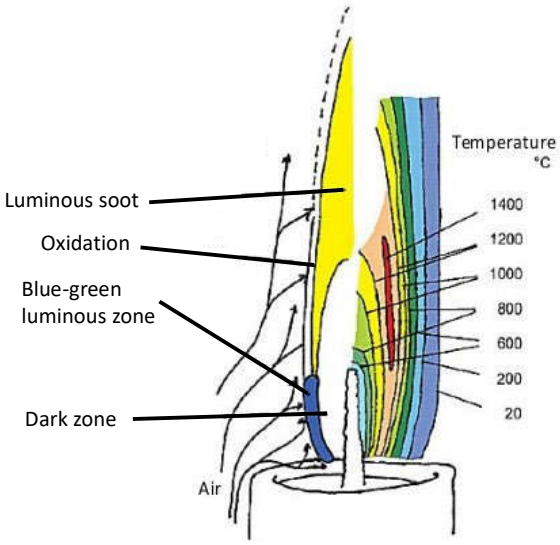
Flame contained in lamp

What is inside of a flame?



Black soot is present in the flame and is heated to yellow hot by the flame.

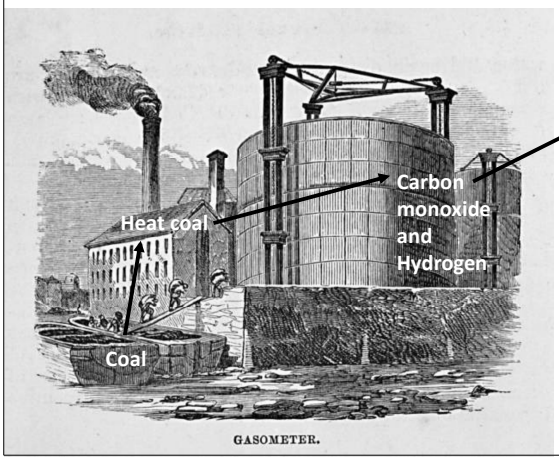
What is the structure of a candle?



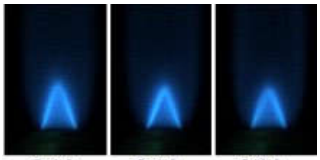
Roth, Klaus. "Chemistry of the Christmas Candle. Part 2." *ChemViews Mag* (2011).

What motivated Faraday?

- Lighting or illumination was the main focus for research into flames.



CO and H_2 burn without a luminous flame.



H_2: CO=30:70 Rich Conditions

Faraday sought out the cause of the soot in coal gas.

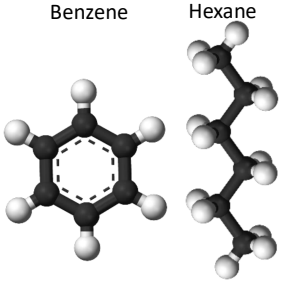
What causes luminosity in coal gas?

- Faraday extracted "Bicarburet of hydrogen" (benzene) from a portable gas for lighting.

Pool flame of benzene is heavily sooting.

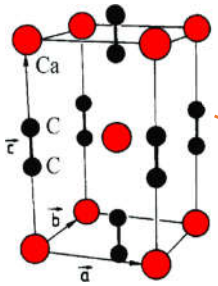
Benzene Hexane

Dark zone in hexane flame is missing in benzene.

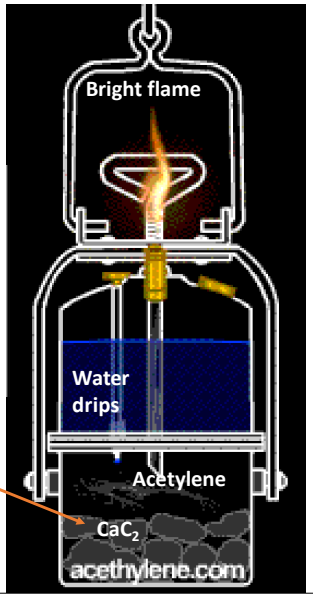


Is there a better gas than coal gas?

- Davy discovered acetylene in 1836 but in 1892 Moisan found a way to make acetylene easily with calcium carbonate (CaC_2)

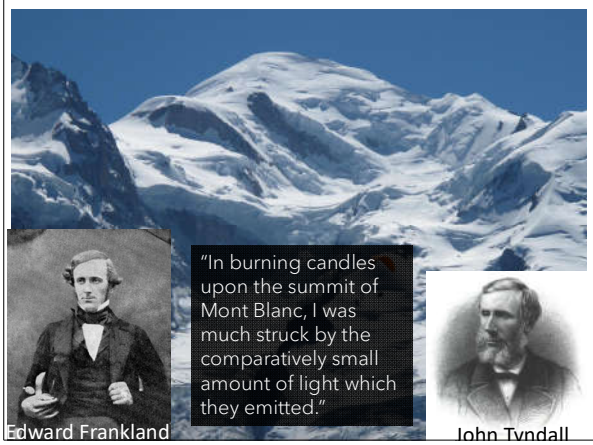


Acetylene

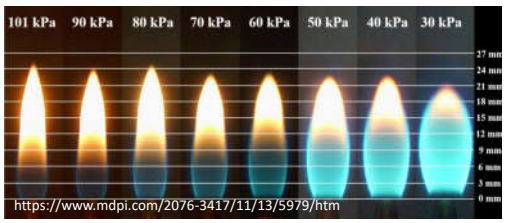


What impact does pressure have on luminosity?

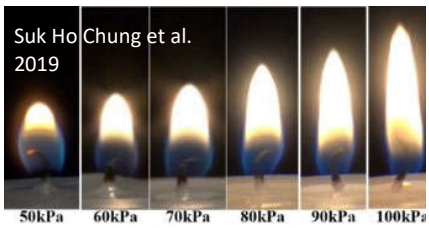
- Frankland and Tyndall lit candles at the top of Mont Blanc the tallest mountain in Europe.



Lower pressures reduces soot production

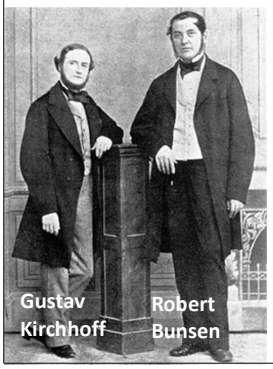


Higher pressures increases soot production



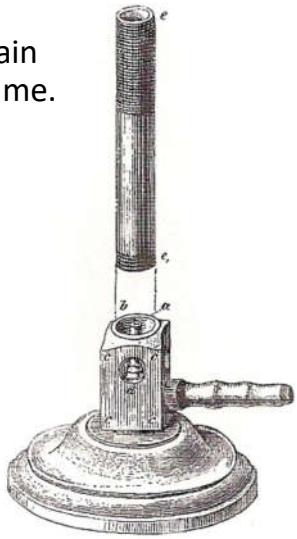
How to prepare a non-luminous flame?

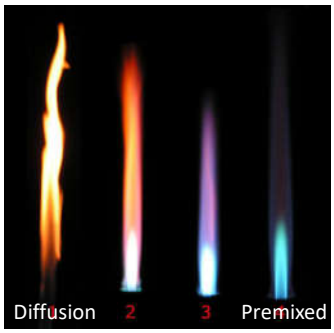
- Bunsen developed a burner to entrain air to produce the first premixed flame.



If the tube…is screwed into the cylinder, and the city gas is allowed to flow into it…, it sucks in so much air through the openings that it burns at the mouth of the tube with a nonluminous, perfectly soot-free flame.

Bunsen 1857

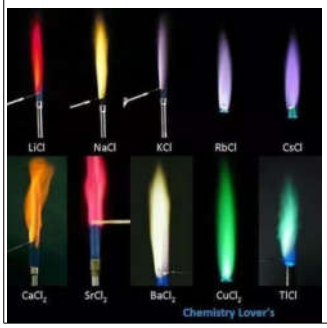


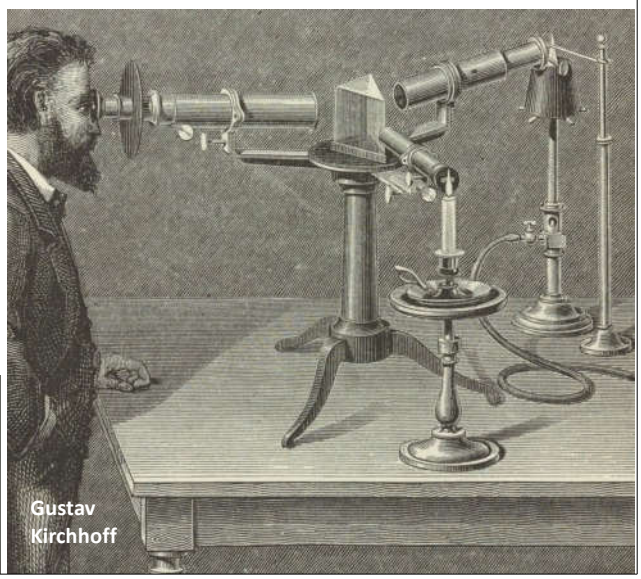


How to prepare a non-luminous flame?

- Kirchhoff and Bunsen used the flame to discover elements using flame emission spectroscopy.

Flame emission spectroscopy

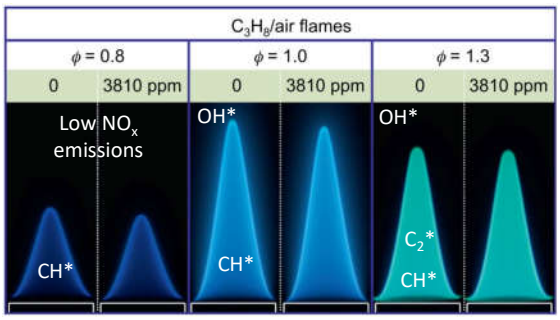




What is the blue-green emission?

- Swan in 1856 found green emissions in fuel rich flames.
- Mulliken in 1927 used early QM to assign C_2*.
- Threshold of $\phi > 2.0$ indicates pyrolysis reaction leads to carbon formation (Street and Thomas 1955)

$$C_mH_n + yO_2 \rightarrow yCO + (n/2)H_2O + (m-y)C_s$$

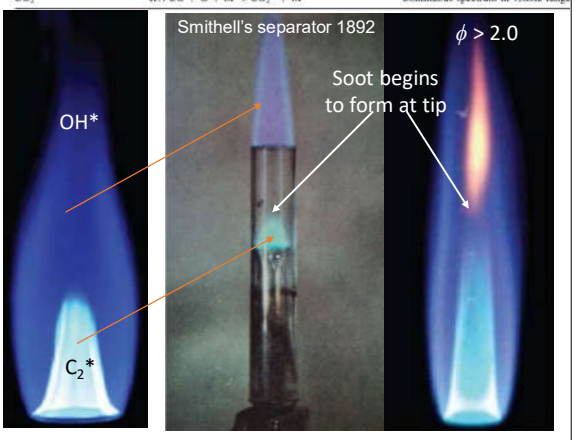


Vu, Tran Manh, et al. *Combustion and Flame* 161.4 (2014): 917-926.

https://doi.org/10.1007/s12046-020-01465-4

Table 1. Formation routes of excited radicals and characteristic wavelengths.

Radical	Reactions	Wavelength (nm)
OH*	R1: CH + O_2 → CO + OH* R2: H + O + M → OH* + M R3: O(H) + OH + H → H_2O + OH*	282.9, 308.9
CH*	R4: C_2H + O_2 → CO_2 + CH* R5: C_2H + O → CO + CH*	387.1, 431.4
C_2*	R6: CH_2 + C → C_2* + H_2	515.8, 516
CO_2*	R7: CO + O + M → CO_2* + M	Continuous spectrum in visible range



https://doi.org/10.1007/s12046-020-01465-4

What was known at the turn of the 20th century?

- Acetylene and Benzene are critical intermediates in soot formation.
- The threshold of soot formation points to pyrolysis and C_2 being present before soot formation.
- Pressure strongly impacts soot formation.
- Some focus on clean burning fuels

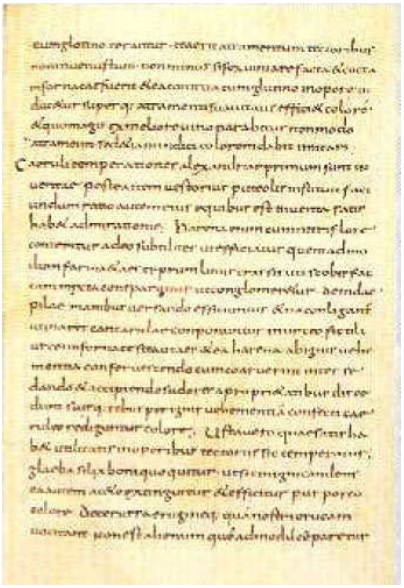
PART B: PIGMENTS

Early art using soot

Soot is used in prehistoric cave paintings
(35,000 -10,000 **BC)**

World's oldest tattoos (Tyrolean iceman, Ötzi) were etched in soot (c. 3,300 BC)

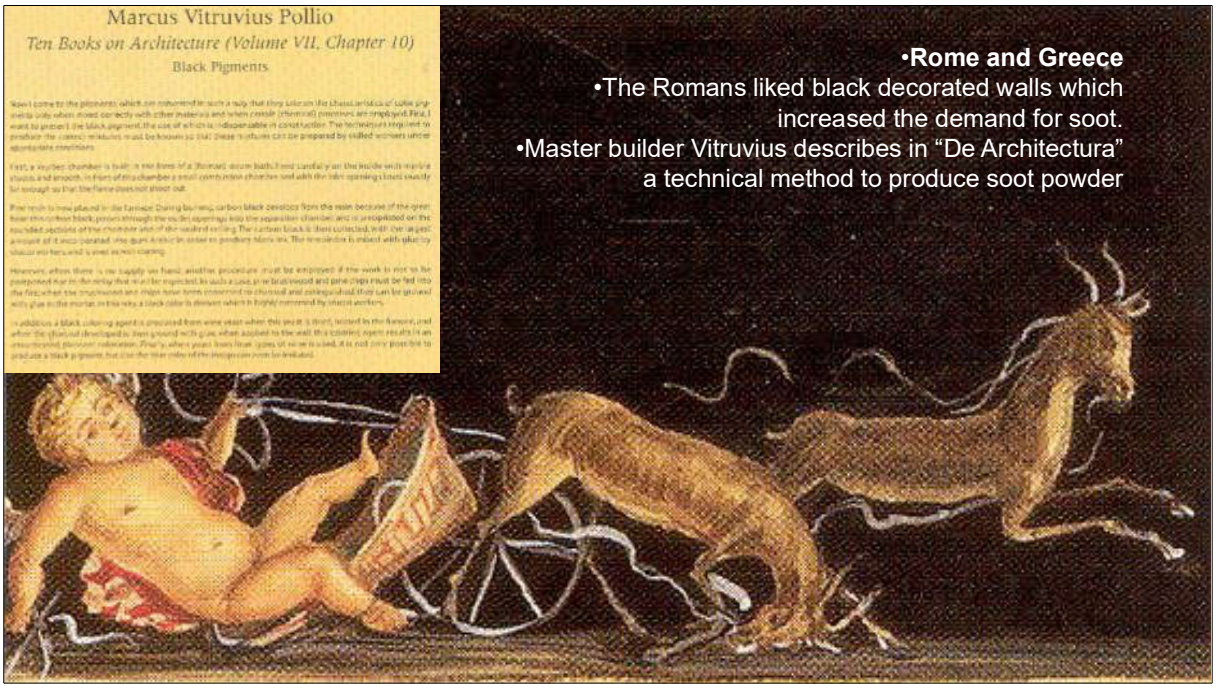
Soot/carbon black in inks



- **Ancient civilizations in China and Egypt** mixed soot into resins, vegetable oil and tar to create colour inks
- Soot produced from an oil lamp brought into contact with a cooled surface from which the soot could then be scraped off and collected as a powder.

Inkstick

Green Duan inkstone from Song Dynasty



•**Rome and Greece**

•The Romans liked black decorated walls which increased the demand for soot.

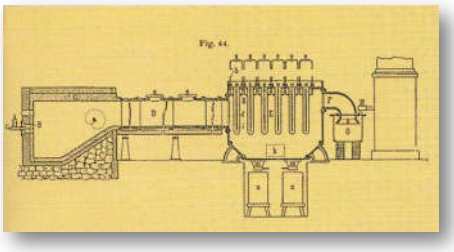
•Master builder Vitruvius describes in “De Architectura” a technical method to produce soot powder

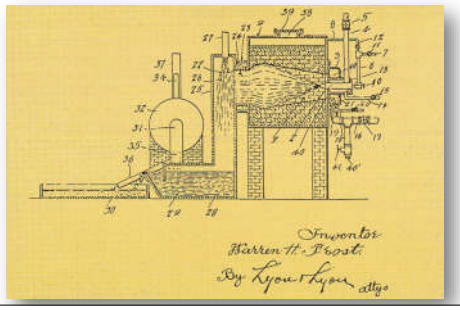
Lamp black process

The **lamp black process** based on resin is refined and many villages in the black forest specialised on carbon black production

Processing pine resin to create pitch, tar and carbon black

Oil furnace



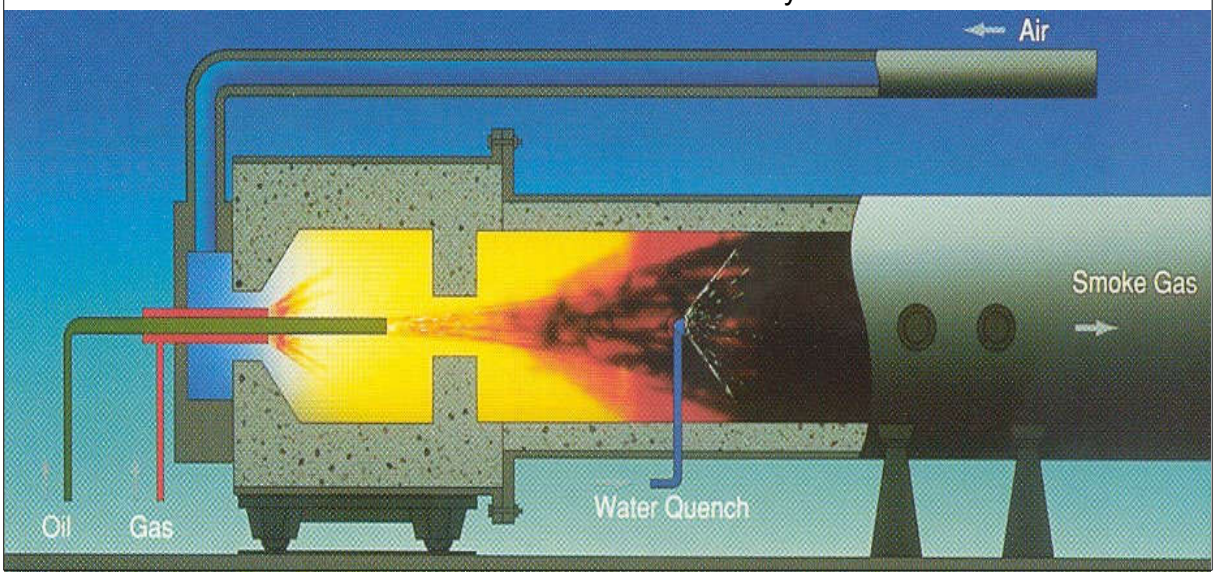


In the 19th century and then more systematically from 1920 onwards attempts were made to produce carbon black in a closed system with mineral oil as feedstock.

The oil furnace method, first commercialised in 1943 is the dominant method producing 98%
of the world wide carbon black production.

Furnace black process

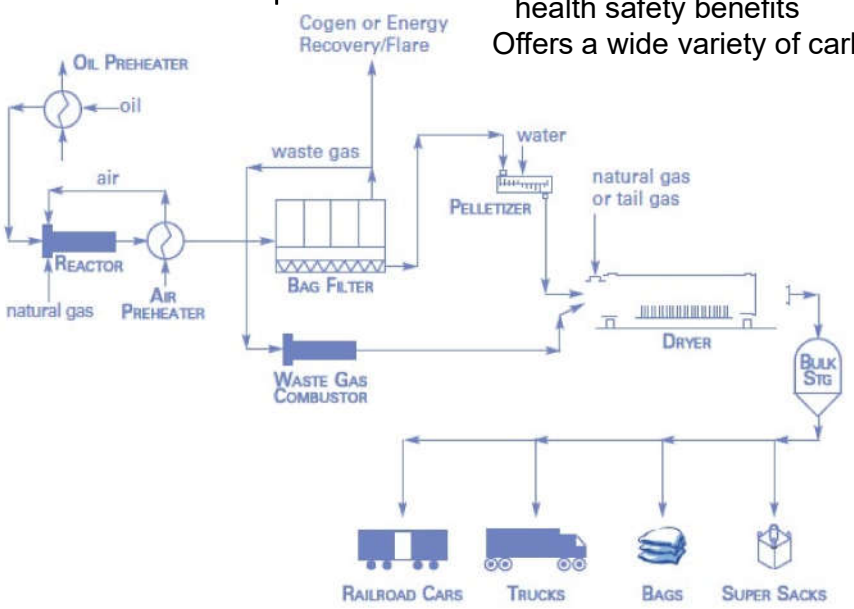
The furnace process offers environmental and health safety benefits
Offers a wide variety of carbon blacks



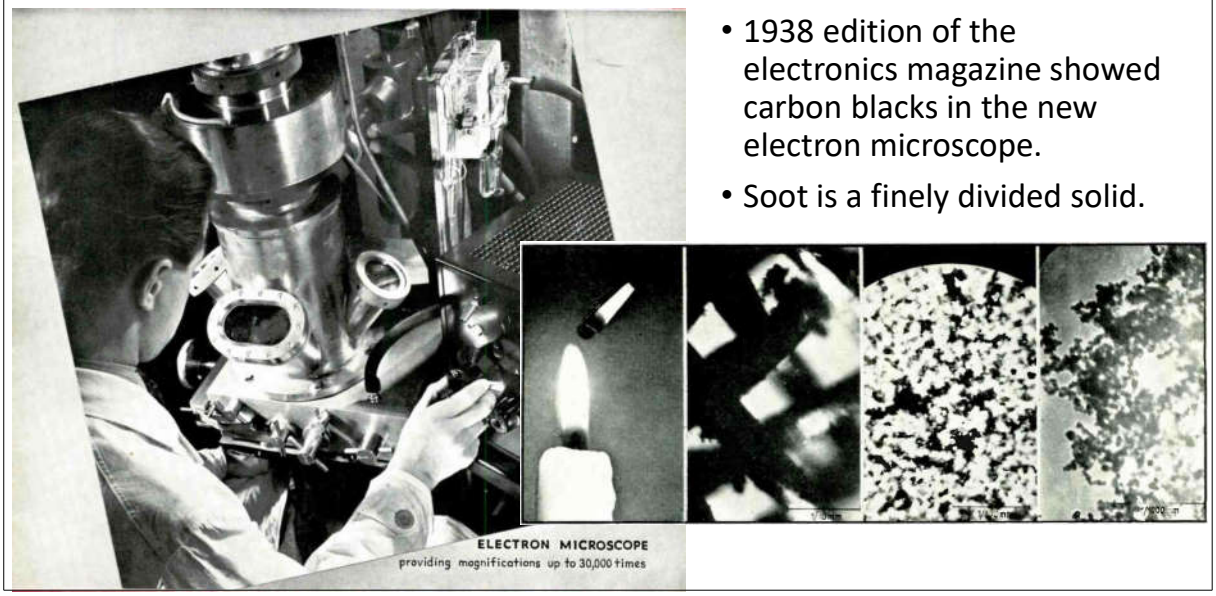
Furnace black process

The furnace process offers environmental and health safety benefits

Offers a wide variety of carbon blacks



What is the structure of soot particles?

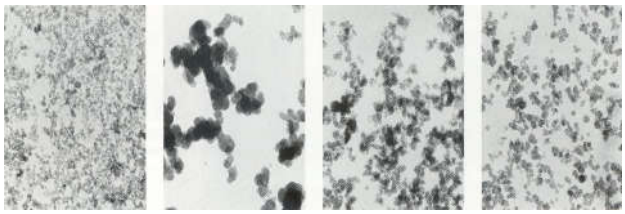


- 1938 edition of the electronics magazine showed carbon blacks in the new electron microscope.
- Soot is a finely divided solid.

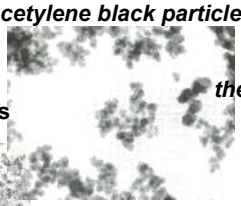
What do the different carbon blacks look like?

Furnace blacks

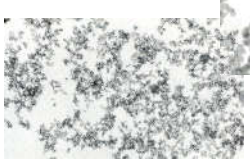
Fine 14 nm | Coarse 50 nm | High structure | Low structure



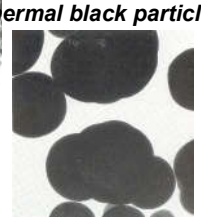
acetylene black particles

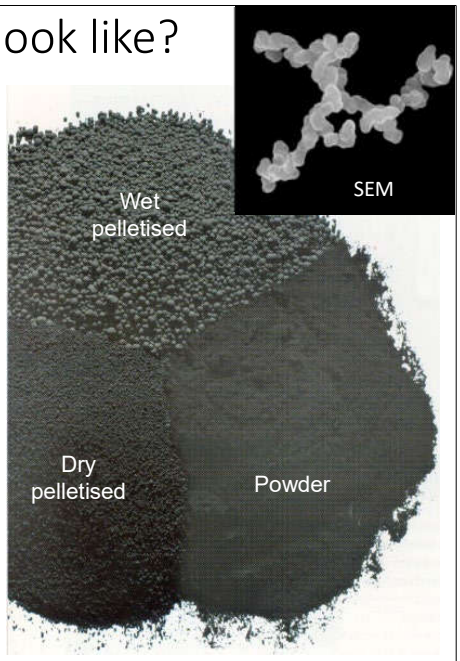


gas black particles



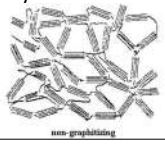
thermal black particles



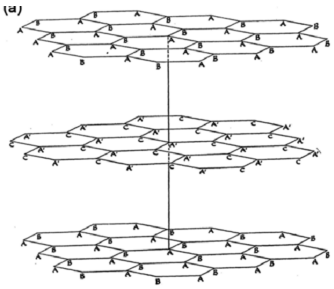


What is inside carbon black particles?

1918 Scherrer determined crystallite size within a powder using X-rays.
1924 Bernal found structure of graphite.
1930s many papers on carbon black's structure.
1939 – Riley reviewed evidence for small graphitic crystallites in carbon black.
1943 Biscoe and Warren assumed a microcrystalline model
1951 – Franklin introduces non-graphitising carbons that have misaligned crystallites and done graphitise.



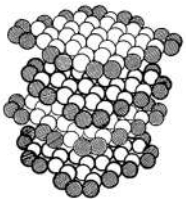
Bernal graphite



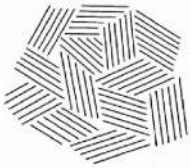
Bernal's 1928 X-ray camera

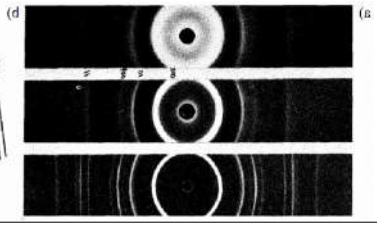
https://collection.sciencemuseumgroup.org.uk/objects/co13268/j-d-bernals-x-ray-diffraction-camera-united-kingdom-1928-x-ray-diffraction-camera

Riley crystallite



Biscoe & Warren microcrystallites

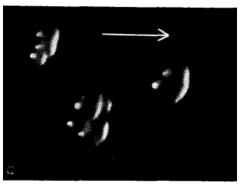




What is inside carbon black particles?

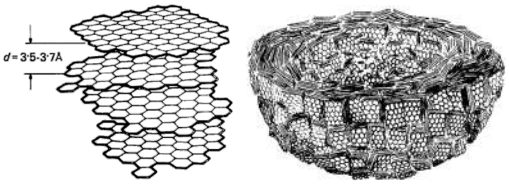
1948 - Hall used dark field electron microscopy to see orientation of crystallites is concentric.
1966 – Heckman revised model
1968 – Heidenreich uses phase contrast in electron microscope to record microstructure of carbon black.
1971 - Marsh imaged different carbon blacks with acetylene blacks being partially graphitised.

Hall's dark field EM

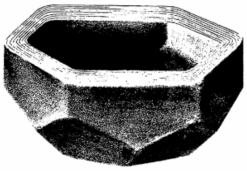


Heckman's carbon black model

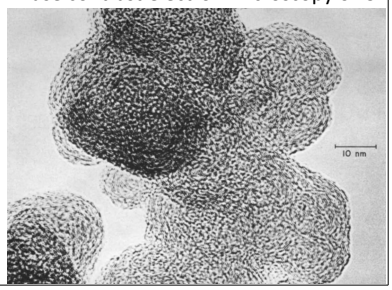
Heidenreich's carbon black model



Heated CB becomes polyhedron



Phase contrast electron microscopy of CB



Carbon blacks

Pigment blacks

- Used for printing inks - particle size and surface determine colour and viscosity
- The coating sector uses jet black – oxidised, fine particles
- **Plastic industry** - fine particles for UV resistance and for anti-static, e.g. power cables, carbon brushes and electrodes
- **Paper industry** - medium size particles – decoration
- **Construction industry** – coarse particles

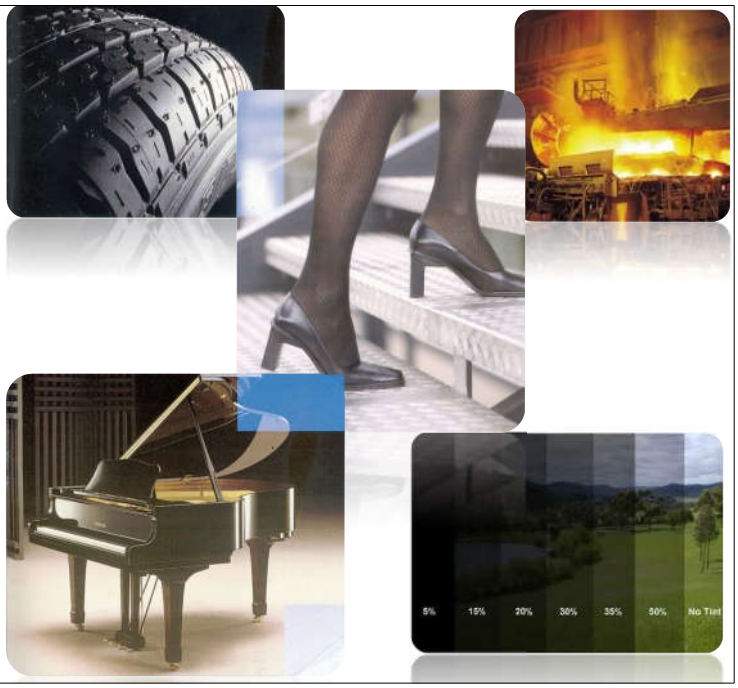
Carbon blacks

Reinforcing and rubber blacks

- Discovered by accident in the 19th century
- Replaced zinc oxide
- Eliminates the stickiness of rubber

- Active blacks
  - E.g.tires - size: 20nm
- Semi-active
  - E.g. floor mates – size:50nm

- Characterised by **size, surface area and after treatment**
- More than **90% of carbon black** for the rubber industry



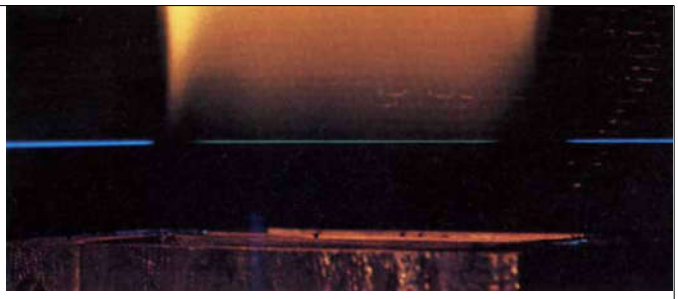
New applications

Fluorescent quantum dots

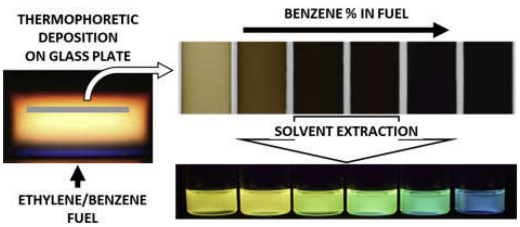
- 1970s nanoparticles detected in electron microscope.
- 1980s UV lasers found fluorescence in flames.

- 2005 Miller suggested stacked aromatics enable excimer states
- 2019 Wang *et al.* determined quantum confinement effect in stacked clusters exp. and theory.

- 2007 Liu, Ye and Mao extracted fluorescent np from candle soot.
- 2019 Ethylene flame and 2020 Benzene controls size.



The Chemistry of Flames William C. Gardiner Jr.



Sirignano, Mariano, Carmela Russo, and Anna Ciajolo. "One-step synthesis of carbon nanoparticles and yellow to blue fluorescent nanocarbons in flame reactors." Carbon 156 (2020): 370-377.

PART C: POLLUTION

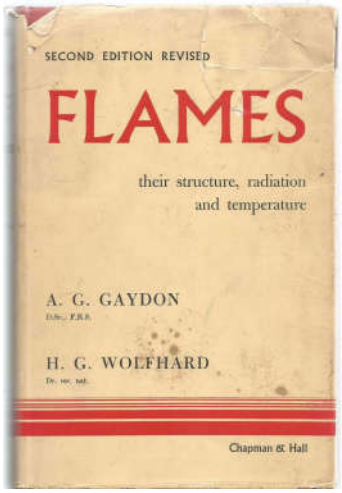
Pollution shines spotlight on combustion

- 1952 Great Smog of London shifted focus to soot reduction with Clean Air Act of 1956 (12k deaths).
- 1950s Haagen-Smit discovered photochemical fog mechanism with nitrous oxides and HC.
- 1963 USA passes Clean Air Act.

Soot symposium Imperial college 1961

1960 – Gaydon and Wolfhards new edition of Flames was just published. Included a significant review of main reaction mechanisms

Focuses on the significant amount of hydrogen present, XRD of crystallites and presence of radicals suggesting polybenzenoid radicals.



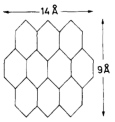


Figure 2. Circumanthracene, $C_{40}H_{16}$

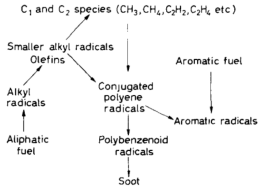


Figure 3. Proposed scheme of reactions

Before this symposium common to refer to carbon formation after this symposium predominately soot formation.

Palmer and Cullis 1965

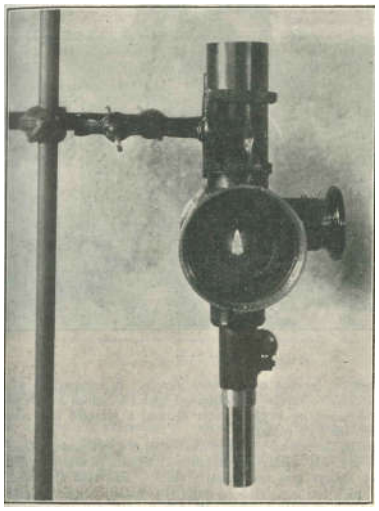
"A major breakthrough in understanding carbon formation will have been achieved when it becomes possible in at least one case to *account for the* ***entire course of nucleation and growth of carbon*** on the basis of a *fundamental knowledge of* ***reaction rates and mechanisms***."

Chemistry and Physics of Carbon, Palmer and Cullis 1965

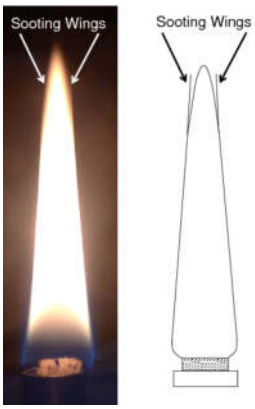
Critical aspect missing is a mechanism for nucleation

Sooting thresholds

1927 – smoke point measured in standard wick-fed diffusion flame.
1983 – Calcote and Manos define the Threshold sooting index that corrects for different burner designs and is dependent on molecular mass.
2007 – Pfefferle et al. developed the Yield Sooting index
2013 – Kraft et al. developed the FURTI method for smoke point.



Kewley and Jackson "Wick fed lamps" (J. Inst. Petr. Techn. 1927 13, 364)

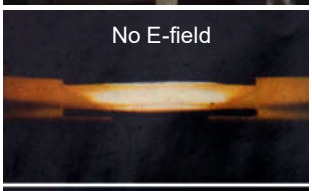


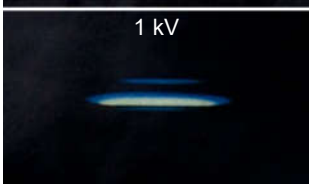
Calcote and Manos 1983 TSI

Electrical control of soot formation

1814 – Brande showed that a flame can be
1957 – Calcote showed chemionisation reactions lead to high concentration of flame ions.
1982 – Hayhurst showed C3H3+ primary ion in sooting flames.
1997 – Fialkov extensive review.

1967 – Weinberg showed reduction in soot formation with an electric field.
2018 – Martin et al. suggest flexoelectricity of curved PAH lead to effects seen.
2022 – Simulation of counterflow flame suggests impact on nucleation (Liu et al. C&F 239)





rich flame

$$CH^* + C_2H_2 \rightarrow C_3H_3^+ + e^-$$

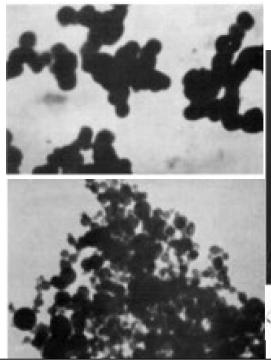
lean flame

$$CH + O \rightarrow CHO^+ + e^-$$

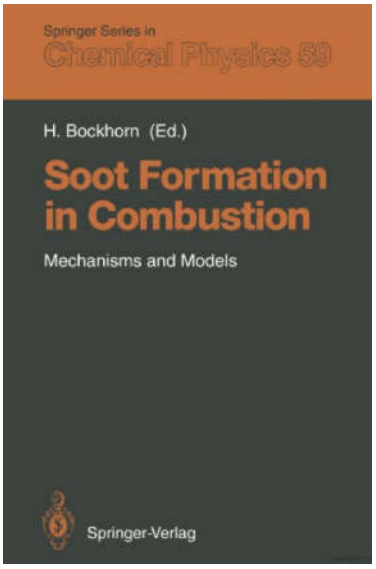
10^9-10^{11} ions cm^{-3}

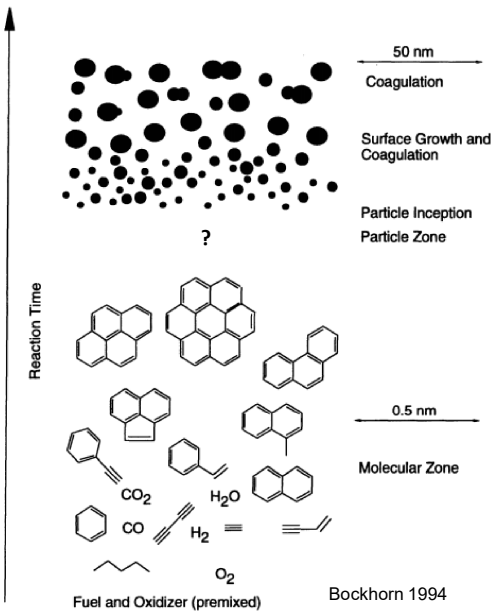
Fialkov 1997

Electron microscopy shows impact on primary particle size



Heidelberg Symposium 1991

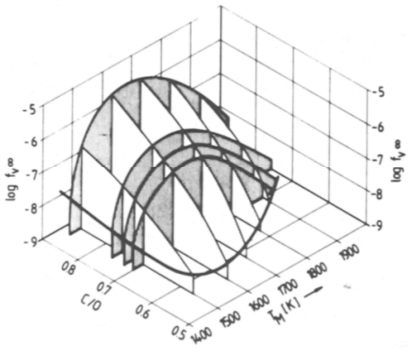


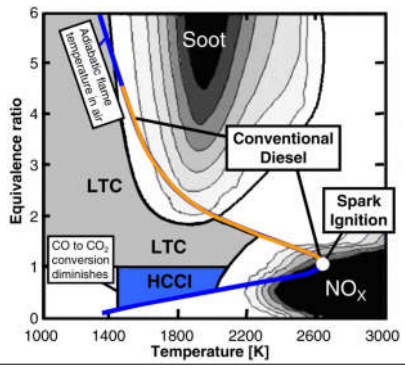


Bockhorn 1994

Heidelberg Symposium 1991 –T and P dependence

1989 – Bohm, Wagner and Weiss show the sooting limit as a function of pressure and richness C/O.
Combined with nitrous oxide kinetic mechanism this has led to the development of the low temperature combustion engines





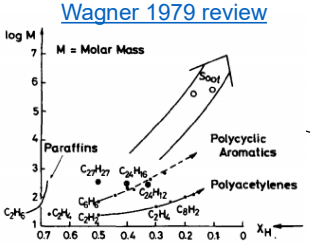
Heidelberg Symposium 1991 – Mass spectroscopy

Wagner and Homann combined the results of mass spectrometry into a plot of hydrogen fraction and the logarithm of molecular mass.

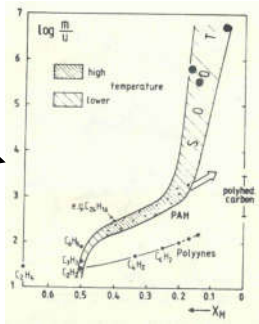
Summary of results

- Polyynes break down to form PAH
- PAH form through the HACA mechanism
- Significant hydrogen concentration in early soot particles.

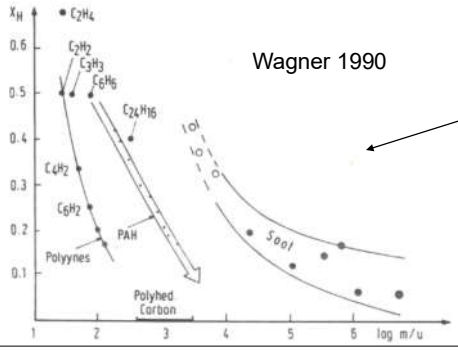
Wagner 1979 review



Homann 1990



Wagner 1990

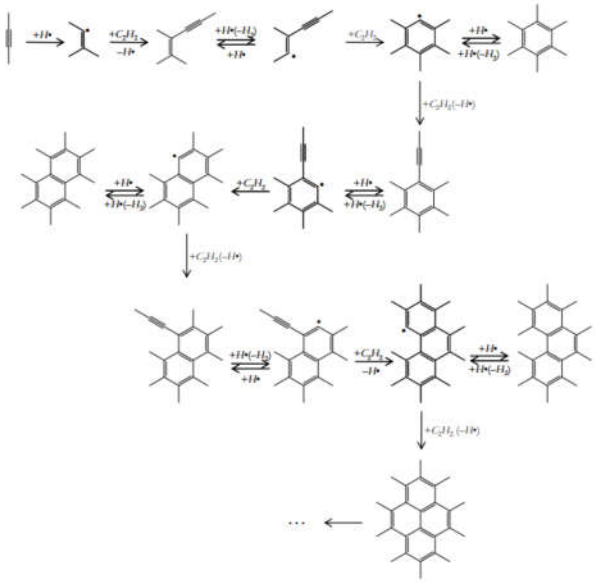


Heidelberg Symposium 1991 – HACA mechanism

1990 – Frenklach and Wang developed HACA mechanism using data from shocktubes.

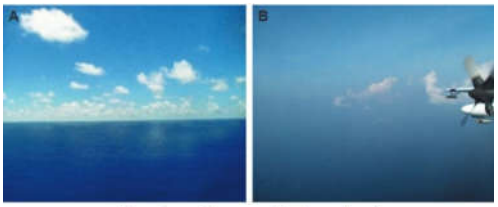
1997 – Kennedy reviewed kinetic models.

Goal was to be able to predict soot emissions from any combustion device.



Impact on the climate

- 1988 – IPCC was formed
- 1997 – Kyoto Protocol
- 2000s saw climate models demonstrate the warming impact of soot on the climate.



Reduction of Tropical Cloudiness by Soot

A. S. Ackerman,[1*] O. B. Toon,[2] D. E. Stevens,[3] A. J. Heymsfield,[4] V. Ramanathan,[5] E. J. Welton[6]

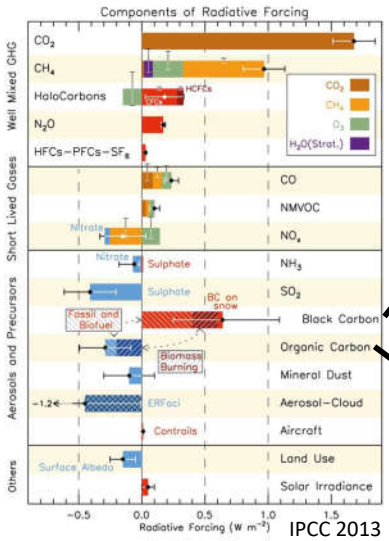
Measurements and models show that enhanced aerosol concentrations can augment cloud albedo not only by increasing total droplet cross-sectional area, but also by reducing precipitation and thereby increasing cloud water content and cloud coverage. Aerosol pollution is expected to exert a net cooling influence on the global climate through these conventional mechanisms. Here, we demonstrate an opposite mechanism through which aerosols can reduce cloud cover and thus significantly offset aerosol-induced radiative cooling at the top of the atmosphere on a regional scale. In model simulations, the daytime clearing of trade cumulus is hastened and intensified by solar heating in dark haze (as found over much of the northern Indian Ocean during the northeast monsoon).

A primary objective of the Indian Ocean Experiment (INDOEX) was to quantify the indirect effect of aerosols on climate through their effects on clouds (*1*). Conventionally, increased aerosol concentrations are expected to increase cloud droplet concentrations, and hence, total droplet cross-sectional area, thereby causing more sunlight to be reflected to space (*2*). Furthermore, model simulations of marine stratocumulus (*3–5*) and observations of ship tracks (*6–8*) suggest that increased aerosol concentrations can enhance cloud water content, physical thickness, and areal coverage by decreasing precipitation. Deep layers of dark (solar-absorbing) haze were observed over much of the tropical northern Indian Ocean in February-March of 1998 and 1999 during INDOEX (*9, 10*). The

[1]NASA Ames Res

12 MAY 2000 VOL 288 SCIENCE www.sciencemag.org

Impact on the climate

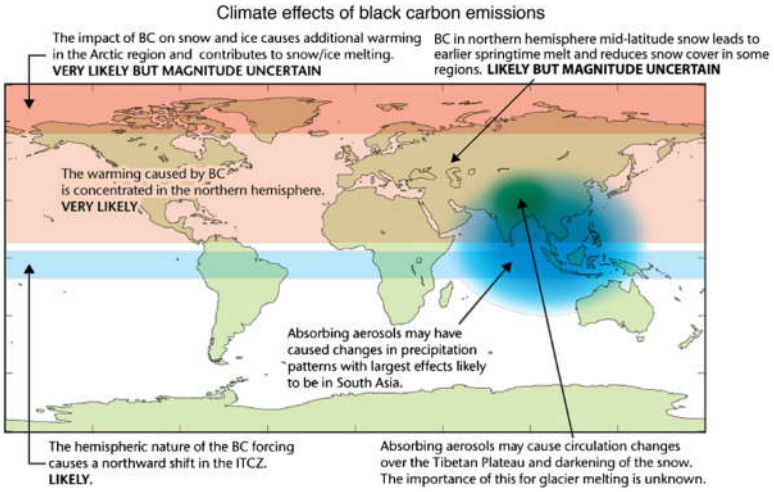


IPCC 2013

Internal combustion engines and furnaces produce black carbon.

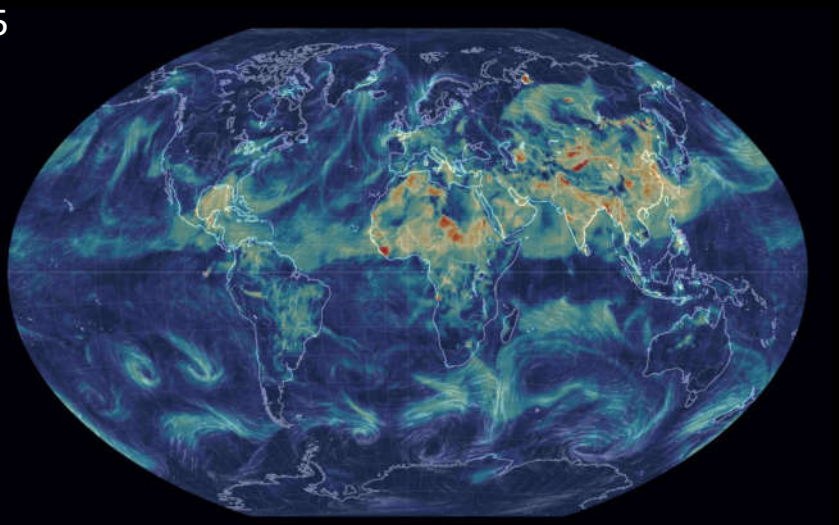
Agricultural fires also produces brown/organic carbon.

Impact on the climate

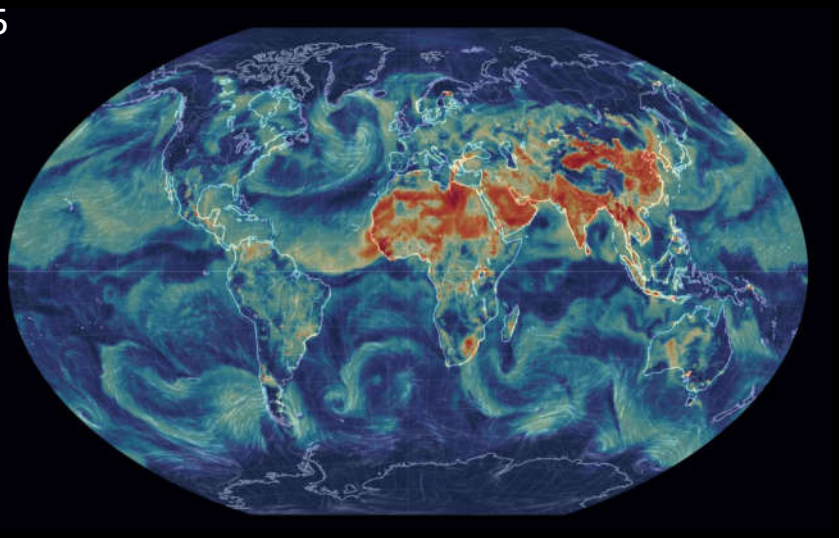


Bond 2013

PM2.5
2020



PM2.5
2019



Six cities study 1993

- Epidemiological studies show a **clear correlation** of **particulate exposure to mortality**
- Soot is a significant portion of ambient particulates as size decrease

The New England
Journal of Medicine

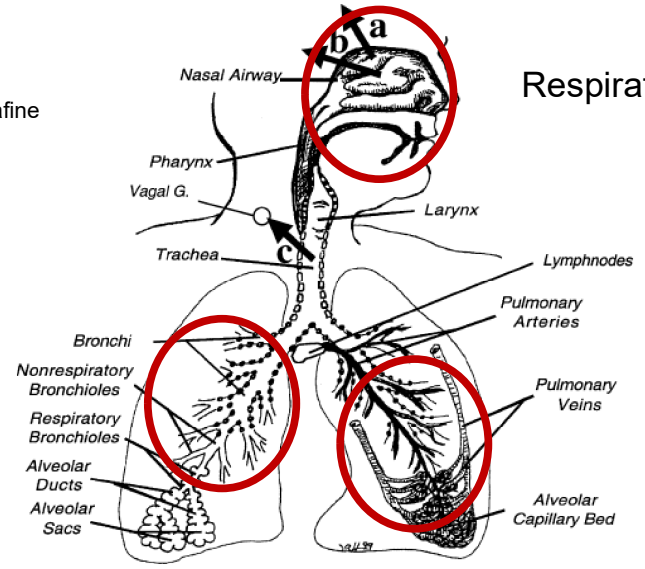
Volume 329 DECEMBER 9, 1993 Number 24

AN ASSOCIATION BETWEEN AIR POLLUTION AND MORTALITY IN SIX U.S. CITIES

Douglas W. Dockery, Sc.D., C. Arden Pope III, Ph.D., Xiping Xu, M.D., Ph.D.,
John D. Spengler, Ph.D., James H. Ware, Ph.D., Martha E. Fay, M.P.H.,
Benjamin G. Ferris, Jr., M.D., and Frank E. Speizer, M.D.

