Accelerated molecular dynamics simulation method development and applications in polymer systems

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN (Dr. rer. Nat.)

von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)
angenommene

DISSERTATION

von

Liu, Modan

Tag der mündlichen Prüfung:	30.10.2020
1. Referent:	Prof. Dr. Wolfgang Wenzel
2. Korreferent:	Prof. Dr. Jörg Schmalian

Erklärung zur Selbstständigkeit

Ich versichere, dass ich diese Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht und die Satzung des KIT zur Sicherung guter wissenschaftlicher Praxis in der gültigen Fassung vom 01.07.2017 beachtet habe.

Karlsruhe, den 22.09.2020, _____

Modan Liu

Als Ansichtsexemplar genehmigt von

Karlsruhe, den 22.09.2020, _____

Prof. Dr. Wolfgang Wenzel

Contents

1. Preface **1**

2. Introduction **3**
2.1. Molecular dynamics 3
2.2. Thesis outline 8

3. Accelerating molecular dynamics via locally softened potentials **11**
3.1. Introduction 11
3.1.1. Accelerated MD simulations via enhanced sampling methods 11
3.1.2. Machine learning 13
3.1.3. Genetic algorithm 15
3.1.4. Efficiency of parameter space exploration and hand designed rules . 16
3.2. Methods 16
3.2.1. The adaptive management of the softened pair-wise potential 17
3.2.2. The transition between the vanilla Lennard-Jones and the softened potential 19
3.3. Acceleration benchmark in a grain boundary movement problem 24
3.3.1. Approaching the quickest shrinking path: the sandwich problem . . 27
3.3.2. Limitations in the selection criterion: the domain shrinking a diamond shaped smaller grain and in a random grid of smaller grains 30
3.3.3. Conservation of thermodynamic order parameters: overall potential energy and Q6 34
3.3.4. Conclusion and outlook 36

4. High-throughput MD simulations of IL diffusion in MOFs **39**
4.1. Introduction 39
4.1.1. Ionic liquids 39
4.1.2. Metal-organic frameworks 40
4.1.3. [BMIM], [NTf_2], and Lithium ions embedded in the HKUST-1 and UiO family MOFs 41
4.2. SURMOF synthesis: layer-by-layer growth of MOF, loading of IL@MOF, and conductivity measurement from experiments 43
4.3. Workflow design for high-throughput IL@MOF MD simulations 47
4.4. Bulk phase [BMIM][NTf_2] versus [BMIM][NTf_2]@HKUST-1 53
4.4.1. Density-dependent bunching and immobilization of confined IL . . . 54
4.4.2. The traffic jam: molecular dyanmics insights from tracer trajectories 55
4.5. IL transport in UiO family MOFs 58
4.5.1. Tunable pore geometries in UiO-66, 67, and 68 and tunable critical IL bunching 58
4.5.2. Two-stages traffic jam: spontaneous symmetry breaking in ions' path finding 59

4.6. Li-doped-IL@HKUST-1 . . . 63
4.6.1. Bi-modal conduction of Li-ion in IL: small size and alternative diffusion routes versus strong ion pairing . . . 63
4.6.2. The Grotthuss-like conduction of Li-ion . . . 64
4.6.3. Fast and well-tempered: the traffic map for Li-ion diffusion . . . 66
4.7. The IL transport in SURMOF vs in MOF pellets . . . 68
4.8. Conclusion . . . 69

5. Coarse-grained molecular dynamics of polymerisation and polymer folding 71
5.1. Introduction . . . 71
5.1.1. Direct laser writing . . . 71
5.1.2. Single-chain nanoparticles . . . 73
5.2. Methods . . . 74
5.2.1. One point for one monomer: Langevin dynamics . . . 75
5.2.2. Dissipative particle dynamics in SCNP folding . . . 76
5.2.3. FENE bonds and SRP potential to prevent bond crossing . . . 78
5.2.4. Stochastic process of dynamic bond recombination . . . 78
5.3. 3D printing a polymer network . . . 79
5.3.1. Coarse-grained building blocks, the reaction scheme, and printing 3D features via voxel stacking . . . 79
5.3.2. Monomer conversion kinetics and efficiency tuning . . . 82
5.3.3. Mechanical strengths characterisation of the printed network . . . 84
5.4. Folding a single-chain nanoparticle . . . 87
5.4.1. Photodimerisation of styrylpyrene containing monomers: in the solution vs on the polymer . . . 87
5.4.2. Intermittent folding features and the photostationary state in the reversible folding . . . 90
5.5. Conclusion . . . 94