Six cities study 1993

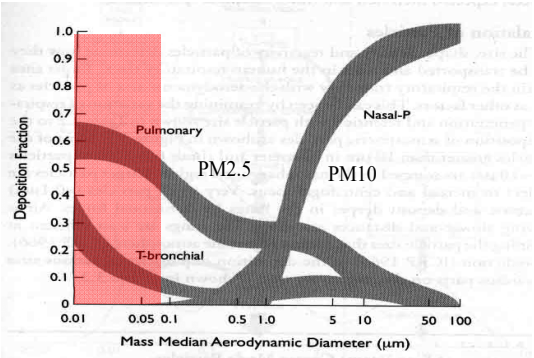
Ultrafine particles
penetrate
deeper into airways
Translocation of ultrafine
particles to brain

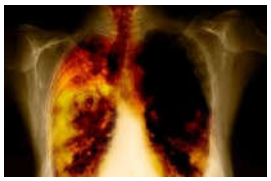


Respiratory tract

Six cities study 1993

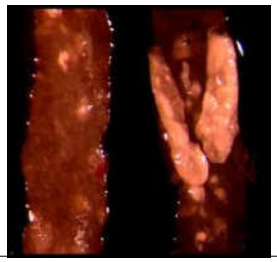
Ultrafine particles
penetrate
deeper into airways
Translocation of ultrafine
particles to brain





1. Lung cancer

2. Asthma

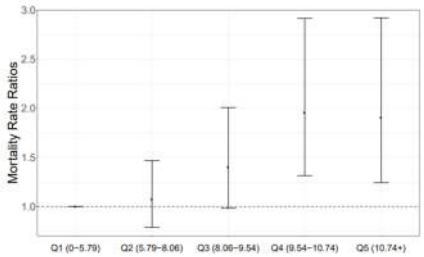


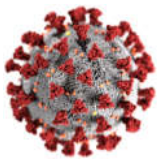
3. Vascular inflammation…

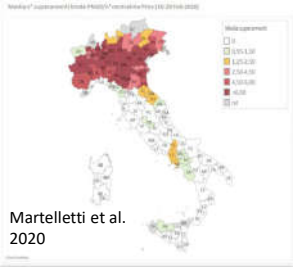
Impact on humans

- SARS correlation with air pollution index in China (Cui et al. 2003)
- Preprint from Harvard shows 1 µg/m³ in $PM_{2.5}$ is associated with an 2-15% increase in COVID-19 death rate? (Wu et al. medRxiv 2020)
- Particulates as carrier of the virus? As with influenza and measles (Setti et al. 2020)
- Is it due to copollutant NOx (Ogen 2020, Martelletti et al. 2020)

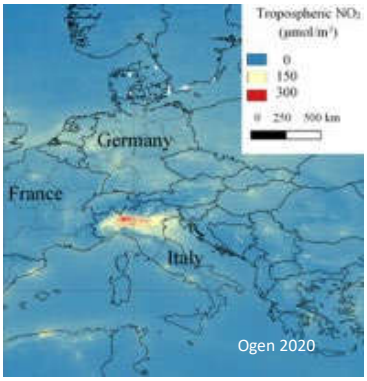
From 66 administrative regions in Italy, Spain, France and Germany, 78% of COVID-19 deaths occurred in the five most polluted regions. (Ogen 2020)





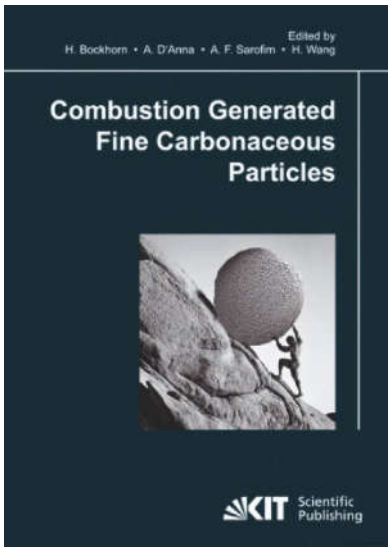


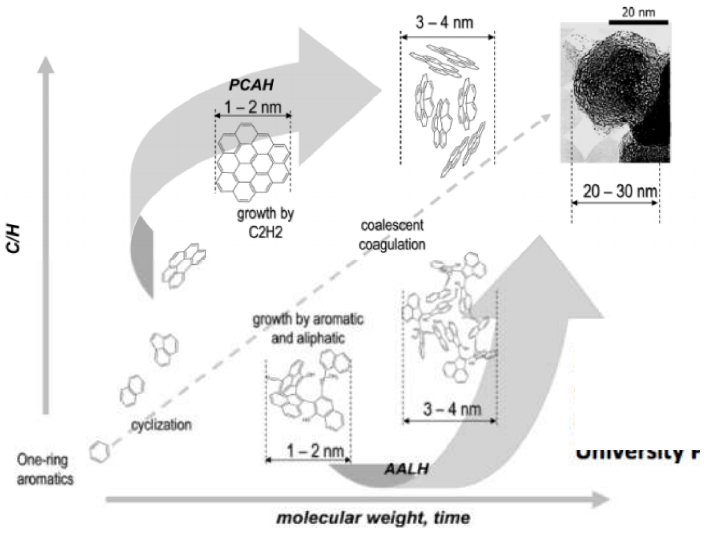
Martelletti et al. 2020



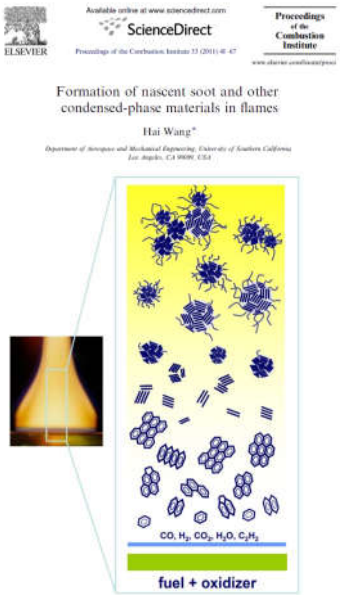
Ogen 2020

Anacapri Workshop 2007

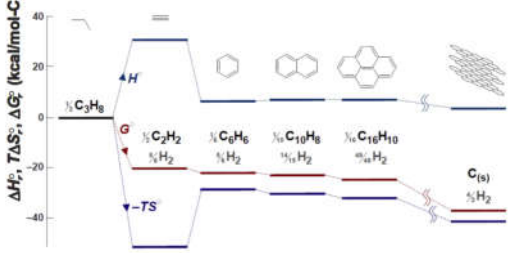




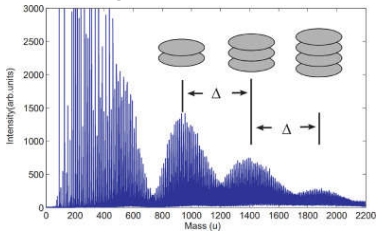
Wang Review 2011 Combustion Symposium



Soot formation is driven by entropy

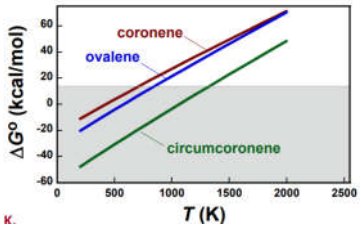


Clustering of aromatics ~500 Da occurs

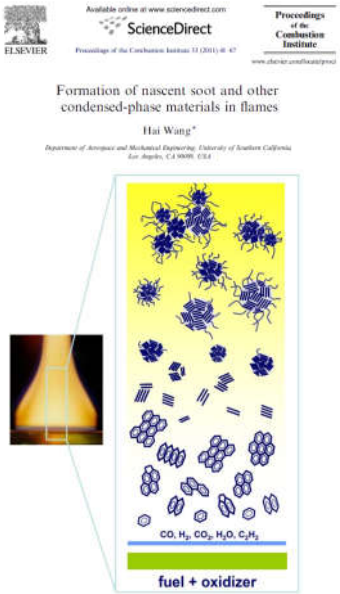


Soot precursors consisting of stacked pericondensed PAHs
Happold, H.-H. Grotheer, M. Aigner 2009

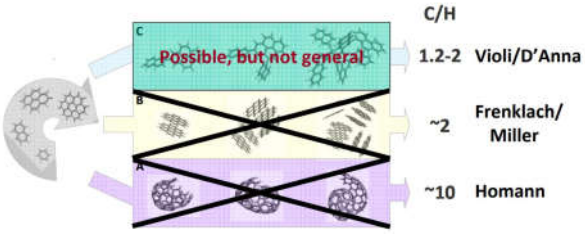
Weak interactions cannot lead to nucleation of PAH ~500 Da.



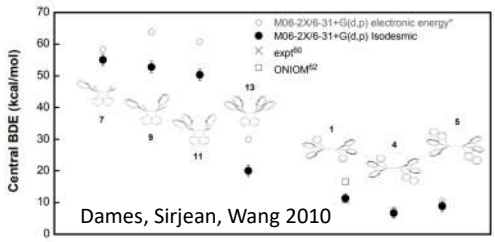
Wang Review 2011 Combustion Symposium



No current mechanism can explain soot formation

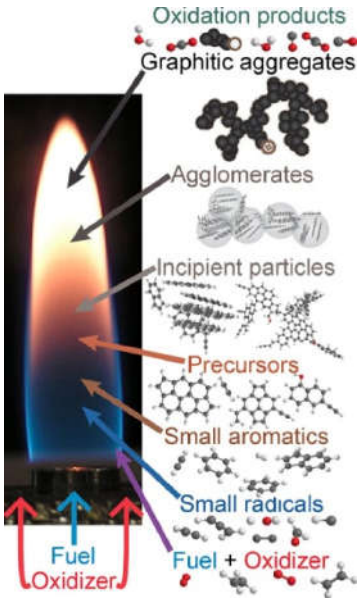


Suggests localised π-radicals could be the key

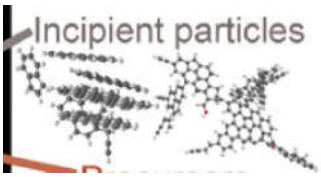


Dames, Sirjean, Wang 2010

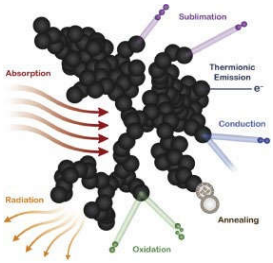
Michelsen 2016 review



Detection of incipient particles is critical

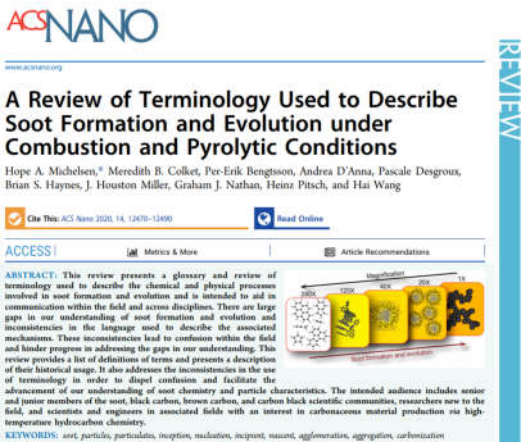


Review of *in situ* experimental techniques



Pressing need for
(1) diagnostics that allow single-shot measurements for turbulent systems,
(2) methods that are feasible for measurements under high and evolving pressure and optically dense conditions, e.g., engine conditions,
(3) techniques that can be applied to rapidly evolving biomass-burning events, i.e., wildfires, and
(4) diagnostics that are non-invasive to the combustion system under study.

Sooting Flame Workshop 2018 Dublin



ACSNANO

REVIEW

A Review of Terminology Used to Describe Soot Formation and Evolution under Combustion and Pyrolytic Conditions

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

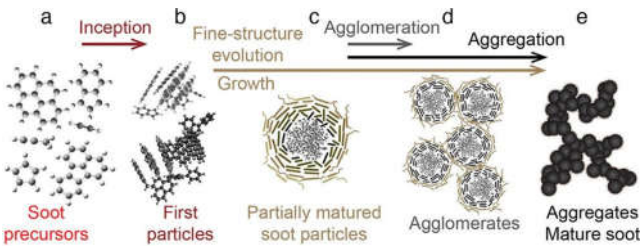
Cite This: ACS Nano 2020, 14, 12470−12490 | Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: This review presents a glossary and review of terminology used to describe the chemical and physical processes involved in soot formation and evolution and is intended to aid in communication within the field and across disciplines. There are large gaps in our understanding of soot formation and evolution and inconsistencies in the language used to describe the associated mechanisms. These inconsistencies lead to confusion within the field and hinder progress in addressing the gaps in our understanding. This review provides a list of definitions of terms and presents a description of their historical usage. It also addresses the inconsistencies in the use of terminology in order to dispel confusion and facilitate the advancement of our understanding of soot chemistry and particle characteristics. The intended audience includes senior and junior members of the soot, black carbon, brown carbon, and carbon black scientific communities, researchers new to the field, and scientists and engineers in associated fields with an interest in carbonaceous material production via high-temperature hydrocarbon chemistry.

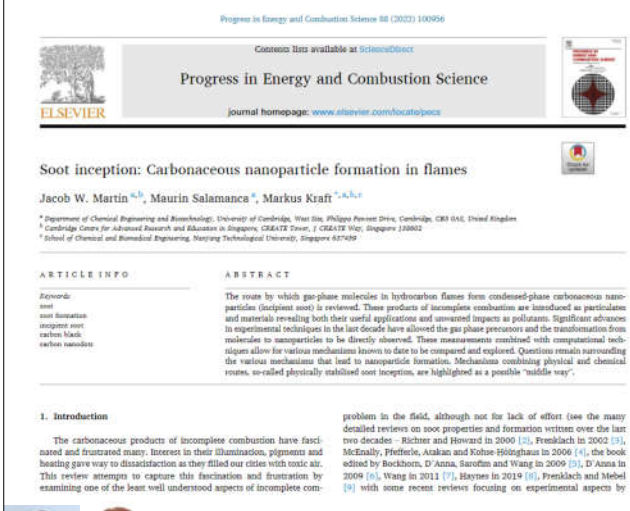
KEYWORDS: soot, particles, particulates, inception, nucleation, incipient, nascent, agglomeration, aggregation, carbonization

Agreement to use the word inception instead of nucleation



Michelsen, H. A., Colket, M. B., Bengtsson, P. E., D'Anna, A., Desgroux, P., Haynes, B. S., ... & Wang, H. (2020). A review of terminology used to describe soot formation and evolution under combustion and pyrolytic conditions. *ACS nano*, *14*(10), 12470-12490.

Soot inception review 2022 Martin, Salamanca and Kraft



Progress in Energy and Combustion Science 88 (2022) 100956

Contents lists available at ScienceDirect

Progress in Energy and Combustion Science

journal homepage: www.elsevier.com/locate/pecs

Soot inception: Carbonaceous nanoparticle formation in flames

Jacob W. Martin [a,b], Maurin Salamanca [a], Markus Kraft [a,a,b,c]

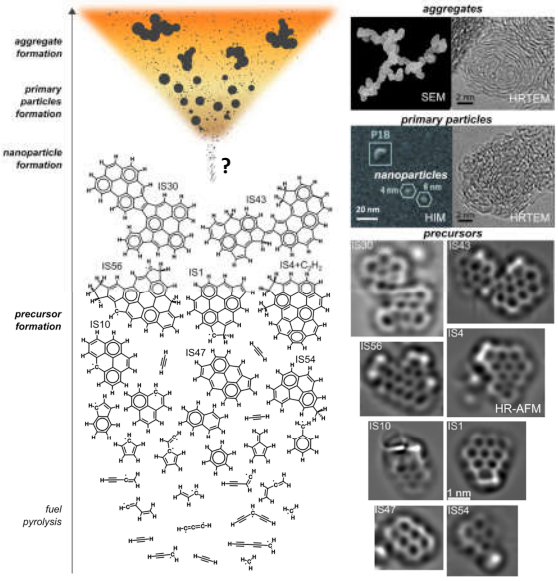
ARTICLE INFO

ABSTRACT

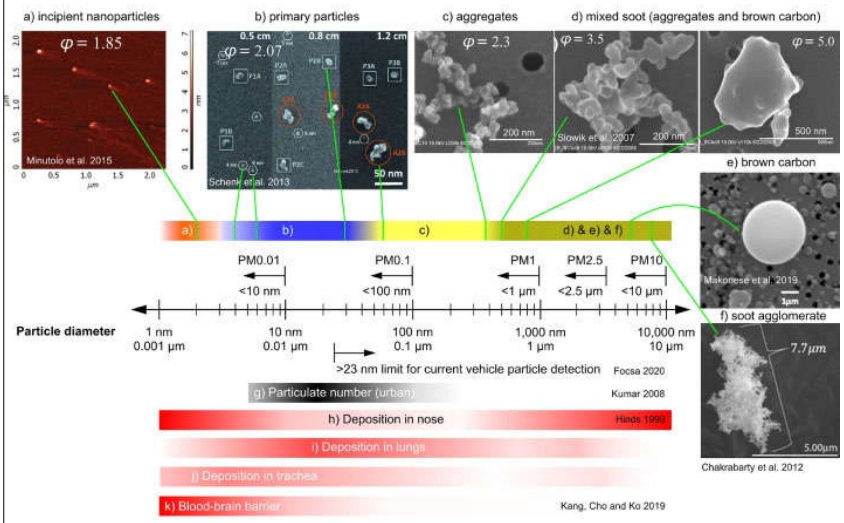
The route by which gas-phase molecules in hydrocarbon flames form condensed-phase carbonaceous nanoparticles (incipient soot) is reviewed. These products of incomplete combustion are introduced as particulates and materials revealing both their useful applications and unwanted impacts as pollutants. Significant advances in experimental techniques in the last decade have allowed the gas phase precursors and the transformation from molecules to nanoparticles to be directly observed. These measurements combined with computational techniques allow for various mechanisms known to date to be compared and explored. Questions remain surrounding the various mechanisms that lead to nanoparticle formation. Mechanisms combining physical and chemical routes, so-called physically stabilized soot inception, are highlighted as a possible "middle way".

1. Introduction

The carbonaceous products of incomplete combustion have fascinated and frustrated many. Interest in their illumination, pigments and heating gave way to dissatisfaction as they filled our cities with toxic air. This review attempts to capture this fascination and frustration by examining one of the least well understood aspects of incomplete combustion in the field, although not for lack of effort (see the many detailed reviews on soot properties and formation written over the last two decades – Richter and Howard in 2000 [2], Frenklach in 2002 [3], McEnally, Pfefferle, Atakan and Kohse-Höinghaus in 2006 [4], the book edited by Bockhorn, D'Anna, Sarofim and Wang in 2009 [5], D'Anna in 2009 [6], Wang in 2011 [7], Haynes in 2019 [8], Frenklach and Mebel [9] with some recent reviews focusing on experimental aspects by

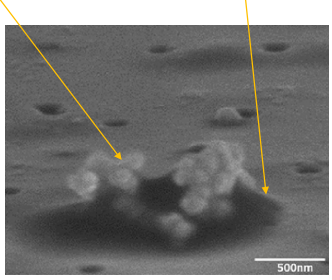


Evidence for nanoparticles <10 nm in flames



Soot aggregate

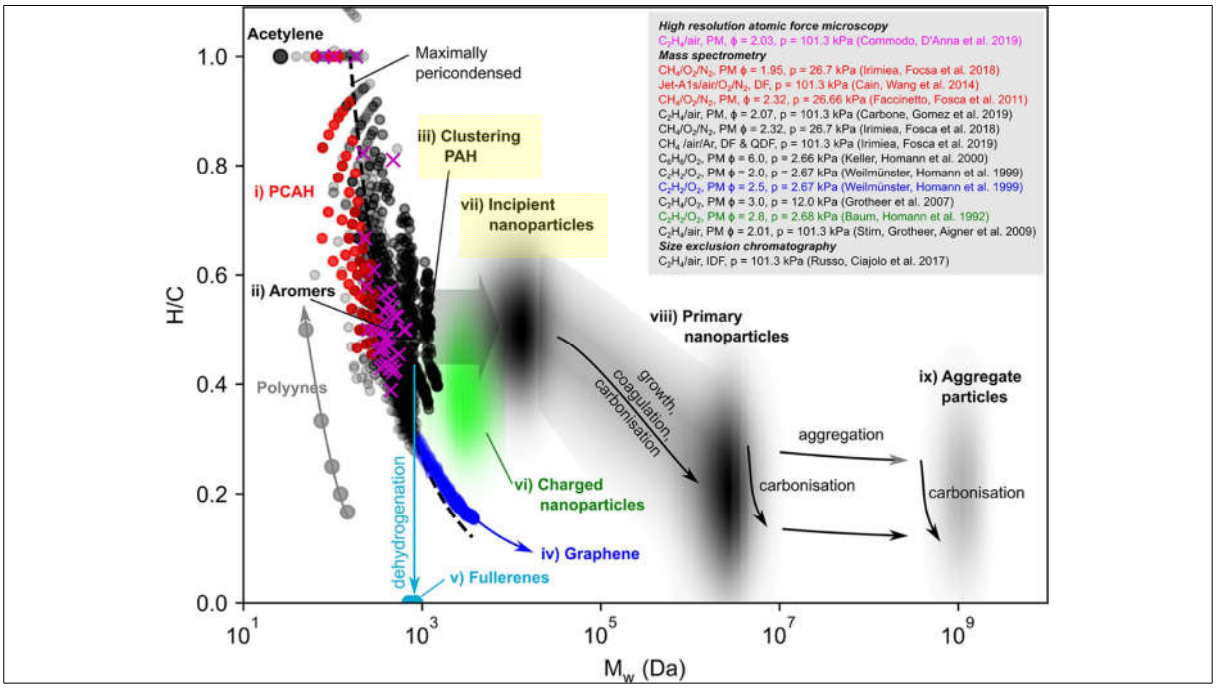
Brown carbon (tar)



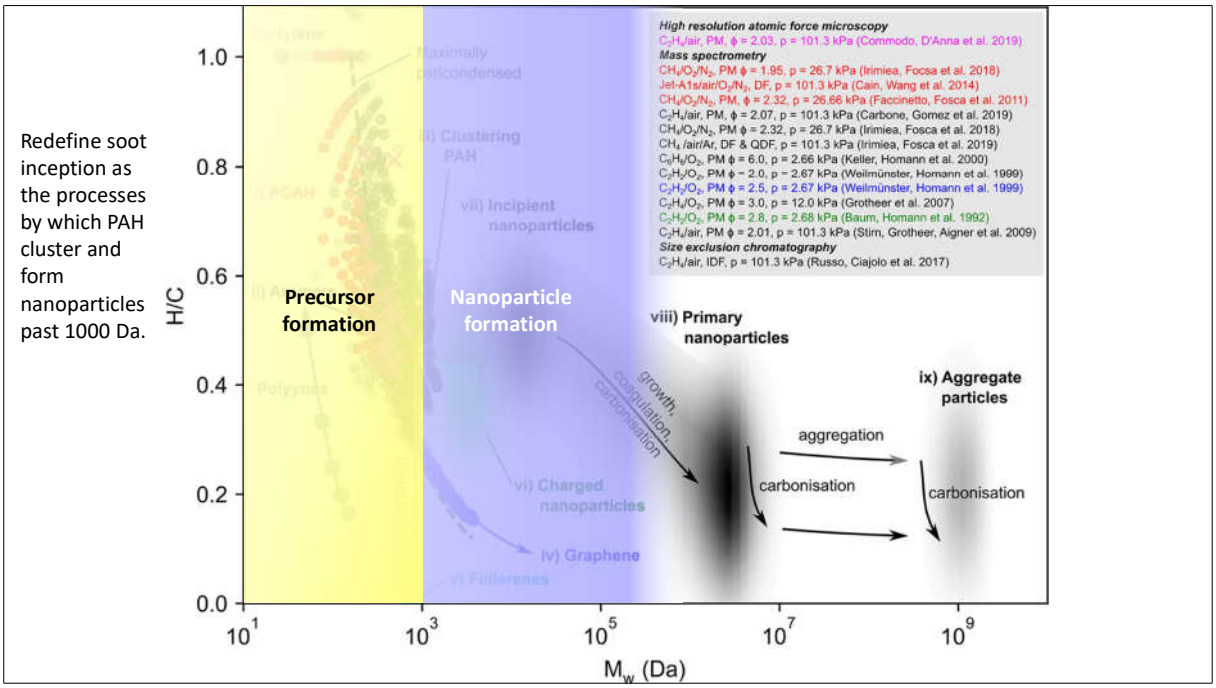
Sharma, Noopur, et al. "Physical properties of aerosol internally mixed with soot particles in a biogenically dominated environment in California." *Geophysical Research Letters* 45.20 (2018): 11-473.

No liquid like particles just a sampling artifact.
Tar is condensing on soot increase the sampling rate and the tar is removed.

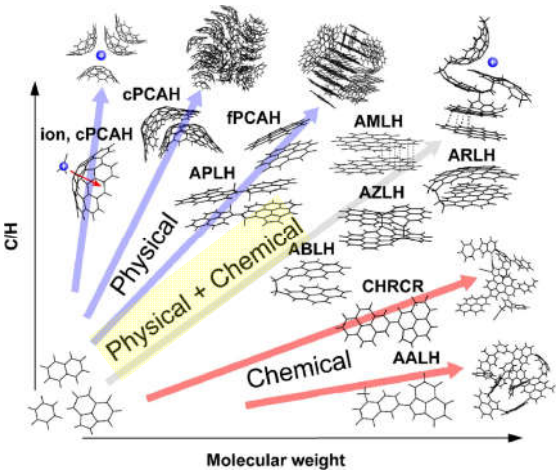
Botero, Maria L., et al. "On the thermophoretic sampling and TEM-based characterisation of soot particles in flames." *Carbon* 171 (2021): 711-722.

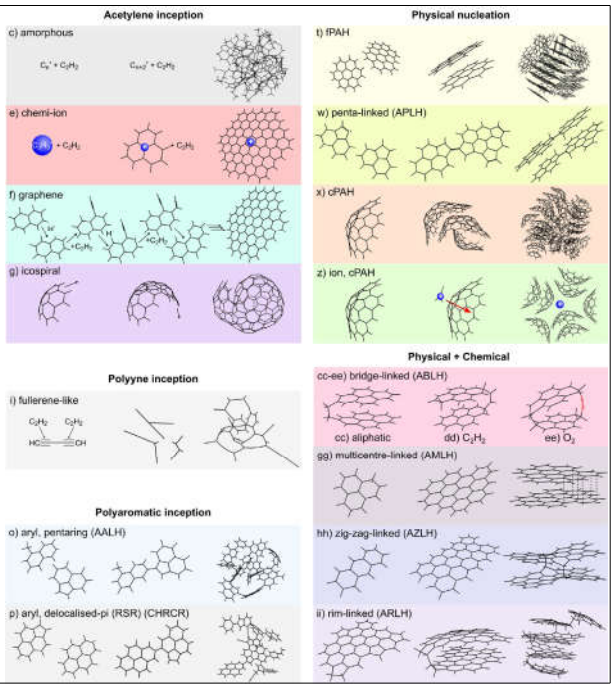


Redefine soot inception as the processes by which PAH cluster and form nanoparticles past 1000 Da.

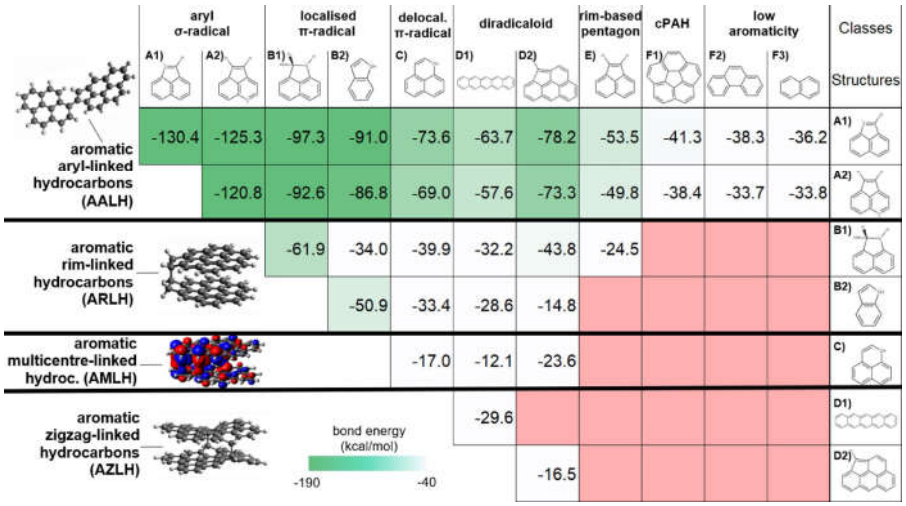


Many old and new mechanisms are discussed and a focus on mechanisms that involve **both physical and chemical interactions** are highlighted.



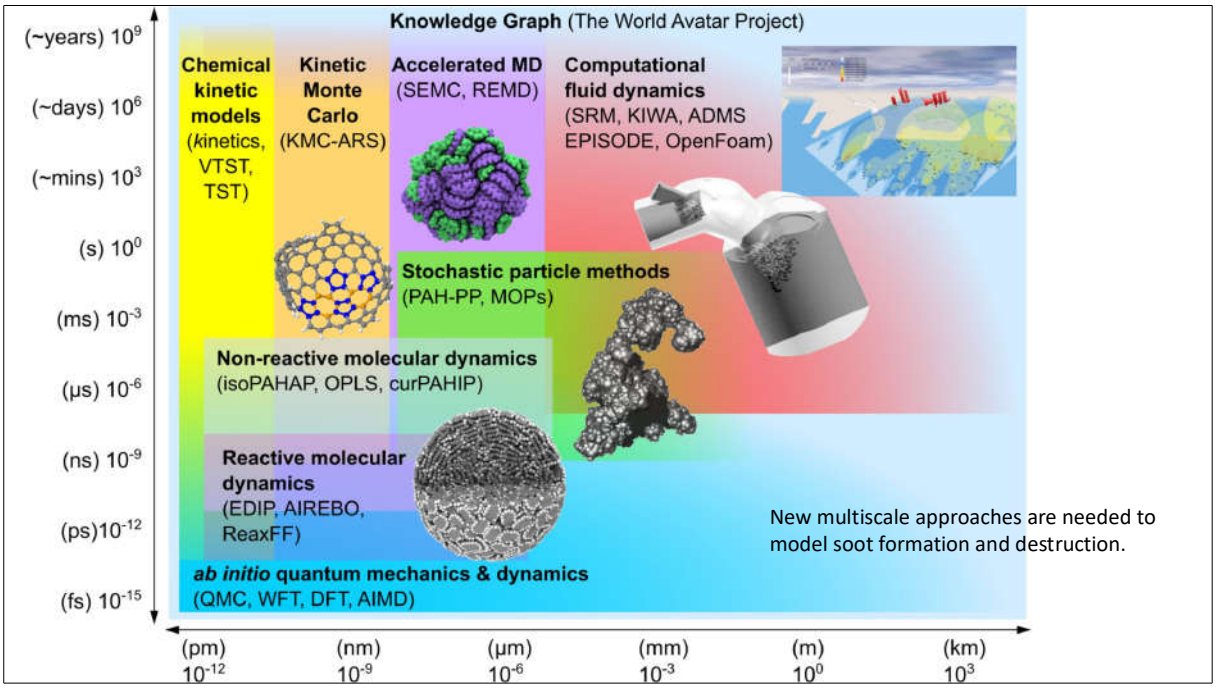


Many old and new mechanisms are discussed and a focus on mechanisms that involve **both physical and chemical interactions** are highlighted.

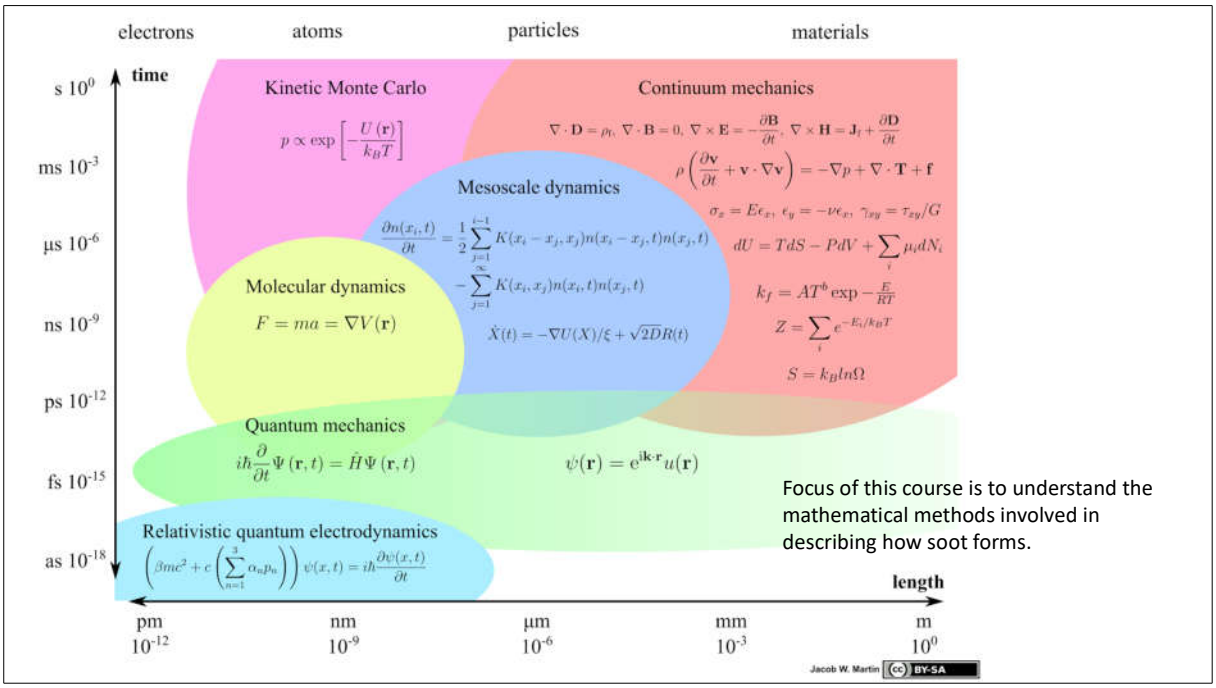


Martin, Jacob W., et al. "Reactivity of polycyclic aromatic hydrocarbon soot precursors: implications of localized π-radicals on rim-based pentagonal rings." *The Journal of Physical Chemistry C* 123.43 (2019): 26673-26682.

Menon, Angiras, et al. "Reactivity of polycyclic aromatic hydrocarbon soot precursors: Kinetics and equilibria." *The Journal of Physical Chemistry A* 124.48 (2020): 10040-10052.



New multiscale approaches are needed to model soot formation and destruction.



Focus of this course is to understand the mathematical methods involved in describing how soot forms.

End of Lecture 1

Soot – Part 2

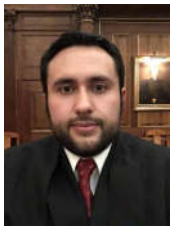
Markus Kraft

Computational Modelling Group Cambridge

Main Contributors:
Dr Jake Martin (Part 1)
Dr Angiras Menon (Part 2)
Dr Laura Pascazio (Part 3)
Dr Gustavo Leon (Part 4)

Jacob Martin

AngirasMenon

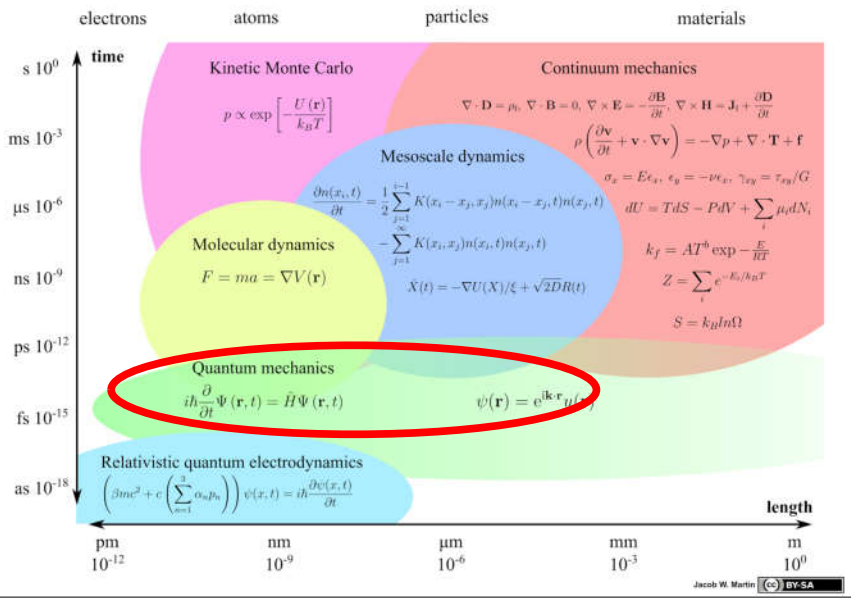


Gustavo Leon

Laura Pascazio

Part 1 Overview
Part 2 Quantum Chemistry
Part 3 Molecular Dynamics
Part 4 Kinetic Monte Carlo
Part 5 Stochastic Particle Methods
Part 6 Application – engine model

Scale Diagramme



Part A: Introduction to Quantum Chemistry

1. Introduction:
   What is Quantum Chemistry? , Some fundamental theory

Part B: Example Calculation in Gaussian

1. Running a DFT calculation in Gaussian

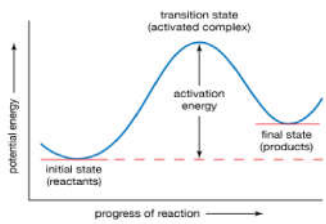
Part C: Introduction to Rate calculations

1. Introduction:
   Potential Energy Surface, Transition State Theory, Partition functions
2. Examples:
   Locating transition states using Gaussian, Using Multiwell with TST.

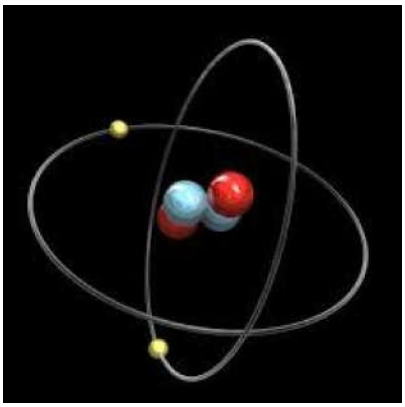
$$H(t)\,|\,\psi(t)\rangle = i\hbar\frac{d}{dt}\,|\,\psi(t)\rangle$$

Gaussian 16

Expanding the limits of computational chemistry



PART A: INTRODUCTION TO QUANTUM CHEMISTRY



What is Quantum Chemistry?

- Electronic Structure Theory: Concerned with the **motion of electrons** within atoms/molecules/complexes. Computes static properties of molecules from **electronic wavefunctions**, and derive **potential energy surfaces** (PES)
- Includes a variety of computational methods: Wavefunction methods, Density Functional Theory (DFT), Multireference methods, Multi-configurational methods, *etc*.
- *Ab Initio* Molecular Dynamics: Calculate the motion of molecules along a potential energy surface - non-equilibrium properties.
- …And much more.

Why do we use Quantum Chemistry?

- More accurate than classical methods (usually)…
- Is able to capture the behavior and movement of electrons (difficult for classical methods).
- Is *ab. Initio* – uses fewer empirical parameters compared to classical methods.

BUT…

- Fundamental equations are significantly complex.
- Calculations are much more expensive.
- Limits size of system that can be studied.

What can Electronic Structure give us?

- Equilibrium geometries – molecular structure
- Orbital energies and shapes – band structure

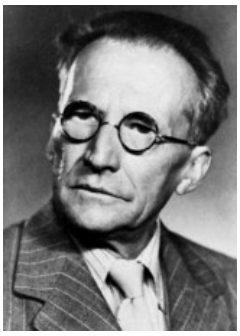
- Vibrational frequencies – for Infrared and Raman spectra
- Excited states, electronic transitions – for UV/Vis and photoelectron spectra
- NMR spectra

- Dipole moment, polarizability, charge densities – molecular interaction parameters

- Barrier heights, reaction paths, reaction rates (when combined with trajectory studies / TST)
- Thermodynamic properties (when combined with statistical mechanics)

The Schrödinger Equation

- Derived his famous equation(s) in 1925, published in 1926, Nobel Prize in 1933.



- Time Dependent Form:

$$i\hbar\frac{\partial}{\partial t}\Psi(r,t) = \hat{H}\Psi(r,t)$$

- Time Independent Form:

$$\hat{H}\Psi = E\Psi$$

The Schrödinger Equation

- For molecules, the Hamiltonian can be written as:

$$\hat{H} = \hat{T}_N(\boldsymbol{R}) + \hat{T}_e(\boldsymbol{r}) + V_{eN}(\boldsymbol{r},\boldsymbol{R}) + V_{NN}(\boldsymbol{R}) + V_{ee}(\boldsymbol{r})$$

Kinetic Energy of Nuclei

Kinetic Energy of Electrons

Electrons – Nuclei Interaction

Nuclei – Nuclei Interaction

Electrons – Electrons Interaction

The Electronic Schrödinger Equation

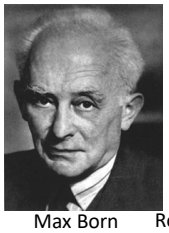
- The Born-Oppenheimer approximation: Nuclei are so heavy compared to electrons, so they are effectively stationary:

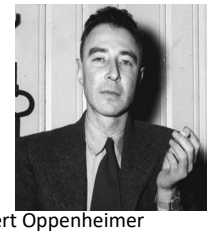
$$\widehat{T_N}(\boldsymbol{R}) \approx 0 \qquad V_{NN}(\boldsymbol{R}) = const.$$

- Once this approximation is made, we are left with the Electronic Schrodinger Equation (ESE):

$$\widehat{H}_{el}\Psi(\mathbf{r},\mathbf{R}) = E_{el}\Psi(\mathbf{r},\mathbf{R})$$

$$\widehat{H}_{el} = \widehat{T}_e(\boldsymbol{r}) + V_{eN}(\boldsymbol{r},\boldsymbol{R}) + V_{ee}(\boldsymbol{r}) + (V_{NN}(\boldsymbol{R}))$$


Max Born


Robert Oppenheimer


Olivia Newton John

- **The goal of most computational quantum chemistry is to solve the Electronic Schrodinger Equation.**

Solving the ESE: Hartree-Fock Theory

- Need to solve for the electronic wavefunction somehow. Typically, the method is to guess.
- The most basic way of doing this that is implemented in quantum chemistry codes is **Hartree – Fock theory.** This is the basis for most more advanced methods.
- How does it work?
  - 1. Use Born-Oppenheimer Approximation.
  - 2. Write the electronic wavefunction as a Slater determinant:
  - 3. Vary the orbitals until the electronic energy is minimized.


John Slater


Douglas Hartree


Vladimir Fock

Slater Determinant

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix}$$

Number of Electrons

Orbitals

Position and spin of Nth electron

Determining the Orbitals: Introducing Basis Sets

- Solving the governing equations in quantum chemistry become easier if the orbitals are represented as a weighted sum of basis functions:

$$\chi_i = \sum_{i=1}^{N} C_{\phi i} \tilde{\chi}_{\phi}$$

- The basis functions describing the molecular orbitals is known as the **basis set.**
- The basis set is an **input** to quantum chemistry codes. **Choice of basis set is often key to getting reasonable results.**
- In most quantum chemistry codes, two types of basis functions are commonly used: atom-centred gaussian-type orbitals, and plane-wave basis sets.

Gaussian-type Orbital Basis Sets

- Can be written as:

$$\tilde{\chi} = N x^a y^b z^c e^{-\zeta r^2}$$

- Use Gaussians to describe atomic orbitals – sum of atomic orbitals gives molecular orbitals.
  - Strengths:
    - Can get decent results with relatively small number of functions
    - Polarization and diffuse functions help describe actual molecular orbitals
    - Can include core electrons in calculations
    - Transition metals described better
  - Weaknesses:
    - Numerically demanding - limited to smaller systems than plane-wave.
    - Atom specific – has to specify functions for every different atom. Molecules and solids are also different.
    - Rarer atoms may not be included.
- **Typically best for gas-phase calculations of molecules.**
- Example software: Gaussian, MOLPRO, GAMESS, NWChem

Plane-wave Basis Sets

- Can be written as:

$$\tilde{\chi} = \frac{1}{\sqrt{\Omega}} e^{-i\boldsymbol{G}.\boldsymbol{r}}$$

- Treat everything like a free-electron gas, and use plane-waves to describe calculation space.
  - Strengths:
    - Naturally orthogonal and periodic
    - Numerically efficient – can study larger systems
    - Not atom-dependent; basis set is identical for gas or solids.
  - Weaknesses:
    - Plane waves cannot describe electrons near nuclei well – need to use pseudopotentials
    - Dependent on the size of your box
    - Very inefficient for non-periodic systems / gas phase

- **Typically best for solid-state calculations (crystals, polymers, periodic systems).**
- Example software: CASTEP, VASP, Quantum ESPRESSO, NWChem

Beyond Hartree – Fock : Density Functional Theory

- Hartree-Fock is an approximation, and unsurprisingly, is often insufficient.
- Main Error: Each electron in the system treats the other electrons as an average potential.
- This induces an error in the calculated energy, called the electron correlation energy.
- As a result Hartree-Fock does very poorly at describing energies of bond-breaking, radicals, transition metals, and more.
- Need more accurate methods : **Density Functional Theory (DFT),** post-Hartree-Fock methods, advanced wavefunction methods.

DFT : What is it?

- Hohenberg and Kohn showed using **magic** (advanced mathematics) that solving for the electron density can give all the necessary ground state properties of a molecule.
- Can write the electronic Hamiltonian in terms of electron density:

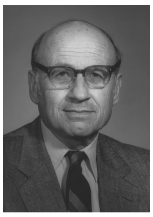
$$\hat{H}_{el} = \hat{T}_e + \hat{V}_{eN} + \hat{V}_{ee}$$

$$\downarrow$$

$$E[p] = T[\rho] + E_{eN}[\rho] + E_{ee}[\rho]$$

- Kohn and Sham then approximated the kinetic energy of an electron density as that of a Slater Determinant that gives the same energy as the electron density.

Pierre Hohenberg



Walter Kohn

Lu Jeu Sham

Kohn-Sham DFT: Introducing Functionals

- The Kohn-Sham DFT gives the energy as:

$$E_{KS-DFT} = T_{Slater}[\rho] + E_{eN}[\rho] + J[\rho] + \boldsymbol{E_{xc}[p]}$$

- All of the terms are known analytically except for the last one, $E_{xc}[\rho]$. This term is called the exchange-correlational term, and is not known *a priori.*
- Typically, one makes some approximation of how to calculate $E_{xc}[\rho]$. These are known as **functionals.**
- The functional is the other major **input** to quantum chemistry codes. **Choice of functional is often key to getting reasonable results.**

Types of Functionals

- Generally functionals are classified as either **pure** functionals or **hybrid** functionals.
- Pure functionals – based on **local** exchange correlation for a free electron gas.
- Two main approximations, the local density approximation (LDA) and the generalized gradient approximation (GGA).

- LDA: $E_{xc}^{LDA}[\rho] = \int \epsilon_{xc}(\rho)\rho(\boldsymbol{r})d^3r$ – value of functional only depends on density at that point. Density is same everywhere. Cheapest but lowest accuracy.

- GGA: $E_{xc}^{GGA}[\rho, \nabla\rho] = \int \epsilon_{xc}(\rho, \nabla\rho)\rho(\boldsymbol{r})d^3r$ – value of functional depends on density and density gradient. Density is not the same everywhere. Better but not that good.

- Hybrid functionals: Improve upon accuracy by mixing GGA/LDA with Hartree-Fock for the electrostatics, and some functions for the correlation and non-local effects: More accurate, but more expensive. **Should use over pure functionals whenever possible.**

PART B: RUNNING CALCULATIONS IN GAUSSIAN 09/16

Gaussian 16

Running a DFT calculation in Gaussian

An example input file for benzene

```
%nprocshared=20
%mem=60GB
%chk=benzexmp.chk
#B3LYP/6-311+g(d,p) Opt Freq

benzexmp

0 1
C          -1.21173        0.47992        0.02427
C          -0.39583        1.68434        0.02427
C           1.05518        1.57995        0.02427
C           1.69029        0.27114        0.02427
C           0.87438       -0.93328        0.02427
C          -0.57663       -0.82889        0.02427
H          -0.86848        2.65837        0.02427
H           1.66240        2.47630        0.02427
H           2.77016        0.19345        0.02427
H           1.34703       -1.90732        0.02427
H          -1.18385       -1.72524        0.02427
H          -2.29160        0.55761        0.02427
```

Number of processors to use, memory, name of checkpoint file (for restarting jobs).

Main job specification: **Functional/Basis Set Type of Job to Run**

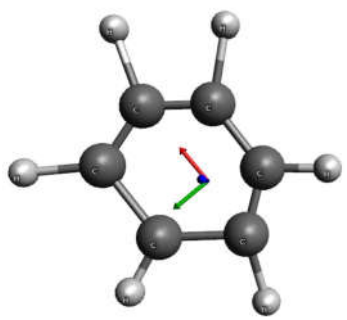
Name of Job (can be anything you want)

Charge of Molecule, Multiplicity of Molecule

Atom list and corresponding xyz coordinates

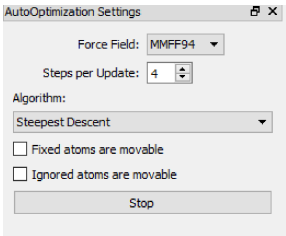
Generating an Input File(1)

- Many molecular editors have in-built features for generating Gaussian input files. I personally use Avogadro (free and open-source) but plenty of others exist (Gaussview, MOLDEN, Chemdraw, etc.)
- First step is to draw molecule freehand in Avogadro, making sure to have the necessary atoms and bonds:

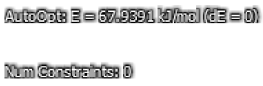


Generating an Input File(2)

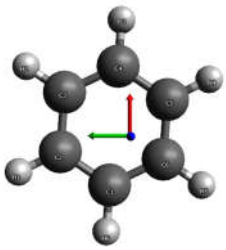
- Then, use one of Avogadro's in-built force fields (click the toolbar E) to make drawing look more like benzene. A poor geometry guess will usually cause DFT to crash, so this should always be done.



Sample Force-field settings



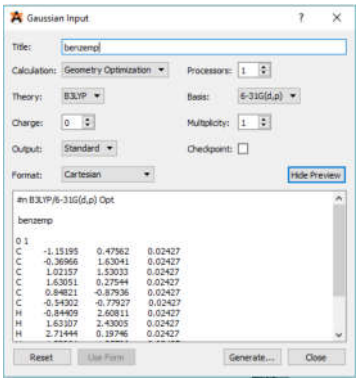
When dE is very small, the force field optimization is done



Gives a reasonable initial geometry

Generating an Input File(3)

- Under Extension in Avogadro, one can generate input files for all sorts of quantum chemistry packages. Here's one in Gaussian format:



- Of course, some manual edits are usually necessary, as Avogadro does not have all the job types, functionals, or basis sets available in it's built-in menu. It generates the coordinates for you which can be left as is.

Choosing a Job Type

- The three main ones:
  - Geometry Optimization (keyword **Opt**): Computes the equilibrium geometry of the input molecule. Can optimize to either a minima (default, for reactants/products) or a saddle point (by writing Opt=(Calcfc,TS), for transition states).
  - Frequencies (keyword **Freq**): Computes the vibrational modes for the input molecule (for IR/Raman spectra). Gaussian will also compute thermochemistry at 298K (enthalpy, entropy, zero-point corrections) when this is specified. Frequencies are necessary for rate constants. **Always run this whenever running an optimization.**
  - Single Point Energy Calculation (keyword **SP**, default if no other keyword provided). Computes the energy for the given input geometry. The input geometry is not changed. Usually, you run this to get the energy from a better functional without having to re-optimize the whole molecule.

Choosing a Basis Set(1)

STO-3G [Hehre69, Collins76]

3-21G [Binkley80a, Gordon82, Pie

6-21G [Binkley80a, Gordon82]

4-31G [Ditchfield71, Hehre72, Ha

6-31G [Ditchfield71, Hehre72, Ha

6-31G†: Gaussian 16 also inclu [Petersson88, Petersson91]. These f functions: e.g., 6-31+G(d'f').

6-311G: Specifies the 6-311G ba Raghavachari80b] (note that the for neutral molecules as well), first transition row, using the s elements in the third row [Binn using the Wachters-Hay basis 6-311G.

cc-pVDZ, cc-pVTZ, cc-pVQZ, cc-pV5Z, cc-pV6Z: Dunning's correlation consistent basis sets [Dunning89, Kendall92, Woon93, Peterson94, Wilson96] (double, triple, quadruple, quintuple-zeta and sextuple-zeta, respectively). These basis sets have had redundant functions removed and have been rotated [Davidson96] in order to increase computational efficiency.

These basis sets include polarization functions by definition. The following table lists the valence polarization functions present for the various atoms included in these basis sets:

Atoms	cc-pVDZ	cc-pVTZ	cc-pVQZ	cc-pV5Z	cc-pV6Z
H	2s,1p	3s,2p,1d	4s,3p,2d,1f	5s,4p,3d,2f,1g	6s,5p,4d,3f,2g,1h
He	2s,1p	3s,2p,1d	4s,3p,2d,1f	5s,4p,3d,2f,1g	*not available*
Li-Be	3s,2p,1d	4s,3p,2d,1f	5s,4p,3d,2f,1g	6s,5p,4d,3f,2g,1h	*not available*
B-Ne	3s,2p,1d	4s,3p,2d,1f	5s,4p,3d,2f,1g	6s,5p,4d,3f,2g,1h	7s,6p,5d,4f,3g,2h,1i
Na-Ar	4s,3p,1d	5s,4p,2d,1f	6s,5p,3d,2f,1g	7s,6p,4d,3f,2g,1h	*not available*
Ca	5s,4p,2d	6s,5p,3d,1f	7s,6p,4d,2f,1g	8s,7p,5d,3f,2g,1h	*not available*
Sc-Zn	6s,5p,3d, 1f	7s,6p,4d,2f,1g	8s,7p,5d,3f,2g,1h	9s,8p,6d,4f,3g,2h,1i	*not available*
Ga-Kr	5s,4p,2d	6s,5p,3d,1f	7s,6p,4d,2f,1g	8s,7p,5d,3f,2g,1h	*not available*

These basis sets may be augmented with diffuse functions by adding the AUG- prefix to the basis set keyword (rather than using the + and ++ notation—see below).

Basis sets of Ahlrichs and coworkers: the SV, SVP, TZV, TZVP keywords refer to the initial formations of the split valence and triple zeta basis sets from this group [Schaefer92, Schaefer94]. The newer redefinitions of these basis sets in [Weigend05, Weigend06] are requested with the keywords Def2SV, Def2SVP, Def2SVPP, Def2TZV, Def2TZVP, Def2TZVPP, Def2QZV, Def2QZVP, Def2QZVPP, and QZVP. Note that Def2SVPP corresponds to the "def2-SV(P)" basis set in [Weigend05]; all other names follow those in the paper with the hyphen removed.

MIDI! of Truhlar and coworkers [Easton96]. The MidiX keyword is used to request this basis set.

EPR-II and EPR-III: The basis sets of Barone [Barone96a] which are optimized for the computation of hyperfine coupling constants by DFT methods (particularly B3LYP). EPR-II is a double zeta basis set with a single set of polarization functions and an enhanced s part: (6,1)/[4,1] for H and (10,5,1)/[6,2,1] for B to F. EPR-III is a triple-zeta basis set including diffuse functions, double d-polarizations and a single set of f-polarization functions. Also in this case the s-part is improved to better describe the nuclear region: (6,2)/[4,2] for H and (11,7,2,1)/[7,4,2,1] for B to F.

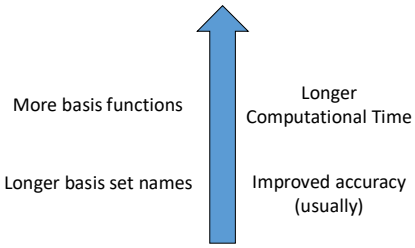
UGBS: The universal Gaussian basis set of de Castro, Jorge and coworkers [Silver78, Silver78a, Moballem86, Moballem87, daCosta87, daSilva89, Jorge97, Jorge97a, deCastro98]. Additional polarization functions may be added by including a suffix to this keyword:

UGBS:nP|V|O

82, Fuentealba83, Stoll84, Fuentealba85, 3, Haeussermann93, Dolg93a, Kuechle94, potentials within Gen basis input. Note ore electrons). OldSDD requests the

Choosing a Basis Set(2)

- The basis set names aren't very helpful, but http://gaussian.com/basissets/ has a list and some description.
- Two very common families in the literature:
  - Pople (named X-YZ(+)G(letters), X-YZW(+)G(letters))
  - Dunning (cc-pVXZ, with X being D,T,Q,5,6).
- For Pople:
  - X is number of functions to describe core electrons,
  - Y,Z, and W are the functions used for the valence electrons.
  - Letters in brackets are for polarization functions (you should always use these),
  - + represent diffuse functions (should really use for optimizing anions, radicals, accurate polarizabilities, dispersion)
- For Dunning, the X represents the angular momentum. The higher the number, the larger the basis set. Can add diffuse functions by using aug-

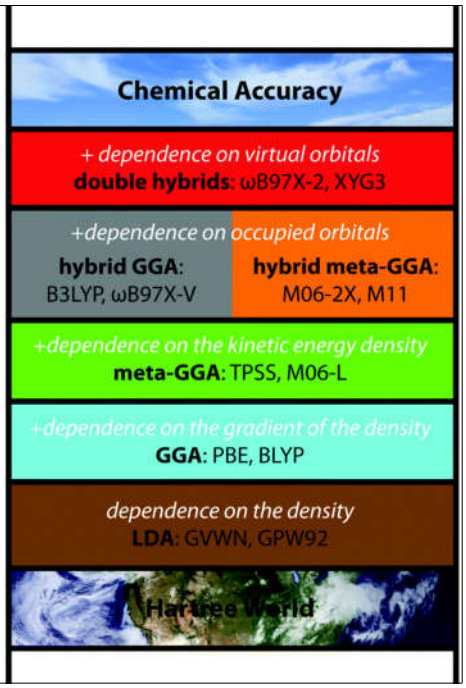


Choosing a Basis Set(3)

- Generally, the longer the name, the larger the basis set.
- Larger basis sets are usually more accurate, but this is not always the case (sometimes get cancellations of errors). Larger basis sets are more computationally demanding.
- Check the literature and see what other basis sets people are using.
- If you are interested in a particular property, try and compare to experiment and use the smallest basis set that gets you sufficiently close.
- Use the same (or very similar) basis set for optimization and frequency calculations on the same molecule.
- Recommendation: If unsure, start with 6-31G(d) or cc-pVDZ (small enough to run quickly, big enough to give general trends. For reasonable accuracy 6-311(+)G(d,p) and cc-pVTZ will often be sufficient.

Choosing a Functional

- Much of the same advice for basis sets applies here. There are so many functionals available and their names don't tell you anything about how to use them.
- Your best friend is the literature here.
- Calculating some property and comparing to higher level calculations/experiment is usually necessary.
- Jacob's Ladder is also a helpful guide.
- My personal suggestion: If in doubt, use B3LYP. For energies use Minnesota functionals. For band gap/optical properties, use HSE06, and for dispersion/interaction use B97D.

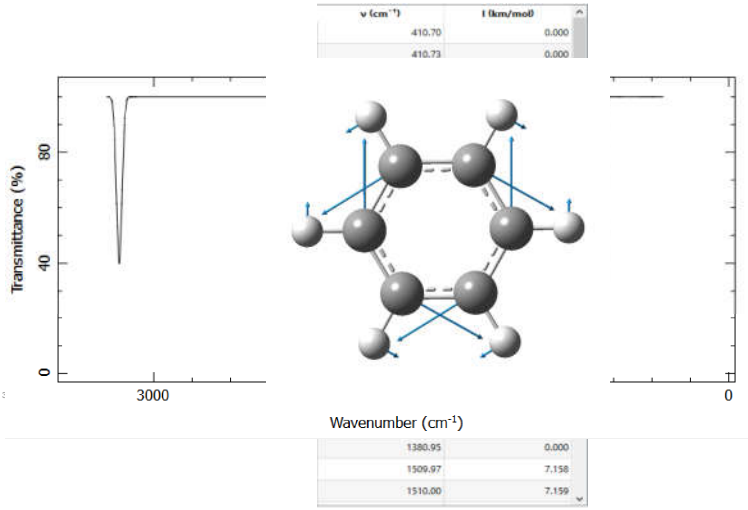


Running Gaussian

- Once the input file is complete, all you need to do is submit it to Gaussian and give it a name for the log file. The log file contains the main results.
- Other results are contained in the formatted checkpoint file which is generated using your checkpoint file by doing formchk <your-file>.chk
- Your job has completed successfully when you see this:

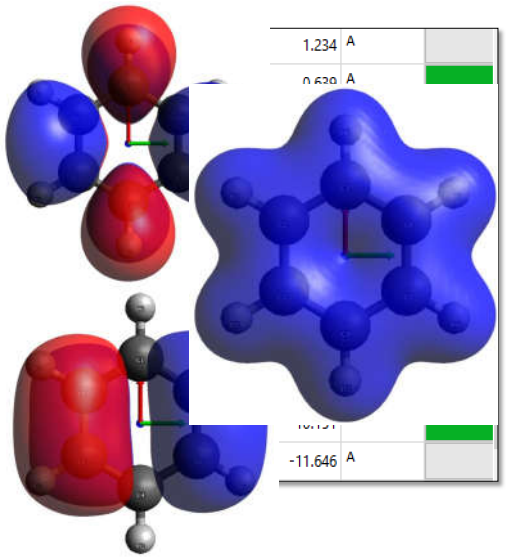
Output of Gaussian Job

- Open the log file in Avogadro. The frequencies should be on the side tool bar. **For a reactant/product, there should be only positive frequencies.**
- Can animate any frequency by clicking on it.
- Show Spectra command will display IR spectra for the calculation:



Output of Gaussian Job(2)

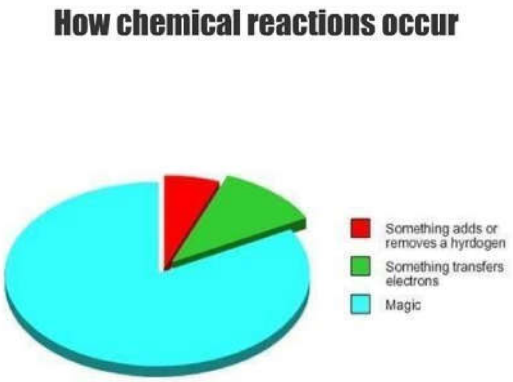
- In the formatted checkpoint file, one can see the energies of the molecular orbitals. Can determine HOMO-LUMO gap,
- Can look at the shape and density of frontier orbitals (HOMO, LUMO, etc).
- Can view van-der-Waals or electron iso-surfaces of electron density – see where the charge is concentrated in your molecule



Some common errors

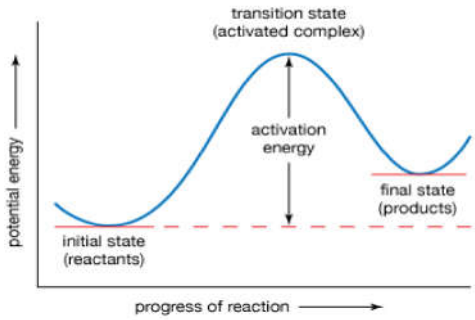
What Gaussian Says	Translation
"Combination of X electrons and Y multiplicity is impossible"	Incorrect Multiplicity. Usually 1 for closed-shell, 2-for singlet radicals, etc.
"Error Termination request processed by link 9999"	The geometry optimization failed. Usually, this is due to the initial guess being poor
"Convergence failure. Run terminated"	The SCF calculation failed to converge. This is usually due to poor initial geometry or poor orbitals.
"Erroneous write"	Out of disk space. The checkpoint files and rwf files can get large, so be careful with storage.
"Error termination in NtrErr: NtrErr Called from FileIO"	Tried to read something from a checkpoint file that isn't there (or the .chk file doesn't exist)
"Wanted an integer as input, found a string as input"	A line (usually job title), is missing in the input file
"End of File in Zsymb"	Missing blank line at end of input file

PART C: INTRODUCTION TO RATE CALCULATIONS

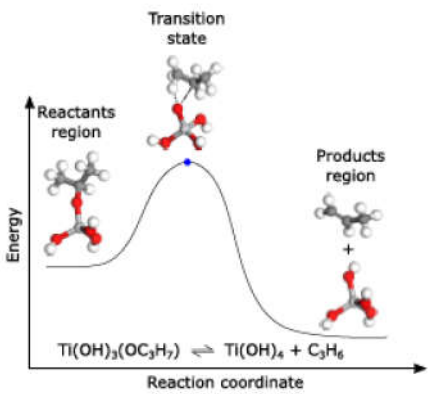


The Potential Energy Surface (PES)

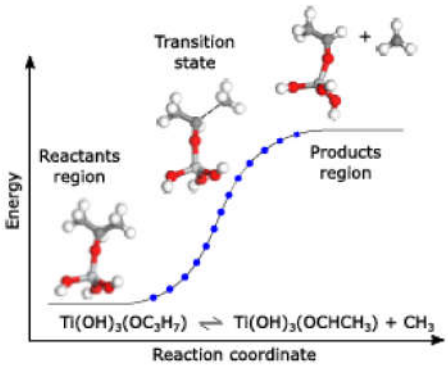
- Describes variation in energy with some arbitrary coordinate.
- For reactions, this is the 'reaction coordinate' – can be a bond length, torsional angle, etc.
- Local minima along the PES correspond to reactants, intermediates, products – stable compounds. These are sometimes called wells.
- Transition states are located at maxima on the PES . They are also called dividing surfaces – separate reactant region from product region.



Example PES's



Reaction with a Barrier



Barrierless Reaction

Getting the Rate: Transition State Theory

- Two main forms: canonical transition state theory and microcanonical transition state theory.

- Canonical transition state theory : Find dividing surface as a function of temperature. The rate can be calculated from:

$$k(T) = \frac{k_B T}{h} \frac{q^{\ddagger}}{q^R} \exp\left(-\frac{E_A^{\ddagger}}{k_B T}\right)$$

- Key assumptions of Transition State Theory:
  - Born-Oppenheimer is valid.
  - Reactants are at equilibrium.
  - No re-crossing – reactants that cross the transition state surface will form product.
  - There is a barrier.
  - Reactants must overcome the barrier to form product.
  - Reactants have sufficient energy to collide by themselves.

- $E_A^{\ddagger}$ is the barrier height, defined as: $E_{TST}(0\mathrm{K}) - \sum E_{reactants}(0\mathrm{K})$.
- $q^{\ddagger}$ is the partition function of the transition state, q^R is the total partition function of the reactants: $q^R = \prod_i q_i$.

Getting the Rate: Evaluating Partition Functions

- The partition functions in the rate constant are the total partition functions. This is composed of different components.

$$q_{tot} = q_{\mathrm{el}} q_{\mathrm{vib}} q_{\mathrm{trans}} q_{\mathrm{rot}}$$

- All of these have standard expressions derived from statistical mechanics:

$$q_{el} = g_0 \qquad q_{trans} = \left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} \frac{k_B T}{P}$$

Ground state degeneracy

$$q_{vib} = \prod_i^{n_{vib}} \frac{e^{-\frac{\beta h \omega_i}{2}}}{1 - e^{-\beta h \omega_i}}$$

Frequencies of Molecule

Molecular Moments of Inertia

$$q_{rot} = \frac{\pi^{\frac{1}{2}}}{\sigma_r} \left(\frac{T^3}{\Theta_A \Theta_B \Theta_C}\right)^{1\backslash 2} \qquad \Theta_j = \frac{h^2}{8\pi^2 I_j k_B}$$

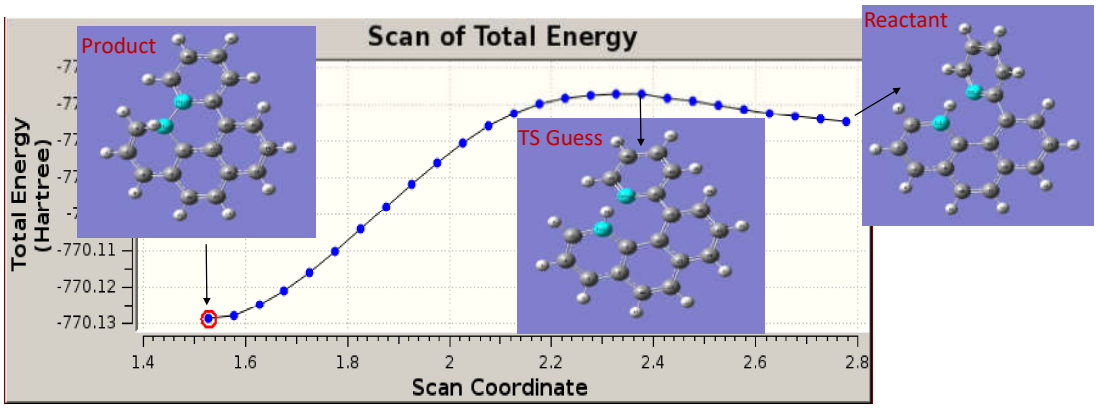
- So, to get the rate, all we need the frequencies and moments of inertia of the reactants and TS and an estimate of the barrier height. These can all be obtained from DFT calculations on optimized reactants and transition states.

Locating Transition States

- This is usually the hardest part of the process as there is no definitive way that will guarantee a successful location of a TS.
- Method 1: Use chemical intuition and make a guess. Then try brute force geometry optimization. (This hardly ever works)
- Method 2: Look in the literature for transition states of a similar class of reactions. E.g. for hydrogen abstractions the TSs tend to look very similar. (Be careful – small differences in reactants can have surprisingly large effects).
- Method 3: Scanning along reaction coordinate. Typically reactions involve bond breaking or forming. Can scan along this distance to look for a maxima. This can be inputted as a transition state guess.
- Method 4: Advanced optimization algorithms. These take the reactant, product and guess transition state geometries and search in between using interpolation. Examples are QST2, QST3 in Gaussian, NEB in NWChem. (Expensive, but are generally robust as long as the provided TS guess is somewhat sensible).

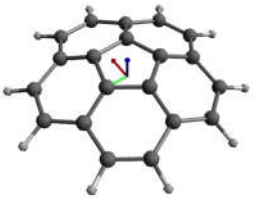
Example of a Coordinate Scan in Gaussian

- Scans are performed using geometry optimizations (keyword **Opt=ModRedundant**). The coordinate for scanning is added to the end of the file. For scanning a long a bond, the syntax is B AtomNumber1 AtomNumber 2 S Nsteps Distance_Increment.
- Example Output:

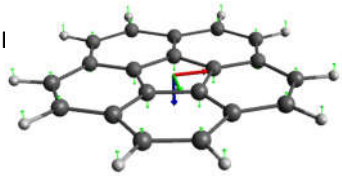


Confirming the Transition State

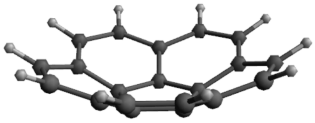
- A successfully found transition state should have only one imaginary frequency (will show as negative in most molecular editors). This frequency should correspond to the reaction in question.
- A ersio



Corannulene



Transition State



Inverted Corannulene

Computing the Rate Constant

- Once the reactants and TS are optimized and have their frequencies calculated, all that is left to do is evaluate the partition functions.
- Fortunately, there are plenty of codes that automate this process (and many more sophisticated calculations than simple canonical TST):
  - **VareCoeF** (previously VariFlex): Developed by Klippenstein *et al.* from Argonne – probably widest capability. Necessary for very high level calculations
  - **Polyrate**: Developed by Truhlar *et al.* from University of Minnesota – Has interfaces to many quantum chemical packages, can do most calculations, but can be difficult to use.
  - **Multiwell**: Developed by Barker *et al.* from University of Michigan – Relatively easy to use, has master equation, Monte Carlo, and canonical/ microcanonical TST options. Some advanced methods not implemented.
  - Reaction Mechanism Generator (RMG): Developed by Green *et al.* from MIT – Includes **Arkane**, a submodule in RMG that can carry out transition state theory and thermodynamic calculations. Python-based, easy to use and read, can extract necessary information from log files directly.

Example calculation in Multiwell

- Canonical TST is implemented in Multiwell's THERMO suite. The input file is described in great detail in the user guide.
- Scripts provided to read Gaussian log file and convert for you.

```
KCAL    MCC
22
200. 250. 298.15 300. 350. 400. 500. 600. 700. 800. 900. 1000. 1100. 1200. 1300. 1400. 1500. 1750. 2000. 2250. 2500. 3000.
3
ctst  'TS-A-B'   275.  209.8293  10.
C2H3O
! Maranzana, Tonachini, et al. 2004
! (blank comment line)
! (blank comment line)
1    1    1
0.0  2
12   'HAR' 'AMUA'
1    hra    122.6869    0.903   -2
1
2    vib    240.1862    0.0    1
3    vib    623.0771    0.0    1
4    vib    639.0296    0.0    1
5    vib    675.0634    0.0    1
6    vib    768.7699    0.0    1
7    vib    773.6569    0.0    1
8    vib   1990.6594    0.0    1
9    vib   3408.1298    0.0    1
10   vib   3497.0541    0.0    1
11   vib   3759.5243    0.0    1
12   top   65.91 12.3740     1    ! K-rotor
```

Energy Units, Concentration Units

Number and list of Temperatures

Number of species (In this case 2 reactants + one TS

Type, Name, Energy at 0K, (magnitude of negative frequency), (reverse barrier)

Empirical Formula

Symmetry No., Number of optical isomers, number of electronic energy levels

Number of vibrations/rotations, vibration units, moment of inertia units

Type, Magnitude of frequency, Anharmonicity value, Degeneracy

Type, First Rotational Constant, Second Rotational Constant, Rotational Symmetry Number

Example calculation in Multiwell(2)

	RATE k(T) = A*exp(E/T) (molecule/cc)**(-1) * sec-1							(CTST) TS-A-B				(REAC) OH_____				(REAC) C2H2_____			
T(K)	k(T)	A(T)	B(T)	DelS(rxn)	DelH(rxn)	DelCp(rxn)	DelG(rxn)	Qelectr	Entropy	Cp	[H(T)-H(0)]	Qelectr	Entropy	Cp	[H(T)-H(0)]	Qelectr	Entropy	Cp	[H(T)-H(0)]
200.00	7.625E-14	1.273E-11	-1024.	-111.38	1.32	-3.20	23.59	2.00	152.81	12.65	2.12	2.73	130.46	7.35	1.48	1.00	133.73	8.50	1.45
250.00	2.153E-13	1.467E-11	-1055.	-111.61	1.17	-2.84	29.07	2.00	155.33	13.98	2.75	2.90	131.64	7.23	1.84	1.00	135.30	9.59	1.91
298.15	4.302E-13	1.675E-11	-1092.	-111.74	1.04	-2.53	34.35	2.00	157.65	15.15	3.48	3.02	132.56	7.16	2.19	1.00	136.72	10.53	2.39
300.00	4.401E-13	1.683E-11	-1093.	-111.74	1.03	-2.52	34.55	2.00	157.63	15.19	3.51	3.02	132.59	7.16	2.20	1.00	136.78	10.56	2.41
350.00	7.475E-13	1.924E-11	-1137.	-111.80	0.91	-2.26	40.04	2.00	159.74	16.21	4.30	3.13	133.38	7.11	2.56	1.00	138.16	11.37	2.96
400.00	1.131E-12	2.190E-11	-1185.	-111.82	0.81	-2.05	45.54	2.00	161.70	17.06	5.13	3.21	134.06	7.07	2.91	1.00	139.46	12.04	3.55
500.00	2.097E-12	2.790E-11	-1294.	-111.80	0.62	-1.75	56.52	2.00	165.21	18.35	6.90	3.34	135.19	7.04	3.62	1.00	141.82	13.07	4.80
600.00	3.288E-12	3.466E-11	-1413.	-111.74	0.45	-1.57	67.50	2.00	168.28	19.32	8.79	3.43	136.12	7.04	4.32	1.00	143.91	13.96	6.15
700.00	4.665E-12	4.205E-11	-1535.	-111.67	0.30	-1.47	78.47	2.00	171.02	20.12	10.76	3.50	136.90	7.07	5.03	1.00	145.79	14.52	7.57
800.00	6.213E-12	4.996E-11	-1668.	-111.60	0.16	-1.41	89.43	2.00	173.49	20.82	12.81	3.56	137.58	7.12	5.74	1.00	147.50	15.11	9.05
900.00	7.901E-12	5.832E-11	-1799.	-111.53	0.02	-1.37	100.39	2.00	175.74	21.47	14.93	3.60	138.15	7.20	6.45	1.00	149.08	15.64	10.59
1000.00	9.718E-12	6.707E-11	-1932.	-111.46	-0.12	-1.34	111.34	2.00	177.82	22.07	17.10	3.64	138.74	7.29	7.18	1.00	150.55	16.13	12.18
1100.00	1.165E-11	7.617E-11	-2065.	-111.40	-0.28	-1.33	122.29	2.00	179.76	22.63	19.34	3.67	139.25	7.39	7.91	1.00	151.91	16.56	13.81
1200.00	1.369E-11	8.561E-11	-2200.	-111.34	-0.38	-1.31	133.23	2.00	181.58	23.14	21.63	3.69	139.72	7.49	8.65	1.00	153.20	16.96	15.49
1300.00	1.583E-11	9.537E-11	-2334.	-111.29	-0.51	-1.30	144.16	2.00	183.30	23.61	23.96	3.71	140.17	7.60	9.41	1.00	154.41	17.32	17.20
1400.00	1.806E-11	1.054E-10	-2470.	-111.23	-0.64	-1.29	155.09	2.00	184.91	24.04	26.35	3.73	140.59	7.70	10.17	1.00	155.56	17.64	18.95
1500.00	2.028E-11	1.158E-10	-2606.	-111.19	-0.77	-1.28	166.01	2.00	186.45	24.43	28.77	3.75	140.99	7.79	10.95	1.00	156.65	17.92	20.73
1750.00	2.663E-11	1.429E-10	-2947.	-111.07	-1.09	-1.26	193.29	2.00	189.97	25.26	34.99	3.78	141.90	8.00	12.92	1.00	159.15	18.51	25.29
2000.00	3.313E-11	1.717E-10	-3291.	-110.98	-1.40	-1.24	220.55	2.00	193.12	26.09	41.39	3.81	142.71	8.18	14.95	1.00	161.39	18.96	29.97
2250.00	4.015E-11	2.020E-10	-3635.	-110.89	-1.71	-1.24	247.79	2.00	195.57	26.88	47.92	3.83	143.45	8.32	17.01	1.00	163.41	19.30	34.75
2500.00	4.754E-11	2.336E-10	-3980.	-110.81	-2.02	-1.24	275.00	2.00	198.56	26.77	54.56	3.85	144.12	8.44	19.10	1.00	165.25	19.56	39.61
3000.00	6.333E-11	3.002E-10	-4668.	-110.67	-2.64	-1.25	329.38	2.00	203.13	27.32	68.09	3.87	145.32	8.62	23.37	1.00	168.49	19.96	49.50

- Output contains rate constant as function of T – can fit this to Arrhenius expression for mechanisms.

Issues with (Canonical) Transition State Theory

- Canonical transition state theory represents the simplest case of a reaction – the high pressure limit with a single transition state.

- When is Transition State Theory inappropriate?
  - Barrierless reactions
  - No or multiple transition states for the reaction
  - Quantum tunneling effects
  - Reactions requiring a third body collision (pressure-dependent)

- When are the Partition Functions inaccurate?
  - Anharmonic effects
  - Hindered rotors
  - Coupling of vibration and rotation
  - Excited States

Beyond Transition State Theory – What to Do?

- For barrierless reactions – Use *Variational* Transition State Theory.

- For pressure-dependent reactions – Can use microcanonical/RRKM theory.

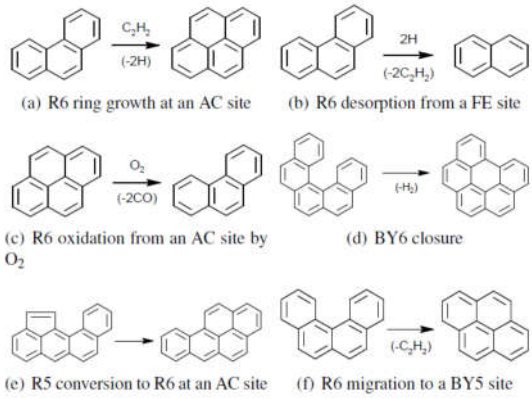
- For multiple transition states – Master Equation simulations.

- For anharmonic frequencies/hindered rotors/Rovibrations – Need to run further calculations on these effects.

- For excited states – Higher level methods than DFT are necessary.

- It can get complicated, but most commercial computational chemistry/rate constant codes can handle all of these too.

PART IV: Application to PAH chemistry

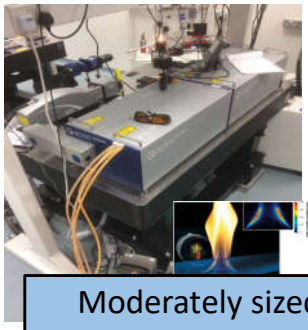


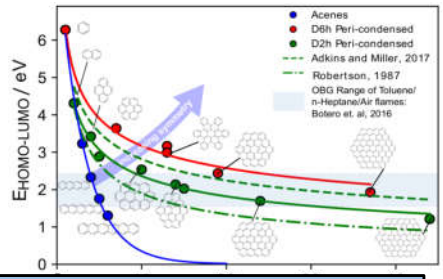
(a) R6 ring growth at an AC site (b) R6 desorption from a FE site

(c) R6 oxidation from an AC site by O_2 (d) BY6 closure

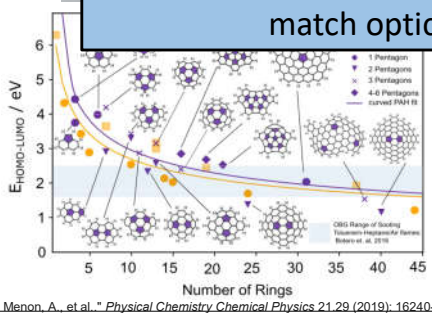
(e) R5 conversion to R6 at an AC site (f) R6 migration to a BY5 site

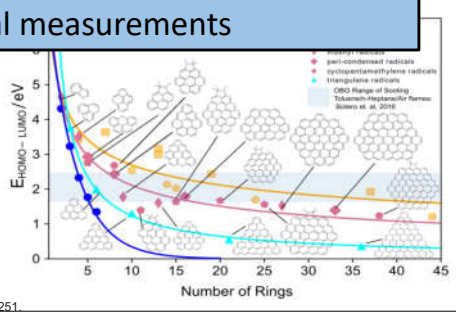
Identifying important PAHs – optical diagnostics





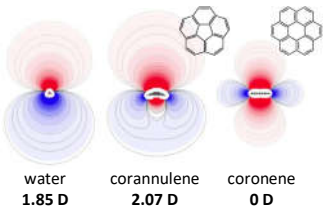
Moderately sized, curved PAHs and radical PAHs match optical measurements

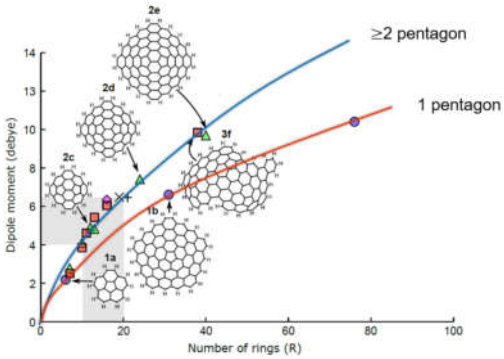




Menon, A., et al.," *Physical Chemistry Chemical Physics* 21.29 (2019): 16240-16251.

Curved PAHs and Flexoelectricity

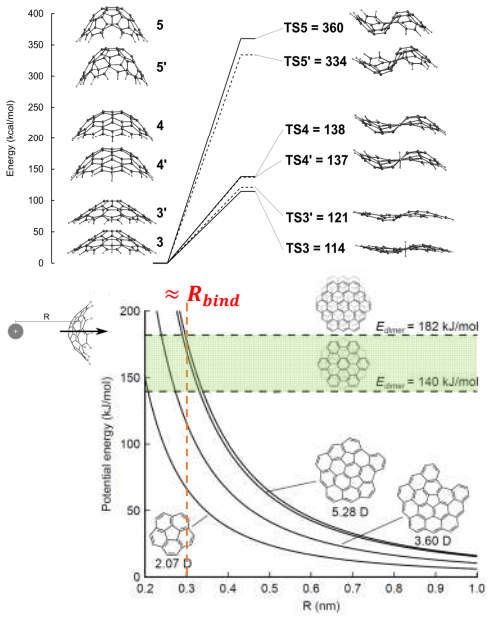


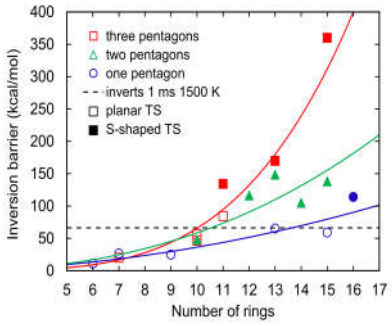


- Curvature in PAHs comes from the presence of five-member rings.
- This causes a dipole moment due to the polarisation of π electrons from the concave to convex surface. The dipole moment increases with size and curvature of the PAHs.

Influence of electric fields

Are curved PAHs stable in flames?





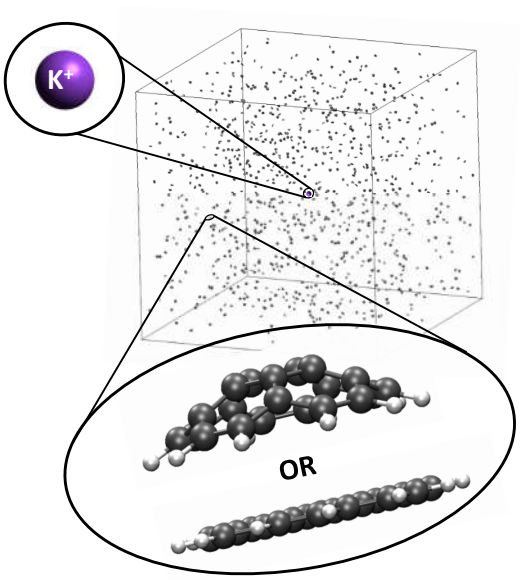
Only small, low curvature PAHs are seen to invert.

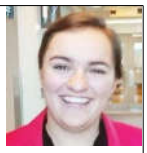
Larger, curved PAHs bind strongly to ions, potentially increasing their stability.

Martin, J. W., Menon, A., Lao, C. T., Akroyd, J., & Kraft, M. (2019). *Combustion and Flame*, *206*, 150-157.

Simulation of PAHs with ions

- New curPAHIP potential
- 500 - 1500 K
- 1000 molecules
- 4 cases:
  1. Planar PAHs
  2. Curved PAHs
  3. Planar PAHs with K^+
  4. Curved PAHs with K^+





Kimberly Bowal

Alston Misquitta

Bowal, K., Martin, J. W., Misquitta, A. J., & Kraft, M. *Comb Sci & Tech* (2019)

Molecular dynamics – Clustering of PAHs

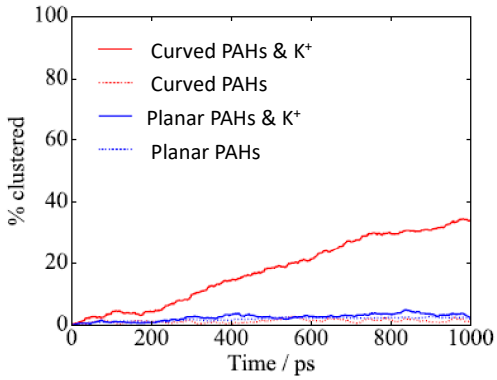
Curved PAHs without K^+

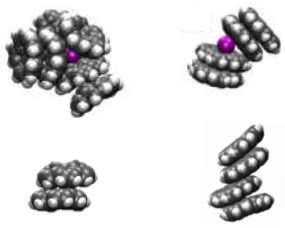
Curved PAHs with K^+

T = 500K

Bowal, K., Martin, J. W., Misquitta, A. J., & Kraft, M. *Comb Sci & Tech* (2019)

Clustering results





Molecular arrangement maximises electrostatic interactions

Curved PAHs and K^+ show greatest ability to form clusters – but mainly at lower temperatures.

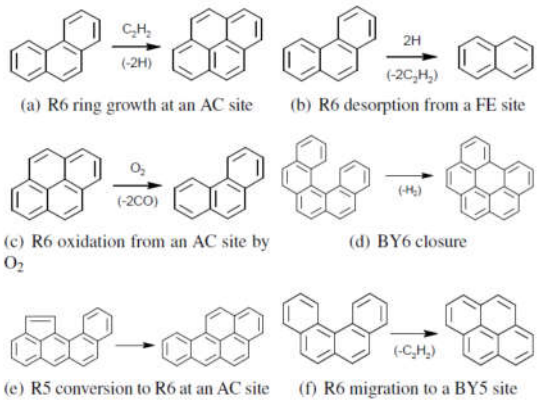
Bowal, K., Martin, J. W., Misquitta, A. J., & Kraft, M. *Comb Sci & Tech* (2019)

Aromatic site modelling reactions of PAHs

List of reactions occurring on different PAH sites.

Rates of processes are modelled using DFT.

KMC simulation of formation and growth of larger PAHs.

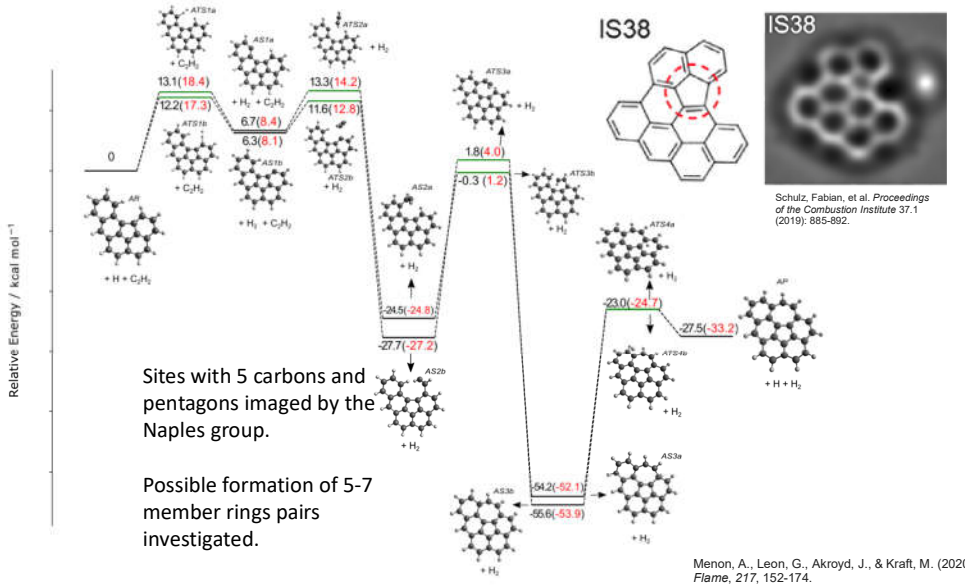


(a) R6 ring growth at an AC site (b) R6 desorption from a FE site

(c) R6 oxidation from an AC site by O_2 (d) BY6 closure

(e) R5 conversion to R6 at an AC site (f) R6 migration to a BY5 site

Deriving new reaction paths for unique PAHs



Schulz, Fabian, et al. *Proceedings of the Combustion Institute* 37.1 (2019): 885-892.

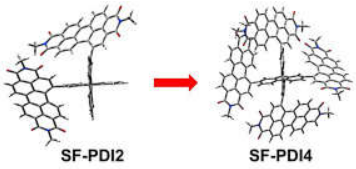
Sites with 5 carbons and pentagons imaged by the Naples group.

Possible formation of 5-7 member rings pairs investigated.

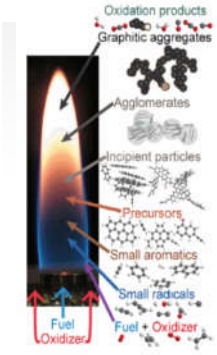
Menon, A., Leon, G., Akroyd, J., & Kraft, M. (2020). *Combustion and Flame*, *217*, 152-174.

Why is Optical Band Gap important?

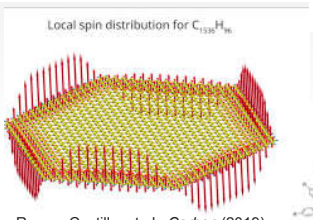
- Understanding what size graphene flakes form nanoparticles and are present in flames.
- Designing molecules for quantum dots, organic solar cells, and molecular electronics.



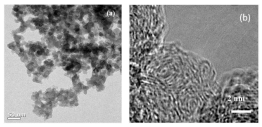
Yang, Lei, et al. *Journal of Power Sources* 324 (2016): 538-546.



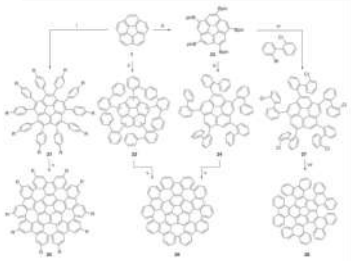
H. Michelsen – CEFRC (2016)



Ramos-Castillo, et al. *Carbon* (2019).

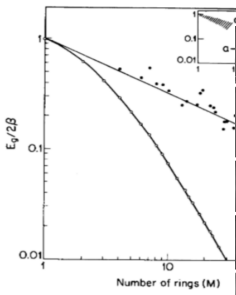


Kim, Dae-Yeong, et al. " *Nanomaterials* 9.5 (2019): 793.

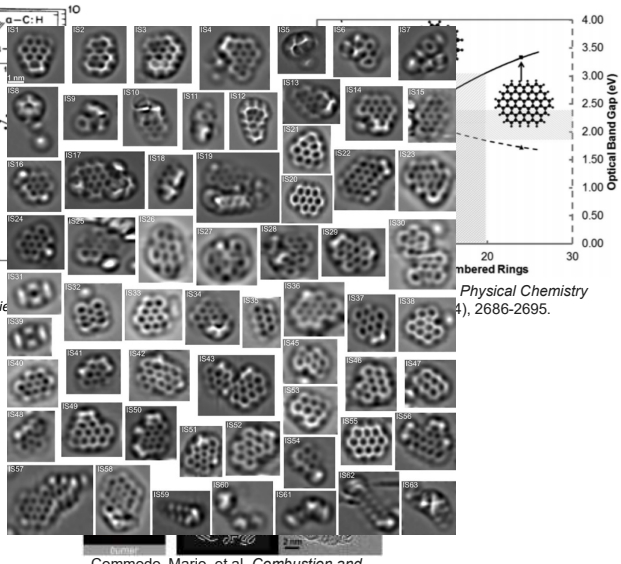


Muzammil and Stuparu. Nature *Communications Chemistry* 2.1 (2019): 58.

Background



Robertson, & O'reilly, *Physical Revie*
(1986).



Physical Chemistry
4), 2686-2695.

John Robertson

Commodo, Mario, et al. *Combustion and Flame* 205 (2019): 154-184.

Martin, Jacob W., et al. *The Journal of Physical Chemistry C* 122.38 (2018): 22210-22215.

Key Questions:

What DFT methods are suitable for computing band gaps?

How is the optical band gap of nanographenes impacted by their structure?

What does this mean for combustion and carbon material applications?

Objectives:

Compute optical band gap for a variety of PAHs

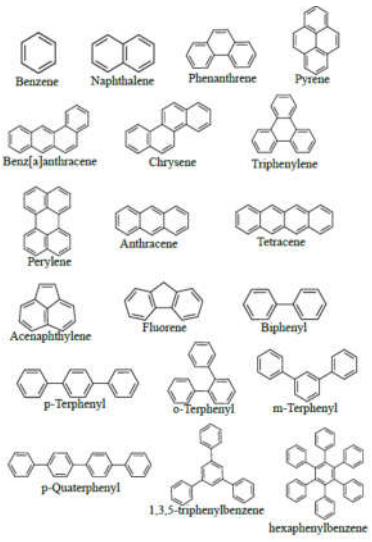
Determine

1) Impact of size, cross-linking, curvature, and radical character on OBG

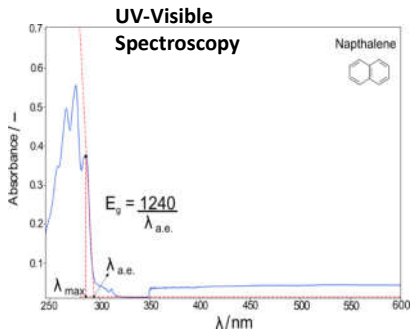
2) What this could mean for tuning carbon nanoparticles

Methodology

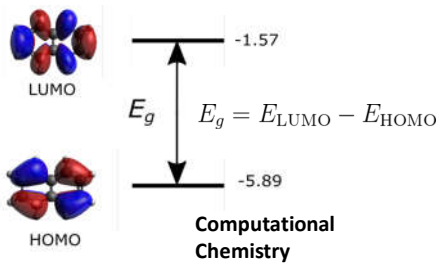
Test Set of Molecules



UV-Visible Spectroscopy

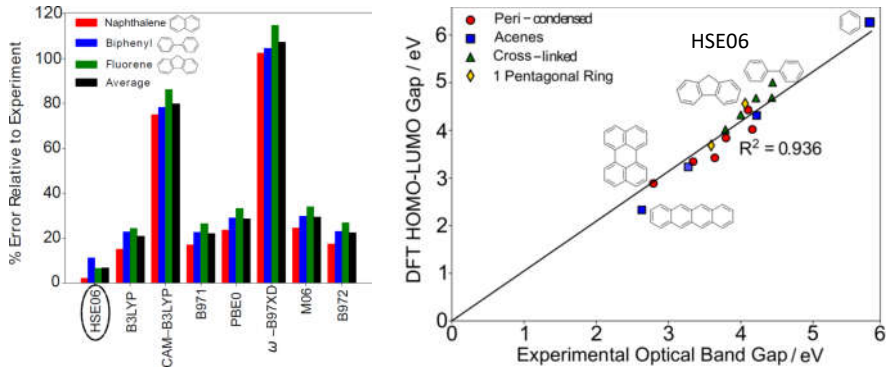


Jochen Dreyer



Computational Chemistry

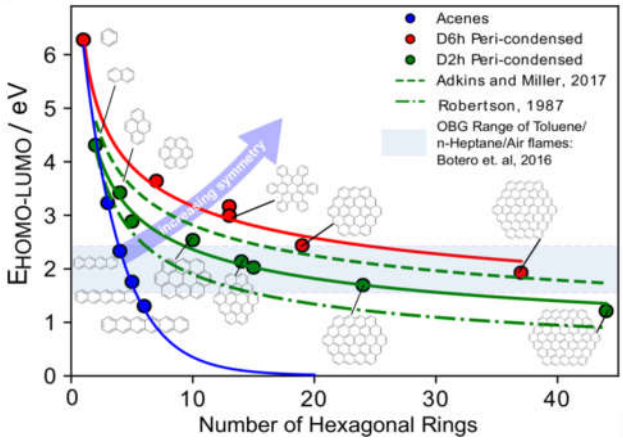
Choosing a computational method



- Tested several popular computational chemistry methods based on density functional theory.

Certain DFT methods can reproduce experimental measurements

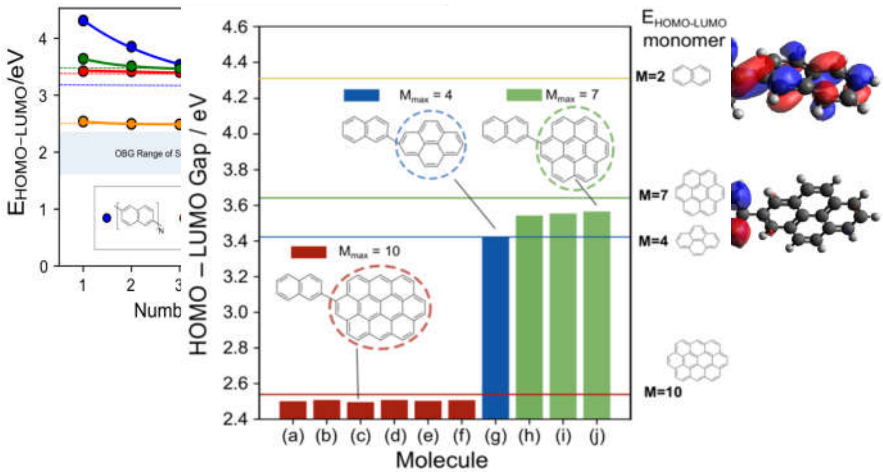
Traditional model nanographenes



- All three groups approach a zero band gap limit
- In flames: 10 – 25 rings in size

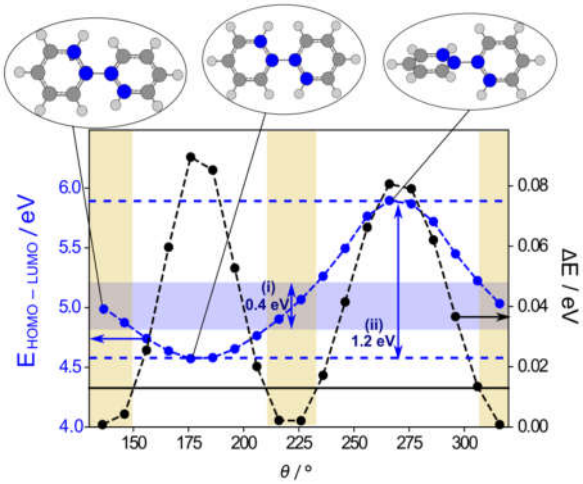
Size and symmetry of nanographene flake is key

Crosslinked Structures



The largest fragment (lowest band gap) is most important.

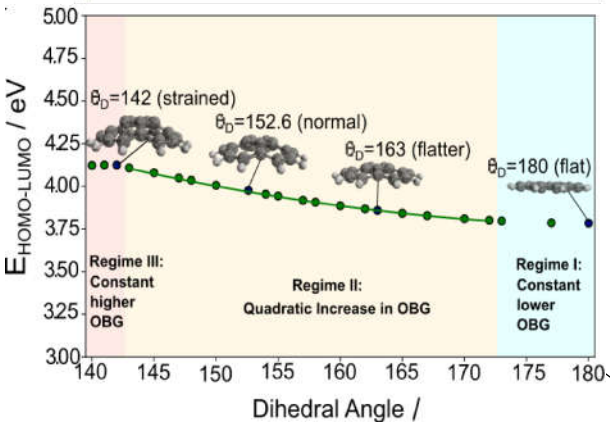
Effect of rotation on optical band gap



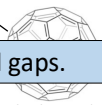
- Formation of cross-link enables rotation between fragments.
- Relaxed potential energy surface scan to check optical properties of different conformers and what energy is required.

Rotation allows tunability of optical band gaps.

Curved vs. planar structures



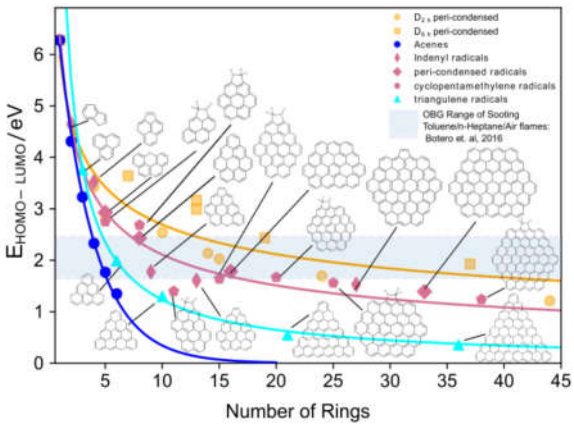
- Curved structures also approach zero-band gap limit.
- More pentagons results in more curvature - higher optical band gap
- Curved nano-graphenes coincide with planar structures for larger sizes.



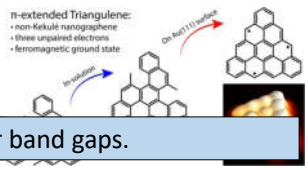
Smaller curved nano-graphenes have larger band gaps.

Rabenau, T., et al. Zeitschrift für Physik B Condensed Matter 90.1 (1993): 69-72.

Radical vs planar structures



- A single radical site does not affect the optical band gap.
- Delocalized radicals (odd carbon number) have significantly lower band gaps at all sizes.
- Multi-radical character reduces optical band gap as well.



π-radical nanographenes have lower band gaps.

Mishra, Shantanu, et al. *Journal of the American Chemical Society* (2019).

Findings

- The optical band gap of nano-graphenes is highly sensitive to the underlying structure – size, symmetry, curvature, and π-radical character are all important and could be useful for tailoring band gap to different applications.
- Other features are less impactful, with cross-linking and hydrogen termination not influencing the optical band gap significantly.
- OBGs observed in flames can be attributed to moderate sized model nano-graphenes, larger curved structures, or smaller π-radical structures.

End of Lecture 2

Soot – Part 3

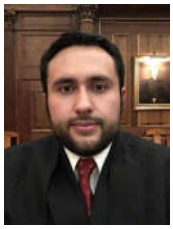
Markus Kraft

Computational Modelling Group Cambridge

Main Contributors:
Dr Jake Martin (Part 1)
Dr Angiras Menon (Part 2)
Dr Laura Pascazio (Part 3)
Dr Gustavo Leon (Part 4)

Jacob Martin

AngirasMenon

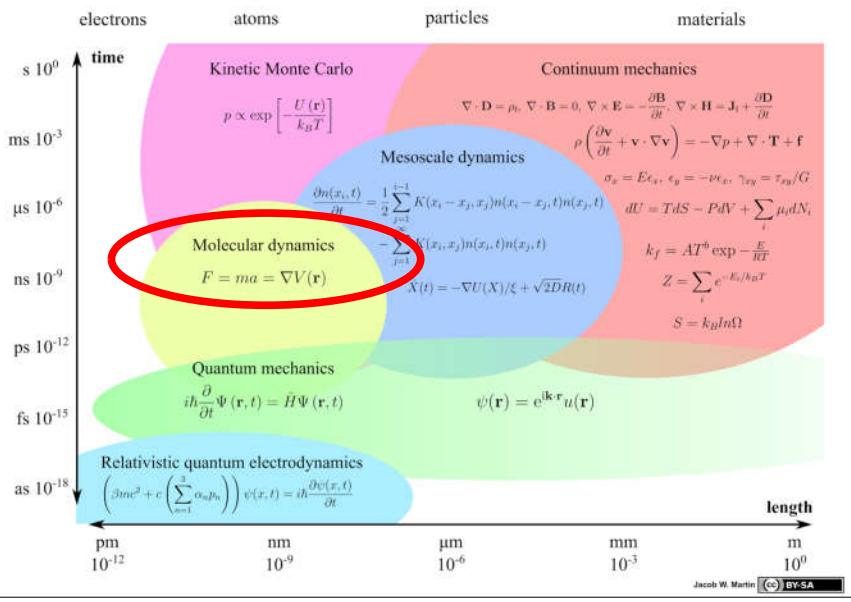


Gustavo Leon

Laura Pascazio

Part 1 Overview
Part 2 Quantum Chemistry
Part 3 Molecular Dynamics
Part 4 Kinetic Monte Carlo
Part 5 Stochastic Particle Methods
Part 6 Application – engine model

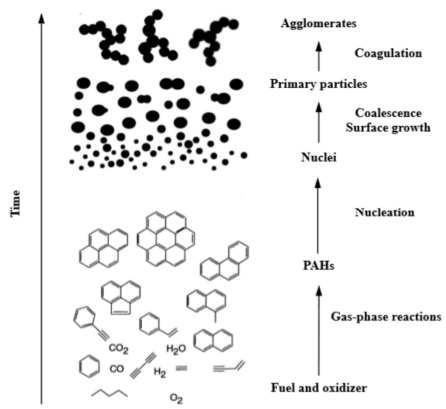
Scale in Simulations

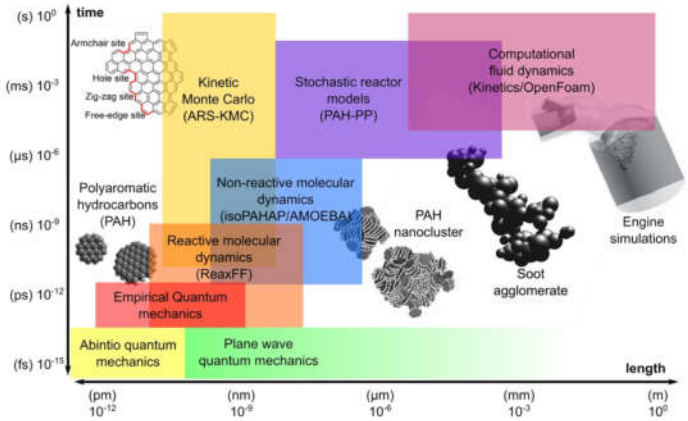


Scale in Simulations

Example: Soot formation in combustion process

http://nznano.blogspot.com/





Quantum Mechanics and Molecular Dynamics

Computational techniques for studying the **time evolution** of a system of interacting atoms at an atomistic level:

QUANTUM MECHANICS	MOLECULAR MECHANICS
Solves **quantum Schrödinger equation** and provide a rigorous description of molecular systems: $H(t)\lvert\psi(t)\rangle = i\hbar\frac{d}{dt}\lvert\psi(t)\rangle$	Solves **Netwon's equations of motion** and uses classical potential energy equations: $\boldsymbol{F} = m\boldsymbol{a}$ $U_{total} = U_{bonded} + U_{non\ bonded}$
Considers atoms as **collections of electron and nuclei**	Considers atoms as **hard spheres**
Time and memory hungry	Computationally efficient
Suitable for **small and medium** sized systems	Suitable for **large** systems

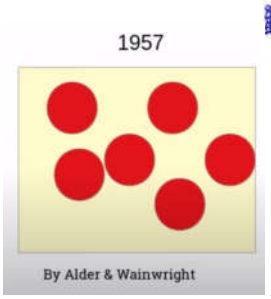
Molecular Dynamics: strengths and weaknesses

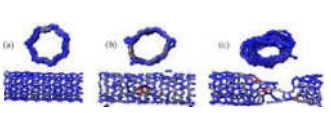
- Why MD?
  - Dynamics: Predict time dependent behaviour
  - Scale: Large collections of interacting particles that cannot be studied by quantum mechanics. MD methods are thousand times faster than quantum chemistry methods
  - Atomistic Model

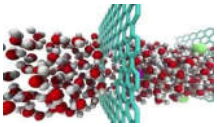
- Limitations:
  - Very large systems
  - Quantum effects
  - Limited by inputs (ex force field)

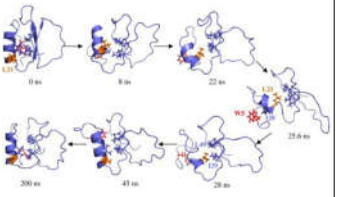
Applications

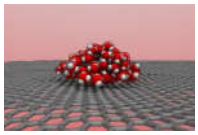
- Adsorption
- Molecular docking
- Protein folding
- Drug-receptor interactions
- Solvation of molecules
- Membranes
- Structure-to-function relationship
- Conformational changes
- Effects of irradiation
- Mechanical failure
- Surface interactions
- Carbon nanotubes, nanodots
- Materials properties
- Crystal structures
- Droplet formation



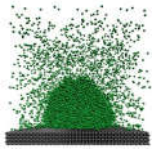


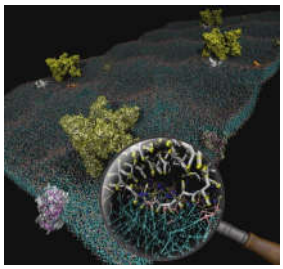






- Biomolecular system fluctuations
- Deformation mechanisms
- Ion transport
- Structure determination
- Vapour phase material growth





MD in Combustion

- Soot nucleation:

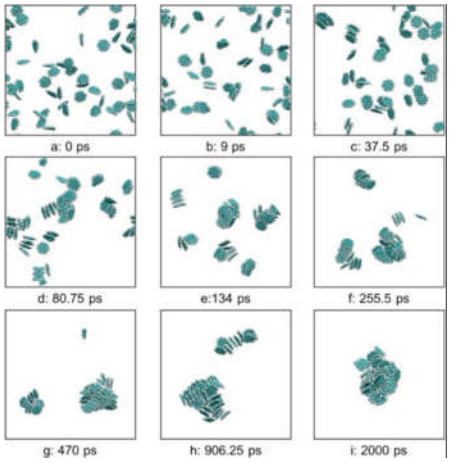
PAH collision, physical nucleation, reactive nucleation

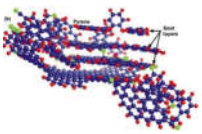
- Soot particle characteristics and properties:

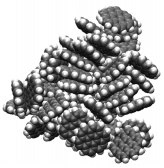
Melting points, mechanical properties

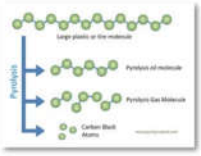
- Pyrolysis
- Particle interactions and growth:

Coalescence and agglomeration









Classic equation of motion

The system can be simulated by solving Newton's equation of motion.