6. Summary 97

Acknowledgements 101

Appendix 103
A. UFF4MOF parametrisation for HKUST-1 . . . 103
B. UFF4MOF parametrisation for UiO-66, 67, and 68 . . . 104
C. OPLS-AA compatible parameters for Li-ion, [BMIM], and [NTf2] . . . 105
D. DPD, FENE, and SRP parametrisation for SCNP folding . . . 107
E. Additional figure for reversible folding of **P1'** into **SCNP1'** . . . 108

Bibliography 109

List of Figures

1.1. COVID-19 virus from TEM imaging and its protease as a digital model. . . 1

2.1. The Lennard-Jones interaction. 6
2.2. Periodic boundary conditions in a 2D representation. 7
2.3. A brief flowchart for molecular dynamics 8

3.1. Enhanced sampling scheme to prevent kinetic trapping in metadynamics-like algorithms. 12
3.2. Incomplete information can be inferred from a lower resolution, pixelated image. 13
3.3. A generic deep learning enabled neuron network. 14
3.4. Operations within an iteration in the genetic algorithm. 15
3.5. Genetic algorithm iteratively optimizes the design of an antenna. 16
3.6. Relative movement between pairs. 17
3.7. Radial distribtuion fuction of an imperfect Lennard-Jones crystal. 18
3.8. Distribution profile of the projected kinetic energy in the an imperfect Lennard-Jones crystal. 19
3.9. Softened potential compared to 12-6 Lennard-Jones. 20
3.10. Flowchart of the accelerated algorithm. 21
3.11. Flowchart of the hybrid voting optimisation of parameter setups. 23
3.12. Workflow of Phoenics algorithm. [31] 24
3.13. The grid search in the parameter space visualised for 3 major components r_{switch}, ϵ, and N_{total}. 25
3.14. Learning curve and accuracy of the deep learning (DL) module. 26
3.15. The grain boundary movement of the rectangular smaller grain. The smaller domain, marked by red color, shrinks in size before eventually vanishes as the dislocation and misalginment of the atoms within the domain is realigned with the base grain. 28
3.16. Time series of the smaller domain size for 7300 unique parameter setups. . . 29
3.17. The distribution of performances of the accleration algorithm. 29
3.18. GBshrinking Phoenics 30
3.19. Snapshot in the process of shrinking a diamond shaped smaller grain. . . . 32
3.20. Snapshot in the process of domain shrinking in a grid of smaller grains of dimensions 20σx20σ. 33
3.21. Q6 orientation order parameter evolution with 10σx10σ smaller grains. . . . 34
3.22. Q6 orientation order parameter evolution with 20σx20σ smaller grains. . . . 34
3.23. Conservation of the overall potential energy in the smaller grain shrinking with respect to various control parameters. 35
3.24. Separate Q6 measurements for the whole system, the base grain, and the smaller grain. 36
3.25. Potential rescaling as the softening method for a generic form of pair-wise interaction. 37