The force acting on each atom *i* of a system composed by *N* atoms is given by:

$$\boldsymbol{F}_i = m_i \boldsymbol{a}_i = m_i \frac{\partial \boldsymbol{v}_i}{t} = m_i \frac{\partial^2 \boldsymbol{r}_i}{\partial t^2} \quad , i = 1 \dots N$$

Newton's equation

If the force is **conservative** (if the potential energy depends only on the position of the atoms), the force is related to the potential energy by:

$$\boldsymbol{F}_i = -\frac{\partial U(\boldsymbol{r}_i, \dots, \boldsymbol{r}_N)}{\partial \boldsymbol{r}_i}$$

Integration Algorithms: Essential Idea

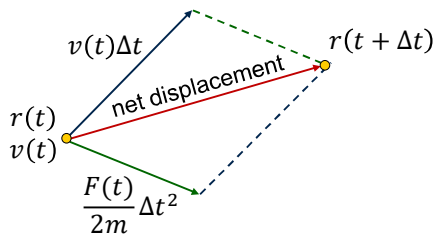
All the integration algorithms assume the positions, velocities and accelerations can be approximated by a Taylor series expansion:

$$r(t+\Delta t) \approx r(t) + r'(t)\Delta t + \frac{1}{2!}r''(t)\Delta t^{\,2} + \cdots + \frac{1}{n!}r^{(n)}(t)\Delta t^{\,n} + O(\Delta t^{\,n+1})$$

$$r(t+\Delta t) \approx r(t) + r'(t)\Delta t + \frac{1}{2!}r''(t)\Delta t^{\,2}$$

$$v(t) = \frac{d\,r(t)}{dt} \qquad a(t) = \frac{d^2\,r(t)}{dt^2} = \frac{F(t)}{m}$$

$$r(t+\Delta t) \approx r(t) + v(t)\Delta t + \frac{F(t)}{2m}\Delta t^{\,2}$$



Integration Algorithms: Essential Idea

Taylor series expansion for velocity around a known point $v(t)$:

$$v(t+\Delta t) \approx v(t) + v'(t)\Delta t + \frac{1}{2!}v''(t)\Delta t^{\,2} + \cdots + \frac{1}{n!}v^{(n)}(t)\Delta t^{\,n} + O(\Delta t^{\,n+1})$$

$$v(t+\Delta t) \approx v(t) + v'(t)\Delta t + \frac{1}{2!}v''(t)\Delta t^{\,2} \qquad \text{where} \qquad v'(t) = a(t) = \frac{F(t)}{m}$$

Taylor series expansion around $\frac{d\,v(t)}{dt}$:

$$v'(t+\Delta t) \approx v'(t) + v''(t)\Delta t$$

Multiplying by $\frac{\Delta t}{2}$:

$$v'(t+\Delta t)\frac{\Delta t}{2} \approx v'(t)\frac{\Delta t}{2} + v''(t)\frac{\Delta t^2}{2} \quad \rightarrow \quad v''(t)\frac{\Delta t^2}{2} = v'(t+\Delta t)\frac{\Delta t}{2} - v'(t)\frac{\Delta t}{2}$$

$$\boldsymbol{v(t+\Delta t)} \approx v(t) + v'(t)\frac{\Delta t}{2} + v'(t+\Delta t)\frac{\Delta t}{2} = \boldsymbol{v(t)} + \frac{\boldsymbol{\Delta t}}{\boldsymbol{2m}}\big(\boldsymbol{F(t)} + \boldsymbol{F(t+\Delta t)}\big)$$

Velocity Verlet Algorithm

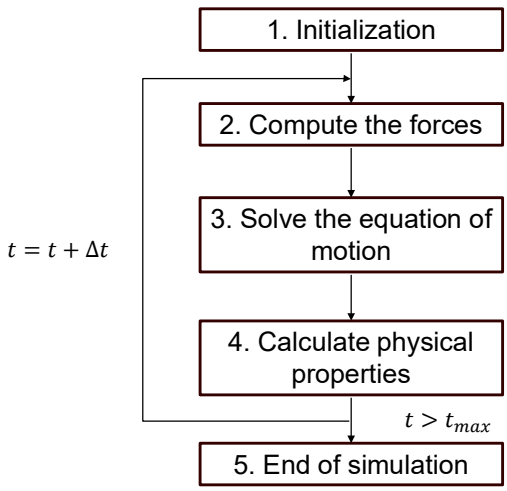
$$\boldsymbol{r_i}(t+\Delta t) = \boldsymbol{r_i}(t) + \boldsymbol{v_i}(t)\Delta t + \frac{\boldsymbol{F_i}(t)}{m_i}\Delta t^{\,2} + O\left(\Delta t^{\,3}\right)$$

$$\boldsymbol{v_i}(t+\Delta t) = \boldsymbol{v_i}(t) + \frac{\Delta t}{2m_i}\left[\boldsymbol{F_i}(t) + \boldsymbol{F_i}(t+\Delta t)\right] + O(\Delta t^{\,3})$$

FAST
ACCURATE
TIME-REVERSIBLE

Other algorithms: Verlet, Leap-frog, Beeman

Key Stages in MD Simulation



Input initial condition: atomic positions $\boldsymbol{r}_i(t_0)$ and velocities $\boldsymbol{v}_i(t_0)$ of all atoms in the system

$$\boldsymbol{F}_i(t) = -\frac{\partial U(\boldsymbol{r}_1(t), \dots, \boldsymbol{r}_N(t))}{\partial \boldsymbol{r}_i}$$

Velocity Verlet:

$\boldsymbol{r}_i(t)$, $\boldsymbol{F}_i(t) \rightarrow \boldsymbol{r}_i(t+\Delta t)$

$\boldsymbol{r}_i(t+\Delta t) \rightarrow \boldsymbol{F}_i(t+\Delta t)$

$\boldsymbol{v}_i(t), \boldsymbol{F}_i(t), \boldsymbol{F}_i(t+\Delta t) \rightarrow \boldsymbol{v}_i(t+\Delta t)$

$$\boldsymbol{r_i}(t+\Delta t) = \boldsymbol{r_i}(t) + \boldsymbol{v_i}(t)\Delta t + \frac{\boldsymbol{F_i}(t)}{m_i}\Delta t^{\,2} + O\left(\Delta t^{\,3}\right)$$

$$\boldsymbol{v_i}(t+\Delta t) = \boldsymbol{v_i}(t) + \frac{\Delta t}{2m_i}\left[\boldsymbol{F_i}(t) + \boldsymbol{F_i}(t+\Delta t)\right] + O(\Delta t^{\,3})$$

Simulation setup

- **STARTING CONFIGURATION:**
  - Atomic positions (x, y, z)
  - Atomic velocities
  - Boundary conditions
- **POTENTIAL ENERGY (FORCE FIELD):**
  - Potential Energy equations
  - Potential Energy parameters
- **SIMULATION PARAMETERS:**
  - Integration algorithm
  - Time step
  - Thermodynamic ensemble
  - T and P and T and P control

CLASSICAL MECHANICS IS DETERMINISTIC: INITIAL STATE AND INTERACTION RULES FULLY SPECIFY THE SYSTEM'S FUTURE

STARTING CONFIGURATION

Initial coordinates

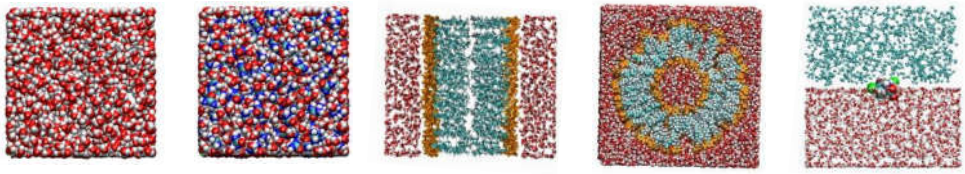
Atomic coordinates:

```
COMPND      Water
HETATM    1  O          1       0.251  -0.360  -0.046  1.00  0.00
HETATM    2  H          1       0.249   0.684   0.231  1.00  0.00
HETATM    3  H          1       0.586  -0.954   0.791  1.00  0.00
TER       4             1
END
```

pdb format

PACKMOL (more complex systems):

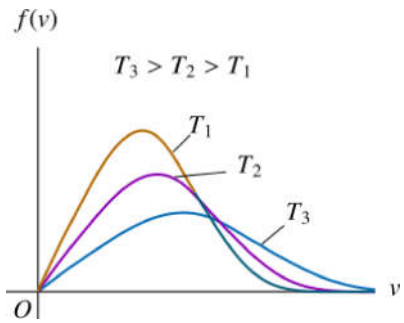
Random positions | Mixture | Layers | Double layered sphere | Interface



Initial velocities

Standard approach is to draw velocities randomly from a **Maxwell-Boltzmann distribution** at the simulation temperature T (assumption of thermal equilibrium)

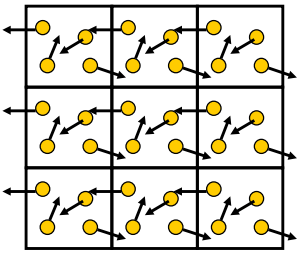
$$f(v_i) = \sqrt{\frac{m_i}{2\pi k_B T}} \exp\left(-\frac{m_i v_i^2}{2 k_B T}\right)$$



In some cases, it is needed to assign them manually, *e.g.* dimerization:

Boundary Conditions

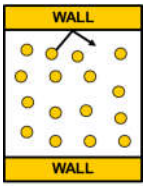
PERIODIC BOUNDARY CONDITIONS



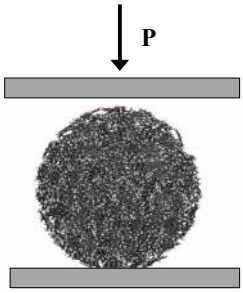
Duplicate the simulation box in one or more directions. An atom moving out of boundary comes from the other side.

- Required to predict and study the properties of a system in bulk (very big or "infinite" system)
- Needed to eliminate surface effects

FIXED BOUNDARY



- Apply external forces (shear deformation)



FORCE FIELDS

Potential energy

The net force acting on each atom in the system is a result of its interactions with all other atoms. The force on an atom i is the negative gradient of a scalar potential energy function:

$$F_i = -\nabla_i U(r_1, r_2, \dots, r_N)$$

The gradient can be computed in three different ways:

- By using a **force field** – Classic Molecular Dynamycs
- By solving the **Schrödinger equation** – Ab initio Molecular Dynamics
- By a combination of both – Quantum Mechanics/Molecular Mechanics (QM/MM)

Force Fields

A **force field** is the collection of **functional forms** and **parameter sets** used to relate the **potential energy** of a system with its **internal coordinates**.

Different types of Force Field:

1. Classical
2. Polarizable
3. Reactive
4. Empirical Valence Bond
5. Coarse-grained
6. Machine Learning

Classical Force Fields

$$U(\boldsymbol{r}) = \sum U_{bonded}(\boldsymbol{r}) + \sum U_{nonbonded}(\boldsymbol{r})$$

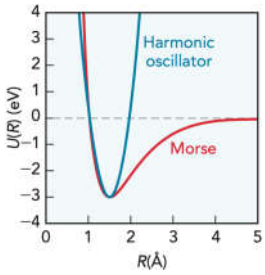
- atoms linked by covalent bonds
- involves 2, 3 and 4 atom interactions

- long-range electrostatic and VdW interactions
- involves interactions between all pairs of atoms

$$\sum U_{bonded} = \sum_{bonds} U_{bond} + \sum_{angles} U_{angles} + \sum_{dihedrals} U_{dihedral} + \sum_{improper} U_{improper}$$

$$= \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 + \sum_{angles} \frac{1}{2} k_a (\theta - \theta_0)^2 +$$

$$+ \sum_{dihedrals} k_\phi [1 + \cos(n\phi - \delta)] + \sum_{improper} \frac{1}{2} k_\xi (\xi_{ijkl} - \xi_0)^2$$



Classical Force Fields

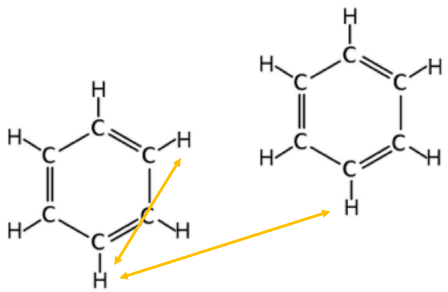
$$U(\boldsymbol{r}) = \sum U_{bonded}(\boldsymbol{r}) + \sum U_{nonbonded}(\boldsymbol{r})$$

- atoms linked by covalent bonds
- involves 2, 3 and 4 atom interactions

- long-range electrostatic and VdW interactions
- involves interactions between all pairs of atoms

$$\sum U_{nonbonded} = \sum U_{VdW} + \sum U_{electr} =$$

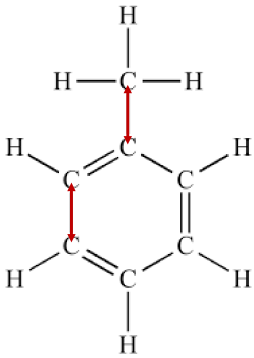
$$= \sum_{i<j}^{atoms} 4\varepsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right) + \sum_{i<j}^{atoms} \frac{1}{4\pi\varepsilon_0} \frac{q_i q_j}{r_{ij}}$$



Classical Force Fields

The atoms of the molecule are classified in different **atom types** to distinguish interactions between the same chemical class of atoms

$$U(\boldsymbol{r}) = \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 + \sum_{angles} \frac{1}{2} k_a (\theta - \theta_0)^2 +$$
$$+ \sum_{dihedrals} k_\phi [1 + \cos(n\phi - \delta)] + \sum_{improper} \frac{1}{2} k_\xi (\xi_{ijkl} - \xi_0)^2 +$$
$$+ \sum_{\substack{atoms \\ i<j}} 4\varepsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right) + \sum_{\substack{atoms \\ i<j}} \frac{1}{4\pi\varepsilon_0} \frac{q_i q_j}{r_{ij}}$$



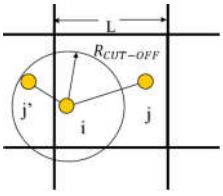
The **parameters** of the force field are different for each atom type

Neighbour list

Evaluating the force is the most computationally demanding part of molecular dynamics.
For each atom a set of lists is generated:

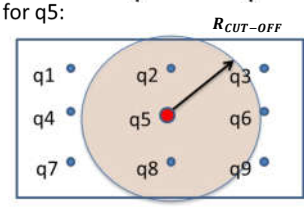
- Fixed (static) lists:
  - Bond list
  - Angle list
  - Dihedral list
- Dynamic list:
  Nonbond list stores non-bond interactions with atoms within cutoff
  $F_{ij} \rightarrow r_{ij} < R_{CUT-OFF}$

PBC:
$R_{CUT-OFF} < L/2$



Nonbond list for q5:

q2, r2
q6, r6
q8, r8
q9, r9



Polarizable Force Field

1. Charge-on-spring (Drude) models
2. Fluctuating-charge models
3. Polarizable point charges models:

Polarization is responsible for (ion/dipole)-induced dipole forces

$$U = \sum_{bonds} U_{bond} + \sum_{angles} U_{angles} + \sum_{dihedrals} U_{dihedral} + \sum_{improper} U_{improper} + \sum_{pairs} U_{VdW}$$
$$+ \sum_{multipole} U_{MPol} + \sum_{polarization} U_{pol}$$

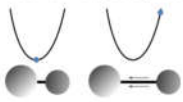
Atomic multipole interactions

Induced dipole polarization

Reactive force fields

- REACTIVE FORCE FIELDS: Enables classical modelling of chemical reactions

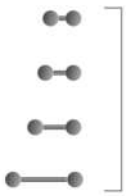
Standard forcefields vs *ReaxFF*



$$BO_{ij}(r_{ij}) = BO_{ij}^{\sigma} + BO_{ij}^{\pi} + BO_{ij}^{\pi\pi}$$

Fixed atomtypes and harmonic potentials: bond breaking impossible, e.g.

$E_{bond} \propto$ *(distance)* ^{2}



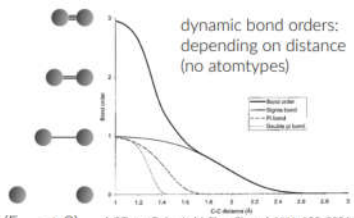
-a triple "bond" will always stay a triple "bond"

- a C_{triple} atom will always be a C_{triple} atom

($E_{bond} \rightarrow \infty$)

Non-harmonic potentials based on bond orders: bond breaking/forming possible, e.g.

$E_{bond} \propto$ *-(bond order)* × exp[*(1 – bond order)*]



dynamic bond orders: depending on distance (no atomtypes)

($E_{bond} \rightarrow 0$)

A.C.T. van Duin et al,J. Phys. Chem. A 2001, 105, 9396-9409.

Reactive force fields

- REAXFF:

$$U = U_{bond} + U_{angles} + U_{torsions} + U_{over} + U_{VdW} + U_{Coulomb} + U_{specific}$$

U_{bond}= energy associated with forming bonds between atoms
U_{angle}= energy associated with 3-body valence angle strain
$U_{torsion}$= energy associated with 4-body torsional angle strain
U_{over}= energy penalty preventing overcoordination of atoms
U_{VdW}= dispersive (van der Waals) contributions
$U_{Coulomb}$= electrostatic contributions
$U_{specific}$= specific energy terms, e.g., lone pair, hydrogen bonding, etc.

C/H/O (2018)		
39 General parameters		
36 Atom parameters	(x3 atoms)	520 parameters
16 Bond parameters	(x6 bonds)	
7 Angles parameters	(x18 angles)	
7 Dihedral parameters	(x21 torsions)	
4 H bond parameters	(x1 H bond)	

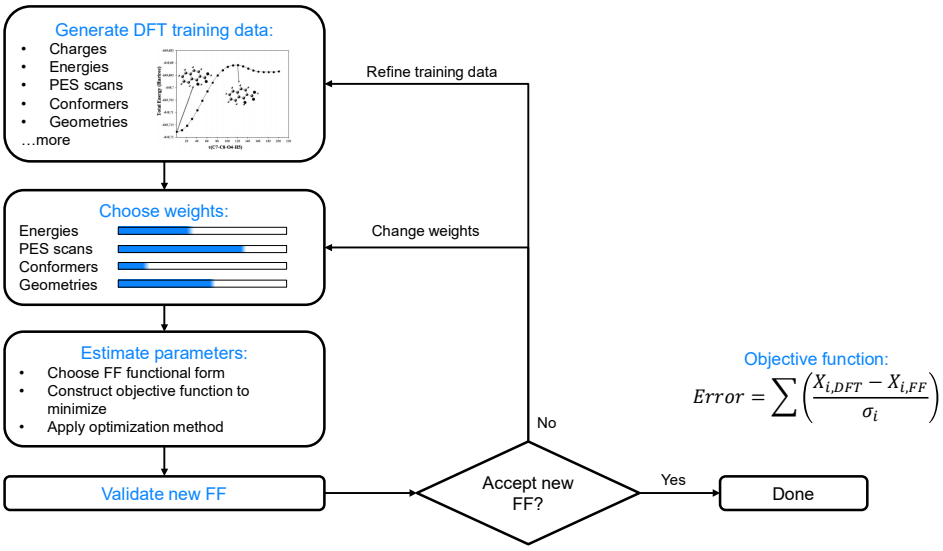
Parametrization

$$U(\boldsymbol{r}) = \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 + \\ + \sum_{angles} \frac{1}{2} k_a (\theta - \theta_0)^2 + \\ + \sum_{dihedrals} k_\phi [1 + \cos(n\phi - \delta)] + \\ + \sum_{improper} \frac{1}{2} k_\xi (\xi_{ijkl} - \xi_0)^2 + \\ + \sum_{i<j}^{atoms} 4\varepsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} \right) + \sum_{i<j}^{atoms} \frac{1}{4\pi\varepsilon_0} \frac{q_i q_j}{r_{ij}}$$

The parameters are typically:

- obtained from quantum mechanical calculations
- Calibrated to experimental data (neutron, X-ray and electron diffraction, NMR, infrared, Raman and neutron spectroscopy, etc.)

Parametrization procedure



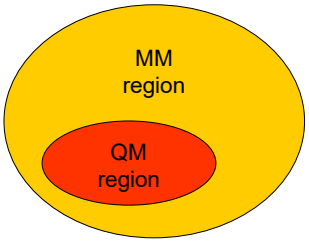
ReaxFF training sets in literature

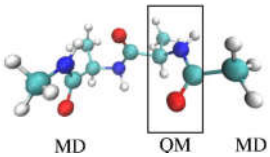
- 2001 (C/H): First parametrization for hydrocarbons. The parameters were derived from quantum chemical calculations on bond dissociation and reactions of small molecules plus heat of formation and geometry data for a number of stable hydrocarbon compounds.
- 2008 (C/H/O): Hydrocarbon oxidation. The force field parameters were determined by combining the original hydrocarbon training set with the QM data for transition states and reaction energies for systems relevant to hydrocarbon oxidation.
- 2015 (C/H – Aromatics): The training set was extended to describe the reactions for thermal decomposition and CC bond formation of large hydrocarbons like graphene and fullerene.
- 2016 (C/H/O): improved description of oxidation of small hydrocarbons and syngas reaction. It includes a number of reactions related to CO and HCO.
- Other versions for specific scopes are reported in literature: simulations of thermal decomposition of RDX and TATP; hydrocarbon chemistry catalyzed by Nickel and Vanadium Oxide

QM/MM

The hybrid **QM/MM** (quantum mechanics/molecular mechanics) approach is a MD method that combines the strengths of *ab initio* QM calculations (accuracy) and MM (speed) approaches, thus allowing for the study of chemical processes.

The chemical systems is partitioned into an electronically important region which requires a quantum chemical treatment and a remainder which only acts in a perturbative fashion and thus admits a classical description.





$$E = E_{QM} + E_{MM} + E_{QM/MM}$$

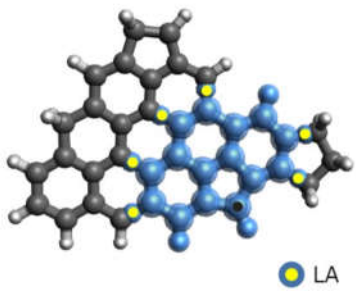
$$E_{QM} = \frac{\langle \psi H \psi \rangle}{\langle \psi \psi \rangle}$$

Interaction between QM and MM

Energy of MM subsystem

QM/MM implementation

The QM/MM division splits the systems along a chemical bond using a so-called link atom LA (hydrogen atom in the QM calculation step). It is not physically present in the MM subsystem, but the forces on it, that are computed in the QM step, are distributed over the two atoms of the bond. The bond length itself is constrained during the computations.



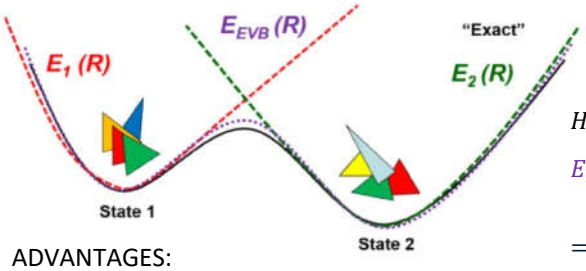
ONIOM approach:

$$E = E_I^{QM} + E_{I+II}^{MM} + E_I^{MM}$$

I: QM subsystem
II: MM subsystem

Empirical Valence Bond Method (EVB)

In the EVB method, simulation of reactive processes is conducted via the coupling of classical FFs, one classical FF for the reactant E_1 and one classical FF for the product E_2 using a coupling term C_{12} between the FFs.



$$H_{\mathrm{EVB}}(R) = \begin{pmatrix} E_1(R) & C_{12}(R) \\ C_{12}(R) & E_2(R) \end{pmatrix}$$

$$E_{\mathrm{EVB}} = \frac{(E_1 + E_2) - \sqrt{(E_1 + E_2)^2 - 4\left(E_1 E_2 - C_{12}^2\right)}}{2}$$

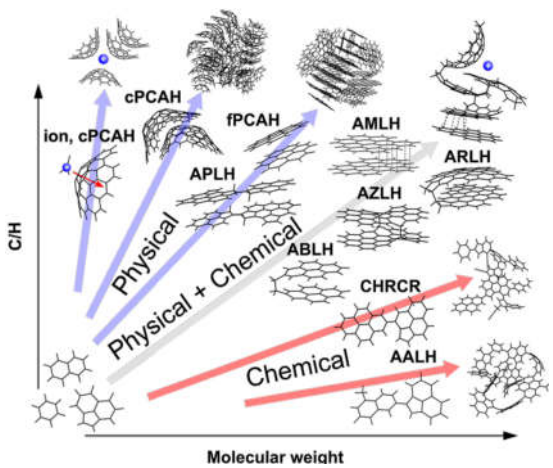
$$C_{12} = A_1 e^{-\left(\frac{E_1 - E_2 - A_2}{A_3}\right)^2} + A_4$$

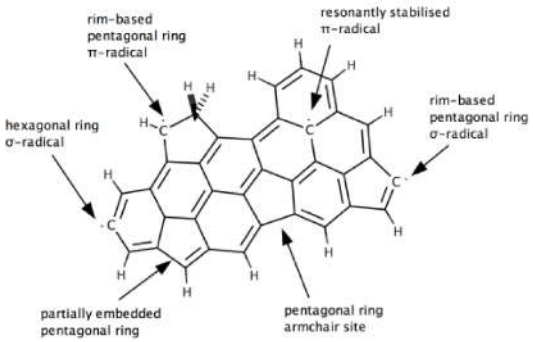
ADVANTAGES:

- Classic FFs are already available in literature
- It requires only bond scans for parameterization
- Easy to parametrize (max 4 parameters per reaction)
- Fast

FORCE FIELDS: APPLICATIONS IN COMBUSTION

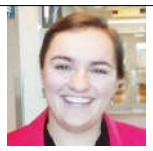
Soot nucleation





Reactive edges of soot precursors.

Basin hopping and potentials

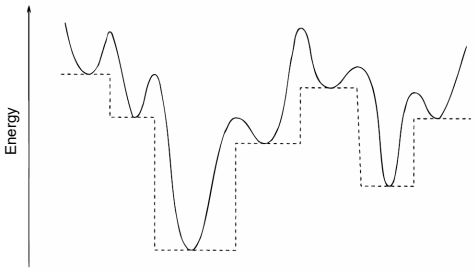

Kimberly Bowal


Peter Grancic


Tim Totton


David Wales

- Finds stable molecular clusters by searching for minima
- Based on potential energy 'landscape'
- Uses Monte-Carlo criterion when 'jumping' between minima




Alston Misquitta

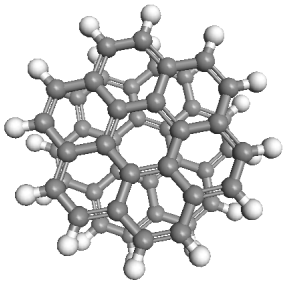
Modelling the internal structure of nascent soot particles Tim Totton, Dwaipayan Chakrabarti, Alston J. Misquitta, Markus Sander, David Wales, and Markus Kraft, *Combustion and Flame* **157**, 909-914, (2010).

Sphere Encapsulated Monte Carlo: Obtaining Minimum Energy Configurations of Large Aromatic Systems
Kimberly L. Bowal, Peter Grancic, Jacob W. Martin, and Markus Kraft, *Journal of Physical Chemistry A* **123**(33), 7303-7313, (2019).

Global minimum clusters

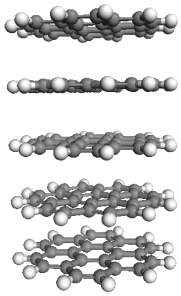
J.Houston Miller

2 Coronene molecules



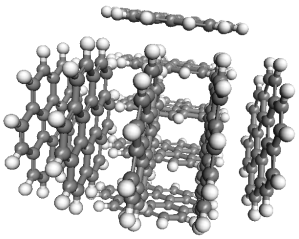
E = -94.90 kJ/mol

5 Coronenes molecules



E = -394.35 kJ/mol

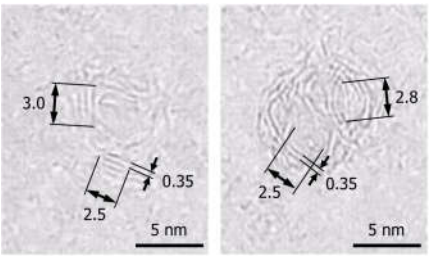
10 Coronene molecules



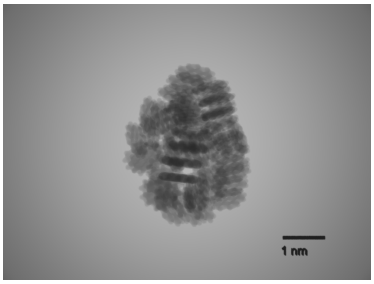
E = -926.42 kJ/mol

Intermolecular potential calculations for polynuclear aromatic hydrocarbon clusters
JD Herdman, JH Miller - The Journal of Physical Chemistry A, 2008

Experimental comparison



Experimental HR-TEM images of an aggregate sampled from a diesel engine. Indicated are length scales of structures within a primary particle (from Mosbach *et al.,* 2009, Combustion and Flame).



A TEM-style projection of a computed cluster of 50 coronene molecules

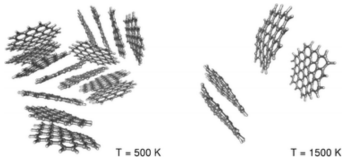
Physical nucleation – flat PAH

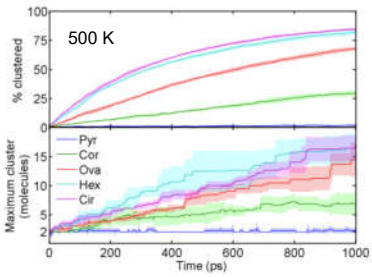
Hypothesis:

- flat PAH are responsible for soot nucleation
- Only physical interaction are important
- No chemical reactions

Force Field:

- Classical force fields
- OPLS-AA and isoPAHAP

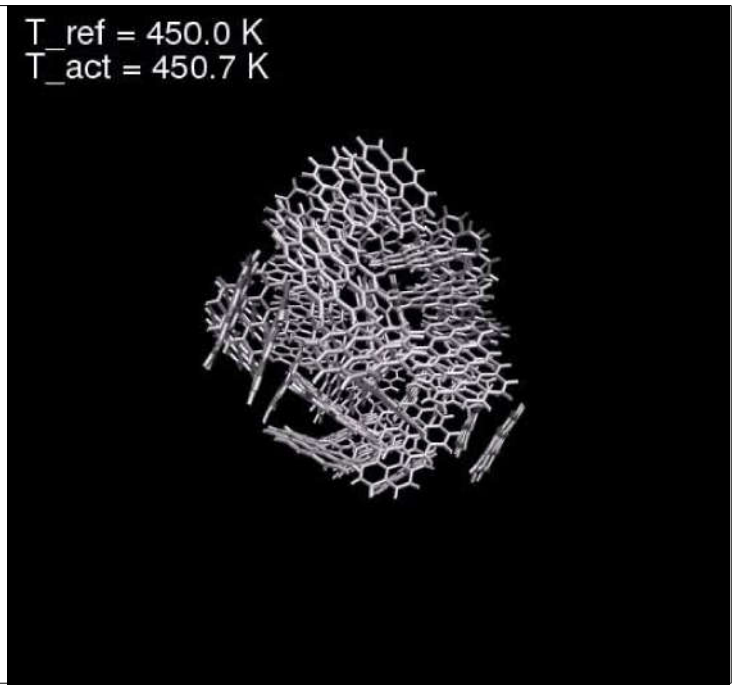




Totton et al. (2012) *Physical Chemistry Chemical Physics* 14, 4081-4096.

Heating a coronene$_{50}$ cluster

- Heated from 50K to 1000K over 20 ns
- Complete evaporation below flame temperatures
- Experimental sublimation temperature ~798K



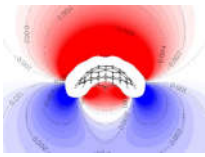
Physical nucleation – curved PAH

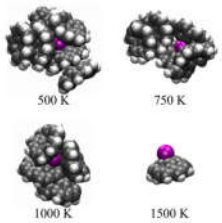
Hypothesis:

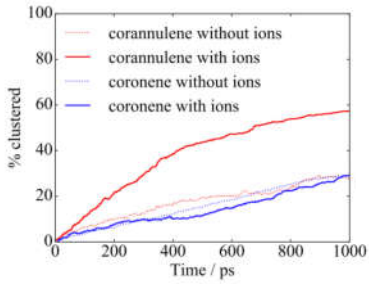
- Interactions between curved PAH and ions are responsible for soot nucleation
- Physical interaction + induced dipole need to be considered
- No chemical reactions

Force Field:

- Polarizable force fields
- AMOEBA or isoPAHAP with dummy atoms







Martin et al. (2019) *Proceedings of the Combustion Institute* 37(1), 1117-1123.

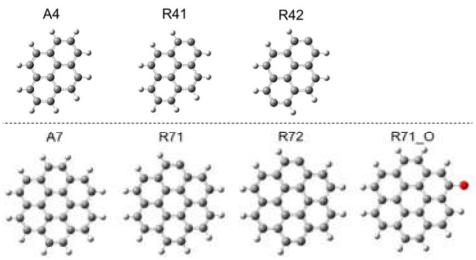
Chemical nucleation – σ-radicals

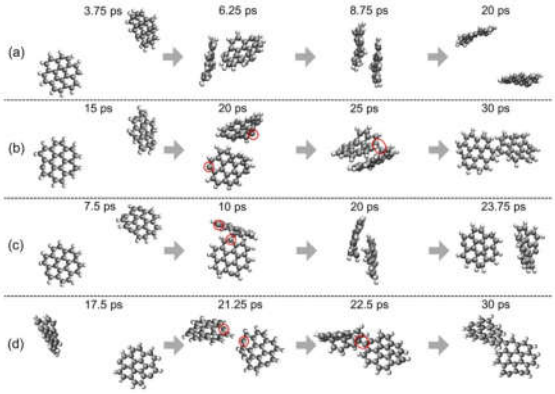
Hypothesis:

- σ-radicals are responsible for soot nucleation
- Chemical reactions between radical sites need to be considered

Force Field:

- Reactive force field
- REAXFF – C/H/O 2016





Mao et al. (2018) *J. Phys. Chem. A* 122, 8701-8708.

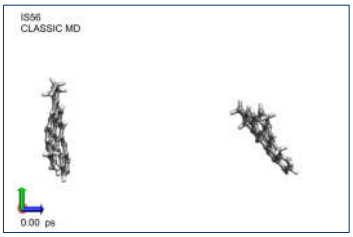
Physical/Chemical nucleation – π-radicals

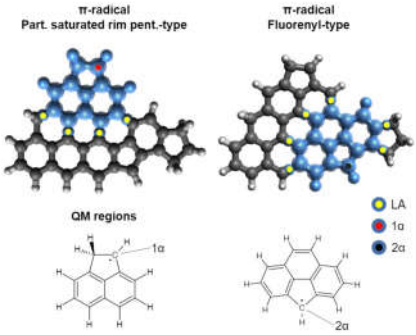
Hypothesis:

- π-radicals are responsible for soot nucleation
- Physical interactions and chemical reactions between radical sites need to be considered

Force Field:

- There is no reactive force field parametrized for reactions between π-radicals
- QMMM





Martin et al., *Journal of the American Chemical Society* **143**(31), 12212-12219, (2021).

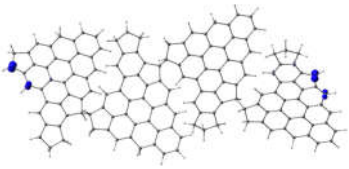
Physical/Chemical nucleation – π-radicals

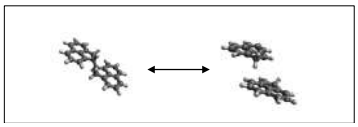
Hypothesis:

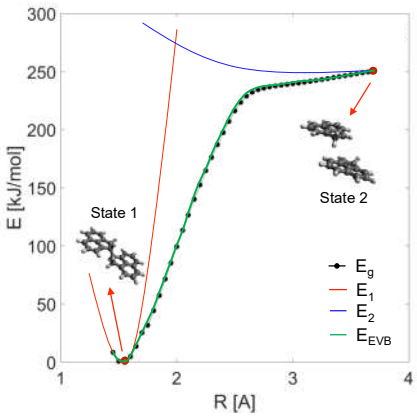
- π-radicals are responsible for soot nucleation
- Physical interactions and chemical reactions between radical sites need to be considered

Force Field:

- There is no reactive force field parametrized for reactions between π-radicals
- QMMM
- EVB method





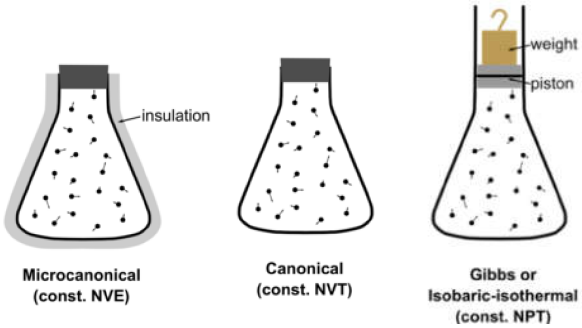


T and P CONTROL: ENSAMBLES, THERMOSTAS AND BAROSTATS

Statistical Ensemble

An **ensemble** is a collection of points in phase space satisfying the condition of particular thermodynamic state.

- Micro-canonical ensemble (NVE): This correspond to an isolated system.
- Canonical Ensemble (NVT): Energy is exchanged with an external heat bath.
- Isobaric-Isothermal Ensemble (NPT): A thermostat and a barostat are needed.



Temperature control

Temperature is related with the kinetic energy (and velocity) of the *N* atoms in our system

$$E_{kin} = \frac{1}{2}\sum_{i=1}^{N} m_i v_i^2 \qquad E_{kin} = \frac{3N}{2} k_B T$$

$$T = \frac{2E_{kin}}{3Nk_B} = \frac{1}{3Nk_B}\sum_{i=1}^{N} m_i v_i^2$$

Thermostats control the temperature by scaling in different ways the velocities.

❖ Velocity scaling thermostat: multiply velocities by a factor λ to obtain desired temperature T_0

- $v_{new}(t) = \lambda\, v(t)$
- $\lambda = \sqrt{\frac{T_0}{T(t)}}$
- Velocities are scaled every time step or every *n* steps

Temperature control

Temperature is related with the kinetic energy (and velocity) of the *N* atoms in our system

$$E_{kin} = \frac{1}{2}\sum_{i=1}^{N} m_i v_i^2 \qquad E_{kin} = \frac{3N}{2} k_B T$$

$$T = \frac{2E_{kin}}{3Nk_B} = \frac{1}{3Nk_B}\sum_{i=1}^{N} m_i v_i^2$$

Thermostats control the temperature by scaling in different ways the velocities.

❖ Berendsen thermostat: system is coupled to a heat bath with temperature T_{bath}

$$v_{new}(t) = \lambda\, v(t)$$

$$\lambda = \sqrt{1 + \frac{\Delta t}{\tau}\frac{(T_{bath} - T)}{T(t)}}$$

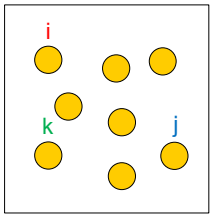
- τ is very large → $\lambda \approx 1$ → no coupling
- τ is very small → strong coupling

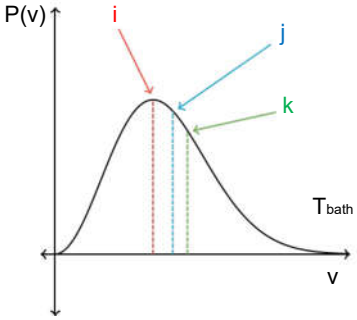
Temperature control

❖ Andersen Thermostat: random collisions of molecules with an imaginary heat bath (randomize velocities). At each time step some atoms are randomly chosen and their velocity are selected from a Maxwell-Boltzmann distribution at the desired temperature:

$$P(v_i) = \sqrt{\frac{m_i}{2\pi k_B T}} \exp\left(-\frac{m_i v_i^2}{2k_B T}\right)$$

The strength of the coupling to the heat bath is specified by a collision frequency *f* which describes the number of times an atom is selected along the simulation





Temperature control

❖ Langevin dynamics: The velocities are corrected by a random force and a constant friction.

The equation of motions are modified as follow:

$$\boldsymbol{F} = m\boldsymbol{a} = -\nabla U(\boldsymbol{r}) - \gamma m\boldsymbol{v} - \sqrt{2\gamma m k_B T}\boldsymbol{R}(t)$$

Standard force

Friction force with coefficient γ

Random force due to stochastic collision

$\gamma = 2 - 5\ ps^{-1}$

$R(t)$ – Gaussian random process

These additional forces are applied every n steps, which are defined by the so-called collision frequency.

Pressure control

The pressure and the volume of a system are related by the virial equation of state:

$$\frac{PV_m}{RT} = 1 + \frac{B}{V_m} + \frac{C}{V_m^2} + \cdots$$

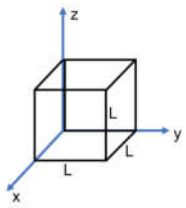
The pressure can be controlled by modifying the volume of the simulation box.

❖ Berendsen barostat – the pressure is weakly coupled to a "pressure bath" and the volume periodically rescaled

$$\mathrm{L}_{new}(t) = \mu\, \mathrm{L}(t)$$

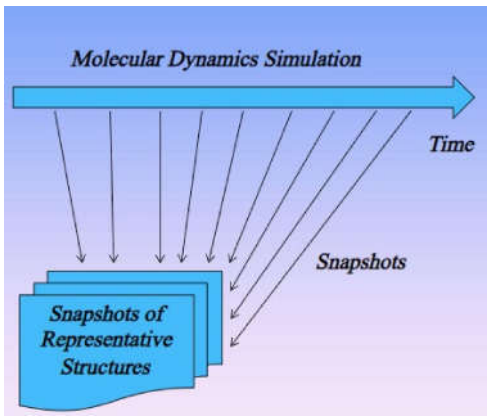
$$\mu = \sqrt[3]{1 - \frac{\Delta t}{\tau}\gamma(P_{bath} - P)}$$

- γ is the thermal compressibility
- τ is the "rise time" that describes the strength of the coupling



MD OUTPUTS

Output: MD trajectory

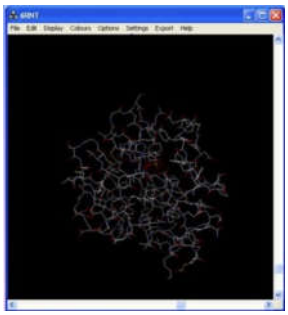


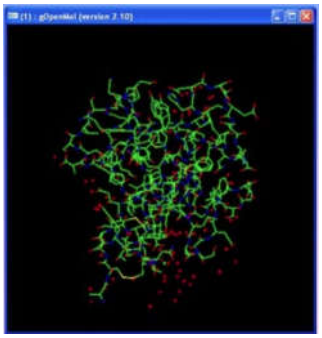
MD trajectory is a file containing positions and velocities of every atom as they vary with time.

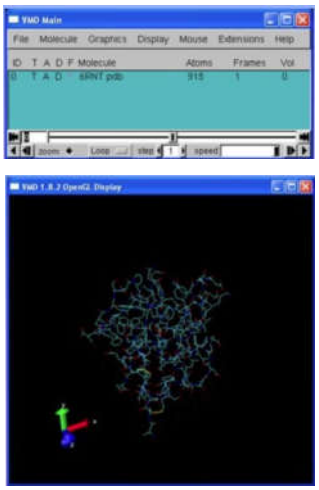
From this trajectory, the average values of properties can be determined.

Visualization software

gOpenMol
Ovito
VMD (Visual MD)
RasMol







Simple statistical quantities

Physical properties are usually a function of the particle coordinates and velocities.
The instantaneous value of a generic physical property *A* at time *t* is:

$$A(t) = f\big(r_1(t), \dots, r_N(t), v_1(t), \dots, v_N(t)\big)$$

Measuring quantities in MD usually means performing time averages of physical properties over the system trajectory:

$$\langle A \rangle = \frac{1}{N_T} \sum_{t=1}^{N_T} A(t)$$

where *t* is an index which runs over the time steps from 1 to the total number of steps N_T.

Ergodic hypothesis

- Ensemble average:

$$\langle A \rangle = \iint A\big(r^N, v^N\big)\rho\big(r^N, v^N\big)\, dr^N dv^N$$

- Time average:

$$\bar{A} = \lim_{T \to \infty} \frac{1}{T} \int_{t_0}^{t_0+T} A\big(r^N(t), v^N(t)\big) dt$$

THE ERGODIC HYPOTHESIS STATES THAT FOR $T \to \infty$

$$\langle \boldsymbol{A} \rangle = \bar{\boldsymbol{A}}$$

So we can compute thermodynamic averages from sufficiently long MD trajectories

Thermodynamic properties

- **Kinetic Energy**

$$\langle E_{kin} \rangle = \left\langle \frac{1}{2} \sum_{i=1}^{N} m_i v_i^2 \right\rangle$$

- **Temperature**

$$T = \frac{2}{3Nk_B} \langle E_{kin} \rangle$$

- **Pressure**

$$PV = Nk_B T + \frac{1}{3} \left\langle \sum_{i=1}^{N} r_i \cdot F_i \right\rangle$$

Other properties

STRUCTURAL PROPERTIES:

- **Radial distribution function:**

particle density of type B at a distance *r* around particles A

$$g_{AB}(r) = \frac{\langle \rho_B(r) \rangle}{\langle \rho_B \rangle_{local}} = \frac{1}{\langle \rho_B \rangle_{local}} \frac{1}{N_A} \sum_{i \in A}^{N_A} \sum_{j \in B}^{N_B} \frac{\delta(r_{ij} - r)}{4\pi r^2}$$

particle density of type B averaged over all spheres around particles A

DYNAMICAL PROPERTIES:

- **Diffusion coefficient:** Green-Kubo relation

$$D = \frac{1}{3N} \int_0^{\infty} \langle v_i(t) \cdot v_i(0) \rangle_{i \in A} dt$$

Velocity autocorrelation function for particle of type A

Common mistakes

- Simulation is **too short** (results are not meaningful, out of thermodynamic equilibrium)
- **Inadequate force field**
- Δt is too large (E not conserved, unstable simulation)
- System is too small (simulation box)
- Cut-off too short (improper treatment of long range interactions)
- Statistical significance: number of trajectories

MD softwares

OPEN SOURCE:

- Gromacs
- Tinker/Tinker-HP
- OpenMM
- NAMD
- LAMMPS
- CP2K
- DL_POLY

COMMERCIAL SOFTWARE:

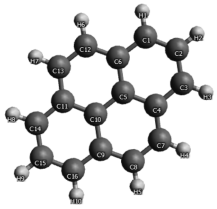
- Amber
- CHARMM

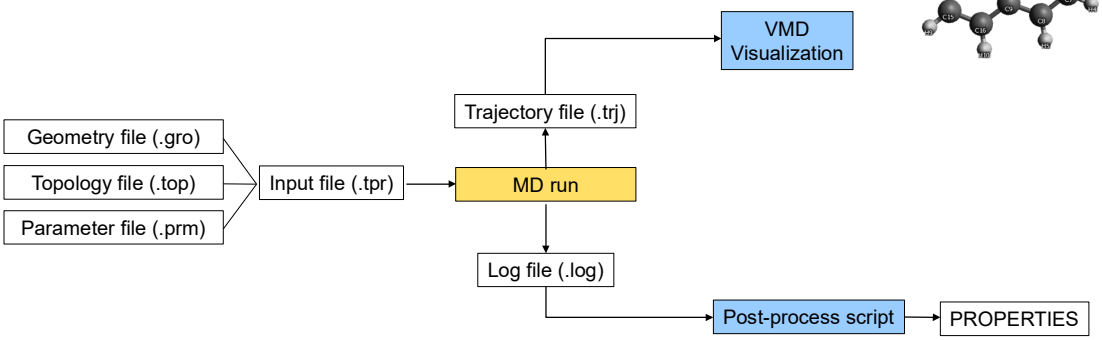
MD with GROMACS

Example

Physical nucleation of soot with pyrene:

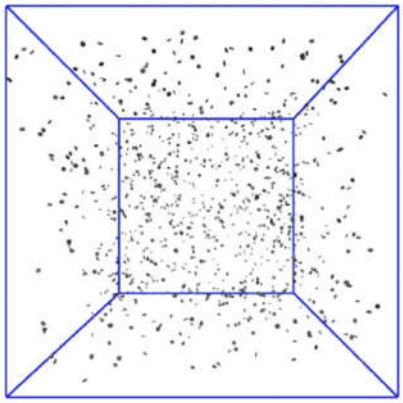
- 1000 PYRENE MOLECULES
- T = 1000 K
- **SOFTWARE: Gromacs**





Geometry file

pyrene_ic.gro



```
PYRENE
260000
    1PYR     C1    1  62.064  65.153 101.343
    1PYR     C2    2  61.956  65.245 101.368
    1PYR     C3    3  62.152  65.107 101.438
    1PYR     C4    4  62.121  65.137 101.379
    1PYR     C5    5  62.000  65.197 101.606
    1PYR     C6    6  61.914  65.247 101.500
    1PYR     C7    7  62.266  65.038 101.420
    1PYR     C8    8  62.352  64.997 101.518
    1PYR     C9    9  62.318  65.030 101.663
    1PYR    C10   10  62.212  65.099 101.688
    1PYR    C11   11  61.962  65.229 101.749
    1PYR    C12   12  62.048  65.186 101.850
    1PYR    C13   13  62.168  65.129 101.814
    1PYR    C14   14  61.784  65.304 101.527
    1PYR    C15   15  61.749  65.343 101.659
    1PYR    C16   16  61.844  65.304 101.765
    1PYR     H1   17  62.124  65.171 101.250
    1PYR     H2   18  61.915  65.307 101.287
    1PYR     H3   19  62.281  65.029 101.314
    1PYR     H4   20  62.428  64.937 101.480
    1PYR     H5   21  62.373  64.987 101.736
    1PYR     H6   22  62.019  65.215 101.949
    1PYR     H7   23  62.223  65.081 101.891
    1PYR     H8   24  61.717  65.313 101.457
    1PYR     H9   25  61.638  65.357 101.696
    1PYR    H10   26  61.825  65.318 101.877
    2PYR     C1   27 114.662 106.265  44.354
    2PYR     C2   28 114.787 106.309  44.334
    2PYR     C3   29 114.631 106.126  44.348
    2PYR     C4   30 114.740 106.032  44.382
    2PYR     C5   31 114.874 106.087  44.375
    2PYR     C6   32 114.891 106.220  44.344
    2PYR     C7   33 114.501 106.085  44.319
.....
```

Number of atoms

Index and name of the molecule

Atom name and number

Atom coordinates

Topology file

```
[ defaults ]
; nbfunc    comb-rule   gen-pairs   fudgeLJ fudgeQQ
1       1       yes     0   0.5

#include "oplsaa.ff/ffbonded.itp"

[atomtypes]
; name      bond_type  mass        charge    ptype  C        A
 oplsca     CA  6      12.01100    0.000      A    1.0      1.0
 oplsha     HA  1       1.00800    0.000      A    1.0      1.0

[nonbond_params]
; i         j       func    C           A
  oplsca    oplsha    1    1.0      1.0

[ moleculetype ]
; Name              nrexcl
Pyrene              3

[ atoms ]
;   nr   type      resnr residue  atom   cgnr     charge       mass
    1    oplsca       1    PYR     C1      1     -0.239     12.011
    2    oplsca       1    PYR     C2      2     -0.239     12.011
    3    oplsca       1    PYR     C3      3      0.235     12.011
    4    oplsca       1    PYR     C4      4     -0.040     12.011
    5    oplsca       1    PYR     C5      5     -0.040     12.011
    6    oplsca       1    PYR     C6      6      0.235     12.011
    7    oplsca       1    PYR     C7      7     -0.239     12.011
    8    oplsca       1    PYR     C8      8     -0.239     12.011
    9    oplsca       1    PYR     C9      9     -0.239     12.011
   10    oplsca       1    PYR    C10     10      0.235     12.011
   11    oplsca       1    PYR    C11     11      0.235     12.011
   12    oplsca       1    PYR    C12     12     -0.239     12.011
```

File containing parameters for bonded interactions

pyrene.top

Topology file

```
[ bonds ]
;  ai    aj  funct
    1     2     1
    1     3     1
    1    17     1
...

[ pairs ]
;  ai    aj  funct
    1     5     1
    1     8     1
    1    10     1
...

[ angles ]
;  ai    aj    ak  funct
    2     1     3     1
    2     1    17     1
    3     1    17     1
...
```

```
...

[ dihedrals ]
;  ai    aj    ak    al  funct
    3     1     2     6     3
    3     1     2    18     3
   17     1     2     6     3
...

[ dihedrals ]
;  ai    aj    ak    al  funct
    2     1     3    17     1    improper_Z_CA_X_Y
    6     2     1    18     1    improper_Z_CA_X_Y
    4     3     7     1     1    improper_Z_CA_X_Y
...

[ system ]
; Name
pyrene

[ molecules ]
; Compound        #mols
Pyrene             1000
```


Parameter file

```
; NVT production

integrator     = md-vv     ; velocity verlet
dt             = 0.001     ; 1 fs
nsteps         = 1000000   ; 1 ns

nstxout        = 1000      ; save output every 1.0 ps
nstvout        = 1000
nstfout        = 1000
nstenergy      = 1000
nstxtcout      = 1000

continuation   = no
constraints    = none

gen-vel        = yes   ; assign velocities from Maxwell distribution
gen-temp       = 1000  ; temperature for Maxwell distribution
gen-seed       = -1    ; generate a random seed

pbc            = xyz

nstlist        = 10    ; 10 fs
rcoulomb       = 9.0   ; short-range electrostatic cutoff (in nm)
rvdw           = 9.0   ; short-range van der Waals cutoff (in nm)

coulombtype    = PME   ; Particle Mesh Ewald for long-range electrostatics
fourierspacing = 0.16  ; grid spacing for FFT
pme_order      = 4     ; cubic interpolation
rcoulomb       = 9     ; Short-range electrostatic cut-off

vdwtype        = User
energygrps     = oplsca oplsha
energygrp_table = oplsca oplsca oplsha oplsha

tcoupl     = nose-hoover   ; thermostat
tau-t      = 0.5           ; time constant, in ps           ;
ref-t      = 1000          ; reference temperature, one for each group, in K
```

Output

- Log file:

```
          Step           Time         Lambda
       1000000     1000.00000        0.00000

Writing checkpoint, step 6000000 at Thu Oct 13 23:51:31 2016


   Energies (kJ/mol)
           Bond  Tab. Bonds NC          Angle    Proper Dih. Ryckaert-Bell.
    1.71958e+05    8.61897e+04    1.17591e+05    9.69476e+03    1.33163e+05
          LJ-14     Coulomb-14        LJ (SR)   Coulomb (SR)   Coul. recip.
    0.00000e+00    1.76593e+05   -1.94184e+04   -1.36995e+05   -4.82148e+04
      Potential    Kinetic En.   Total Energy  Conserved En.    Temperature
    4.90561e+05    4.48802e+05    9.39362e+05    7.12142e+05    9.99628e+02
 Pressure (bar)
    1.17419e-01
```


Output

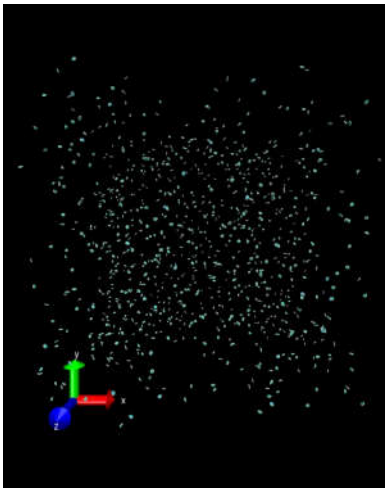
- Log file:

```
<======  ###############  ==>
<====  A V E R A G E S  ====>
<==  ###############  ======>

Statistics over 1000001 steps using 10001 frames

 Energies (kJ/mol)
           Bond  Tab. Bonds NC          Angle    Proper Dih. Ryckaert-Bell.
    1.67140e+05    8.58071e+04    1.13310e+05    9.23079e+03    1.28346e+05
          LJ-14     Coulomb-14        LJ (SR)   Coulomb (SR)   Coul. recip.
    0.00000e+00    1.76605e+05   -1.94947e+04   -1.36994e+05   -4.82151e+04
      Potential    Kinetic En.   Total Energy  Conserved En.    Temperature
    4.75736e+05    4.48975e+05    9.24711e+05    6.12277e+05    1.00001e+03
 Pressure (bar)
    2.32399e-01

 Total Virial (kJ/mol)
    1.46163e+05    5.77300e+01    6.40120e+01
    5.77278e+01    1.46380e+05   -3.45783e+00
    6.40155e+01   -3.44560e+00    1.45924e+05

 Pressure (bar)
    2.31423e-01    4.91329e-03    4.12217e-04
    4.91344e-03    2.22955e-01   -6.14399e-03
    4.11984e-04   -6.14480e-03    2.42817e-01

 Epot (kJ/mol)        Coul-SR          LJ-SR        Coul-14          LJ-14
 oplsca-oplsca    4.27991e+04   -1.42722e+04    1.86964e+05    0.00000e+00
 oplsca-oplsha   -4.15405e+05   -7.80988e+03   -5.05093e+04    0.00000e+00
 oplsha-oplsha    2.35612e+05    2.58736e+03    4.01509e+04    0.00000e+00
```

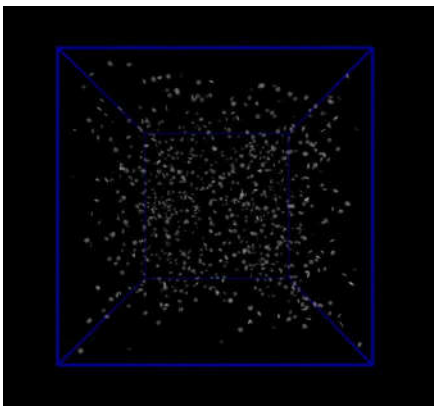
Output

prd.trr

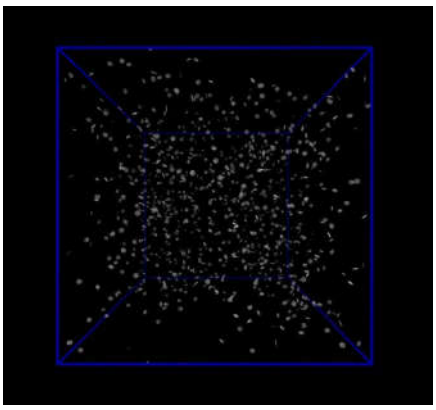
VMD →



MD – circumcoronene ($C_{54}H_{18}$)



500K



1500K

End of Lecture 3

Soot – Part 4

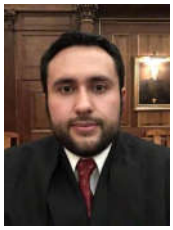
Markus Kraft

Computational Modelling Group Cambridge

Main Contributors:
Dr Jake Martin (Part 1)
Dr Angiras Menon (Part 2)
Dr Laura Pascazio (Part 3)
Dr Gustavo Leon (Part 4)

Jacob Martin

AngirasMenon

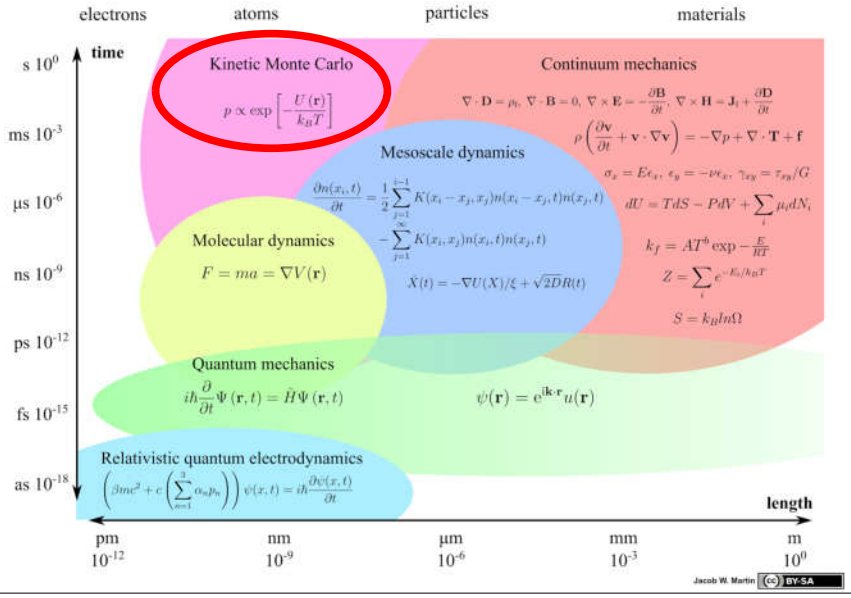


Gustavo Leon

Laura Pascazio

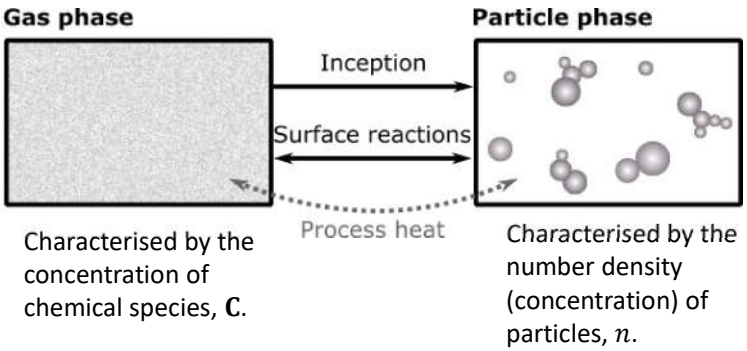
Part 1 Overview
Part 2 Quantum Chemistry
Part 3 Molecular Dynamics
Part 4 Kinetic Monte Carlo
Part 5 Stochastic Particle Methods
Part 6 Application – engine model

Scale Diagramme



Models for aerosol systems

Phases in aerosol systems



Gas Phase

Contains N_{sp} chemical species such that, $\mathbf{C} = \left(C_1, \ldots, C_{N_{\mathrm{sp}}}\right)$, represents their concentrations, and has a Temperature, T.

The concentration of each species k changes in time, t, and space, ω, due to reactions, particle processes and flow[9],

$$\frac{\delta C_k}{\delta t} + \frac{\delta}{\delta \omega} \cdot \underbrace{\left[u_{\mathrm{g}}(\mathbf{C}, T) C_k\right.}_{\text{Advective flow}} \underbrace{\left. - D_k(\mathbf{C}, T) \frac{\delta C_k}{\delta \omega}\right]}_{\text{Diffusive flow}} = \underbrace{\mathcal{W}(\mathbf{C}, T) + \mathcal{G}(\mathbf{C}, T, n)}_{\text{Particle processes}}$$

Particle Phase

Formed by collisions between gas phase (precursor) species.	**Inception**
Interacts with the gas-phase via surface chemical reactions.	**Surface processes**
Collisions between particles result in interconnected aggregates.	**Coagulation**
Surface reactions produce more rounded (spherical) structures.	**Sintering and coalescence**

Governed by a *Population Balance Equation* (PBE).

Number density

For a particle of type $x \in \mathcal{E}$, at position $\omega \in \Omega$, the number density (concentration) of particles is $n(x, \omega)$.

The number density evolves as particles experience different particle processes.

Can be used to calculate *integral* properties of the particle ensemble (particle size distribution).

Population Balance Equation

$$\frac{\delta n}{\delta t} + \frac{\delta}{\delta x} \cdot \underbrace{[S_x(\mathbf{C}, T, x)n]}_{\text{Surface reactions}} + \frac{\delta}{\delta \omega} \cdot \underbrace{\left[u_{\mathrm{P}}(\mathbf{C}, T, x)n - D_{\mathrm{P}}(\mathbf{C}, T, x)\frac{\delta n}{\delta \omega}\right]}_{\text{Particle flow}} =$$

$$\underbrace{I(\mathbf{C}, T, x)}_{\text{Inception}} + \underbrace{\int_{\varepsilon} K(T, x-y, y)n(x-y)n(y)dy - \int_{\varepsilon} K(T, x, y)n(x)n(y)dy}_{\text{Coagulation}}$$

Casper Lindberg

Astrid Boje

Particle models

A detailed particle model for polydisperse aggregate particles Casper Lindberg, Manoel Y. Manuputty, Edward K. Y. Yapp, Jethro Akroyd, Rong Xu, and Markus Kraft, *Journal of Computational Physics* **397**, 108799, (2019).

Stochastic population balance methods for detailed modelling of flame-made aerosol particles Astrid Boje and Markus Kraft, *Journal of Aerosol Science* **159**, 105895, (2022).

Type space

Space used for mathematically representing particles.

A general particle, P_q, can have multivariate states with different coordinates[9]:

$$P_q \begin{cases} \text{spatial (external) coordinates} & \text{physical space,} \quad \omega \in \Omega \\ \text{property (internal) coordinates} & \text{type space,} \quad x \in \mathcal{E} \end{cases}$$

An ensemble containing N_P particles has P_q, with $q \in \{1, \dots, N_\mathrm{P}\}$.

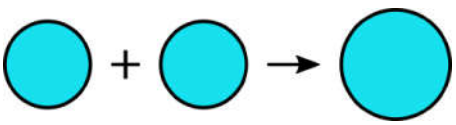
Low dimensional models

Coalescent sphere

Particles are represented as spheres of constant composition and density.
One internal coordinate: Mass[9].

$$P_q = P_q(\eta_q)$$

$$d_p(P_q) = \left(\frac{6}{\pi}\frac{m(P_q)}{\rho}\right)^{\frac{1}{3}}$$



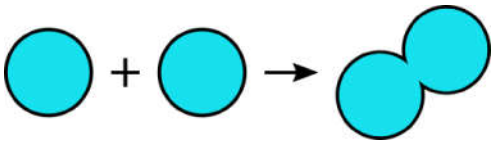
Implies instantaneous coalescence after coagulation events.

Low dimensional models

Surface-volume models

Two internal coordinates: Mass and Surface area.
Surface area is conserved after coagulation events.

$$P_q = P_q(\eta_q, A_q)$$



Fractal dimension needs to be assumed. Typically, ~ 1.8.
Surface processes affect the total area of the particle.

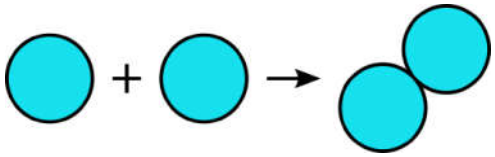
Detailed particle models

A *particle aggregate*, P_q, is defined by its constituent *primary particles*, p_i, where i ∈ $1, \dots, n_q$, and their connectivity $\mathbf{C}_q$, with elements C_{ij}.

$$P_q = P_q\left(p_1, \dots, p_{n_q}, \mathbf{C}_q\right)$$

$$C_{ij} = \begin{cases} 1 & \text{if } p_i, p_j \text{ are adjacent} \\ 0 & \text{if } p_i, p_j \text{ are not adjacent} \end{cases}$$

$$p_i = p_i(v_i)$$



Connectivity is retained but fractal dimension is still assumed.

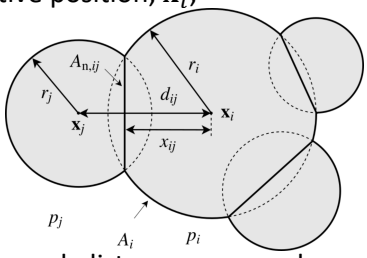
These *overlapping-sphere* models have been successfully used to study aggregation with surface growth in soot[10] and silica aerosols[11].

Detailed particle models

Detailed particle model by Lindberg et al.[12] describes particles by their chemical composition, η_i, radius, r_i, and relative position, $\mathbf{x}_i$,

$$P_q = P_q\left(p_1, \dots, p_{n_q}, \mathbf{C}_q\right)$$

$$p_i = p_i(\eta_i, r_i, \mathbf{x}_i)$$

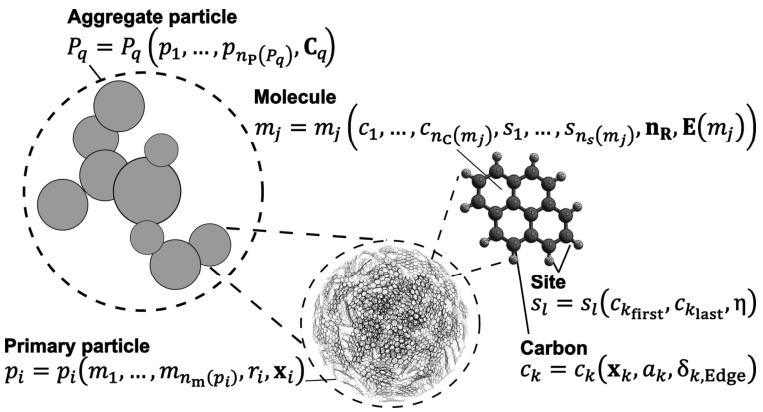


Centre-to-centre separation, d_{ij}, or centre-to-neck distance, x_{ij}, can be obtained from geometry.

No need to assume fractal dimension.

Detailed particle models

Detailed particle model by Leon[13] that describes carbonaceous particles by containing the chemical species that conform them.



Two *layers* in the model[13]:

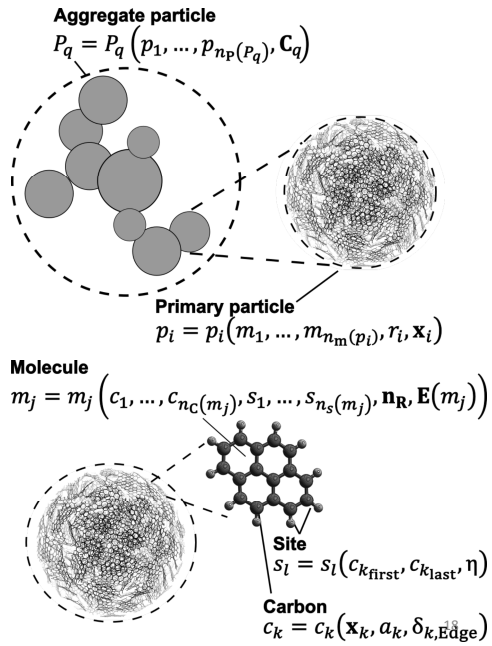
1. Particles and aggregates

   Each primary particle, p_i, contains n_m molecules, radius, r_i, and relative position, $\mathbf{x}_i$.

   Similar description to previous slides.

2. Molecules (PAHs)

   A general molecule, m_j, contains n_C carbon atoms (c_k), n_s sites (s_l), $\mathbf{n_R}$ rings, and an Edge connectivity matrix $\mathbf{E}(m_j)$.



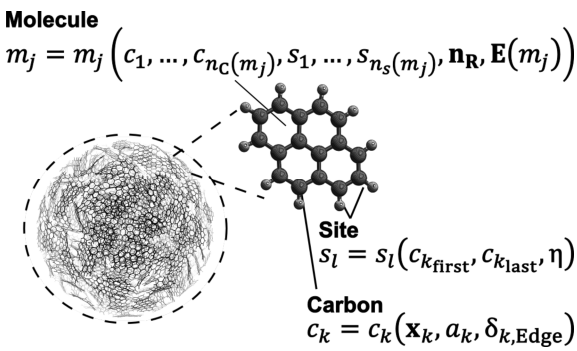
The list of carbon atoms and the list of sites can be written as:

$C(\mathrm{m_j}) = \left(c_1, \dots, c_{n_\mathrm{C}}\right)$ and

$S(\mathrm{m_j}) = \left(s_1, \dots, s_{n_\mathrm{S}}\right)$.

The number of rings in a molecule is:

$$\mathbf{n_R} = \begin{bmatrix} n_\mathrm{R6} \\ n_\mathrm{R5} \\ n_{\mathrm{R5_{emb}}} \\ n_\mathrm{R7} \end{bmatrix}$$



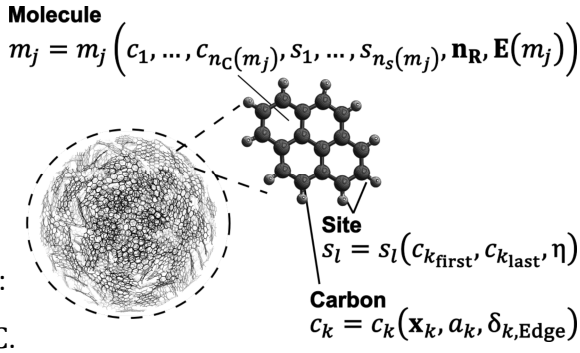
The Edge connectivity matrix $\mathbf{E}(m_j)$ has elements:

$$E_{k_1,k_2} = \begin{cases} 1, & \text{if } c_{k_1} \text{and } c_{k_2} \text{ are connected,} \\ 0, & \text{otherwise,} \end{cases}$$

where $k_1, k_2 \in I_\mathrm{Edge}$, the set of edge carbon atom indices

A carbon atom, c_k, contains its spatial coordinates, $\mathbf{x}_k$, the attached third atom, a_k, typically carbon (C) or hydrogen (H) and a binary variable $\delta_{k,\mathrm{Edge}}$:

$$\delta_{k,\mathrm{Edge}} = \begin{cases} 1, & \text{on the edge,} \\ 0, & \text{otherwise,} \end{cases}$$

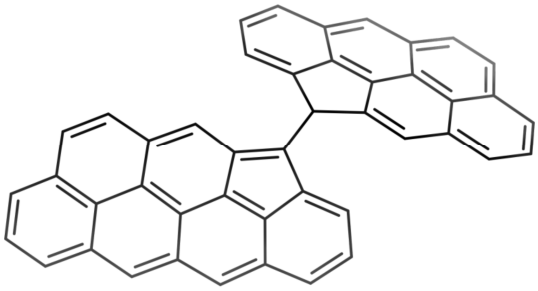


A site, s_l, contains two carbons, $c_{k,\mathrm{first}}$ and $c_{k,\mathrm{last}}$ that are part of a reactive site:

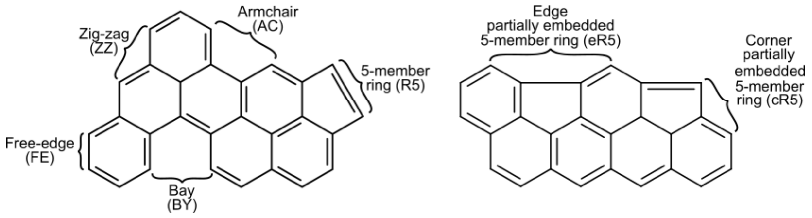
$k_\mathrm{first}, k_\mathrm{last} \in I_\mathrm{Edge}$ and $a_{k_\mathrm{first}}, a_{k_\mathrm{last}} \neq \mathrm{C}$.

The site type, η, is determined by the shortest path between $c_{k,\mathrm{first}}$ and $c_{k,\mathrm{last}}$ in the graph defined by $\mathbf{E}(m_j)$ and has values $\eta \in \{\mathrm{FE}, \mathrm{ZZ}, \mathrm{AC}, \dots\}$

Crosslinking inside the same molecule is possible with this description



Sites



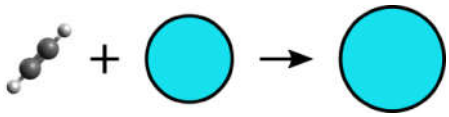
Can have different length (number of carbons)
Different site types depending on the embedding of five-member rings
Edge (R5), Edge partially-embedded (eR5), Corner partially-embedded (cR5)

Molecular (PAH) processes

Particle layer

Particle radius is modified by molecular processes (e.g., acetylene addition).
The locality of the addition is not tracked.

$$p_i(m_1, \dots, m_j, \dots, m_{n_\mathrm{m}}, r_i, \mathbf{x}_i) + \mathrm{A_{gas}} \rightarrow p_i^* \left(m_1, \dots, m_j^*, \dots, m_{n_\mathrm{m}}, r_i^*, \mathbf{x}_i\right)$$



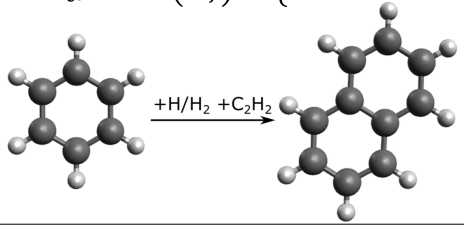
Molecular layer

Molecule is modified as well as carbon atoms and sites.
Additional carbon atoms and sites are added/removed.
For example: $n_{\mathrm{c_{New}}}$ carbon atoms are added, two existing carbon atoms are modified.

$$m_j \left(C(\mathrm{m_j}), S(\mathrm{m_j}), \mathbf{n_R}, \mathbf{E}(m_j)\right) + \mathrm{A_{gas}} \rightarrow m_j^* \left(C^* \left(\mathrm{m_j^*}\right), S^* \left(\mathrm{m_j^*}\right), \mathbf{n_R^*}, \mathbf{E}^* \left(m_j^*\right)\right)$$

$$C(\mathrm{m_j}) = \{c_1, \dots, c_u, c_{u+1}, \dots, c_{n_\mathrm{c}}\} \rightarrow C^* \left(\mathrm{m_j^*}\right) = \left\{c_1, \dots, c_u^*, c_{u+1}^*, \dots, c_{n_\mathrm{c}}, c_{1_\mathrm{New}}, \dots, c_{n_{\mathrm{c_{New}}}}\right\}$$

$c_u(\mathbf{x}_u, \mathrm{H}, 1)$
$c_{u+1}(\mathbf{x}_u, \mathrm{H}, 1)$



$c_u^*(\mathbf{x}_u, \mathrm{C}, 1)$
$c_{u+1}^*(\mathbf{x}_u, \mathrm{C}, 1)$

Molecular layer

Molecule is modified as well as carbon atoms and sites.

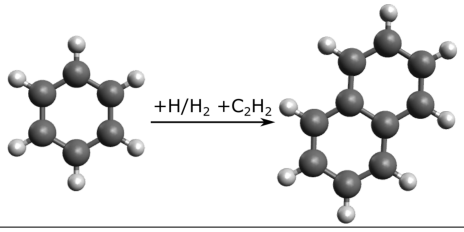
Additional carbon atoms and sites are added/removed.

For example: n_{sNew} sites are added, three existing sites are modified.

$$S(\mathrm{m_j}) = \{s_1, \dots, s_{v-1}, s_v, s_{v+1}, \dots, s_{n_\mathrm{s}}\} \rightarrow S^*\left(\mathrm{m_j^*}\right) = \{s_1, \dots, s_{v-1}^*, s_v^*, s_{v+1}^*, \dots, s_{n_\mathrm{s}}, S_\mathrm{New}\}$$

$$S_\mathrm{New} = \left\{s_{1_{New}}, \dots, s_{n_{s_{New}}}\right\}$$

$$s_{v-1}\left(c_{v-1_\mathrm{first}}, c_u, \mathrm{FE}\right)$$
$$s_v\left(c_u, c_{u+1}, \mathrm{FE}\right)$$
$$s_{v+1}\left(c_{u+1}, c_{v+1_\mathrm{last}}, \mathrm{FE}\right)$$

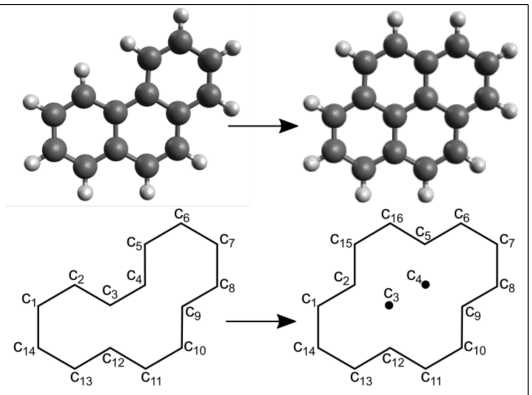


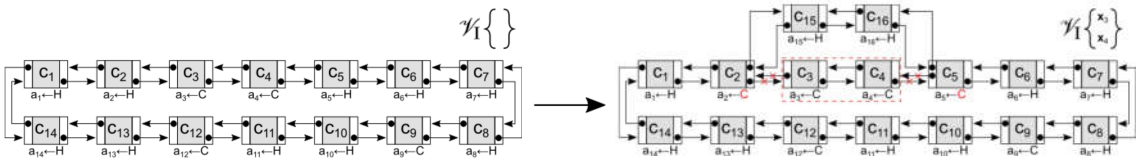
$$s_{v-1}^*\left(c_{v-1_\mathrm{first}}, c_{v-1_\mathrm{last}}, \mathrm{ZZ}\right)$$
$$s_v^*\left(c_{v_\mathrm{first}}, c_{u+1}, \mathrm{FE}\right)$$
$$s_{v+1}^*\left(c_{v+1_\mathrm{first}}, c_{v+1_\mathrm{last}}, \mathrm{ZZ}\right)$$

Molecular layer

Two carbon atoms are moved from L_c to $\mathscr{V}_\mathrm{I}$.

Important to notice that these carbon atoms are *disconnected* from the rest and only their coordinates are stored.

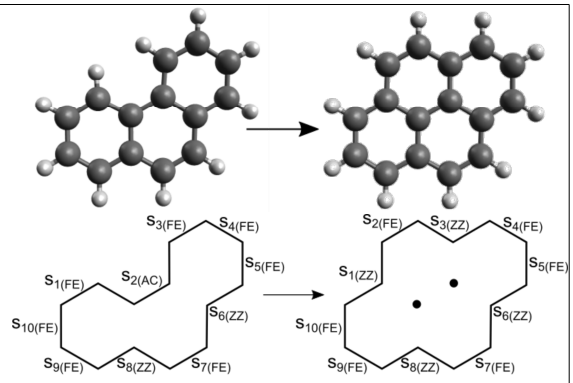


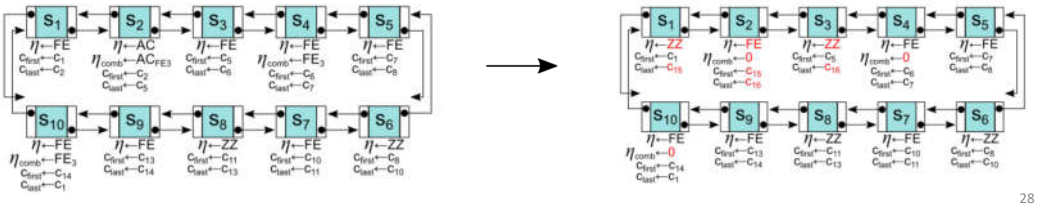


Molecular layer

Sites from L_S are modified. Principal and combined site types are updated, as well as the site connectivity.

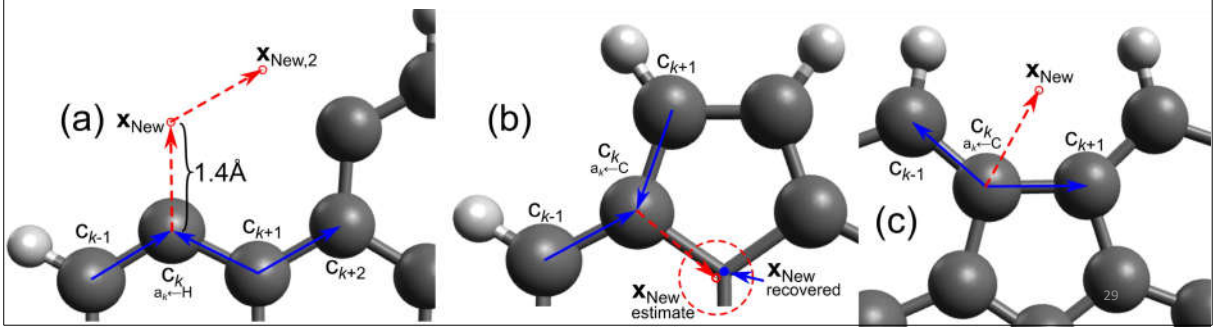
In this case no sites were added/removed.





Molecular layer

The positions of new carbon atoms can be obtained using geometry and the positions of neighbouring carbon atoms. Edge carbon atoms coming from the internal structure are searched in a similar matter. A position for the atom is estimated using geometry and selecting the right coordinates from $\mathscr{V}_I$ by minimising the distance to this position.



Structure Optimisation

After certain situations it may be desirable to optimise the molecular structure to better represent the geometry of the studied material.

In the model, this is done using OpenBabel[16] with the following steps:

1. Pass carbon atom coordinates from L_c and $\mathscr{V}_I$ to an OpenBabel *molecule*. Pass coordinates for possible hydrogen atom locations if $a_k \leftarrow \mathrm{H}$.
2. Use the OpenBabel internal routines to assume the connectivity of the molecule.
3. Verify that the guessed connectivity represents L_c correctly, otherwise fix it.
4. Solve a wrong connectivity (next slide).
5. Perform a structure optimisation. The model uses the MMFF94 forcefield[17] with an energy convergence criteria of 10^{-6} and 4000 steps.
6. Return the carbon atom coordinates to L_c and $\mathscr{V}_I$.

Structure Optimisation

The connectivity can be fixed looking for the hybridisation of carbon atoms with the following logic:

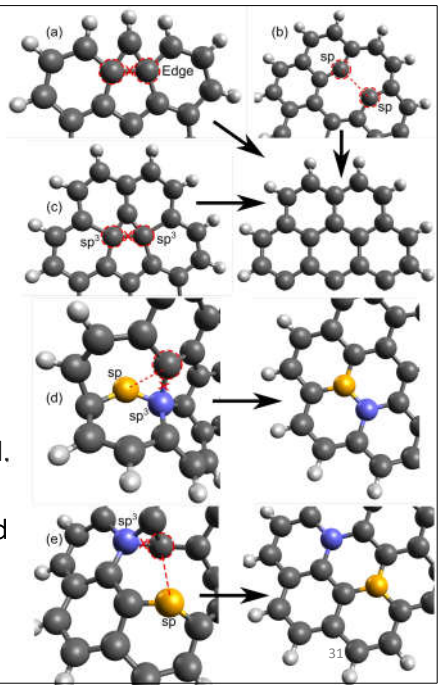
a) Remove bonds from edge atoms not defined in L_c.

b) Remove the likely additional bond from sp^3 atoms.

c) Add the likely missing bond from two sp atoms.

d) sp3 atoms bonded to sp atoms are probably wrong. Look for the closest neighbour and exchange a bond.

e) sp3 atoms nearby sp atoms are probably wrong. Look for the closest neighbour (of the sp3 atom) and exchange a bond.



Process Rates

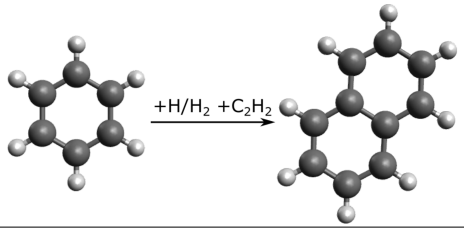
Typically[18], the process rate for a process i can be calculated as:

$$R_i = k_i f_i N_{\mathrm{Site},\eta} C_{\mathrm{A_{gas}}}$$

$$k_{\mathrm{i}} = A_i T^{n_i} \exp\left(-\frac{E_{\mathrm{A}_i}}{RT}\right)$$

f_i is the radical site fraction

$N_{\mathrm{Site},\eta}$ is the number of sites of type η, that process i affects

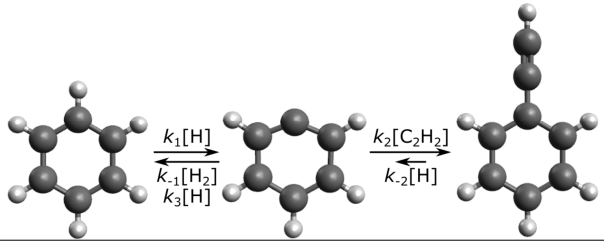


Process Rates

The radical site fraction, f_i, can be estimated using the steady-state approximation.

$$f_i = \frac{k_1 C_{\mathrm{H}}}{k_{-1} C_{\mathrm{H_2}} + k_2 C_{\mathrm{C_2H_2}} + k_3 C_{\mathrm{H}}}$$

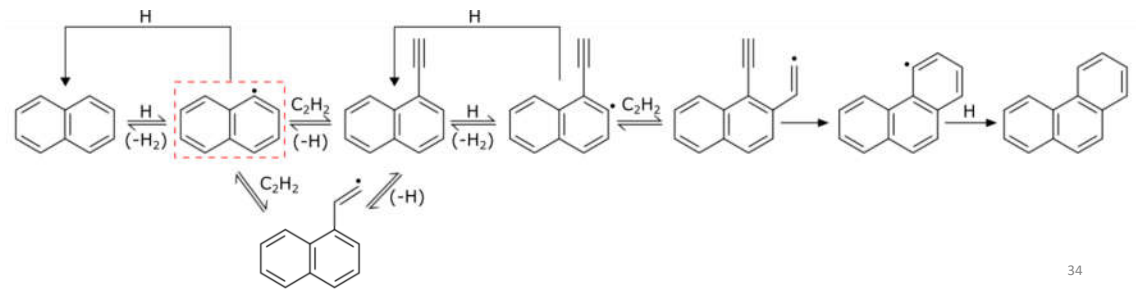
Assumes $k_{-2} C_{\mathrm{H}} \ll k_1 C_{\mathrm{H}}$ and $k_{-1} C_{\mathrm{H_2}} + k_2 C_{\mathrm{C_2H_2}} + k_3 C_{\mathrm{H}} \gg k_1 C_{\mathrm{H}} + k_{-2} C_{\mathrm{H}}$



Process Rates

However, several additional steps contribute to the growth of PAHs.

Assuming f_i for species in the red box assumes that acetylene addition is the rate-limiting step. <u>Not always the best approximation</u>.



Process Rates

Another approximation[19]:

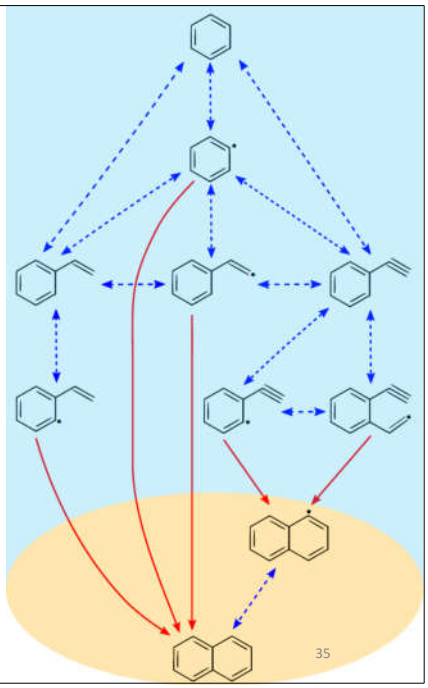
Assume <u>all</u> species on the right follow a steady-state (ss) approximation with <u>red arrows</u> used to calculate process rate.

$$M_{ss} c_{ss} = b_{ss}$$

c_{ss}: vector of ss radical species

M_{ss}: matrix of reactions that produce ss species from other ss species

b_{ss}: vector of reactions that produce ss species from non-ss species



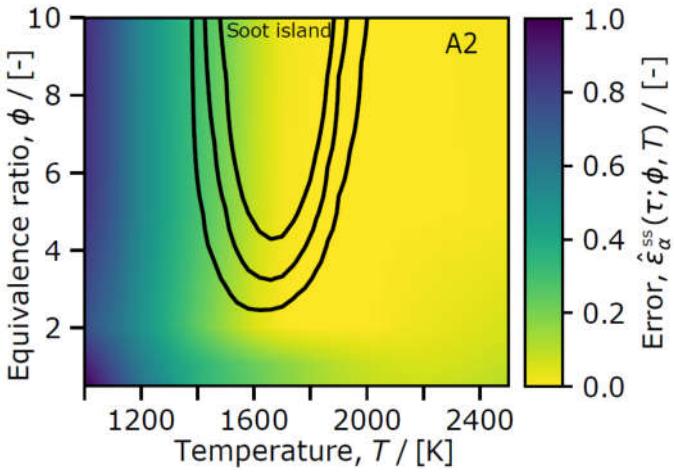
Steady-state approximation

Good at high temperatures

Good at long residence times

(not enough build-up of radical species)

Significant errors on regions where soot processes are important[19].



Process Rates

Yet another approximation:

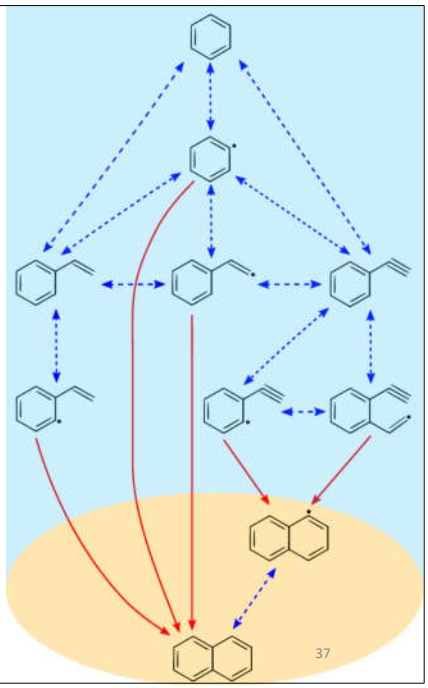
Assume species in <u>blue</u> follow a partial-equilibrium (peq) approximation with <u>red arrows</u> used to calculate process rate.

$$\mathrm{M_{peq}c_{peq}} = \mathrm{b_{peq}}$$

$\mathrm{c_{peq}}$: vector of peq radical species

$\mathrm{M_{peq}}$: matrix of reactions that produce peq species from other peq species

$\mathrm{b_{peq}}$: vector of reactions that produce peq species from non-peq species



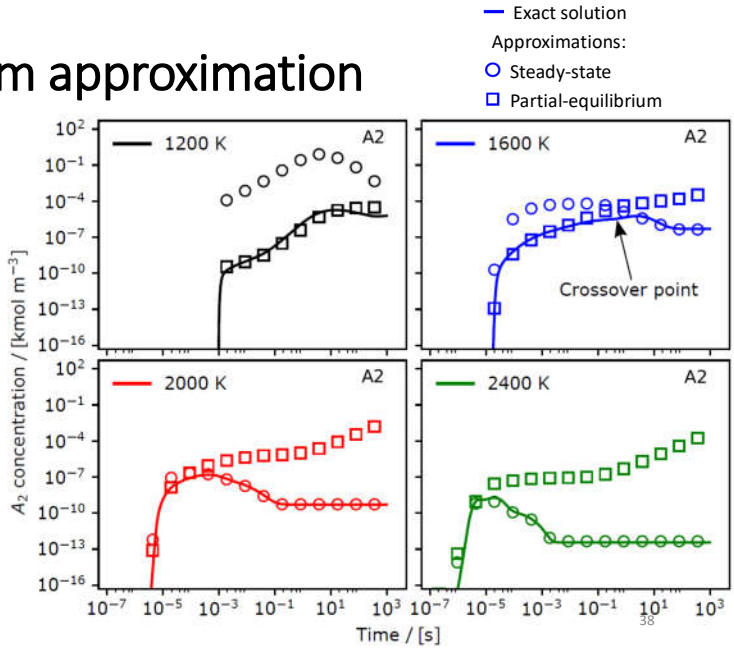
Partial-equilibrium approximation

Good at low temperatures

Good at short residence times

Overestimates process rates at high temperatures.

Crossover point: Happens where the reverse rate becomes as large as the partial-equilibrium rate. Steady-state approximation becomes more accurate.

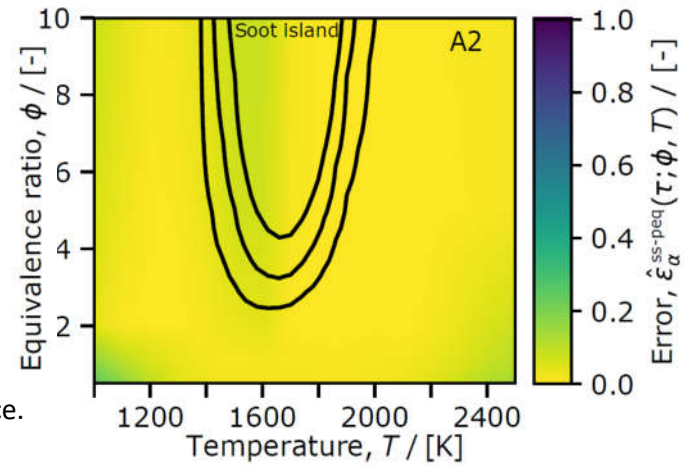


Combined SS-PEQ approximation[19]

1. Calculate $R^{\mathrm{ss}}_{\alpha\to\beta}$ and $R^{\mathrm{peq}}_{\alpha\to\beta}$
2. Calculate $R^{\mathrm{ss-peq}}_{\alpha\to\beta}$

$$R^{\mathrm{ss-peq}}_{\alpha\to\beta} = \begin{cases} R^{\mathrm{peq}}_{\alpha\to\beta} & \text{if } R^{\mathrm{peq}}_{\alpha\to\beta} > R^{\mathrm{ss}}_{\beta\to\alpha}, \\ R^{\mathrm{ss}}_{\alpha\to\beta} & \text{otherwise,} \end{cases}$$

Good performance in full T/ϕ space.



Molecular Processes

The model[13] uses 227 individual reaction rate coefficients (including and reverse rate coefficients).

Table 1: Elementary reaction rate coefficients

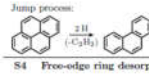
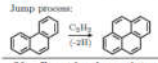
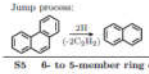
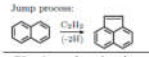
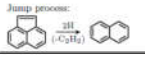
No.	Reactions	$k = AT^n \exp(-E_A/R_{gas}T)$ * A	n	E_A	References
	Hydrogen abstraction from six-member rings				
1	$C_xR_6{-}H + H \rightarrow C_xR_6^{\bullet} + H_2$	4.570×10^{08}	1.880	14.839	[16]
-1	$C_xR_6^{\bullet} + H_2 \rightarrow C_xR_6 + H$	1.690×10^{04}	2.620	4.559	[16]
2	$C_xR_6{-}H + OH \rightarrow C_xR_6^{\bullet} + H_2O$	5.190×10^{03}	3.040	3.675	[7]
-2	$C_xR_6^{\bullet} + H_2O \rightarrow C_xR_6 + OH$	5.590×10^{00}	3.573	8.659	[7]
3	$C_xR_6^{\bullet} + H \rightarrow C_xR_6{-}H$	4.170×10^{13}	0.150		[5]
	Hydrogen abstraction from five-member rings				
4	$C_xR_5{-}H + H \rightarrow C_xR_5^{\bullet} + H_2$	4.890×10^{09}	1.508	19.862	[6]
-4	$C_xR_5^{\bullet} + H_2 \rightarrow C_xR_5 + H$	5.068×10^{04}	2.445	4.520	[6]
5	$C_xR_5{-}H + OH \rightarrow C_xR_5^{\bullet} + H_2O$	5.190×10^{03}	3.040	3.675	[7]
-5	$C_x^{\bullet} + H_2O \rightarrow C_x + OH$	5.590×10^{00}	3.573	8.659	[7]
6	$C_xR_5^{\bullet} + H \rightarrow C_xR_5{-}H$	6.080×10^{12}	0.270		[22]
	Hydrogen addition to five-member rings				
7	$C_xR_5H{-}C_xR_5H + H \rightarrow C_xR_5H_2{-}C_xR_5H^{\bullet}$	5.400×10^{11}	0.450	1.820	[22]
-7	$C_xR_5H_2{-}C_xR_5H^{\bullet} \rightarrow C_xR_5H{-}C_xR_5H + H$	3.015×10^{11}	0.450	−33.367	[22]
8	$C_xR_5H_2{-}C_xR_5H^{\bullet} + H \rightarrow C_xR_5H{-}C_xR_5H + H_2$	2.000×10^{12}			[22]
	Armchair growth				
9	$C_xR_6^{\bullet} + C_2H_2 \rightarrow C_xR_6{-}R_6 + H$	1.190×10^{22}	−2.450	18.890	[4]
10	$C_xR_6^{\bullet} + C_2H_2 \rightarrow C_xR_6{-}R_6 + H$	1.060×10^{14}	−0.490	8.204	[4]
11	$C_xR_6^{\bullet} + C_2H_2 \rightarrow C_xR_6{-}C_2H + H$	4.240×10^{14}	0.025	33.080	[4]
12	$C_xR_6^{\bullet} + C_2H_2 \rightarrow C_xR_6{-}C_2H + H$	7.640×10^{-2}	3.950	16.495	[4]
	Free-edge desorption to produce an armchair				
-9	$C_xR_6{-}R_6 + H \rightarrow C_xR_6^{\bullet} + C_2H_2$	5.465×10^{30}	−3.657	86.240	[4, 11]
-10	$C_xR_6{-}R_6 + H \rightarrow C_xR_6^{\bullet} + C_2H_2$	4.868×10^{22}	−1.697	75.550	[4, 11]
	Free-edge ring growth and desorption				
13	$C_xR_6^{\bullet} + C_2H_2 \rightarrow C_xR_6{-}C_2H_2$	1.910×10^{61}	−14.600	28.610	[9]
-13	$C_xR_6{-}C_2H_2 \rightarrow C_xR_6^{\bullet} + C_2H_2$	2.499×10^{69}	−16.430	71.290	[9]
14	$C_xR_6^{\bullet} + C_2H_2 \rightarrow C_xR_6{-}C_2H + H$	1.100×10^{31}	−4.830	26.620	[9]
-14	$C_xR_6{-}C_2H + H \rightarrow C_xR_6^{\bullet} + C_2H_2$	2.542×10^{37}	−6.213	37.610	[9]
15	$C_xR_6^{\bullet} + C_2H_2 \rightarrow C_xR_6{-}C_2H_3^{\bullet}$	1.360×10^{75}	−18.400	40.880	[9]
-15	$C_xR_6{-}C_2H_3^{\bullet} \rightarrow C_xR_6^{\bullet} + C_2H_2$	4.055×10^{82}	−20.120	79.400	[9]
16	$C_xR_6^{\bullet} + C_2H_3 \rightarrow C_xR_6{-}C_2H_3$	6.000×10^{12}			[11]
-16	$C_xR_6{-}C_2H_3 \rightarrow C_xR_6^{\bullet} + C_2H_2$	8.216×10^{23}	−2.162	119.100	[11]
17	$C_xR_6^{\bullet} + C_2H_4 \rightarrow C_xR_6 + C_2H_3$	9.450×10^{-3}	4.470	4.472	[11]
-17	$C_xR_6 + C_2H_3 \rightarrow C_xR_6^{\bullet} + C_2H_4$	2.316×10^{-2}	4.416	6.709	[11]
18	$C_xR_6^{\bullet} + C_4H_4 \rightarrow C_xR_6{-}R_6 + H$	1.260×10^{04}	2.610	1.434	[11]
-18	$C_xR_6{-}R_6 + H \rightarrow C_xR_6^{\bullet} + C_4H_4$	1.130×10^{16}	0.754	66.940	[11]
19	$C_xR_6 + C_2H_3 \rightarrow C_xR_6{-}C_2H_3 + H$	1.870×10^{97}	1.470	5.533	[11]
-19	$C_xR_6{-}C_2H_3 + H \rightarrow C_xR_6 + C_2H_3$	2.042×10^{14}	−0.221	10.410	[11]
20	$C_xR_6{-}C_2H_3 \rightarrow C_xR_6{-}C_2H_2 + H$	3.010×10^{14}	0.340	111.255	[11]
-20	$C_xR_6{-}C_2H_2 + H \rightarrow C_xR_6{-}C_2H_3$	2.184×10^{11}	0.722		[11]
21	$C_xR_6{-}C_2H_3 + H \rightarrow C_xR_6{-}C_2H_2 + H_2$	6.350×10^{04}	2.750	11.649	[11]
-21	$C_xR_6{-}C_2H_2 + H_2 \rightarrow C_xR_6{-}C_2H_3 + H$	2.509×10^{01}	3.375	2.604	[11]
22	$C_xR_6{-}C_2H_3 + OH \rightarrow C_xR_6{-}C_2H_2 + H_2O$	6.550×10^{-2}	4.200	−0.860	[11]
-22	$C_xR_6{-}C_2H_2 + H_2O \rightarrow C_xR_6{-}C_2H_3 + OH$	6.705×10^{-4}	4.613	6.162	[11]

Molecular Processes

The coefficients are used to calculate the process rate for 30 molecular processes.

Processes include HACA growth and desorption on different sites, bay closures (dehydrogenation), migration of five-member rings[20], formation of seven-member rings[21] and oxidation of five-, six- and seven-member rings[22].

Table 2: Kinetic Monte Carlo jump processes.

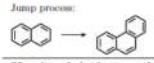
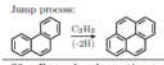
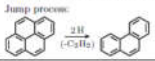
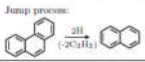
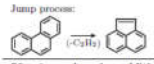
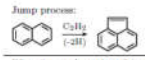
Process [Reference]	Parent site
S1 Free-edge ring growth [1]	**Free-edge (FE)**
Jump process: 	Rate: * $\big(k_{18}[C_xR_6^{\bullet}][C_2H_2] + k_{28}[C_xR_6{-}C_2H_2^{\bullet}][C_2H_2] + k_{38}[C_xR_6{-}C_2H][C_2H_2] + k_{39}[C_xR_6{-}C_2H^{\bullet}][C_2H_2] + k_{40}[C_xR_6{-}C_2H^{\bullet}][C_2H_2] + k_{41}[C_xR_6{-}C_2H^{\bullet}][C_2H_2] + k_{42}[C_xR_6{-}C_2H^{\bullet}][C_2H_2] + k_{43}[C_xR_6{-}C_2H_2][C_2H_2] + k_{44}[C_xR_6(C_2H)(C_2H_2)][H]\big)[C_{FE}]$
S2 Armchair ring growth [1]	**Armchair (AC)**
Jump process: 	Rate: $(k_9 + k_{10})\left(\frac{k_2[H]+k_3[OH]}{k_{-1}[H_2]+k_2[H]+k_{-2}[H_2O]+k_3+k_{10}+k_{11}+k_{12}[C_2H_2]}\right)[C_2H_2][C_{AC}]$
S3 Free-edge desorption to an armchair [4, 11]	**Free-edge adjacent to non-free-edges (FE_{HACA})**
Jump process:	Rate: $(k_{-9} + k_{-10})[C_{FE_{HACA}}]$
S4 Free-edge ring desorption [1]	**Free-edge with two adjacent free-edges (FE3)**
Jump process: 	Rate: $\big(k_{-18} + k_{-24} + k_{-28} + k_{-33} + k_{-34} + k_{-38}\big)[H][C_{FE3}] + \left(\frac{(k_2[H]+k_3[OH])(k_{-39}+k_{-40}+k_{-41})}{k_{-1}[H_2]+k_2[H]+k_{-2}[H_2O]}\right)[C_{FE3}]$
S5 6- to 5-member ring conversion at armchair [1]	**Armchair next to FE3 (AC_{FE3})**
Jump process: 	Rate: $k_{57}\left(\frac{k_2[H]+k_3[OH]}{k_{-1}[H_2]+k_2[H]+k_{-2}[H_2O]+k_{57}}\right)[C_{AC_{FE3}}]$
S6 5-member ring addition [1]	**Zig-zag (ZZ)**
Jump process: 	Rate: $(k_{64} + k_{65})\left(\frac{k_2[H]+k_3[OH]}{k_{-1}[H_2]+k_2[H]+k_{-2}[H_2O]+k_{64}+k_{65}+k_{66}+k_{67}[C_2H_2]}\right)[C_2H_2][C_{ZZ}]$
S7 5-member ring desorption [1]	**5-member ring (R5)**
Jump process: 	Rate: $\left(\frac{k_{68}(k_2[H]+k_3[OH])}{k_{-1}[H_2]+k_2[H]+k_{-2}[H_2O]+k_{68}} + \frac{k_{69}[H](k_{69}+k_{70})}{k_{-5}[H]+k_{69}[H]+k_{70}[C_2H_2]}\right)[C_{R5}]$

Molecular Processes

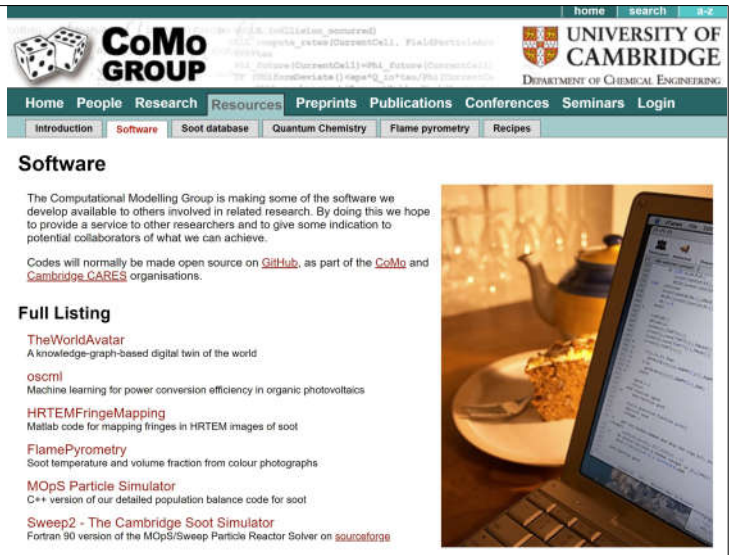
Each molecular process includes the addition or removal of carbon atoms, hydrogen atoms, number of rings and sites.

The model also includes the transformations to existing sites and rules for structure optimisation.

Table 3: Transformations for each Kinetic Monte Carlo jump process.