4.1. Room temperature ionic liquid [BMIM][NTf_2] compared against a common salt of NaCl. . . . 40
4.2. Scanning electron microscope (SEM) image of a seed during MOF growth. . 41
4.3. The ionic liquid pair of [BMIM][NTf_2]. . . . 42
4.4. Schematic representation of layer-by-layer assembly of SURMOF. . . . 43
4.5. Side-view and top-views of the scanning electron microsope (SEM) images of the SURMOF HKUST-1. . . . 44
4.6. Scheme of the [BMIM][NTf_2]@HKUST-1 system. . . . 45
4.7. Infrared reflection absorption spectra of loading the IL into HKUST-1. . . . 46
4.8. QCM and EDX quantification of IL loading in the MOF. . . . 46
4.9. Nyquist plots for the impedance measurements. . . . 47
4.10. Workflow of high-throughput IL@MOF simulations. . . . 48
4.11. HKUST-1 and UiO MOF unitcells and atom type designations. . . . 50
4.12. Li-ion, [BMIM], and [NTf_2] in the IL and the corresponding atom type designations. . . . 51
4.13. Side-view and cross section patterns of the bulk phase IL transport under intense external electric field. . . . 54
4.14. Side-view and cross section patterns of the embedded IL transport under external electric field. . . . 54
4.15. Homogeneous flow with low IL loading and bunched layer with high IL loading. 56
4.16. Molar conductivity of [BMIM][NTf_2]@HKUST-1. . . . 56
4.17. Tracer trajectories of [BMIM] in [BMIM][NTf_2]@HKUST-1. . . . 57
4.18. View along the ion transport channel along the (110) direction in the UiO-67. 59
4.19. Mobility of [BMIM][NTf_2] in UiO-66, 67, and 68 in the (110) direction. . . . 59
4.20. Side-view of the ion tranpsort channel along the (001) direction in the UiO-67. 60
4.21. Mobilities of the IL transport along the (100) and (010) directions in the UiO MOFs. . . . 61
4.22. Single ion trajectories of bunched IL in UiO-66 with 60% loading and partial bunching in UiO-68 at 100% loading. . . . 62
4.23. Single ion trajectories of partial bunching of IL in UiO-68 at 80% loading and 100% loading. . . . 62
4.24. Conductivity of [BMIM][NTf_2] versus [$Li_{0.2}BMIM_{0.8}$][NTf_2] embedded in SURMOF HKUST-1. . . . 64
4.25. Conductivity of [$Li_{0.2}BMIM_{0.8}$][NTf_2]@HKSUT-1 from MD simulations. . . . 65
4.26. The localization of Li-ion near anion complexes. . . . 66
4.27. The Grotthuss-like conduction of Li-ion. . . . 67
4.28. The trajectories of individual [BMIM] vs Li-ions. . . . 68
4.29. Conductivity of [$Li_{0.2}BMIM_{0.8}$][NTf_2] in HKUST-1 pellets. . . . 69

5.1. The 3D printed sculpture on a human hair, the two-photon absorption, and the printing voxels. Image sources: 3D printed sculpture by Jonty Hurwitz, Karlsruhe Institute of Technology, and the two-photon absorption image by Steve Ruzin and Holly Aaron, UC Berkeley. . . . 72
5.2. Folding a polymer coil into SCNP. . . . 74
5.3. Showcase of 3D printing the KIT logo in a thin 3D layer. . . . 76
5.4. The PETA monomer used in the 3D polymer printing. . . . 79
5.5. Reaction schemes in the 3D polymer printing. . . . 80
5.6. Snapshots during the 3D priniting of a pyramidal structure. . . . 82
5.7. Short vs long exposure on the monomer conversion degree. . . . 83
5.8. The comprehensive effect of laser power and exposure time on the monomer conversion degree. . . . 84

5.9. Polymeric cubes printed from voxels with aspect ratio of 1:1:2 with variable laser powers. . . . 85
5.10. Stress-strain relationships of the printed cubes with respect to laser power. 86
5.11. Stress-strain relationships of the printed cubes with respect to exposure time. 86
5.12. Projected strians during the bending of printed cylindrical rods with variable aspect ratios. . . . 87
5.13. Photocycloaddition and reversion in the covalent bond recombinations of styrylpyrene. . . . 88
5.14. Experimental setup of the SCNP folding. . . . 89
5.15. DPD coarse-graining scheme of the monomers and the THF solvent. . . . 89
5.16. The synthesis and characterisation of the polymer with mixed MMA and **M1** units. . . . 91
5.17. The UV/vis spectra and the number of open bonds of folding **P1** into the photostationary state (d-f). . . . 92
5.18. The schematic representation, UV/vis spectra, and the number of open bonds of partially unfolding **SCNP1** into the photostationary state (a-c), and folding of **P1** into the photostationary state (d-f). . . . 92
5.19. The ratio of open bonds and the radius of gyration during folding **P1** into **SCNP1**. . . . 95

6.1. A schematic representation of simulation and experiment methods in a multi-scale hierarchy. . . . 97
E.1. The ratio of open bonds during folding **P1'** into **SCNP1'**. . . . 108
E.2. The ratio of radius of gyration during folding **P1'** into **SCNP1'**. . . . 108

List of Tables

3.1. List of parameters for accelerated molecular dynamics with softened potentials. 22
3.2. Top 10 performing parameter setups in the rectangular grain boundary shrinking with DL-GA-rules sampling. . . . 30
3.3. Top 10 performing parameter setups in the rectangular grain boundary shrinking with Phoenics sampling. . . . 31

4.1. Pore specifications of the SURMOFs of HKUST-1 and UiO-66, 67, and 68. . 42
4.2. Statistics of number concentrations at 100% loading of IL@MOF. . . . 52