Process	Parent site	Transformation: Atoms / Rings	Transformation: Sites	Structure optimisation
S1 Free-edge ring growth				
Jump process: 	Free-edge (FE)	+4 C +2 H +1 R6	parent site ← FE +2 Free-edges +1 C to adjacent sites	-
S2 Armchair ring growth				
Jump process: 	Armchair (AC)	+2 C +1 R6	parent site ← FE +1 C to adjacent sites	Before: If $d_{site} < 2.6$ Å After: If $d_{i,i+1} > 1.7$ Å
S3 Free-edge desorption to an armchair				
Jump process: 	Free-edge (FE_{HACA}) not next to RFE[a]	-2 C -1 R6[b]	parent site ← AC[b] -1 C to adjacent sites	Before: If $d_{i,i+1} < 1.2$ Å If $d_{i,i+1} > 1.6$ Å
S4 Free-edge ring desorption				
Jump process: 	Three adj. free-edges (FE3) non-adjacent to RFE, eR5, eR5[a]	-4 C -2 H -1 R6	parent site ← FE -2 Free-edges -1 C to adjacent sites	-
S5 6- to 5-member ring conversion at armchair				
Jump process: 	Armchair next to FE3 (ACyra) No R5 or eR5 to other side[a] No crosslink	-2 C -2 H -1 R6 +1 R5	parent site ← R5 -3 Free-edges +R5 to adjacent sites -1 C to adjacent sites	After: Every time
S6 5-member ring addition				
Jump process: 	Zig-zag (ZZ)	+2 C +1 R5	parent site ← R5 +1 C to adjacent sites +R5 to adjacent sites	Before: If $d_{site} < 2.1$ Å After: If $d_{i,i+1} > 1.8$ Å
S7 5-member ring desorption				
Jump process: 	5-member ring (R5)	-2 C -1 R5	parent site ← ZZ -1 C to adjacent sites -R5 to adjacent sites	-

Software



MOpS Particle Simulator
C++ version of our detailed population balance code for soot

Model insights

$T = 1500\ \mathrm{C}$

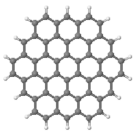
$x_{\mathrm{H}} = 10^{-2}$

$x_{\mathrm{H_2}} = 10^{-1}$

$x_{\mathrm{C_2H_2}} = 10^{-1}$

$x_{\mathrm{O_2}} = 10^{-4}$

$x_{\mathrm{O}} = 10^{-2}$



0.0 ms

Jmol

$T = 1500\ \mathrm{C}$

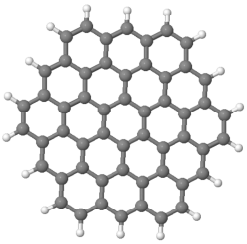
$x_{\mathrm{H}} = 10^{-2}$

$x_{\mathrm{H_2}} = 10^{-1}$

$x_{\mathrm{C_2H_2}} = 10^{-1}$

$x_{\mathrm{O_2}} = 10^{-3}$

$x_{\mathrm{O}} = 10^{-3}$



0.0 ms

$T = 1500\ \mathrm{C}$

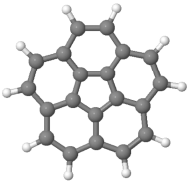
$x_{\mathrm{H}} = 10^{-2}$

$x_{\mathrm{H_2}} = 10^{-1}$

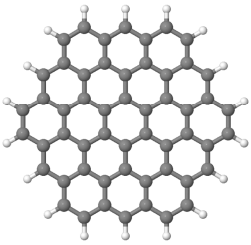
$x_{\mathrm{C_2H_2}} = 10^{-1}$

$x_{\mathrm{O_2}} = 10^{-1}$

$x_{\mathrm{O}} = 10^{-8}$



0.0 ms



0.0 ms

$T = 1500\ \mathrm{C}$

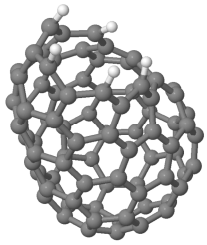
$x_{\mathrm{H}} = 10^{-2}$

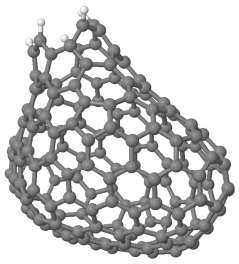
$x_{\mathrm{H_2}} = 10^{-1}$

$x_{\mathrm{C_2H_2}} = 10^{-1}$

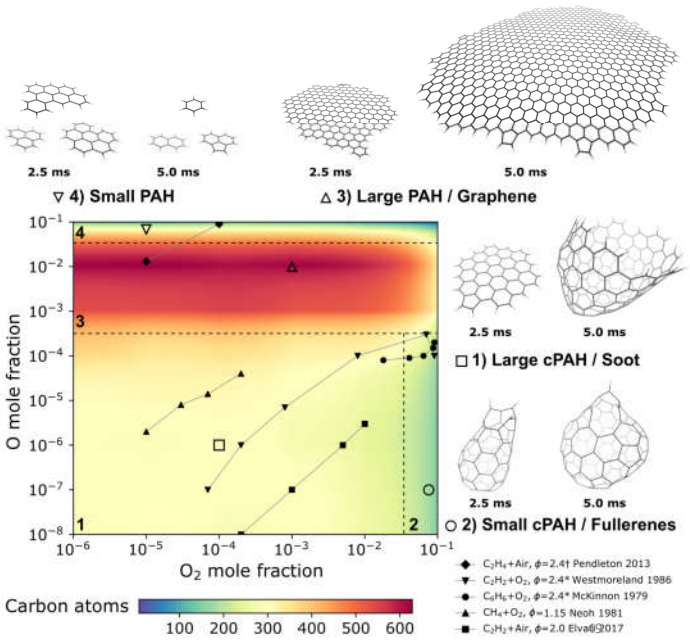
$x_{\mathrm{O_2}} = 10^{-1}$

$x_{\mathrm{O}} = 10^{-8}$



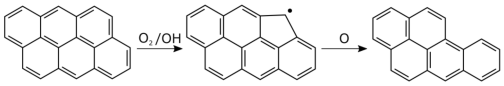


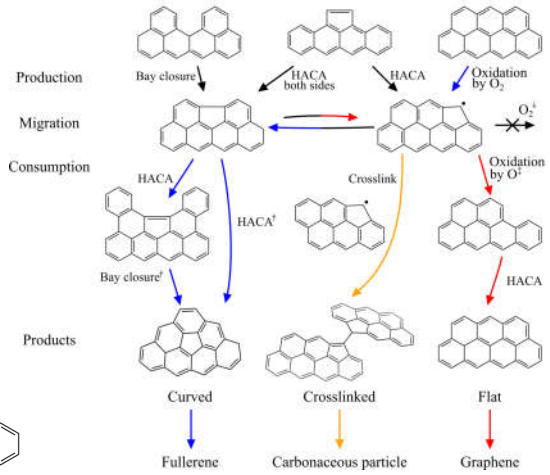
Regions where oxygen produces different products: Graphene, fullerenes, and soot[44].



Molecular and atomic oxygen remove carbon atoms from different sites: Molecular oxygen produces five-member rings and atomic oxygen removes them[22].

Different products can be obtained by combining growth and oxidation[44].





Temperature dependance

$x_{\mathrm{H}} = 10^{-2}$

$x_{\mathrm{H_2}} = 10^{-1}$

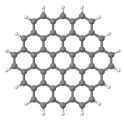
$x_{\mathrm{C_2H_2}} = 10^{-1}$

$x_{\mathrm{O_2}} = 10^{-4}$

$x_{\mathrm{O}} = 10^{-2}$

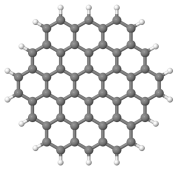
0.0 ms

$T = 1500$ C



0.0 ms

$T = 2000$ C



0.0 ms

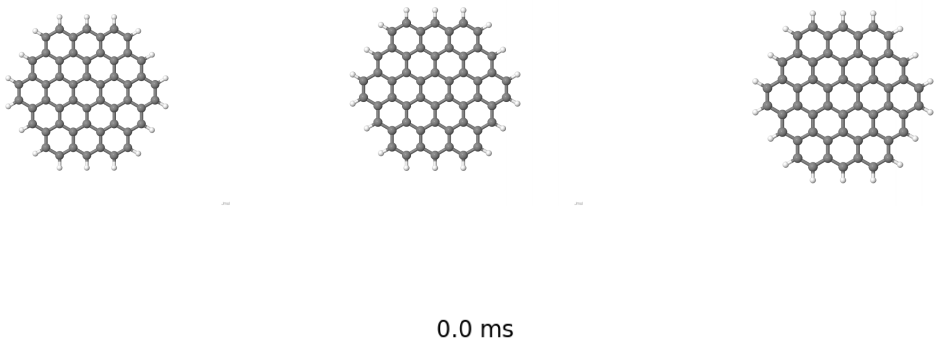
$T = 2500$ C

$x_{\mathrm{H}} = 10^{-2}$

$x_{\mathrm{H_2}} = 10^{-1}$

$x_{\mathrm{C_2H_2}} = 10^{-1}$

Temperature dependance – no oxygen



$T = 1500$ C $\quad$ $T = 2000$ C $\quad$ $T = 2500$ C

$T = 1500$ C

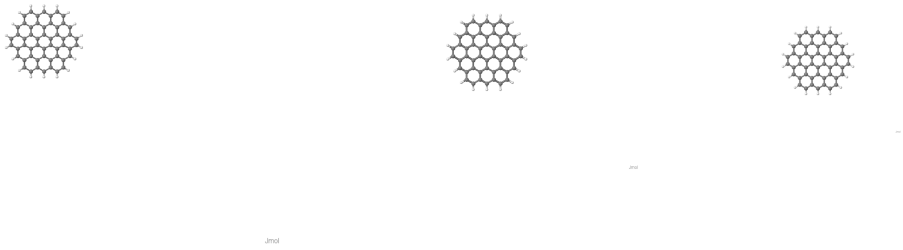
$x_{\mathrm{H}} = 10^{-2}$

$x_{\mathrm{H_2}} = 10^{-1}$

$x_{\mathrm{O_2}} = 10^{-4}$

$x_{\mathrm{O}} = 10^{-2}$

C_2H_2 dependance



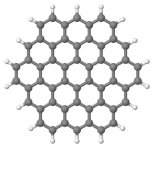
$x_{\mathrm{C_2H_2}} = 0.1$ $\quad$ $x_{\mathrm{C_2H_2}} = 0.075$ $\quad$ $x_{\mathrm{C_2H_2}} = 0.05$

$T = 1500\ \mathrm{C}$

$x_{\mathrm{H}_2} = 10^{-1}$

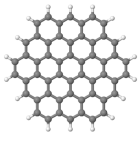
$x_{\mathrm{C}_2\mathrm{H}_2} = 0.1$

Hydrogen dependance



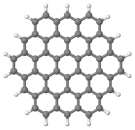
0.0 ms

$x_{\mathrm{H}} = 0.1$



0.0 ms

$x_{\mathrm{H}} = 0.01$



0.0 ms

$x_{\mathrm{H}} = 0.001$

End of Lecture 4

Soot – Part 5

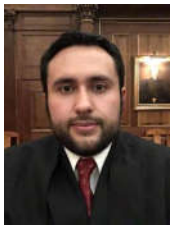
Markus Kraft

Computational Modelling Group Cambridge

Main Contributors:
Dr Jake Martin (Part 1)
Dr Angiras Menon (Part 2)
Dr Laura Pascazio (Part 3)
Dr Gustavo Leon (Part 4)

Jacob Martin

AngirasMenon

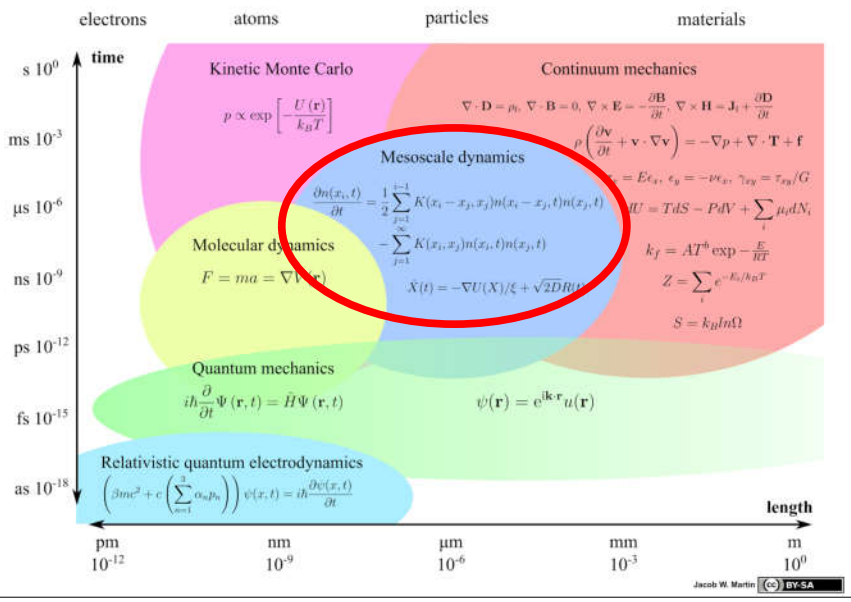


Gustavo Leon

Laura Pascazio

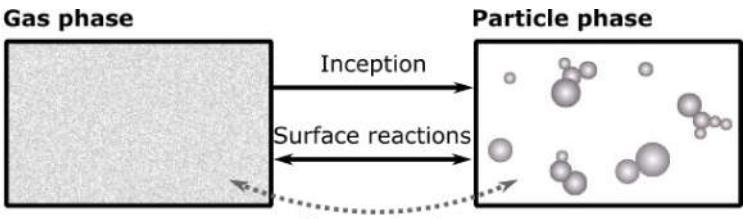
Part 1 Overview
Part 2 Quantum Chemistry
Part 3 Molecular Dynamics
Part 4 Kinetic Monte Carlo
Part **5 Stochastic Particle Methods**
Part 6 Application – engine model

Scale Diagramme



Models for aerosol systems

Phases in aerosol systems



Characterised by the concentration of chemical species, $\mathbf{C}$.

Characterised by the number density (concentration) of particles, n.

Gas Phase

Contains N_{sp} chemical species such that, $\mathbf{C} = \left(C_1, \dots, C_{N_{\mathrm{sp}}}\right)$, represents their concentrations, and has a Temperature, T.

The concentration of each species k changes in time, t, and space, ω, due to reactions, particle processes and flow[9],

$$\frac{\delta C_k}{\delta t} + \frac{\delta}{\delta \omega} \cdot \underbrace{\left[u_{\mathrm{g}}(\mathbf{C}, T) C_k\right.}_{\text{Advective flow}} \underbrace{\left. - D_k(\mathbf{C}, T) \frac{\delta C_k}{\delta \omega}\right]}_{\text{Diffusive flow}} = \underbrace{\mathcal{W}(\mathbf{C}, T) + \mathcal{G}(\mathbf{C}, T, n)}_{\text{Particle processes}}$$

Particle Phase

Formed by collisions between gas phase (precursor) species.	**Inception**
Interacts with the gas-phase via surface chemical reactions.	**Surface processes**
Collisions between particles result in interconnected aggregates.	**Coagulation**
Surface reactions produce more rounded (spherical) structures.	**Sintering and coalescence**

Governed by a *Population Balance Equation* (PBE).

Number density

For a particle of type $x \in \mathcal{E}$, at position $\omega \in \Omega$, the number density (concentration) of particles is $n(x, \omega)$.

The number density evolves as particles experience different particle processes.

Can be used to calculate *integral* properties of the particle ensemble (particle size distribution).

Population Balance Equation

$$\frac{\delta n}{\delta t} + \frac{\delta}{\delta x} \cdot \underbrace{[S_x(\mathbf{C}, T, x)n]}_{\text{Surface reactions}} + \frac{\delta}{\delta \omega} \cdot \underbrace{\left[u_{\mathrm{P}}(\mathbf{C}, T, x)n - D_{\mathrm{P}}(\mathbf{C}, T, x)\frac{\delta n}{\delta \omega}\right]}_{\text{Particle flow}} =$$

$$\underbrace{I(\mathbf{C}, T, x)}_{\text{Inception}} + \underbrace{\int_{\mathcal{E}} K(T, x-y, y)n(x-y)n(y)dy - \int_{\mathcal{E}} K(T, x, y)n(x)n(y)dy}_{\text{Coagulation}}$$


Casper Lindberg Astrid Boje

Particle models

A detailed particle model for polydisperse aggregate particles Casper Lindberg, Manoel Y. Manuputty, Edward K. Y. Yapp, Jethro Akroyd, Rong Xu, and Markus Kraft, *Journal of Computational Physics* **397**, 108799, (2019).

Stochastic population balance methods for detailed modelling of flame-made aerosol particles Astrid Boje and Markus Kraft, *Journal of Aerosol Science* **159**, 105895, (2022).

Type space

Space used for mathematically representing particles.

A general particle, P_q, can have multivariate states with different coordinates[9]:

$$P_q \begin{cases} \text{spatial (external) coordinates} & \text{physical space,} \quad \omega \in \Omega \\ \text{property (internal) coordinates} & \text{type space,} \quad x \in \mathcal{E} \end{cases}$$

An ensemble containing N_P particles has P_q, with $q \in \{1, \dots, N_\mathrm{P}\}$.

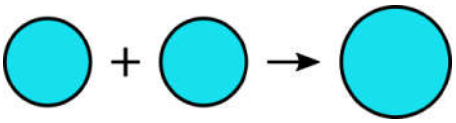
Low dimensional models

Coalescent sphere

Particles are represented as spheres of constant composition and density.
One internal coordinate: Mass[9].

$$P_q = P_q(\eta_q)$$

$$d_p(P_q) = \left(\frac{6}{\pi}\frac{m(P_q)}{\rho}\right)^{\frac{1}{3}}$$



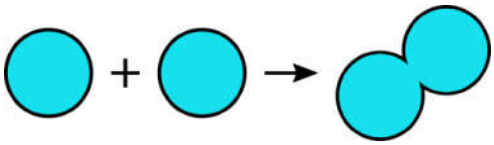
Implies instantaneous coalescence after coagulation events.

Low dimensional models

Surface-volume models

Two internal coordinates: Mass and Surface area.
Surface area is conserved after coagulation events.

$$P_q = P_q(\eta_q, A_q)$$



Fractal dimension needs to be assumed. Typically, ~ 1.8.
Surface processes affect the total area of the particle.

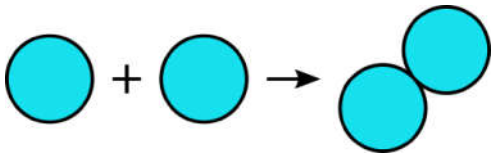
Detailed particle models

A *particle aggregate*, P_q, is defined by its constituent *primary particles*, p_i, where i ∈ $1, \dots, n_q$, and their connectivity $\mathbf{C}_q$, with elements C_{ij}.

$$P_q = P_q\left(p_1, \dots, p_{n_q}, \mathbf{C}_q\right)$$

$$C_{ij} = \begin{cases} 1 & \text{if } p_i, p_j \text{ are adjacent} \\ 0 & \text{if } p_i, p_j \text{ are not adjacent} \end{cases}$$

$$p_i = p_i(v_i)$$



Connectivity is retained but fractal dimension is still assumed.

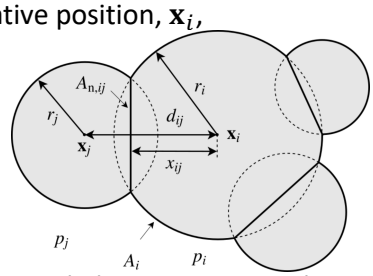
These *overlapping-sphere* models have been successfully used to study aggregation with surface growth in soot[10] and silica aerosols[11].

Detailed particle models

Detailed particle model by Lindberg et al.[12] describes particles by their chemical composition, η_i, radius, r_i, and relative position, $\mathbf{x}_i$,

$$P_q = P_q\left(p_1, \dots, p_{n_q}, \mathbf{C}_q\right)$$

$$p_i = p_i(\eta_i, r_i, \mathbf{x}_i)$$

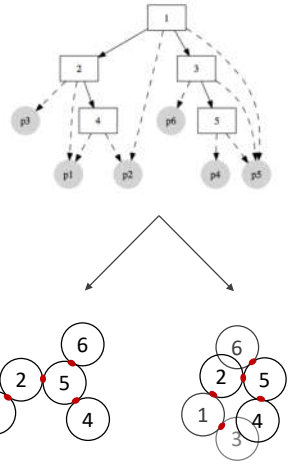


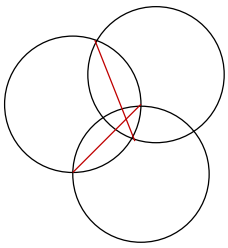
Centre-to-centre separation, d_{ij}, or centre-to-neck distance, x_{ij}, can be obtained from geometry.

No need to assume fractal dimension.

Key model assumptions

- Binary tree structure determines connectivity.
- Necks are circular *i.e.* no double overlaps.





Derived properties

Centre to neck distance $$x_{ij} = \frac{d_{ij}^2 - r_j^2 + r_i^2}{2d_{ij}}$$

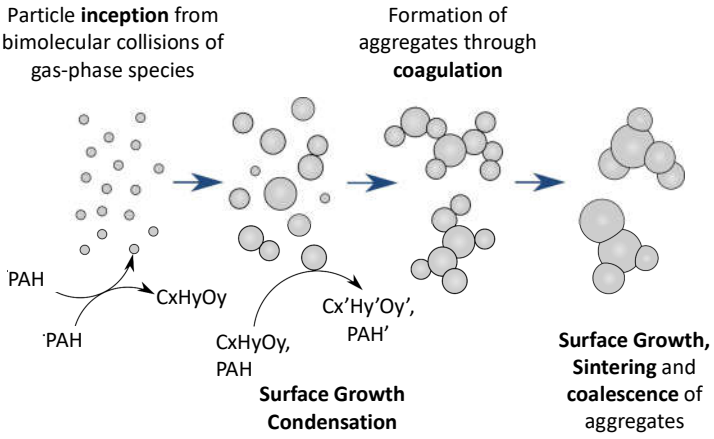
Primary volume $$v_i = V_{\text{sphere}}(r_i) - \sum_j V_{\text{cap}}\left(r_i, x_{ij}\right)$$

$$v_i = \frac{4}{3}\pi r_i^3 - \frac{1}{3}\pi \sum_j \left(2r_i^3 + x_{ij}^3 - 3r_i^2 x_{ij}\right)$$

Free surface area $$A_i = \frac{\partial v_i}{\partial r_i} = 4\pi r_i^2 - 2\pi \sum_j \left(r_i^2 - r_i x_{ij}\right)$$

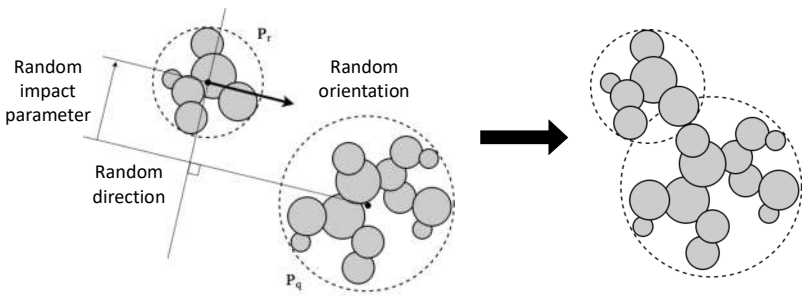
Neck area $$A_{\text{n},ij} = \frac{\partial v_i}{\partial x_{ij}} = \pi\left(r_i^2 - x_{ij}^2\right)$$

Particle processes



Coagulation

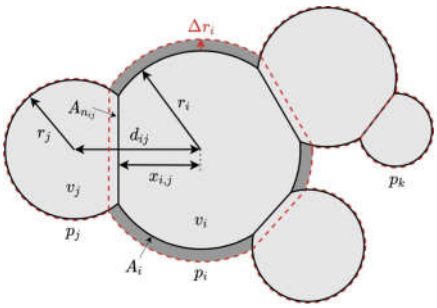
- Ballistic cluster-cluster aggregation with a random impact parameter used to determine point of contact.



- Assumes that the particle size is smaller than the mean free path.

Jullien R., **(1984),** Transparency effects in cluster-cluster aggregation with linear trajectories. *J. Phys. A: Math. Gen*, *17*, L771-L776.

Condensation/Surface growth



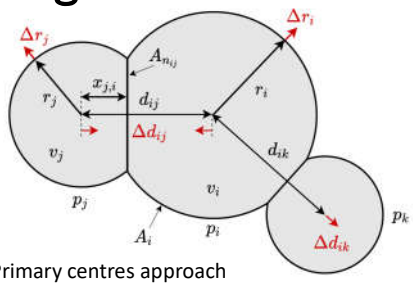
Mass is added to the free surface of a primary, A_i

The particle volume increase by $\Delta V(P_q)$

The primary radius increases:

$$\frac{dr_i}{dt} = \left(\frac{1}{A_i}\right)\frac{dV(P_q)}{dt}$$

Sintering



Primary centres approach each other due to **grain boundary diffusion:**

$$\frac{\mathrm{d}x_{j,i}}{\mathrm{d}t} = -\frac{r_{j,0}^4}{\tau_{j,0}A_{\mathrm{n}_{ij}}}\left(\frac{1}{r_j - x_{j,i}} - \frac{1}{R_{ij}}\right)$$

Conservation of mass/volume:

- Primary radius increases to account for change in overlap.
- Separation of other neighbours also increases.

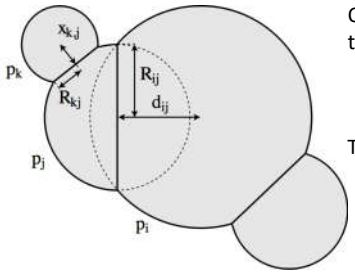
$$\frac{\mathrm{d}v_j}{\mathrm{d}t} = \frac{\partial v_j}{\partial r_j}\frac{\mathrm{d}r_j}{\mathrm{d}t} + \sum_k \frac{\partial v_j}{\partial x_{j,k}}\frac{\mathrm{d}x_{j,k}}{\mathrm{d}t} = 0$$

Characteristic sintering time:

$$\tau_{j,0} = 1.458\times10^{19} r_{j,0}^4 T \exp\left(\frac{-258\mathrm{kJmol}^{-1}}{RT}\right)\mathrm{s}$$

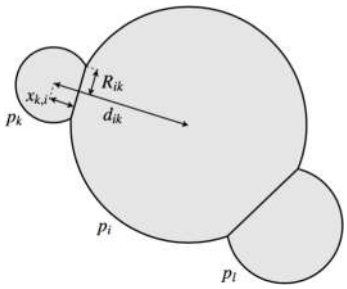
Eggersdorfer, M. L., Kadau, D., Herrmann, H. J. & Pratsinis, S. E., (**2012**). Aggregate morphology evolution by sintering: Number and diameter of primary particles. *J. Aerosol Sci.*, *46*, 7–19.

Coalescence



Once sufficiently sintered primaries are assumed to coalesce.

The smaller primary is merged into the larger primary.

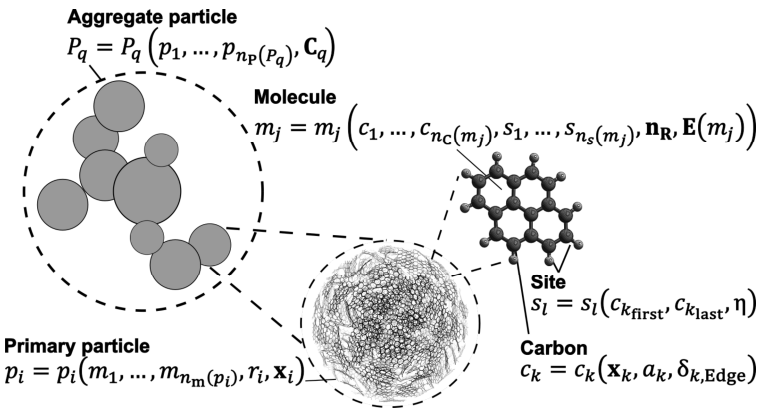


Particle connectivity is updated preserving the neck radii of neighbours.

$$R_{ik} = R_{kj}$$

Detailed particle models

Detailed particle model by Leon[13] that describes carbonaceous particles by containing the chemical species that conform them.



Two *layers* in the model[13]:

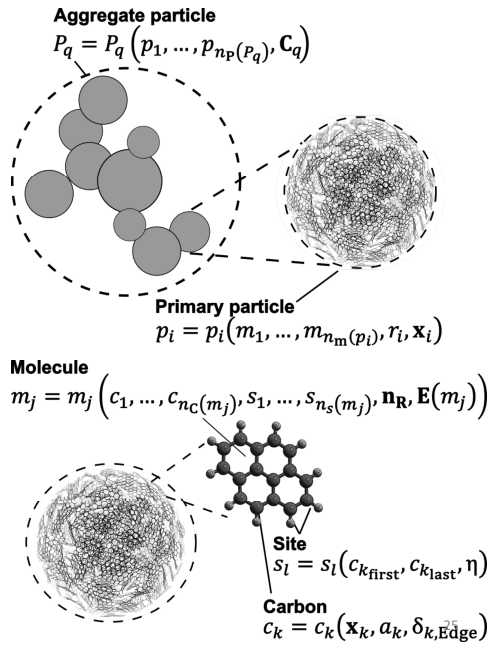
1. Particles and aggregates

   Each primary particle, p_i, contains n_m molecules, radius, r_i, and relative position, $\mathbf{x}_i$.

   Similar description to previous slides.

2. Molecules (PAHs)

   A general molecule, m_j, contains n_C carbon atoms (c_k), n_s sites (s_l), $\mathbf{n_R}$ rings, and an Edge connectivity matrix $\mathbf{E}(m_j)$.



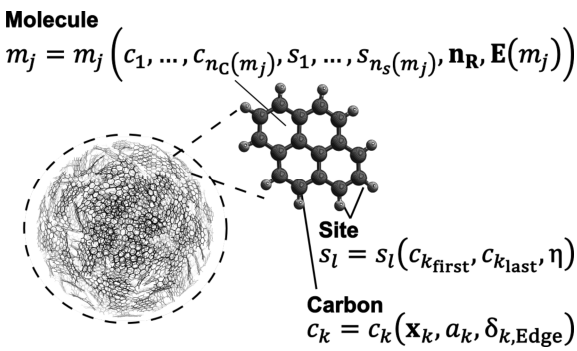
The list of carbon atoms and the list of sites can be written as:

$C(\mathrm{m_j}) = \left(c_1, \dots, c_{n_\mathrm{C}}\right)$ and

$S(\mathrm{m_j}) = \left(s_1, \dots, s_{n_\mathrm{S}}\right)$.

The number of rings in a molecule is:

$$\mathbf{n_R} = \begin{bmatrix} n_\mathrm{R6} \\ n_\mathrm{R5} \\ n_{\mathrm{R5_{emb}}} \\ n_\mathrm{R7} \end{bmatrix}$$



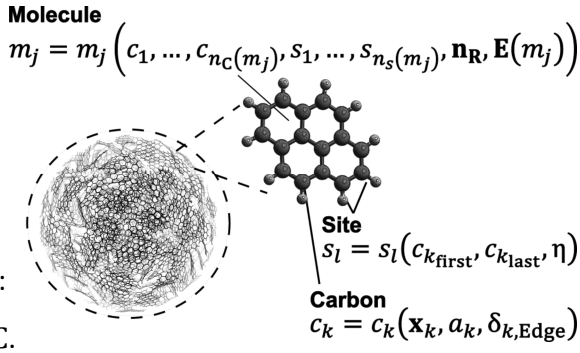
The Edge connectivity matrix $\mathbf{E}(m_j)$ has elements:

$$E_{k_1,k_2} = \begin{cases} 1, & \text{if } c_{k_1} \text{and } c_{k_2} \text{ are connected,} \\ 0, & \text{otherwise,} \end{cases}$$

where $k_1, k_2 \in I_\mathrm{Edge}$, the set of edge carbon atom indices

A carbon atom, c_k, contains its spatial coordinates, $\mathbf{x}_k$, the attached third atom, a_k, typically carbon (C) or hydrogen (H) and a binary variable $\delta_{k,\mathrm{Edge}}$:

$$\delta_{k,\mathrm{Edge}} = \begin{cases} 1, & \text{on the edge,} \\ 0, & \text{otherwise,} \end{cases}$$



A site, s_l, contains two carbons, $c_{k,\mathrm{first}}$ and $c_{k,\mathrm{last}}$ that are part of a reactive site:

$k_\mathrm{first}, k_\mathrm{last} \in I_\mathrm{Edge}$ and $a_{k_\mathrm{first}}, a_{k_\mathrm{last}} \neq \mathrm{C}$.

The site type, η, is determined by the shortest path between $c_{k,\mathrm{first}}$ and $c_{k,\mathrm{last}}$ in the graph defined by $\mathbf{E}(m_j)$ and has values $\eta \in \{\mathrm{FE}, \mathrm{ZZ}, \mathrm{AC}, \dots\}$

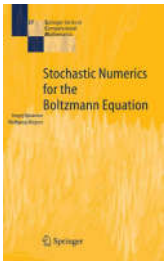
Direct Simulation Monte Carlo algorithm

Stochastically Weighted Particle Models

Andreas *Eibeck* and Wolfgang *Wagner*. SIAM J. Sci. Comput. 22-3 (2000), pp. 802-821, https://doi.org/10.1137/S1064827599353488

Sergej Rjasanow, Wolfgang Wagner, Stochastic Numerics for the Boltzmann Equation, Springer Berlin, Heidelberg, https://doi.org/10.1007/3-540-27689-0

The linear process deferment algorithm: A new technique for solving population balance equations, RIA Patterson, J Singh, M Balthasar, M Kraft, JR Norris
SIAM Journal on Scientific Computing 28 (1), 303-320,2006



Stochastic weighted particle methods for population balance equations RIA Patterson, W Wagner, M Kraft, Journal of Computational Physics 230 (19), 7456-7472

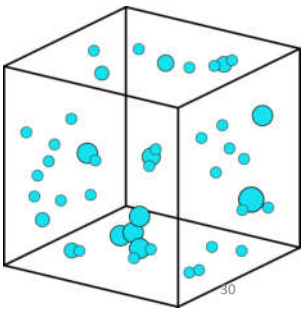
Sample volume

The particle phase is simulated using a finite ensemble of *computational particles,* P_q, $q = 1, \dots, N \leq N_{\max}$ in a sample volume, V_{smp}.

Computational particles represent the concentration of physical particles with a statistical weight, w_i.

Sample volume maps between real and simulated particles.

$$\frac{\#\text{real particles}}{m^3} = \frac{1}{V_{\text{smp}}} \sum_{i=1}^{N} w_i \,, \qquad \text{where } w_i = 1 \text{ in DSA}$$



Initialising simulation

Two parameters need to be defined to start the simulation: V_{smp} and N_{max}

However, V_{smp} needs to be scaled for changes in Temperature and Pressure.

Alternatively, selecting the maximum number density, $M_{0,\mathrm{max}}$, and N_{max} allows initialising the sample volume as:

$$V_{\mathrm{smp},0} = \frac{N_{\mathrm{max}}}{M_{0,\mathrm{max}}}$$

Stochastic numerical method

Waiting time between events, τ, is drawn from an exponential distribution specified by the total process rate:

$$R_{\mathrm{total}} = \sum_{i=1}^{N_{\mathrm{process}}} R_i$$

$$\tau \sim \exp(R_{\mathrm{total}})$$

Probability of a process j, with rate, R_j, being selected for an event is proportional to its relative rate:

$$\mathbb{P}(j) \propto \frac{R_j}{R_{\mathrm{total}}}$$

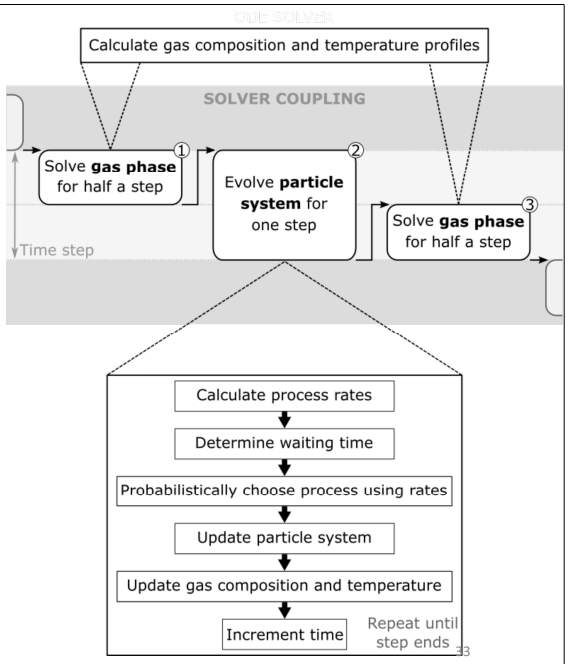
Phase coupling

Operator splitting has been widely used to solve the chemical flow equations coupled to the PBE.

Each equation is decomposed into a sum of its parts and solved separately.

Notable examples are the *Predictor-Corrector* algorithm[29], and the *Strang splitting* algorithm[30].

Accuracy increases as splitting times are decreased but computational time increases.



Phase decoupling

Another methodology to solve the PBE:

1. Solve the chemical flow equations with a simpler particle model (e.g., Method of Moments) to correct the source and sink terms for relevant chemical species.
2. Solve the detailed particle model stochastically along a Lagrangian trajectory of the chemical flow solution.

Valid only for dilute systems: Particle processes do not affect the gas phase solution significantly.

This methodology has been widely used for titania[31,32] and carbonaceous nanoparticles[33,34].

Phase decoupling

A species from the chemical flow equations is used as a *transfer species* for the stochastic model (e.g., pyrene). The number density of transfer species in the model:

$$N_{\mathrm{dens}}^{\mathrm{transf}} = \frac{N^{\mathrm{transf}}}{N\,V_{\mathrm{smp}}}$$

Number density of transfer species in the chemical flow solution:

$$N_{\mathrm{gas}}^{\mathrm{transf}} = C_{\mathrm{A}^{\mathrm{transf}}} N_{\mathrm{A}} V_{\mathrm{smp}}$$

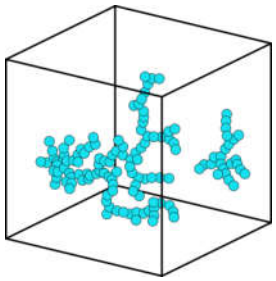
The model follows the concentration of the transfer species:

$$\mathrm{If}\begin{cases} N^{\mathrm{transf}} < N_{\mathrm{gas}}^{\mathrm{transf}}, & \text{Add } N_{\mathrm{gas}}^{\mathrm{transf}} - N^{\mathrm{transf}} \text{ transfer species particles.} \\ N^{\mathrm{transf}} > N_{\mathrm{gas}}^{\mathrm{transf}}, & \text{Remove } N^{\mathrm{transf}} - N_{\mathrm{gas}}^{\mathrm{transf}} \text{ transfer species particles.} \end{cases}$$

Simulation issues

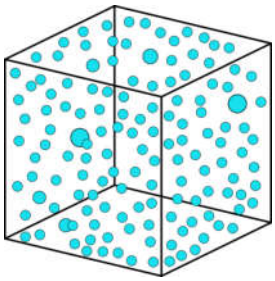
Choosing N_{max} and V_{smp} determines the total number of particles to be used in the simulation. However, two problems arise:

Ensemble depletion
(large coagulation rates)



$N \to 0$

Ensemble saturation
(large inception rates)



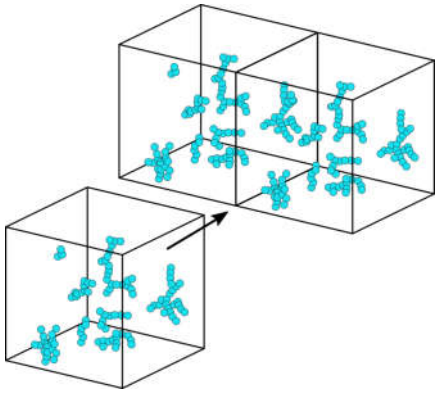
$N \to N_{\mathrm{max}}$

Ensemble doubling

$$\text{If } \frac{N}{N_{\max}} < 0.5, \qquad V_{\text{smp, New}} = 2V_{\text{smp}}, \qquad N_{\text{New}} = 2N$$

Fixes the problem of ensemble depletion while keeping constant number density[35].

Duplicating the particles does not cause issues if $N_{\max}$ is sufficiently large.



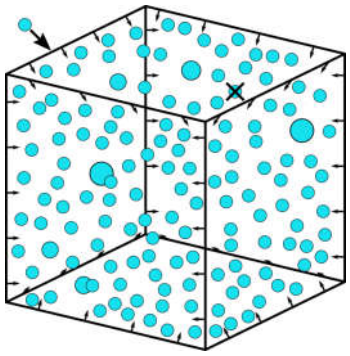
Ensemble contraction

$$\text{If } N = N_{\max}, \qquad N_{\text{New}} = N - 1, \qquad V_{\text{smp, New}} = V_{\text{smp}} \frac{N-1}{N},$$

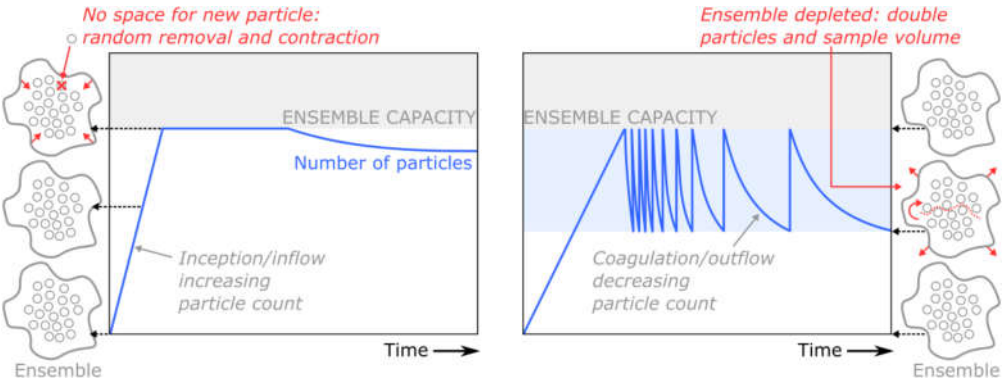
One particle is removed randomly avoiding ensemble saturation[36].

However, excessive contractions increase statistical error and reduce average size of particles in ensemble.

Better to restart simulation with smaller V_{smp} or larger $N_{\max}$ to avoid excessive contractions.



Ensemble filling



Doubling and contracting algorithms get activated the first time that $N \geq N_{\max}$[9].

Statistical accuracy

Statistical accuracy of these methods is controlled by the number of computational particles used, $N_{\max}$, (resolution of the PSD), and the number of different repeat runs, L.

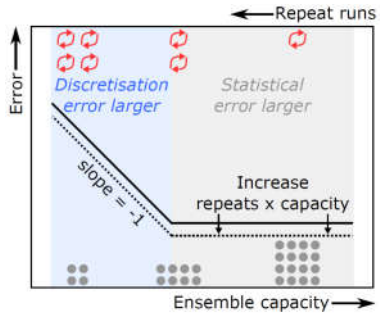
Theoretical convergence[9]. For a given function ξ:

$$\lim_{V_{\text{smp}} \to \infty} \frac{1}{V_{\text{smp}}} \sum_{i=1}^{N} w_i \, \xi(x_i) = \int \xi(x) n(x) dx$$

Numerical convergence tests

Keeping $(N_{\max} \times L)$ while increasing $N_{\max}$ and comparing against a reference solution gives the following behaviour:

Ideal capacity and repeat runs found!



Efficient Simulation Techniques

Software

MOpS Particle Simulator
C++ version of our detailed population balance code for soot

Stochastically weighted algorithm

Similar to the DSMC but the particle statistical weight, w_i, is used to conserve particle number density and other properties such as mass.

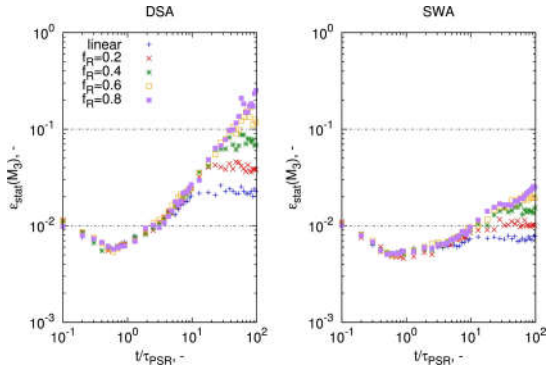
It uses *weight transfer functions* to define coagulation rates.

Special care is needed to demonstrate that the event frequency and symmetry of the real coagulation process is represented[37].

SWA Statistical accuracy

Theoretical work by Patterson et al.[38] and DeVille et al.[39] has indicated that convergence to the correct behaviour is possible for certain formulations.

Menz et al.[40] demonstrated that weighted particle methods can be used to reduce systematic errors.



Coagulation rates

The system coagulation rate, R_{coag}, scales with N^2. Remembering that:

$$K_{\mathrm{fm}}(P_i, P_j) = 2.2\sqrt{\frac{\pi k_{\mathrm{B}} T}{2}\left(\frac{1}{m(P_i)} + \frac{1}{m(P_j)}\right)}\left(d_{\mathrm{c}}(P_i) + d_{\mathrm{c}}(P_j)\right)^2, \text{then}$$

$$R_{\mathrm{coag}}^{\mathrm{fm}} = \frac{1}{2V_{\mathrm{smp}}}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\left[2.2\sqrt{\frac{\pi k_{\mathrm{B}} T}{2}\left(\frac{1}{m(P_i)} + \frac{1}{m(P_j)}\right)}\left(d_{\mathrm{c}}(P_i) + d_{\mathrm{c}}(P_j)\right)^2\right]$$

This double summation can be very expensive to compute.

Majorant kernels

The majorant technique[42] relies on finding an upper bound on the true coagulation rate that is less expensive to calculate. For example:

$$\widehat{K}_{\mathrm{fm}}(P_i, P_j) = 4.4\sqrt{\frac{\pi k_{\mathrm{B}} T}{2}}\left(\frac{1}{\sqrt{m(P_i)}} + \frac{1}{\sqrt{m(P_j)}}\right)\left(d_{\mathrm{c}}(P_i) + d_{\mathrm{c}}(P_j)\right)^2,$$

then

$$R_{\mathrm{coag}}^{\mathrm{fm}} \leq \hat{R}_{\mathrm{coag}}^{\mathrm{fm}} = \frac{2.2}{V_{\mathrm{smp}}}\sqrt{\frac{\pi k_{\mathrm{B}} T}{2}}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\left[\frac{d_{\mathrm{c}}(P_i)^2}{\sqrt{m(P_i)}} + \frac{d_{\mathrm{c}}(P_i)^2}{\sqrt{m(P_j)}} + \frac{d_{\mathrm{c}}(P_j)^2}{\sqrt{m(P_i)}} + \frac{d_{\mathrm{c}}(P_j)^2}{\sqrt{m(P_j)}}\right]$$

Majorant kernels

$$R_{\mathrm{coag}}^{\mathrm{fm}} \leq \hat{R}_{\mathrm{coag}}^{\mathrm{fm}} = \frac{2.2}{V_{\mathrm{smp}}}\sqrt{\frac{\pi k_{\mathrm{B}} T}{2}}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\left[\frac{d_{\mathrm{c}}(P_i)^2}{\sqrt{m(P_i)}} + \frac{d_{\mathrm{c}}(P_i)^2}{\sqrt{m(P_j)}} + \frac{d_{\mathrm{c}}(P_j)^2}{\sqrt{m(P_i)}} + \frac{d_{\mathrm{c}}(P_j)^2}{\sqrt{m(P_j)}}\right]$$

$$R_{\mathrm{coag}}^{\mathrm{fm}} \leq \hat{R}_{\mathrm{coag}}^{\mathrm{fm}} = \frac{2.2}{V_{\mathrm{smp}}}\sqrt{\frac{\pi k_{\mathrm{B}} T}{2}}\left[(N-2)\sum_{i=1}^{N}\frac{d_{\mathrm{c}}(P_i)^2}{\sqrt{m(P_i)}} + \frac{\sum_{i=1}^{N} d_{\mathrm{c}}(P_i)^2}{\sum_{i=1}^{N}\sqrt{m(P_i)}}\right],$$

which is easier to compute.

Majorant kernels

Since the majorant coagulation rate is larger than the *real* coagulation rate, the majorant will over-predict the frequency of events.

This is corrected by using *fictitious jumps*. A coagulation event only occurs successfully with probability equal to the ratio of the true kernel to the majorant kernel:

$$\mathbb{P}(\text{success}) = \frac{K(P_i, P_j)}{\widehat{K}(P_i, P_j)}, \qquad \text{the real kernel is only computed for a particle pair.}$$

Linear process deferment algorithm[43]

An efficient way to simulate particle processes when *linear processes* dominate.

A *linear process* refers to processes that are independent of the concentration of particles (e.g., surface processes).

Linear processes are deferred and only resolved at the end of the DSA time step.

In the case of a coagulation event the linear processes are performed on the affected particles before they coagulate.

Benefit: Orders of magnitude reduction in computation time, without significant accuracy loss.

Algorithm 1: Direct simulation algorithm with particle doubling.

Input: State of the system Q_0 at time t_0, Stop time t_f, Transfer species A^{transf}
Output: State of the system Q_f at stop time t_f.
$t_i \leftarrow t_0$, $Q_i \leftarrow Q_0$
while $t_i < t_f$ **do**
  Calculate the number of transfer species molecules from the gas phase:

$$N_{\text{gas}}^{\text{transf}} = \lfloor C_{A^{\text{transf}}} N_A V_{\text{sample}} \rfloor$$

  while $N_{\Xi}^{\text{transf}} \neq N_{\text{gas}}^{\text{transf}}$ **do** /* Compare with transfer species in the ensemble */
    if $N_{\Xi}^{\text{transf}} < N_{\text{gas}}^{\text{transf}}$ **then**
      Add transfer species (as single molecule primary particles) to the ensemble
      if $N_P(Q) > N_{\text{max}}$ **then**
        Uniformly remove a particle
        Contract the ensemble
      end
    else
      Remove transfer species from the ensemble
    end
  end
  Calculate total rate of non-deferred processes:

$$R_{\text{tot}}^{\text{non-def}}(Q) = R_{\text{eq}}^{\text{tr}}(Q) = \frac{1}{2}\sum_{q,p,i}^{N_P} K(P_q, P_i).$$

  Calculate an exponentially distributed waiting time:

$$\tau = \frac{\ln U}{R_{\text{tot}}^{\text{non-def}}(Q)},$$

  where U is a uniformly distributed random variable, $U \in [0, 1]$
  Select two particles P_q and P_i
  Calculate the majorant kernel for two particles: $K_{\text{eq}}^{\text{tr}}(P_q, P_i)$
  Perform deferred processes for P_q and P_i
  Calculate the true kernel for the two particles: $K_{\text{eq}}^{\text{tr}}(P_q, P_i)$
  Perform particle-particle process with probability

$$P = \frac{K_{\text{eq}}^{\text{tr}}(P_q, P_i)}{K_{\text{eq}}^{\text{tr}}(P_q, P_i)},$$

  if $N < N_{\text{min}}$ **then**
    for $i \in \{1, \ldots, N_P(Q)\}$ **do** /* Double the ensemble */
      Create a new particle $P_{N_{\text{new}}+i} \leftarrow P_i$
    end
    Double sampling volume
  end
  $Q \leftarrow Q_i$
  $t \leftarrow t + \tau$
end
for $q \in \{1, \ldots, N_P(Q)\}$, $i \in \{1, \ldots, n_p(P_q)\}$, $j \in \{1, \ldots, n_m(p_i)\}$ **do** /* LPDA */
  Perform deferred (PAH) processes for all molecules in all particles
  (See Algorithm 2)
end
$Q \leftarrow Q_i$
$t \leftarrow t_f$

Algorithm 2: Deferred (PAH) processes for a single molecule according to LPDA.

Input: State of molecule Q_0 at t_0, Final time t_f
Output: State of molecule Q_f at t_f
$t \leftarrow t_0$, $Q \leftarrow Q_0$
while $t < t_f$ **do**
  Calculate total rate of surface growth (deferred) processes:

$$R_{\text{tot}}^{\text{def}}(Q) = \sum_j R_j(Q), \qquad \text{where } j \in \{S1, \ldots, S30\} \text{ Table 3.4,}$$

  Calculate an exponentially distributed waiting time:

$$\tau = \frac{\ln U}{R_{\text{tot}}^{\text{def}}(Q)},$$

  where U is a uniformly distributed random variable, $U \in [0, 1]$
  Select a process j with probability

$$P(j) = \frac{R_j}{R_{\text{tot}}^{\text{def}}}.$$

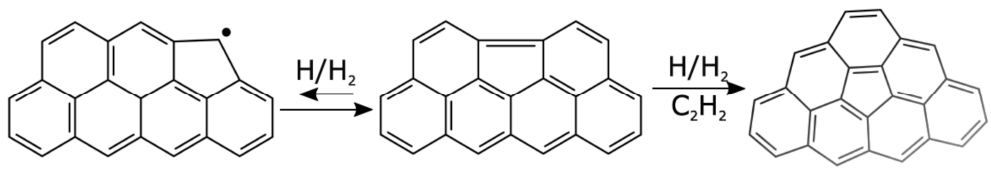
  Uniformly select a site s_k to perform process j, such that η_k or $\eta_{\text{comb},k}$ are the same site type as required by the parent site type for process j
  /* Perform Jump Process j */
  As stated by Process j (Table 3.5):
  Optimise the structure if necessary
  Add or remove atoms
  Add or remove rings
  Add or remove sites
  Transform existing sites
  Optimise the structure if necessary
  $Q \leftarrow Q_i$
  $t \leftarrow t + \tau$
end
$Q \leftarrow Q_i$
$t \leftarrow t_f$

These two algorithms represent the previous discussed simulation techniques for Direct simulation Monte Carlo

Deferred update migration algorithm[13]

The migration of partially-embedded five-member rings is an extremely fast process[20]. The movement of the ring behaves as a one-dimensional random walk around the molecular edge.

The migration is typically ended by the embedding of the five-member ring or by its desorption.



Deferred update migration algorithm

Migration processes modify the molecular structure significantly but do not add/remove atoms.

Updating the structure can be an expensive operation (more if it involves optimizing the structure).

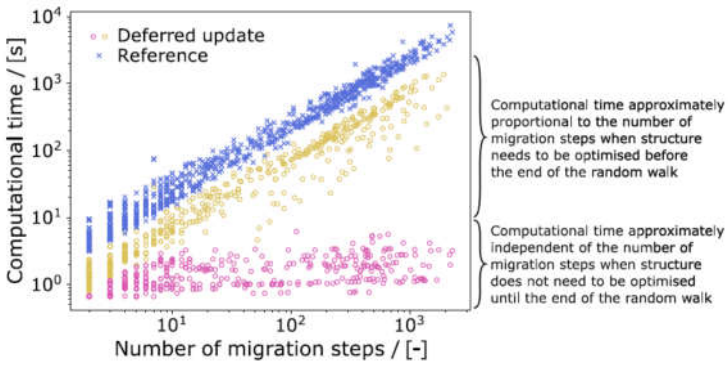
The algorithm *freezes* the structure (L_c) while updating the site types (so the process rates are updated). After a non-migration process is sampled, the structure is updated for the migration and then for the non-migration process.

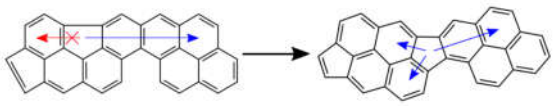
There can be several migrating five-member rings in the algorithm.

Deferred update migration algorithm

Reduces computational time by at least one order of magnitude.

Migration processes that affect sites on both sides of the molecule still require updating the structure.





The algorithm uses a set of random walkers $\mathscr{W}$, where a random walker w_k is represented as:

$$w_k = (i_k + n_k)$$

i_k → initial location of w_k

n_k → number of steps taken by w_k

The algorithm adds and removes a site after the random walker goes around the corner of the structure.

Algorithm 3: Deferred (PAH) processes for a single molecule with deferred update migration algorithm.

Input: State of molecule Q_0 at t_0, Final time t_f
Output: State of molecule Q_0 at t_f
$Q \leftarrow Q_0$
$t \leftarrow t_0$
deferredMigration $\leftarrow False$
while $t < t_f$ **do**
  Calculate total rate of surface growth (deferred) processes:
  $$R_{tot}^{def}(Q) = \sum_j R_j(Q), \qquad \text{where } j \in \{S1, \ldots, S30\} \text{ Table 5.4.}$$
  Calculate an exponentially distributed waiting time:
  $$\tau \leftarrow \frac{\ln U}{R_{tot}^{def}(Q)}$$
  where U is a uniformly distributed random variable, $U \in [0, 1]$
  Select a process j with probability
  $$P(j) = \frac{R_j}{R_{tot}^{def}}$$
  Uniformly select a site s_k to perform process j such that s_k corresponds to the principal site type of process j
  /* Perform Jump Process j */
  if $j \in$ MIGRATION **then** /* A migration process */
    if deferredMigration $= False$ **then** /* Defer migration process */
      deferredMigration $\leftarrow True$
      Initialise set of random walkers, $\mathscr{W} = \{w_1, w_2, \ldots, w_{N_k}\}$, /* One per migrating site */ where each random walker is represented as a pair $w_m = (i_m, n_m)$ defined by the initial location of the migrating site i_m and the net number of steps taken during the random walk n_m, such that the current location of migrating site k, described by walker $m(k)$, is $k \leftarrow i_m + n_m$.
    end
    Look up index of migrating site
    $m \leftarrow m(k)$
    Select migration direction
    Select the landing site for the migration, $s_l \in \{s_{k+1}, s_{k-1}\}$.
    $n_m \leftarrow n_m \pm 1$.
    if go around corner **then** /* Site moves around the corner */
      Modify structure – only affected sites
      Add carbon atom to site s_{l_m}. /* Site now contains six-member ring */
      Remove carbon atom from s_k. /* Site now contains five-member ring */
      Remove site next to s_l. Add site next to s_{l_m}. /* Update number of sites */
      Update walker with new migrating site data, $w_m \leftarrow (l_m - 1, n_m - 0)$.
    end
    Update sites /* Rates are accurate */
    Update the site type of s_k (remove partially-embedded five-member ring). /* Update site list */
    Update the site type of s_l (add partially-embedded five-member ring). /* Update site list */
    Update neighbouring sites. /* Sites affect their neighbours */
  else /* Not a migration process */
    if deferredMigration **then**
      Modify structure – update the location of all migration sites
      Add carbon atoms and new sites, update site types for all initial sites s_{i_m} in $\mathscr{W}$.
      Remove carbon atoms and neighbour sites, update site types for all final sites $s_{i_m+n_m}$ in $\mathscr{W}$.
      Update neighbouring sites. /* Sites affect their neighbours */
      Optimise structure /* After all migration processes */
      deferredMigration $\leftarrow False$
    end
    Perform non-migration jump process
    $Q \leftarrow Q$.
  end
  $t \leftarrow t + \tau$
end
$Q \leftarrow Q$.
$t \leftarrow t_f$

Hybrid methods

Hybrid particle models

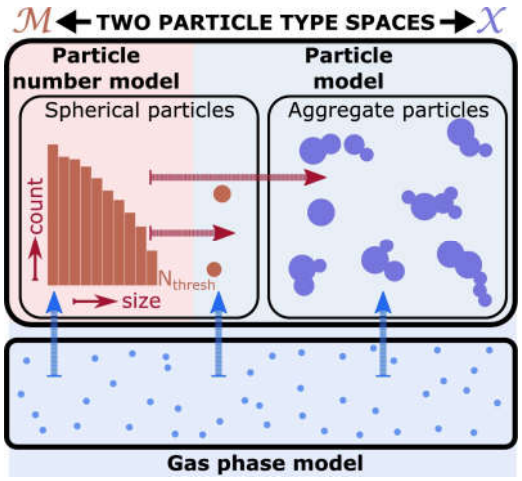
Useful for conditions with large inception rates.

Particles smaller than a threshold, N_{thresh}, are stored as spherical particles with a stochastic weight.

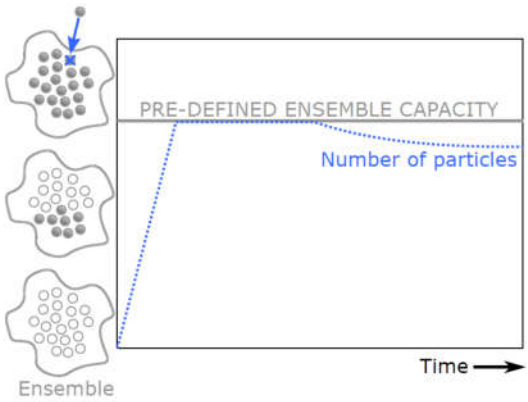
Particles larger than this are stored in a detailed particle model.

This allows using reasonably sized ensembles without losing the details of aggregate particles

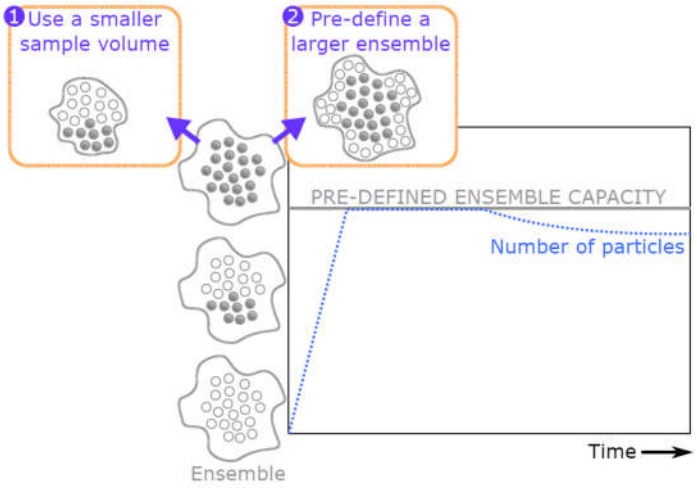
The model by Boje et al. has been used to study titania production in industrial conditions[41].



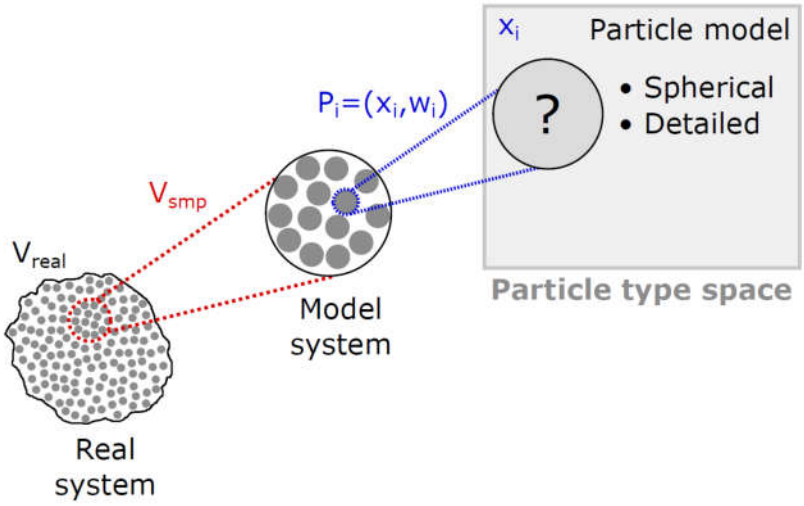
Large ensembles are required to resolve the PSD



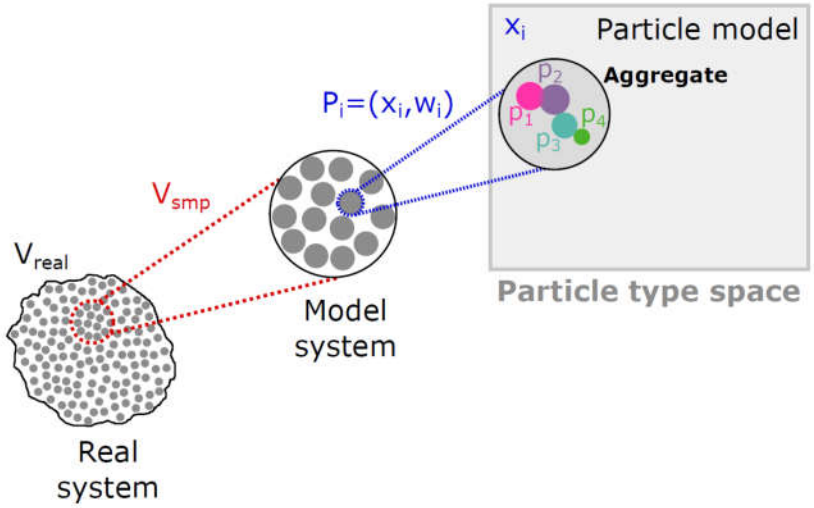
Large ensembles are required to resolve the PSD



Particle complexity varies...



Particle complexity varies...



... we can exploit this by splitting the type space

i.e. use a hybrid **particle-number** and **detailed particle** model[1]

For a particle P:

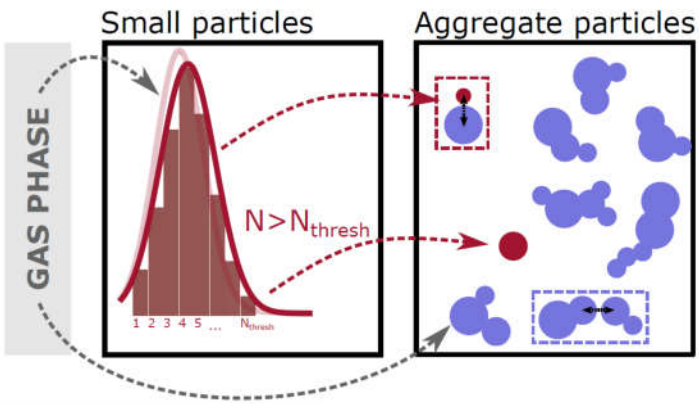
$P \in \mathcal{M}$
- small particles

$P \in \mathcal{X}$
- particles above a threshold size
- structurally complex particles

[1] Boje *et al.* J. Comput. Phys., 389:189–218, 2019.

... we can exploit this by splitting the type space

i.e. use a hybrid **particle-number** and **detailed particle** model[1]



[1] Boje *et al.* J. Comput. Phys., 389:189–218, 2019.

Particle process treatment with a hybrid particle model

We have two discrete particle systems

$y_i \in z_{\mathcal{M}}(t),\ i = 1, \ldots, N_{\text{thresh}}$ $\qquad$ $x_i \in z_{\mathcal{X}}(t),\ i = 1, \ldots, N(t)$

Inception

Increment counter for particles with η_{inc} monomers:

$$N_{\eta_{\text{inc}}} \leftarrow N_{\eta_{\text{inc}}} + 1$$

Particle process treatment with a hybrid particle model

We have two discrete particle systems

$y_i \in z_{\mathcal{M}}(t),\ i = 1, \ldots, N_{\text{thresh}}$ $\qquad$ $x_i \in z_{\mathcal{X}}(t),\ i = 1, \ldots, N(t)$

Surface reaction

Small particles, space $\mathcal{M}$:

Alter counters for particles with η_i, $\eta_i + \eta_{\text{add}}$ monomers:

$$N_{\eta_i+\eta_{\text{add}}} \leftarrow N_{\eta_i+\eta_{\text{add}}} + 1$$
$$N_{\eta_i} \leftarrow N_{\eta_i} - 1$$

Particle process treatment with a hybrid particle model

We have two discrete particle systems

$y_i \in z_{\mathcal{M}}(t),\ i = 1, \ldots, N_{\text{thresh}}$ $\qquad$ $x_i \in z_{\mathcal{X}}(t),\ i = 1, \ldots, N(t)$

Surface reaction

Small particles, space $\mathcal{M}$:

Alter counters for particles with η_i, $\eta_i + \eta_{\text{add}}$ monomers:

$$N_{\eta_i+\eta_{\text{add}}} \leftarrow N_{\eta_i+\eta_{\text{add}}} + 1$$
$$N_{\eta_i} \leftarrow N_{\eta_i} - 1$$

Large particles, space $\mathcal{X}$:

Add $\eta_i + \eta_{\text{add}}$ monomers to primary particle p_i:

$$p_i(\eta_i, r_i, \mathbf{z}_i) \leftarrow p_i(\eta_i + \eta_{\text{add}}, r_i, \mathbf{z}_i)$$

Particle process treatment with a hybrid particle model

We have two discrete particle systems

$y_i \in z_{\mathcal{M}}(t),\ i = 1, \ldots, N_{\text{thresh}}$ $\qquad$ $x_i \in z_{\mathcal{X}}(t),\ i = 1, \ldots, N(t)$

Coagulation

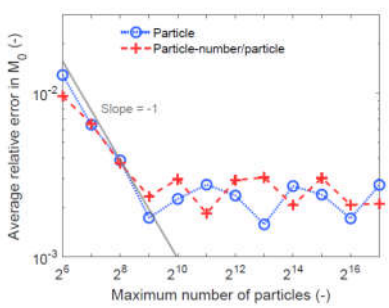
Numerical rate includes terms from both systems:

$$R_{\text{coag}} = \frac{1}{2V_{\text{smp}}} \left[\sum_{i=1}^{N(t)} \sum_{\substack{j=1 \\ j \neq i}}^{N(t)} K(x_i, x_j) + \sum_{i=1}^{N_{\text{thresh}}} \sum_{\substack{j=1: \\ j \neq i \text{ if } N_i < 2}}^{N_{\text{thresh}}} K(y_i, y_j) N_i N_j \right]$$

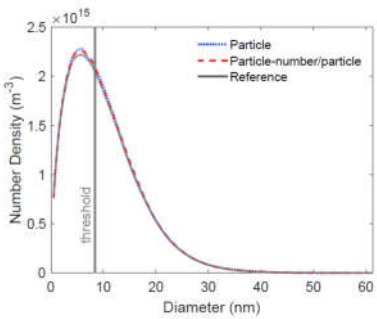
$$+ \frac{1}{V_{\text{smp}}} \sum_{i=1}^{N(t)} \sum_{j=1}^{N_{\text{thresh}}} K(x_i, y_j) N_j.$$

Adapting DSA to allow particle-number and particle models

Hybrid approach is **exact** for a univariate primary particle model



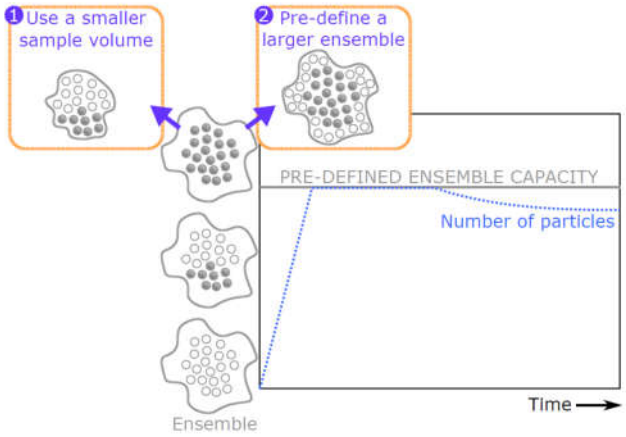
Convergence study



Particle size distribution (KDE)

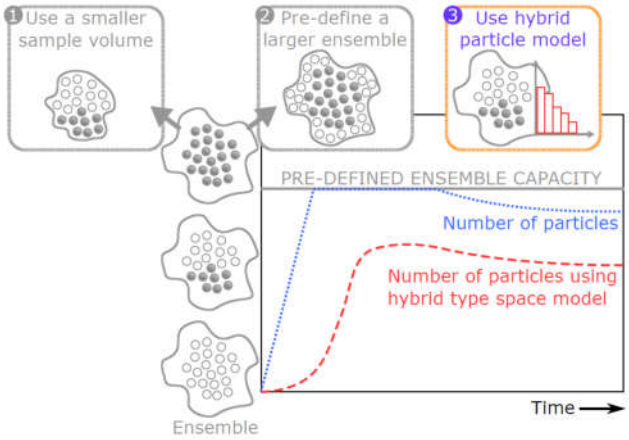
Boje *et al.* J. Comput. Phys., 389:189–218, 2019.

Hybrid method is robust to changes in inception rate and poor choice of numerical parameters



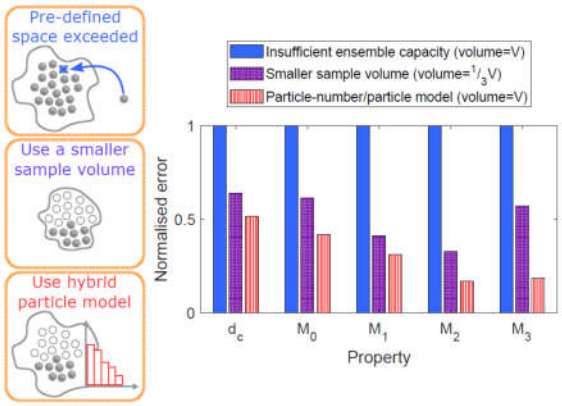
Boje *et al.* J. Comput. Phys., 389:189–218, 2019.

Hybrid method is robust to changes in inception rate and poor choice of numerical parameters



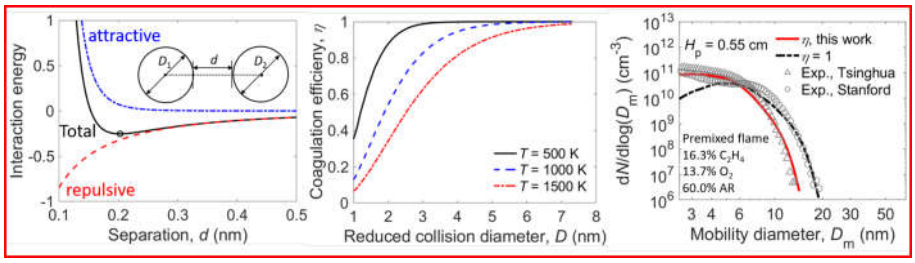
Boje *et al.* J. Comput. Phys., 389:189–218, 2019.

Hybrid method reduces the error for a given cost



Boje *et al.* J. Comput. Phys., 389:189–218, 2019.

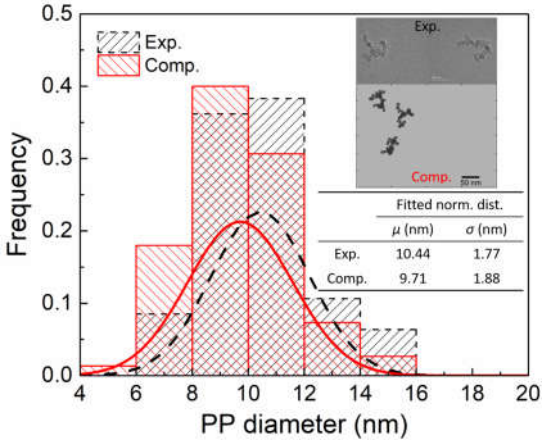
SPM - Particle evolution



- The **interaction energy between two spherical particles was derived from the L-J potentials of the constituent atoms of the two particles**.
- A coagulation efficiency model for soot was proposed based on the interaction energy between the colliding partners and their kinetic energy.

On the coagulation efficiency of carbonaceous nanoparticles
Dingyu Hou, Diyuan Zong, Casper Lindberg, Markus Kraft, and Xiaoqing You, Journal of Aerosol Science 140, 105478, (**2019**).

SPM - Particle evolution



Numerical simulations of soot aggregation in premixed laminar flames
Neal M. Morgan, Markus Kraft, **Michael Balthasar**, David Wong, **Michael Frenklach**, and Pablo Mitchell, Proceedings of the Combustion Institute 31(1), 693-700, (**2007**).

Simulation of primary particle size distributions in a premixed ethylene stagnation flame
Dingyu Hou, Casper Lindberg, Mengda Wang, Manoel Y. Manuputty, Xiaoqing You, and Markus Kraft, Combustion and Flame 216, 126-135, (**2020**).

End of Lecture 5

Soot – Part 6

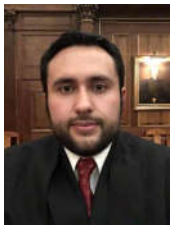
Markus Kraft

Computational Modelling Group Cambridge

Main Contributors:
Dr Jake Martin (Part 1)
Dr Angiras Menon (Part 2)
Dr Laura Pascazio (Part 3)
Dr Gustavo Leon (Part 4)

Jacob Martin

AngirasMenon



Gustavo Leon

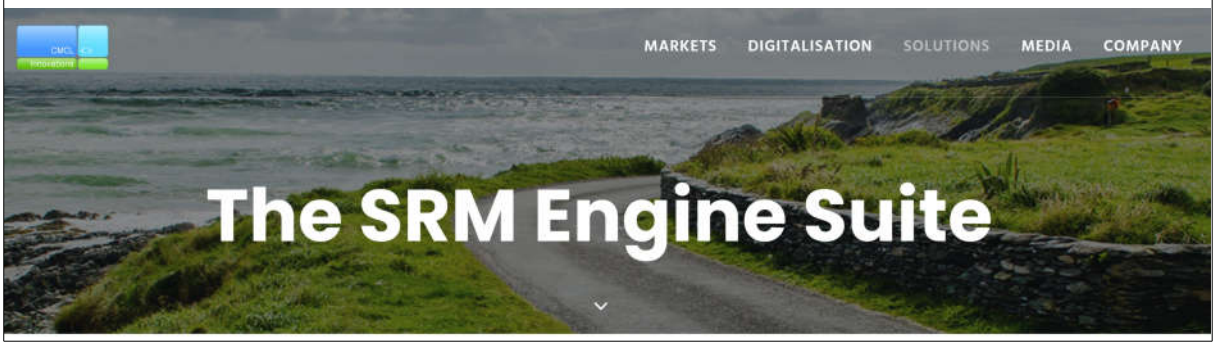
Laura Pascazio

Part 1 Overview
Part 2 Quantum Chemistry
Part 3 Molecular Dynamics
Part 4 Kinetic Monte Carlo
Part 5 Stochastic Particle Methods
Part 6 Application and Outlook

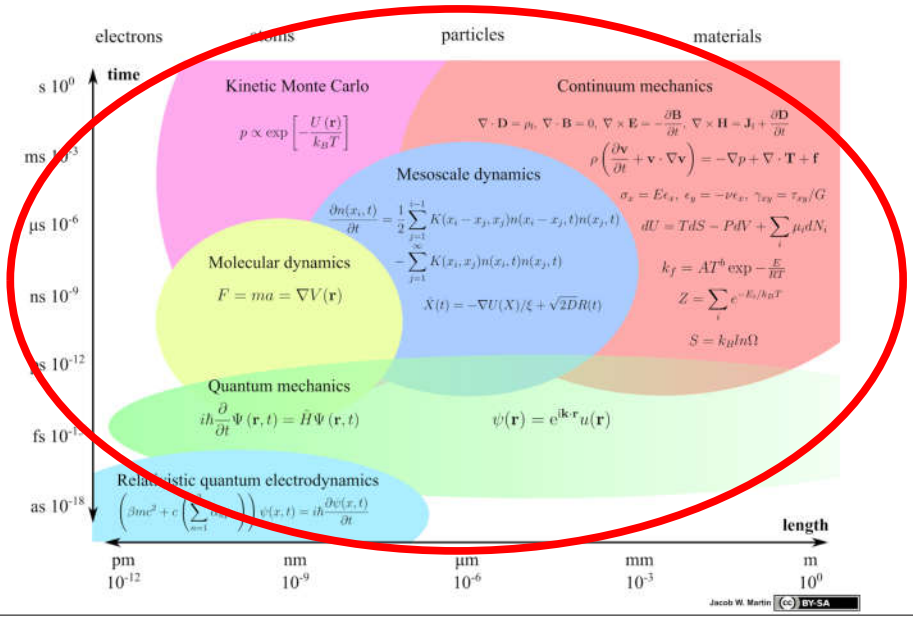
Soot in engines

Sebastian Mosbach

Towards a detailed soot model for internal combustion engines Sebastian Mosbach, Matthew S. Celnik, Abhijeet Raj, Markus Kraft, Hongzhi R. Zhang, Shuichi Kubo, and Kyoung-Oh Kim, *Combustion and Flame* **156**(6), 1156-1165, (2009).



Scale Diagramme

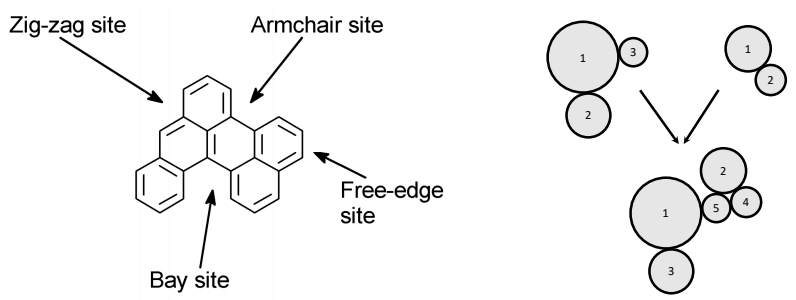


Soot model: site-counting

Describe soot particles by $9+N$ dimensional state space (ARS-SC-PP model):

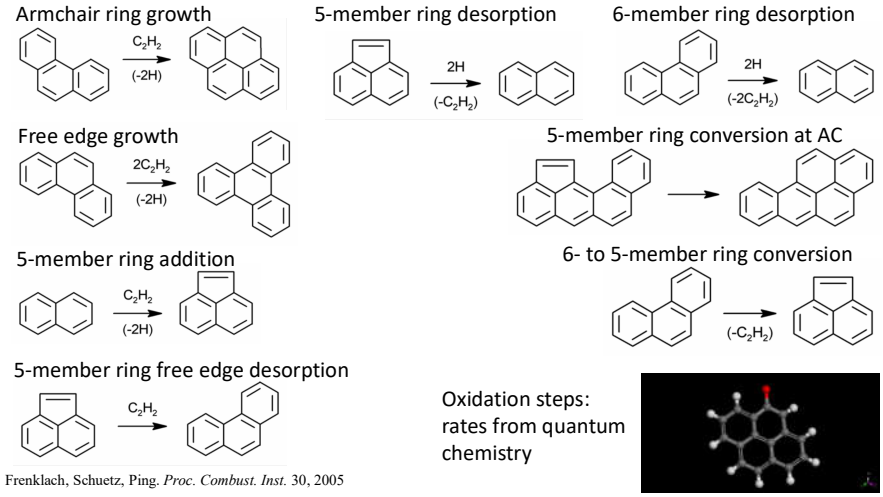
$$E = (C,\ H,\ S_a,\ N_{ed},\ N_{zz},\ N_{ac},\ N_{bay},\ N_{R5},\ N_{PAH},\ PP_{(1\text{-}N)})$$

PP = primary particle list



Towards a detailed soot model for internal combustion engines S Mosbach, MS Celnik, A Raj, M Kraft, HR Zhang, S Kubo, KO Kim, 2009, Combustion and Flame 156 (6), 1156-1165

PAH reaction steps



Frenklach, Schuetz, Ping. *Proc. Combust. Inst.* 30, 2005

Engine model: SRM

Stochastic Reactor Model (SRM)

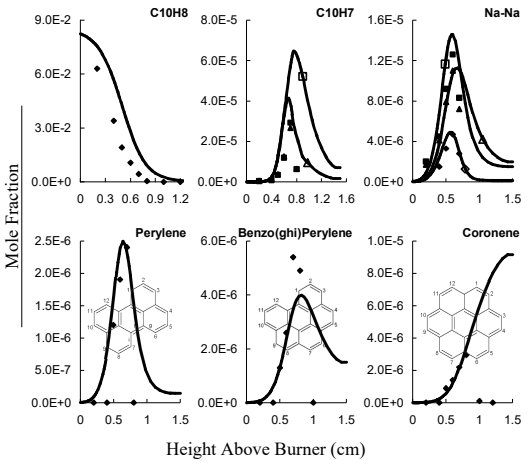
$$\frac{\partial}{\partial t}\mathcal{F}(\psi;t) = \underbrace{-\sum_{j=1}^{S+1}\frac{\partial}{\partial\psi_j}\left[G_j(\psi)\mathcal{F}(\psi;t)\right]}_{\text{chemical reactions, volume change}} + \underbrace{\sum_{j=1}^{S+1}\frac{\partial}{\partial\psi_j}\left[\frac{C_\phi}{2\tau}\left(\psi_j - \langle\psi_j\rangle\right)\mathcal{F}(\psi;t)\right]}_{\text{IEM mixing}} -$$

$$\underbrace{-\frac{\dot{V}}{V}\mathcal{F}(\psi;t)}_{\text{piston movement}} \underbrace{-\frac{1}{h}\left[U(\psi_{S+1}+h)\mathcal{F}(\psi_1,\ldots,\psi_S,\psi_{S+1}+h;t) - U(\psi_{S+1})\mathcal{F}(\psi;t)\right]}_{\text{heat transfer}}$$

- Detailed chemical kinetics → Chemical mechanism: PRF + small aromatics (extended by H. R. Zhang) 208 species, 1002 reactions
- Turbulent mixing
- Convective heat transfer
- Computationally cheap (1-2 CPU-hrs/cycle)

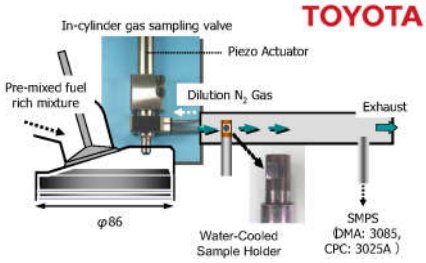
PAHs in gas-phase chemistry

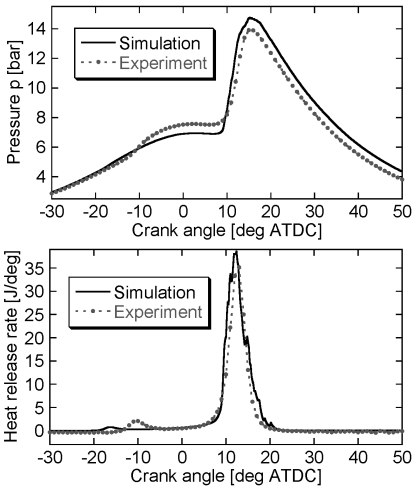
- Hongzhi R. Zhang
- Before: PRF+NOx, 157 species
- After: PRF+NOx+ variety of PAHs and highly unsaturated HCs, 208 species
- Validation against fuel-rich flame experiments



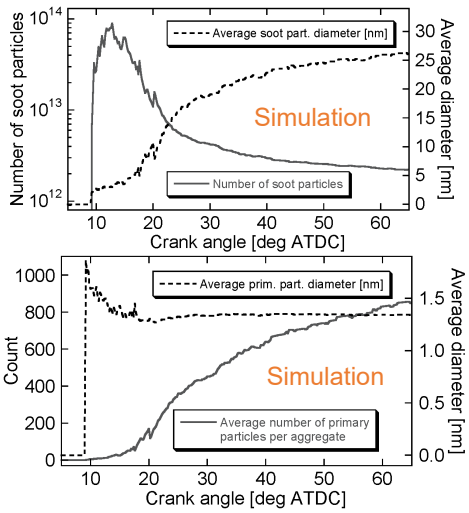
Soot in engines!

- HCCI, n-heptane
- Compression ratio 12
- Equivalence ratio 1.93
- Throttled, 20% EGR

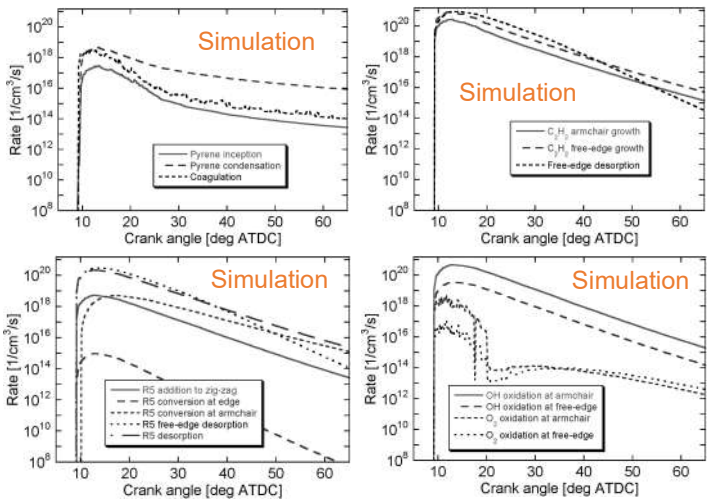




Averaged soot quantities

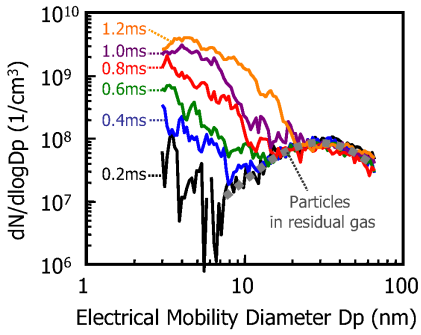


Rates of soot processes

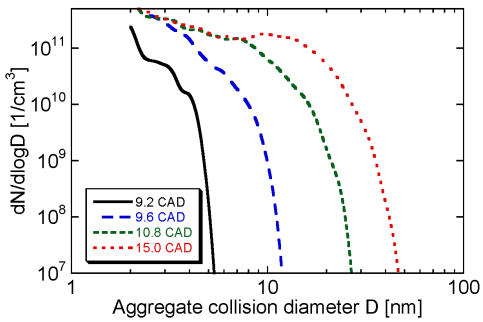


Aggregate size distributions (I)

Experiment

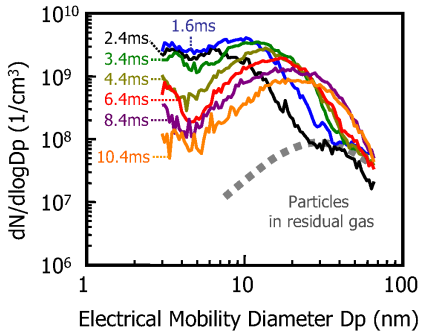


Simulation

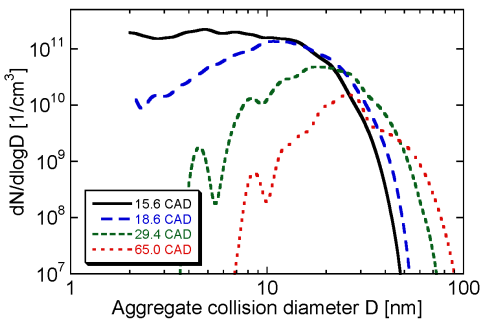


Aggregate size distributions (II)

Experiment

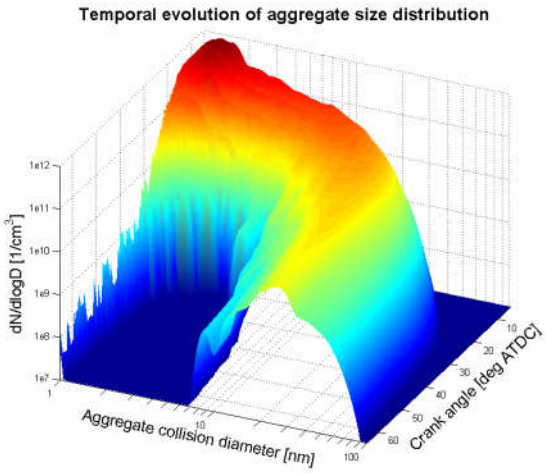


Simulation

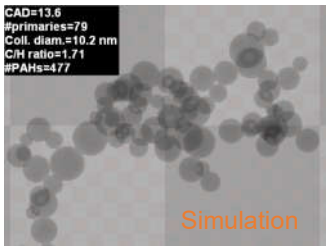


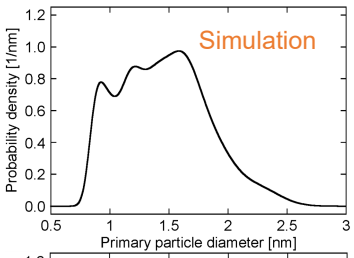
Aggregate size distributions (III)

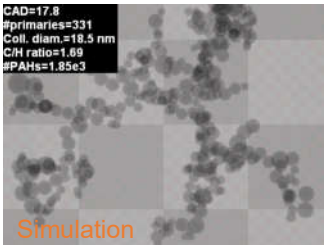
Simulation

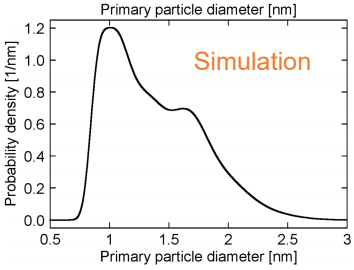


Sampled aggregates (I)

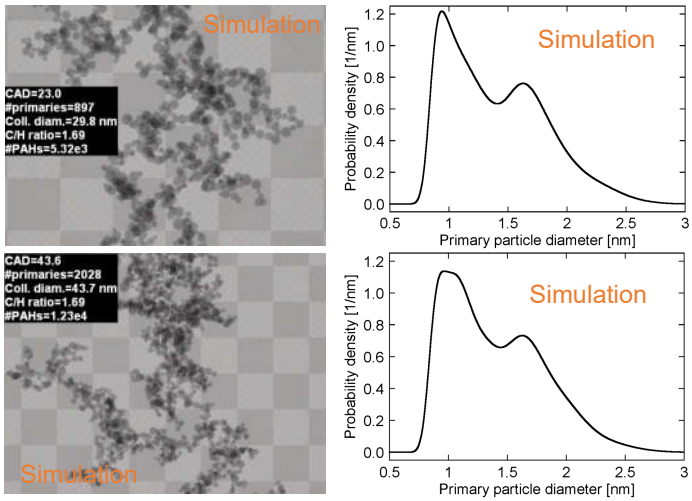






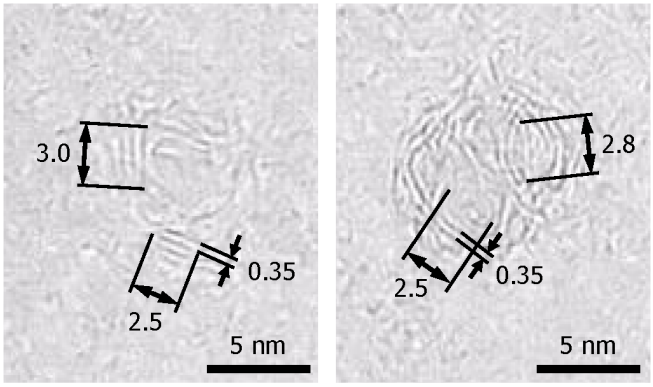


Sampled aggregates (II)



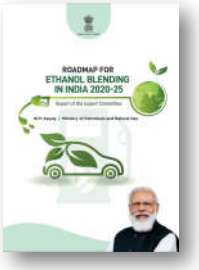
Sampled aggregates (III)

Experiment, sampled at ~16 CAD ATDC



Emissions from sustainable fuels

- Oxygenated fuels as alternative fuels or additive to conventional fuels for internal combustion engine (ICE)
  - Help mitigate the impacts of well-to-wheel emissions to environments
  - 5% to 10% inclusion of ethanol as per the current fuel quality standards



- Roadmap for Ethanol Blending in India 2020-2025
  - A target of 20 % blending of ethanol in gasoline (petrol) and 5 % blending of biodiesel in Diesel fuel by 2030

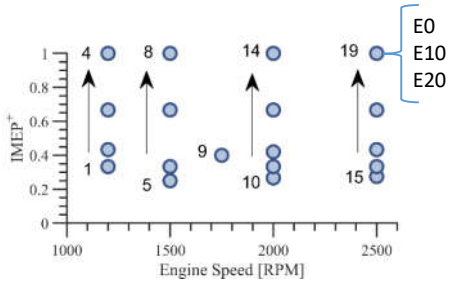
Geely GDI engine Tsinghua University

Image Source: global.geely.com/intec/g-power/

- Study the effects of ethanol-blended fuel on GDI engines
  - Through digital workflow & measurements
  - Combustion characteristics
  - Gaseous and particulate emissions
- Three ethanol-blended fuels
  - Pure gasoline (E0)
  - 10% ethanol-blended gasoline (E10)
  - 20% ethanol-blended gasoline (E20)

Geely Engine Specifications				Measurements	
Volume	1 L	Bore length	73.4 mm	Fuels	E0 – E20
Cylinder	3	Stroke length	78.6 mm	Engine speed	1250 to 2500 RPM
Compression Ratio	9.6	Con rod length	136 mm	Engine Load	Low to full

Operating Conditions

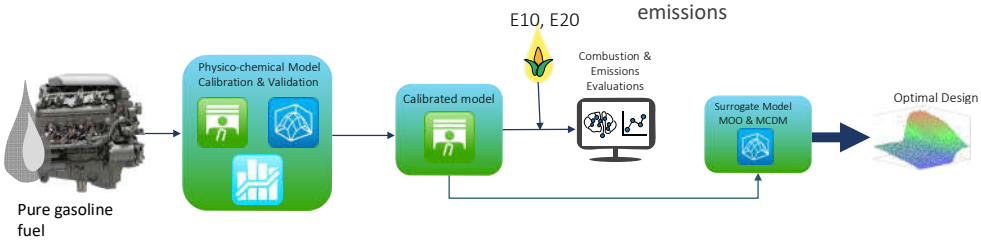


•**Measurement**

- 3 sets of measurement data with 19 operating conditions
- In total 57 sets of data with pressure and emissions

•**Modelling approach**

- Physico-chemical model with SRM Engine Suite
  - E0 data for model calibration
  - MoDS-SRM Engine Suite workflow
- Model validation
  - E10 and E20 data – adjusting ethanol content
  - In-cylinder pressure profile, NOx, uHC and PN emissions



SRM Agent

Engine

Process conditions:

$$\xi = (\xi_1, \ldots, \xi_M)$$

Engine speed
Engine Load
Inlet temperature
Inlet pressure
Spark timing
Fuel, ethanol content
E0, E10, E20

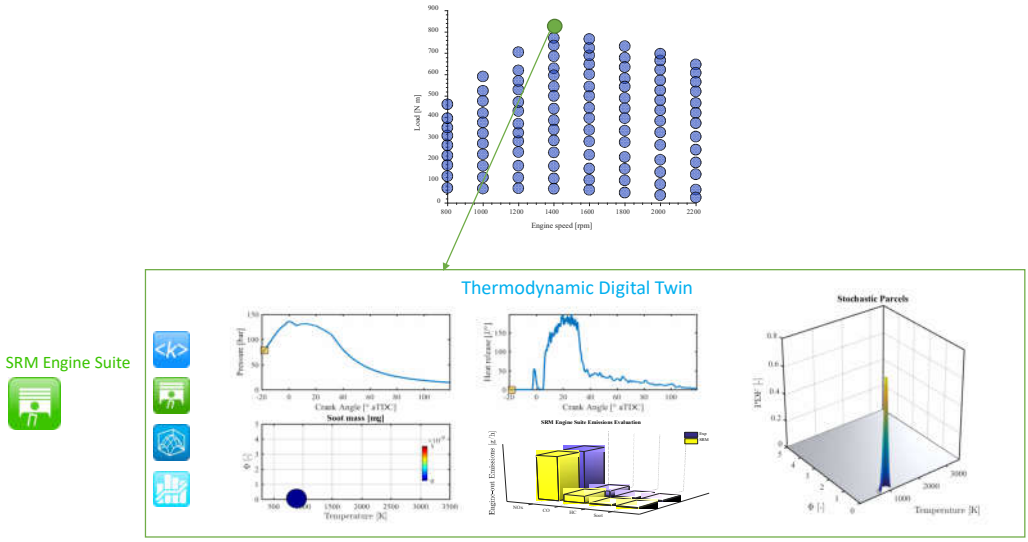
Responses:

$$\eta^{\exp} = (\eta_1^{\exp}, \ldots, \eta_L^{\exp})$$

In-cylinder pressure
NOx
uHCs
PN

Series of experiments at different process conditions: $\xi^{(n)}, \eta^{\exp,(n)}$

SRM Agent

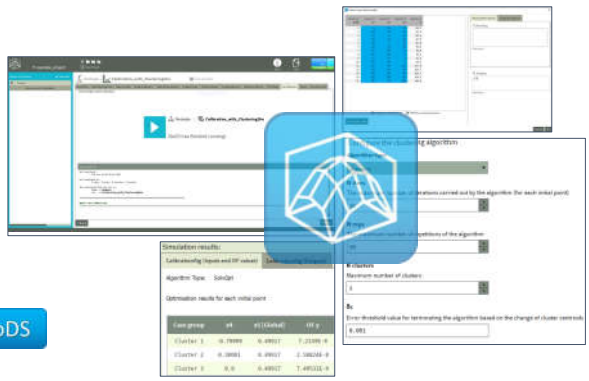


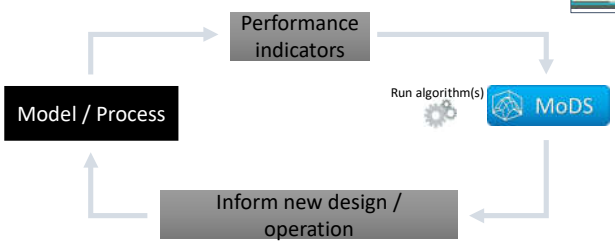
MoDS Agent

MoDS is a highly flexible model development application that uses statistical algorithms to automate common tasks

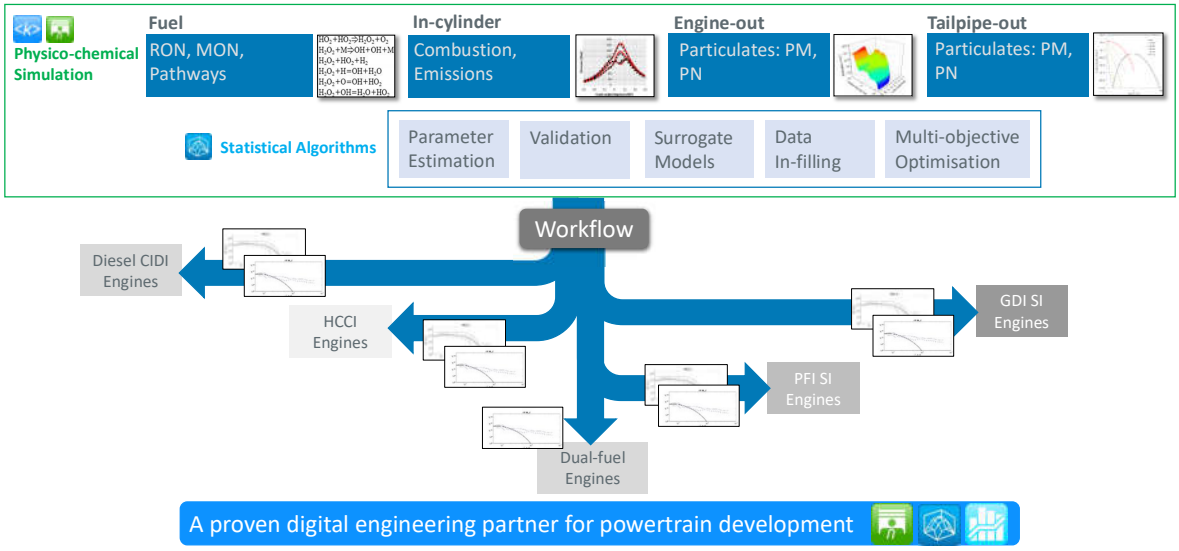
Purpose-built workflows
- Model calibration / parameter estimation
- Model validation
- Surrogate generation and evaluation
- Multi-objective optimisation
- Global sensitivity analysis

Currently backend only
- Intelligent DoE / space filling
- Local sensitivity analysis
- Uncertainty propagation

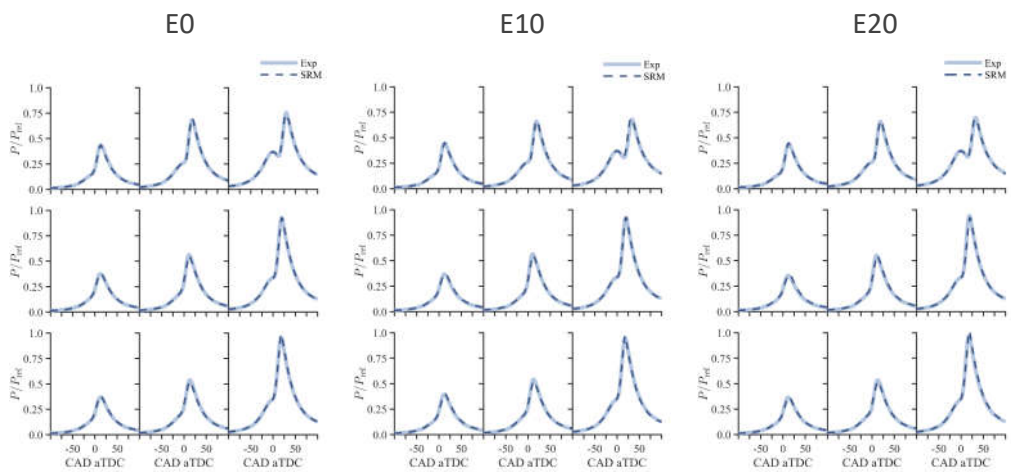




Workflow for Calibration and Validation

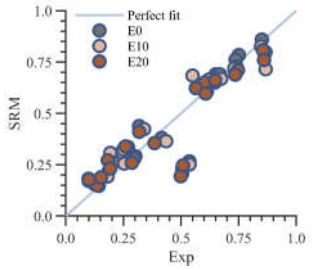


In-cylinder Pressure Profiles

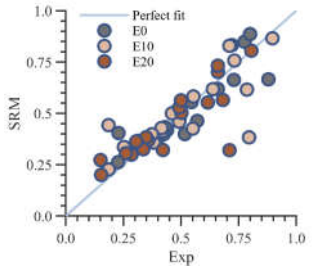


Gaseous and Particulates Emissions

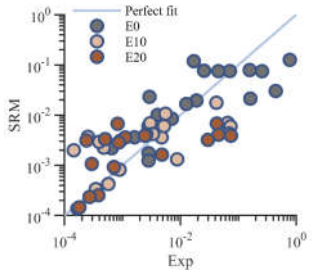
NOx



uHC

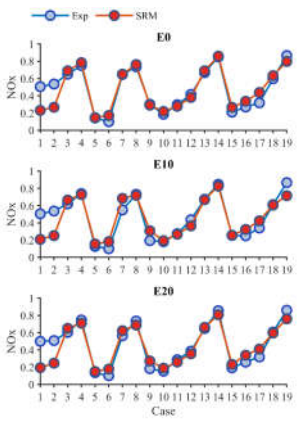


PN

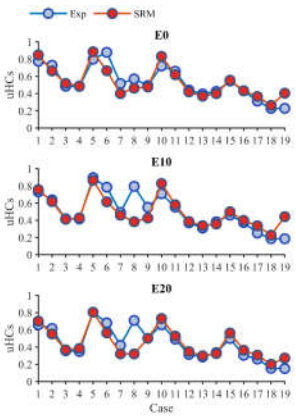


Gaseous and Particulates Emissions

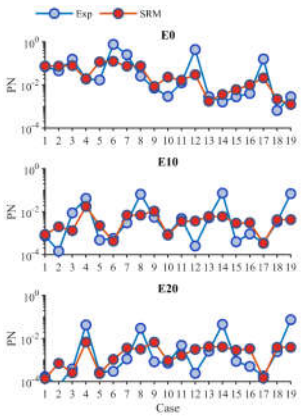
NOx



uHC

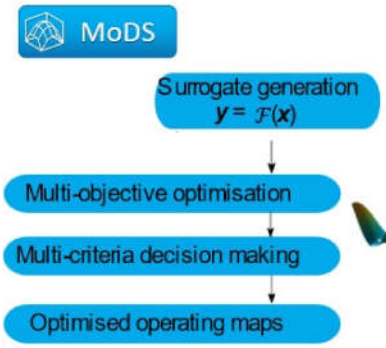


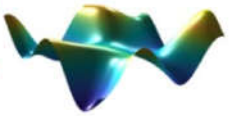
PN

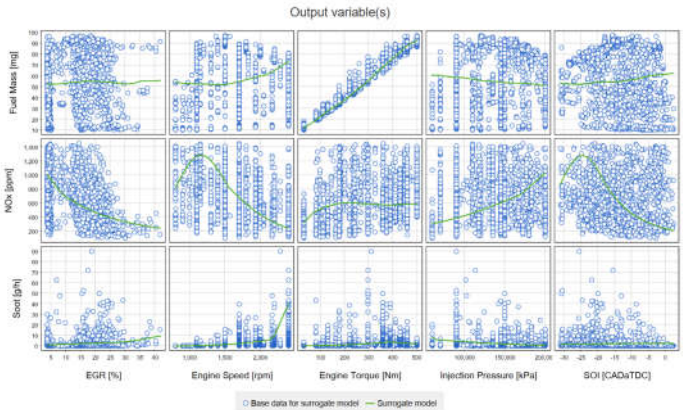


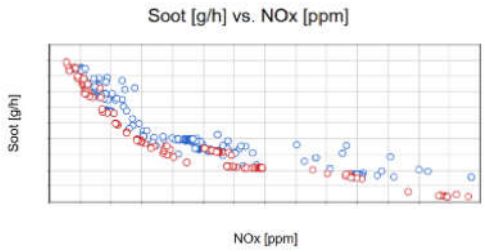
Optimal operation

Maintain the **power output at each operating points** while **minimising the emissions output.**
What is the
best spark timing
and
best inlet manifold pressure?







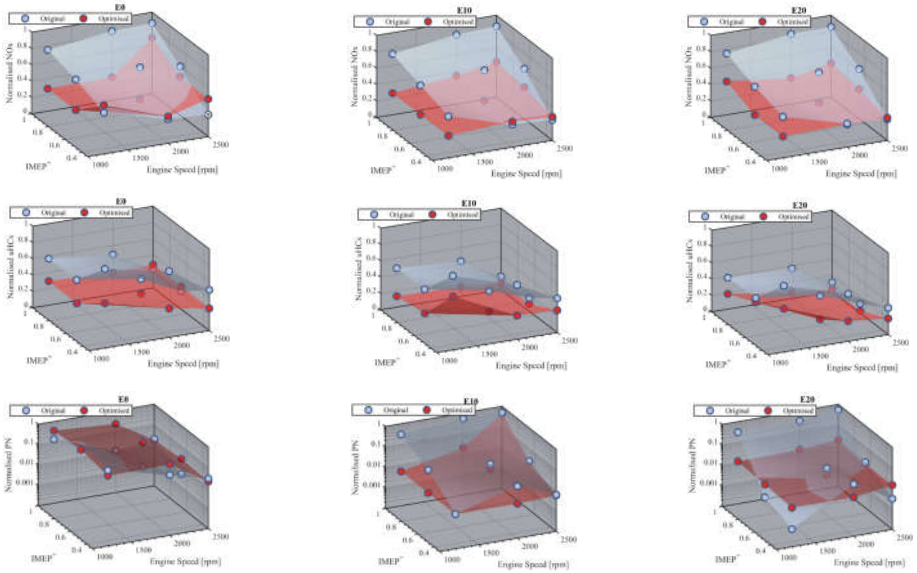


Optimal operation

NOx

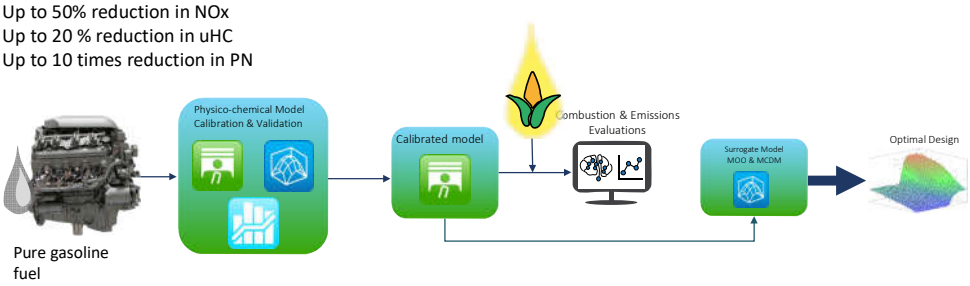
uHC

PN

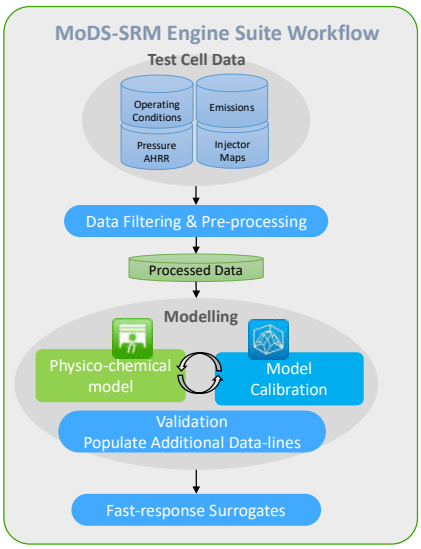


Summary

- The SRM Engine Suite is calibrated based on **pure** gasoline fuel
- The SRM Engine Suite is capable of **accurately** capturing the effects of ethanol blended fuels
  - In-cylinder pressure profiles, NOx and PN emissions
  - uHCs with modified factors within the workflow
- The mean residual error for the simulated in-cylinder pressure is **within 5 %**
- The model prediction results show a **good agreement** as compared to the measurement data in all emissions
  - No clear trend in NOx emission as function of ethanol content in the fuel
  - Strong evidence that increasing ethanol content in gasoline fuel leads to a reduction in uHC and PN emissions
- The integrated workflow for MOO-MCDM is applied to optimise input conditions in order to achieve **further reduction** in emissions
  - Up to 50% reduction in NOx
  - Up to 20 % reduction in uHC
  - Up to 10 times reduction in PN



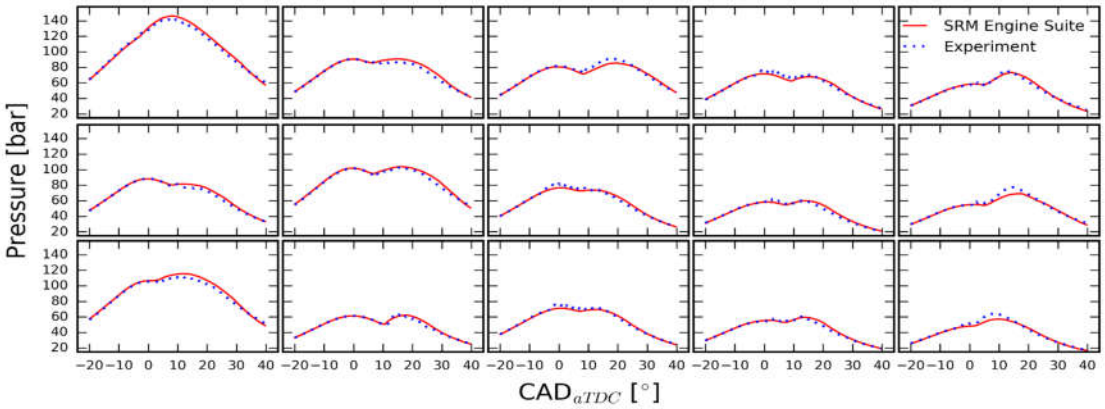
Non-Road Use Case: Heavy Duty Engine



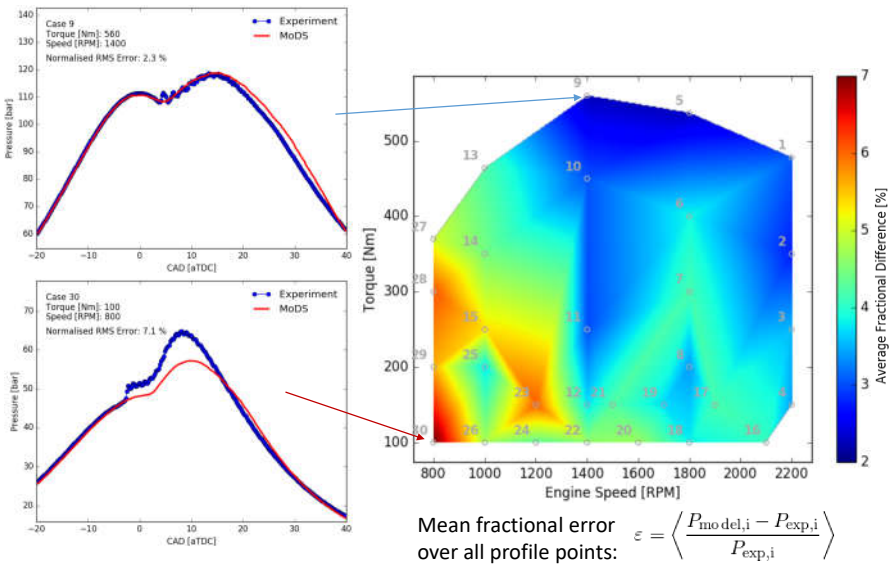
Use Case: C4.4 ACERT Industrial Engine
4.4 L Turbocharged Aftercooled (TA) IC engine
Source: www.cat.com

- Heavy duty CIDI engine
- 280 measurement data
- Identify and filter inconsis tent data
- Model calibration: 30%
- Model validation: 70%
- Pressure and emissions

Non-Road Use Case: Heavy Duty Engine

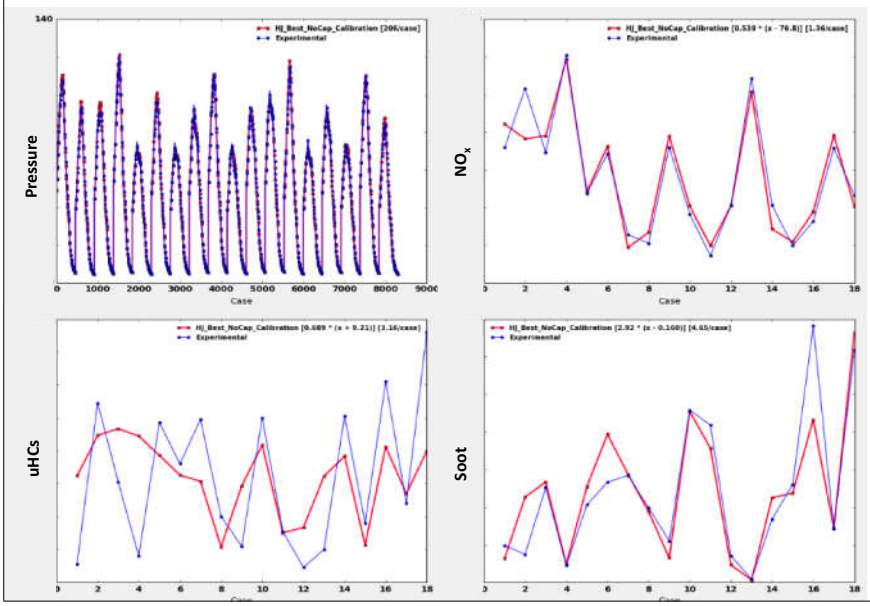


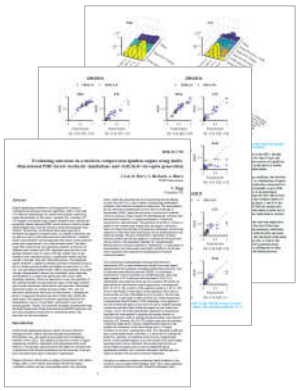
Non-Road Use Case: Heavy Duty Engine



Mean fractional error over all profile points: $\varepsilon = \left\langle \frac{P_{\mathrm{model,i}} - P_{\mathrm{exp,i}}}{P_{\mathrm{exp,i}}} \right\rangle$

Non-Road Use Case: Heavy Duty Engine





For more details: SAE Paper 2018-01-1739

Soot in a world model - The World Avatar TWA

Soot in a world model - The World Avatar TWA

What is required?

- *Digitalisation*
- *(chemical) Knowledge*
- *Interoperability*

- *Holistic view – **no system boundaries***

the World Avatar Project
From Platform to Knowledge Graph:

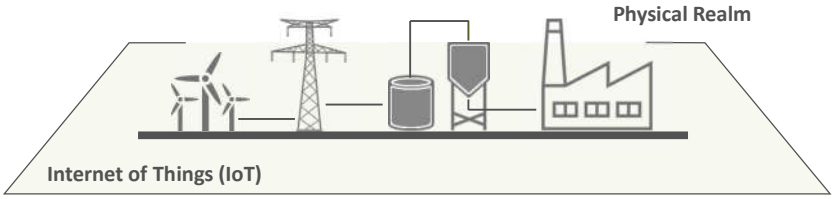
Context: From Digitisation to Digitalisation

Context: From Digitisation to Digitalisation

IoT

Internet of Things

- **Inter-connecting** numerous devices/sensors
- **Ubiquitous computing** by smart devices

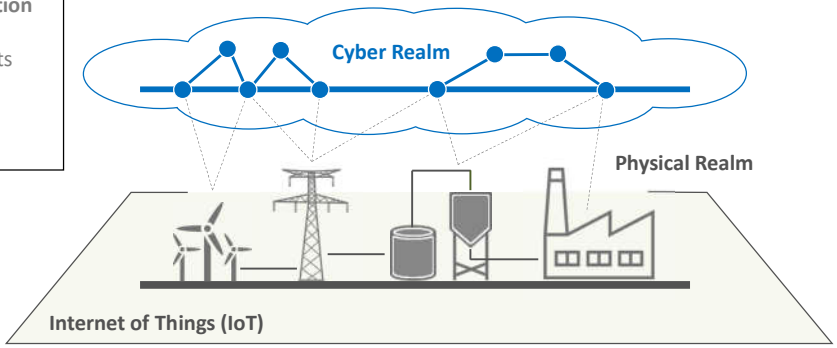


Context: From Digitisation to Digitalisation

CPS

Cyber-physical System

- **Higher level of integration** between physical and computational elements
- Deeply **intertwined physical and software** components

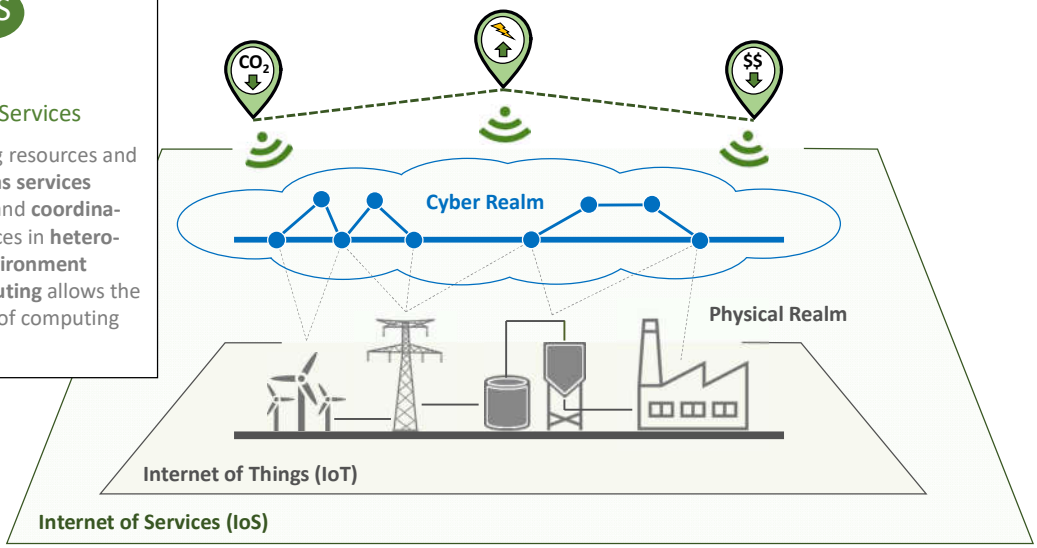


Context: From Digitisation to Digitalisation

IoS

Internet of Services

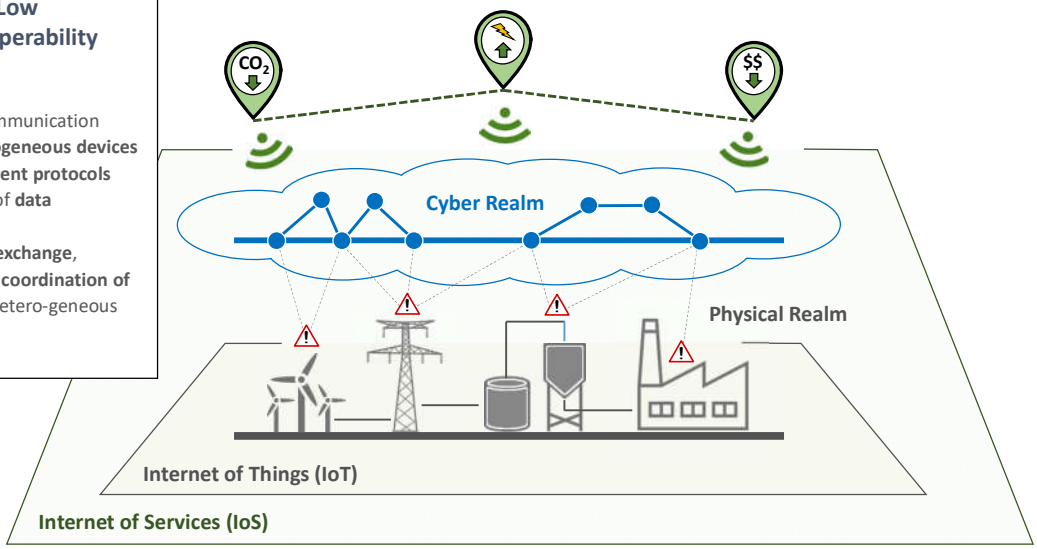
- Representing resources and capabilities **as services**
- **Integration** and **coordination** of services in **heterogeneous environment**
- **Cloud computing** allows the outsourcing of computing tasks



Context: Current Challenges for Digitalisation

Low Interoperability

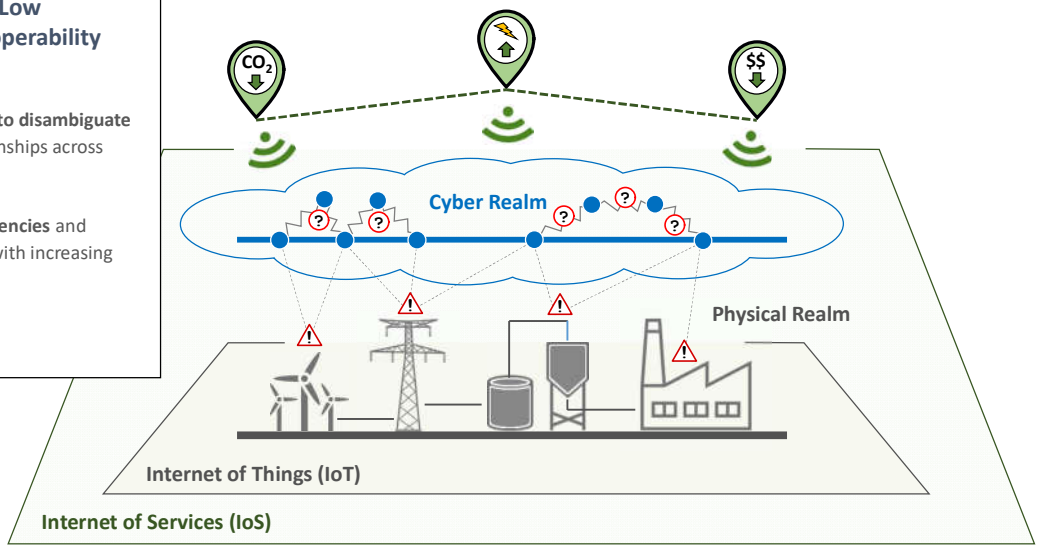
- Obstacles in communication between **heterogeneous devices** caused by **different protocols** and **ambiguity** of **data**
- High friction in **exchange**, **integration** and **coordination of data** between hetero-geneous services in IoS



Context: Current Challenges for Digitalisation

Low Interoperability

- **Lack capability to disambiguate** complex relationships across domains
- Surging **dependencies** and switching cost with increasing amount of data



The Solution

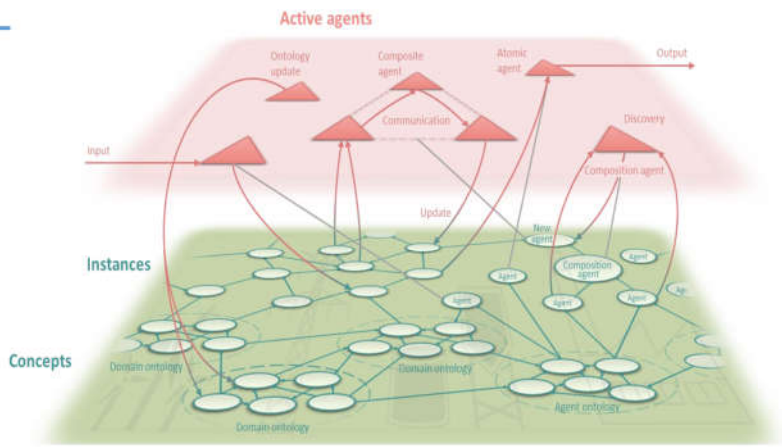
Knowledge Graphs

The Universal Digital Twin – the World Avatar Project

*formerly known as the **J-Park Simulator***

Dynamic Knowledge Graph

Distributed, directed graph

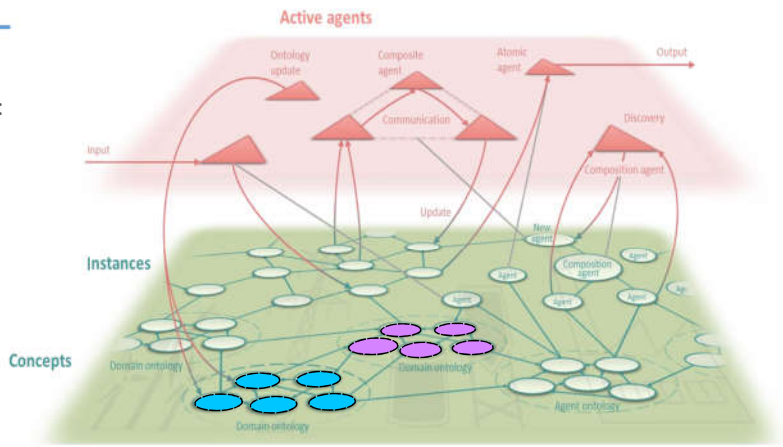


J-Park Simulator: An ontology-based platform for cross-domain scenarios in process industry. Computers & Chemical Engineering,131:106586, 2019.

Dynamic Knowledge Graph

Distributed, directed graph based on:

- **ontological** representation of data

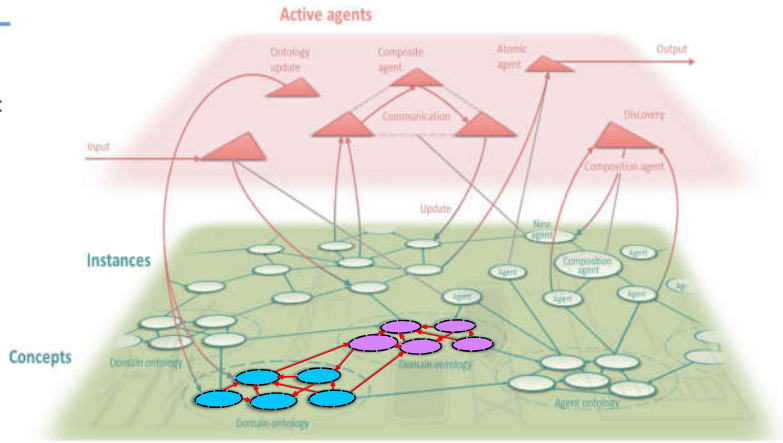


J-Park Simulator: An ontology-based platform for cross-domain scenarios in process industry. Computers & Chemical Engineering,131:106586, 2019.

Dynamic Knowledge Graph

Distributed, directed graph based on:

- **ontological** representation of data
- the concept of **Linked Data**

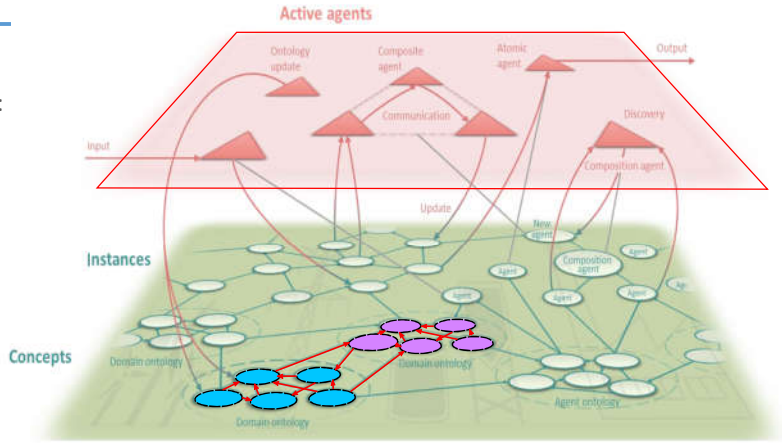


J-Park Simulator: An ontology-based platform for cross-domain scenarios in process industry. Computers & Chemical Engineering,131:106586, 2019.

Dynamic Knowledge Graph

Distributed, directed graph based on:

- **ontological** representation of data
- the concept of **Linked Data**
- contains an **ecosystem of agents**



J-Park Simulator: An ontology-based platform for cross-domain scenarios in process industry. Computers & Chemical Engineering,131:106586, 2019.

Dynamic Knowledge Graph

Distributed, directed graph

... is **unambiguous**

... is **connected**

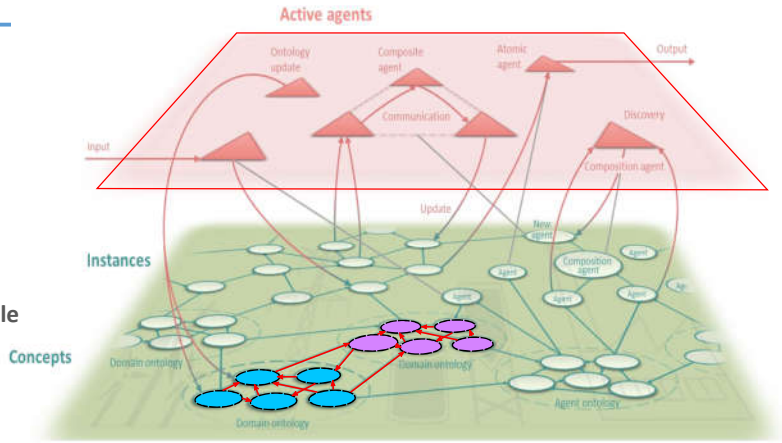
... is **scalable**

... is **distributed**

... is **accessible**

... is **multi-domain interoperable**

... is **evolving** in time



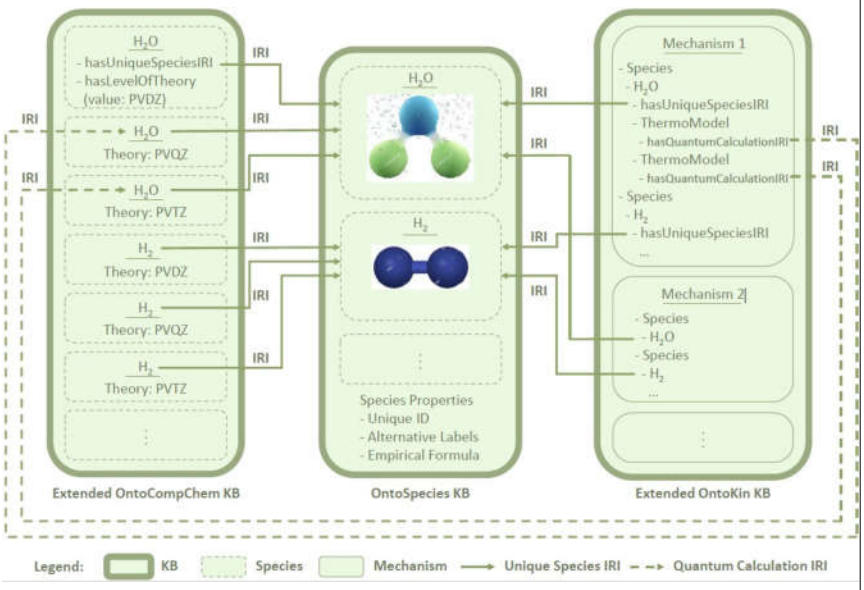
J-Park Simulator: An ontology-based platform for cross-domain scenarios in process industry. Computers & Chemical Engineering,131:106586, 2019.

elementary chemistry

OntoKin mechanisms and **OntoCompChem** calculations are linked to the appropriate entry in **OntoSpecies**.

Enables identification of the same species in different mechanisms through **OntoSpecies**.

Can use **OntoCompChem** calculations to update information on entities in **OntoKin** by mutual links through **OntoSpecies**.



Use cases

Potential energy surfaces
e.g. Automatic creation of Potential energy surfaces
OntoPESScan

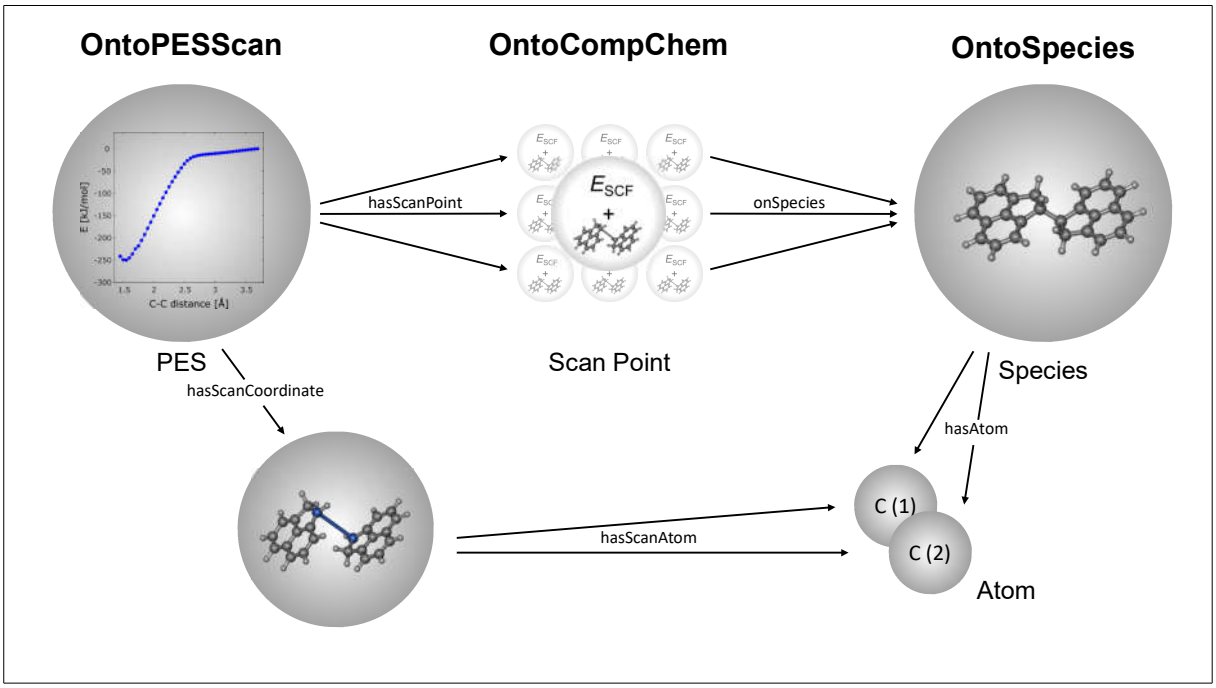
- **Potential Energy Surfaces (PES)** and their exploration are crucial to computational kinetics and force field development.
- **Development of OntoPESScan,** a new ontology to describe computational scans of a variety of potential energy surfaces (bond, angle, torsion).
- **Linked to OntoSpecies** to uniquely define molecules and fragments along the surface.
- Unambiguously define scan coordinates for a PES.
- **Linked to OntoCompChem** so each point along the surface has an associated calculation.
- **Retrieval** of energies and geometries via OntoCompChem to enable **automated fitting of forcefields for molecular dynamics.**

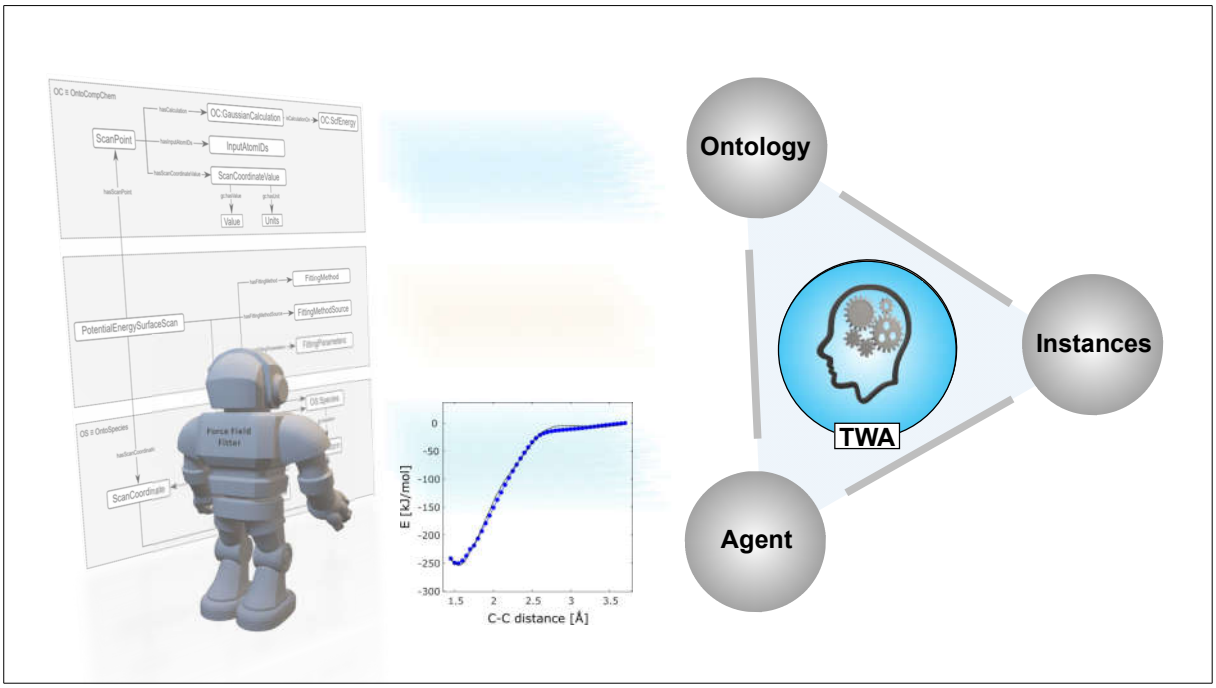
C2H5OH -> CH3+CH2OH - C-C bond scan

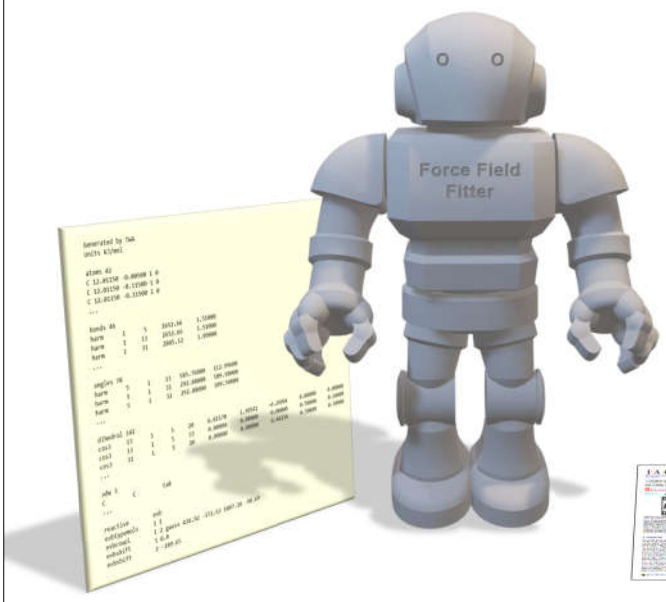
C2H5OH - C-O-H angle scan

cis-ClCHCHCl -> trans-ClCHCHCl - dihedral angle scan (Cl-C-C-Cl)

```
OntoPESScanIndividual:
    hasUniqueSpecies:
OntoSpeciesIRI
    hasScanCoordinate:
        Atom_1, Atom_2 etc…

    hasScanPoints:
        - (1)
            job:
OntoCompChemJobIRI,
            scanCoordinateValue:
                value: 1.5
                unit: Angstrom
        - (2) ...
```


OntoPESScan
OntoCompChem
OntoSpecies
E [kJ/mol]
C-C distance [Å]
hasScanPoint
onSpecies
PES
Scan Point
Species
hasScanCoordinate
hasAtom
C (1)
C (2)
hasScanAtom
Atom


Ontology
Instances
Agent
TWA
E [kJ/mol]
C-C distance [Å]



Molecular Dynamics

J. Am. Chem. Soc., **2021**, 143, 12212–12219.

Inference with Ontologies

- OWL (**W**eb **O**ntology **L**anguage)
  - a very expressive language to describe ontologies
  - e.g. “sameAs”, “equivalentProperty”, “cardinality”, symmetric / reflexive / transient relations
- SWRL (**S**emantic **W**eb **R**ule **L**anguage)
- Semantic Reasoner, e.g.
  - to infer new facts
  - to proof data consistency
  - to apply business rules

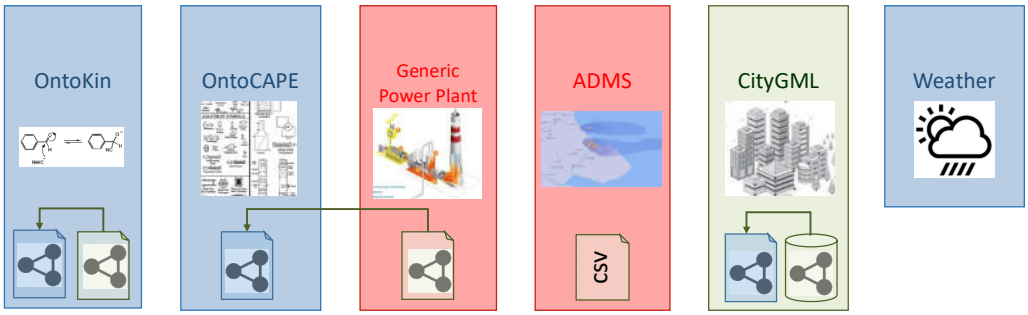
No silver bullet

- Use of ontologies and semantic technologies is no silver bullet
  - There might be trade-offs
  - Efficient use depends on requirements and on the nature of the problem
- Ontologies and semantic technologies complement rather than replace existing data management infrastructures
  - Integration with heterogeneous data sources possible

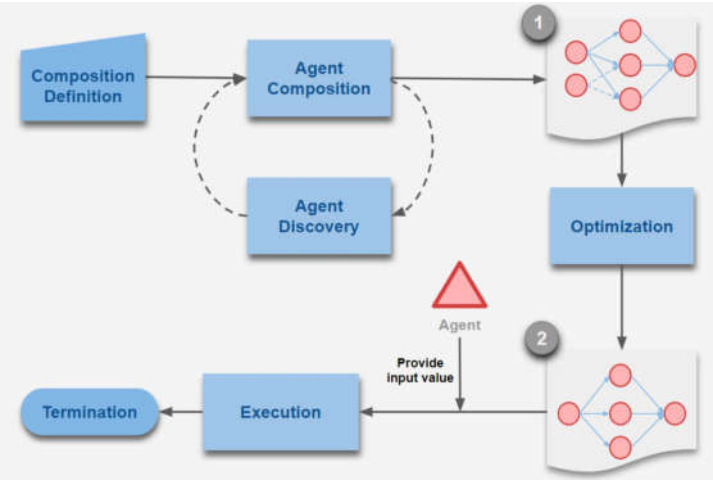
Agent Types

- **Type 0** agents operate at the **real-world boundary** of JPS
  - e.g. input agents (from sensors, users), output agents (to actuators, browser)
- **Type 1** agents query, calculate, and **update** the knowledge graph
  - e.g. querying OntoKin knowledge base, calculating and updating emission stream, simulating atmospheric dispersion
- **Type 2** agents **restructure** the knowledge graph
  - e.g. restructuring the heat waste network of Jurong Island
- **Type 3** agents **integrate** new ontologies and linked open data
  - e.g. integrating DBpedia with the help of natural language queries
- **Type 4** agents **create new agents** and provide higher-level services
  - e.g. composed agents, scenario-based analysis, optimization on superstructures

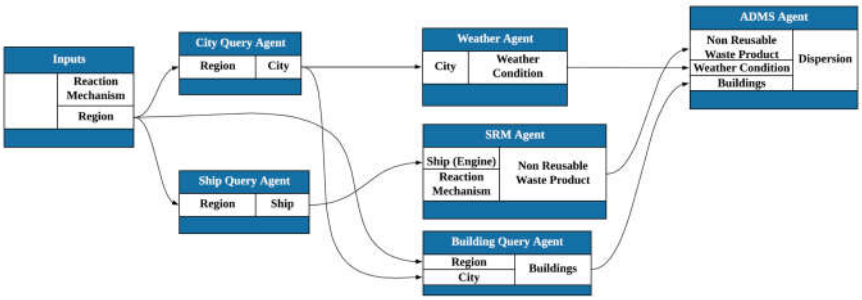
Emissions - power plants and ships



Semantic Agent Composition Framework



Semantic Web Service Composition for Cross-Domain Dispersion Scenario



City Query Agent

City Query Agent	
Region	City

Selection of region on google map

Google Geocoding API

Google Maps Platform

Name of City "Berlin"

DBpedia

URI of City e.g. http://dbpedia.org/resource/Berlin

Building Query Agent

Building Query Agent	
Region	Buildings
City	

Pre-conversion from CityGML data to OntoCityGML and loaded into triple store.

OGC Making location count.

CityGML

OntoCityGML

Selection of region on google map

URI of City e.g. http://dbpedia.org/resource/Berlin

Triple Store

URI of building instances with the region e.g. http://www.theworldavatar.com/kb/deu/berlin/buildings/3920_5818.owl#BuildingBLDG_00030009006684ea

Weather Agent

Weather Agent	
City	Weather Condition

URI of City e.g.
http://dbpedia.org/resource/Berlin

Real-time weather data for selected city which includes:

1. Temperature
2. Wind Speed
3. Wind Direction
4. Cloud Cover
5. Relative Humidity
6. Precipitation

Ship Query Agent

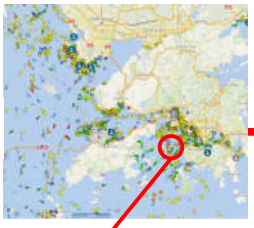
Ship Query Agent	
Region	Ship

Web scrapping is executed periodically to acquire updated AIS data for ships within pre-defined regions. Acquired AIS data is stored in PostgreSQL database.

Selection of region on google map

PostgreSQL Database

URI of ship instances with the region e.g.
http://www.theworldavatar.com/JPS_SHIP/rest/ships/477996210



AIS DATA	
AIS Type	Passenger ship
Flag	Hong Kong
Destination	CENTRAL-CHEUNG CHAU
ETA	-
IMO / MMSI	- / 477995507
Callsign	VRS4424
Length / Beam	-
Current draught	2.2 m
Course / Speed	229.1° / 14.8 kn
Coordinates	22.26892 N/114.08425 E
Last report	Sep 24, 2019 02:40 UTC

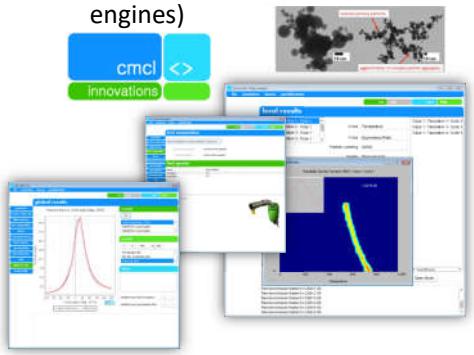
SRM Engine Suite Agent

URI of ship instances with the region e.g. http://www.theworldavatar.com/JPS_SHIP/rest/ships/477996210

URI of the reaction mechanism e.g. http://www.theworldavatar.com/kb/ontokin/Toluene.owl#ReactionMechanism_187077735769001

Stochastic Reactor Model

(commercial software for the simulation of exhaust emission from internal combustion engines)

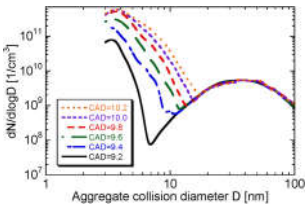


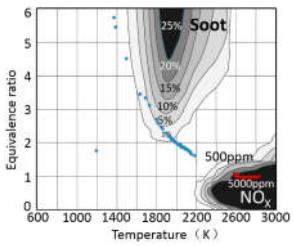
Individual simulated exhaust emission from ship which is updated in individual ship's funnel URI e.g. http://www.theworldavatar.com/kb/ships/Chimney-1.owl#WasteStreamOfChimney-1

Engine Combustion (SRM) Agent

Stochastic Reactor Model (SRM) to simulate in-cylinder combustion:

- Fuel injection
- Turbulent mixing
- Detailed chemistry
- Soot formation



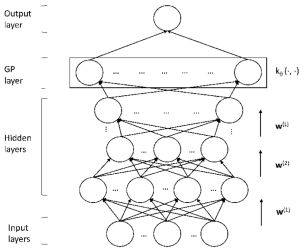


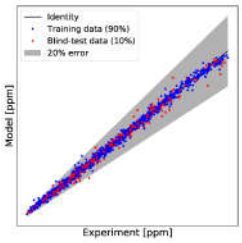
Model outputs engine performance as well as emissions of:

- CO, CO_2
- unburnt hydrocarbons
- NO, NO_2
- Particulate matter (size distributions)

Speed Load Map Agent

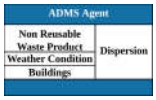
- Fast-response surrogate model suitable for real-time simulations
- Purely data-driven, either experimental or simulated data
- Deep Kernel Learning (DKL), combination of deep neural network and Gaussian process





- NO_x and soot emissions as function of engine operating conditions, *i.e.* engine speed (in RPM) and load/torque (in Nm)
- Empirical model to relate vessel speed (in knots) to engine operating parameters

ADMS Agent



Ship's funnel URI e.g. http://www.theworldavatar.com/kb/ships/Chimney-1.owl#WasteStreamOfChimney-1

Building URI e.g. http://www.theworldavatar.com/kb/deu/berlin/buildings/3920_5818.owl#BuildingBLDG_00030009006684ea

Complex Terrain Profile

Real-time weather data

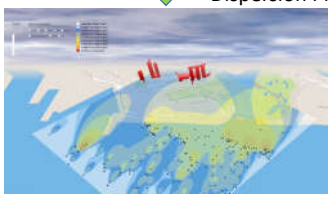
Selected Region (coordinates)

Background concentration

ADMS

(commercial software: advanced dispersion model to model the air quality impact)

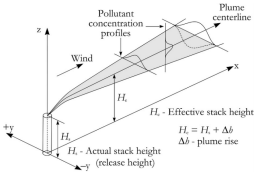
Dispersion Profile



ADMS Agent

Simulate atmospheric dispersion with **ADMS-5** from CERC

Current model is based on a **Gaussian Plume Dispersion**



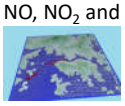
$$\bar{c}(x,y,z) = \frac{Q}{2\pi\sigma_y\sigma_z\bar{u}} \exp\left(\frac{-y^2}{2\sigma_y^2}\right)\left(\exp\left(\frac{-(z-h)^2}{2\sigma_z^2}\right) + \exp\left(\frac{-(z+h)^2}{2\sigma_z^2}\right)\right)$$

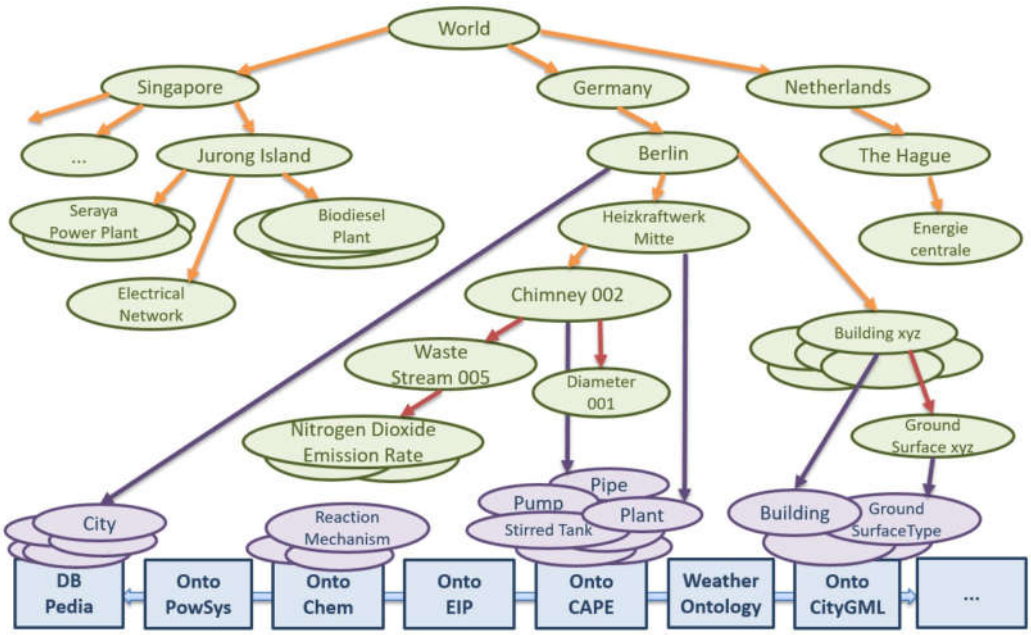
CERC

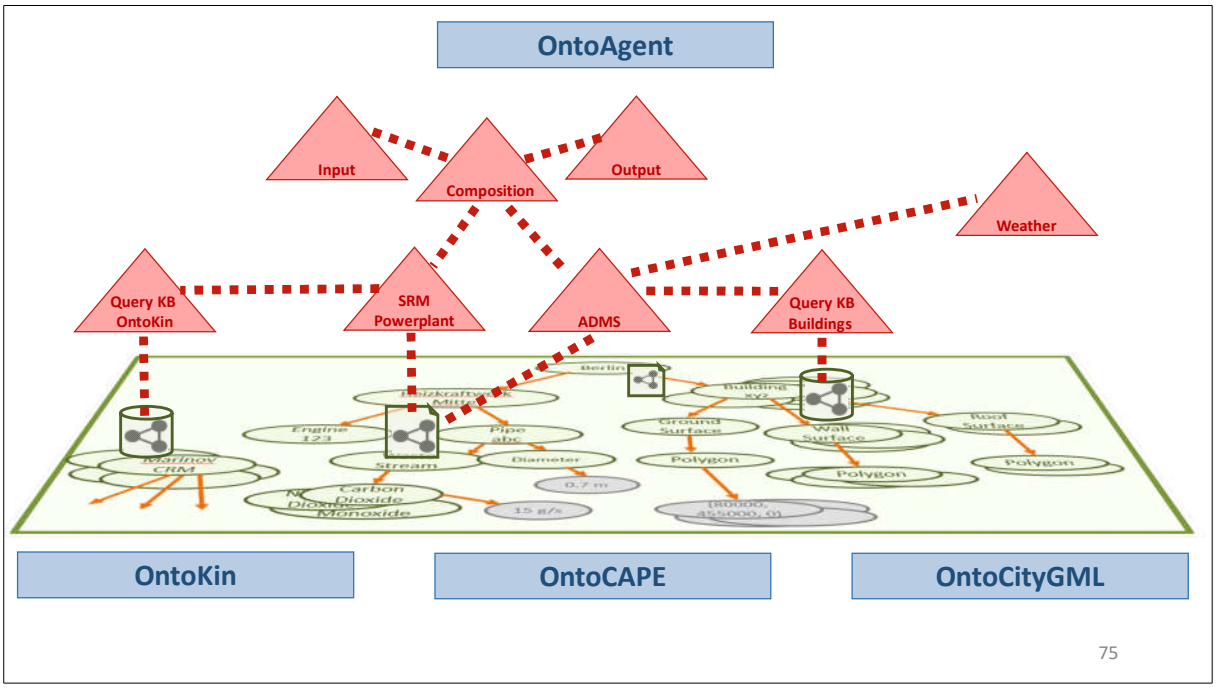
Model also considers the effects of:

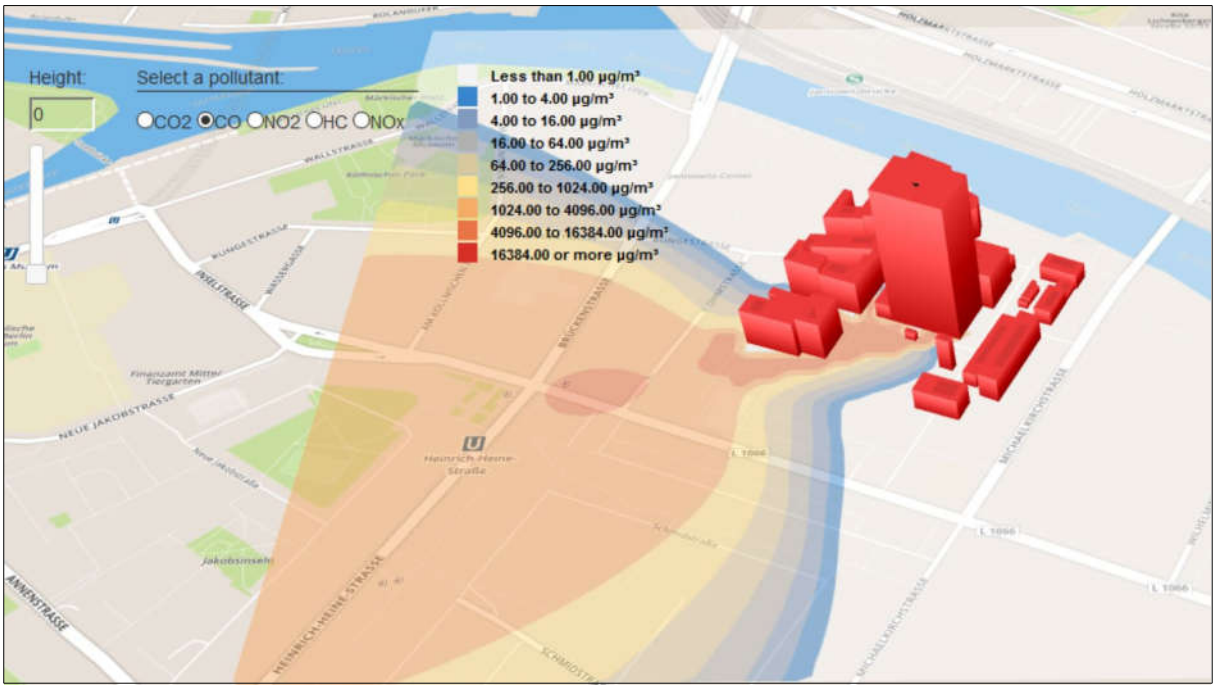
1. Wet deposition
2. Atmospheric chemistry between NO, NO_2 and O_3
3. Buildings
4. Complex terrain
5. Plume visibility

$$NO_2 + h\nu \rightarrow NO + O_3$$
$$NO + O_3 \rightarrow NO_2$$

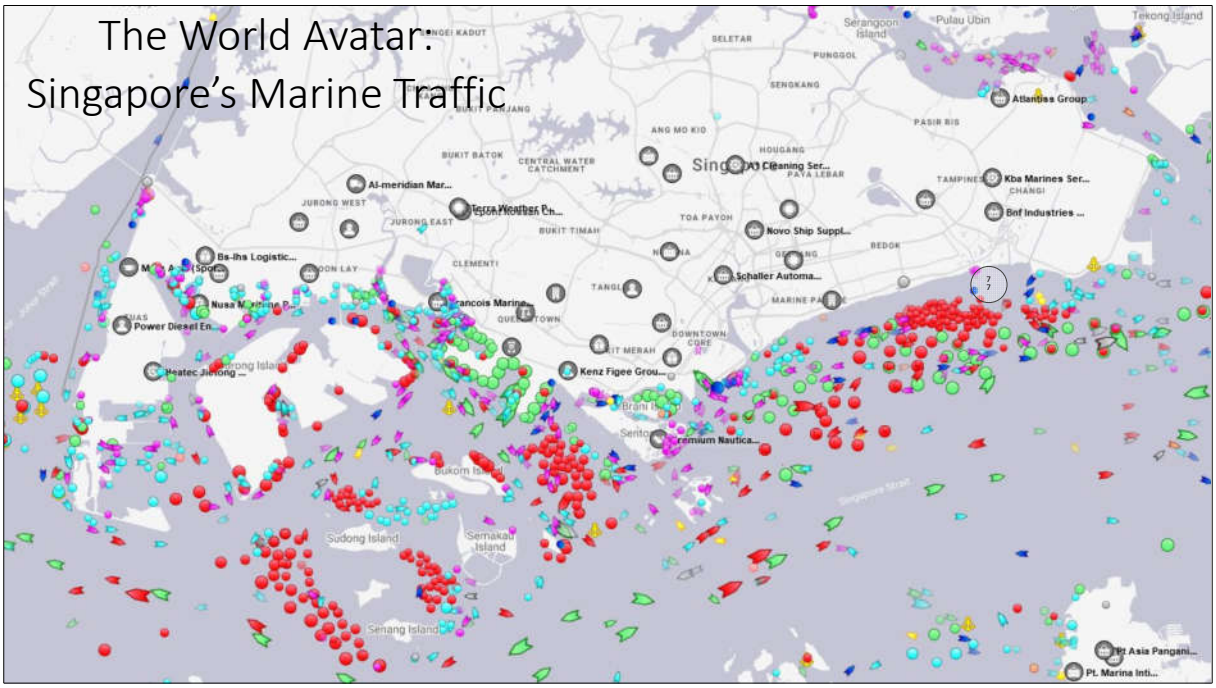




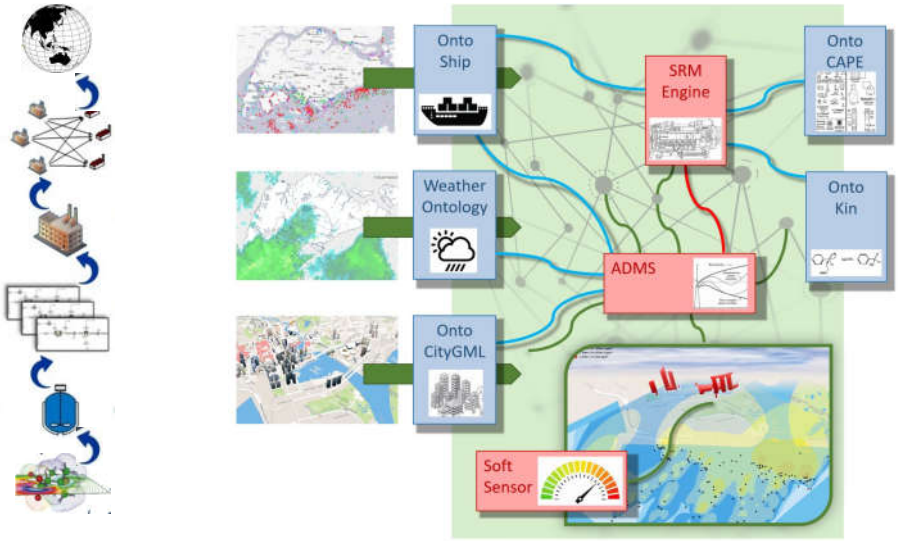

OntoAgent
Input
Composition
Output
Weather
Query KB OntoKin
SRM Powerplant
ADMS
Query KB Buildings
Engine 123
Pipe abc
Stream
Diameter
0.7 m
Carbon Dioxide Monoxide
15 g/s
Ground Surface
Polygon
Wall Surface
Roof Surface
Polygon
Polygon
OntoKin
OntoCAPE
OntoCityGML


Height:
0
Select a pollutant:
CO2 CO NO2 HC NOx
Less than 1.00 µg/m³
1.00 to 4.00 µg/m³
4.00 to 16.00 µg/m³
16.00 to 64.00 µg/m³
64.00 to 256.00 µg/m³
256.00 to 1024.00 µg/m³
1024.00 to 4096.00 µg/m³
4096.00 to 16384.00 µg/m³
16384.00 or more µg/m³

The World Avatar: Singapore's Marine Traffic



Cross-Domain Dispersion Scenario



End of Lecture 6