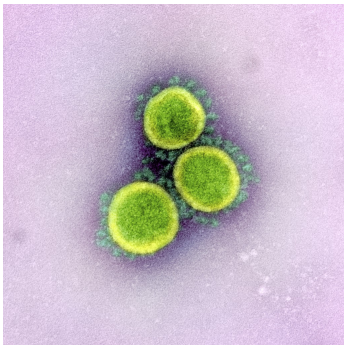
5.1. Coarse-grained particle types assignment in the simulation of direct laser writing. . . . 75
5.2. Summary of control paramters in the 3D polymer printing. . . . 81
5.3. Statistic of polymer systems **P1** and **P1'**. x, and y correspond to the number of styrylpyrene containing **M1** units, and the base MMA units. . . . 88
5.4. Bond recombination parameters of the SCNP folding . . . 93

A.1. UFF4MOF parametrisation of HKUST-1. . . . 103
B.2. UFF4MOF parametrisation of UiO family MOFs. . . . 104
C.3. OPLS-AA compatible parametrisation of Li, [BMIM], and [NTf2] I. . . . 105
C.4. OPLS-AA compatible parametrisation of Li, [BMIM], and [NTf2] II. . . . 106

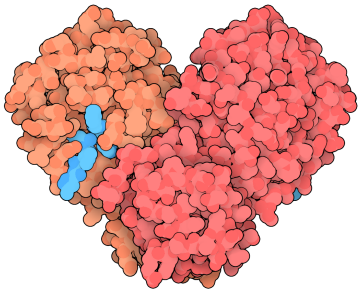
D.5. DPD, FENE, and SRP parameterization of the SCNP folding 107

1. Preface

As we are bogged in the outbreak of COVID-19, or SARS-CoV-2, since the spring of 2020 [1], there is no greater demand than a fast computer simulation to help with the race against the coronavirus in saving human lives. Personally, this is my second experience of a coronavirus outbreak following the vivid memories of cities lock-down, school-shutdown, the smell of disinfectant sprays, the sight of everybody behind face-masks, and a suddenly increased enthusiasm for physical exercise, when SARS hit China back in 2003. During the writing of this thesis, our understanding of COVID-19 has been constantly evolving. From the first few days immediately after the discovery of the new virus, it has been possible to isolate and capture the individual virion experimentally, shown in Figure 1.1a. After the isolation of the virus, a 3D reconstruction for the molecular geometries based on the imaging had been accomplished, while the sequence analysis for the RNA building blocks made it possible to establish a comprehensive digital model for the virus. [2] In general for the coronavirus class, the RNA virus translates itself into reproducing not only more building blocks to spread the disease but also the very building tools. One of the key enzymes that attracted principle attention is the protease, represented in Figure 1.1b, as a target protein for drug development. [3]



(a) Transmission electron microscope (TEM) image of the COVID-19.* The separation of the virus from its host cell is achieved within days' time since the discovery of the disease, before a collection of analysis tools can provide information as its RNA sequence and the conformation of the protein it synthesizes. The diameter of the virus is approximately 100nm.



(b) The protease of COVID-19 represented as a digital model. [4] The viral protease is an essential enzyme for cleaving the precursor viral polyprotein and thence functionalised. If deactivated by the protease inhibitor, which preoccupies the active site on the enzyme, shown in light blue, the virus can no longer reproduce.

Figure 1.1.: COVID-19 virus from TEM imaging and its protease as a digital model.
* Image license: CC-O (public domain; ID: 2294135), original source: US Government departement, the National Institute of Allergy and Infection Diseases.

In an effort to disrupt the reproduction of such RNA virus or disable its functionalities of causing symptoms, molecular dynamics is crucial in understanding how these building

blocks and building tools interact with each other. Similar strategies apply also to other drug design applications, e.g. development of the inhibitor against the malicious protein folding which triggers Alzheimer's. [5] Understanding of molecule dynamics in an atomistic world can be also applied to applications in inorganic materials as alloys [6] and semiconductors [7], and organic materials as polymers, discussed in the following chapters, while analogies can be even drawn further to active materials [8], and collective behaviours [9], bridging the gap between the atomic level insights to the world of statistics.

Despite of the fact that molecular dynamics experts today possess the greatest computing power ever since its inauguration, computing resources are nevertheless inadequate. Again in the case of COVID-19, experiment reveals that the virus consists of a single RNA strand bound to nucleocapsid protein to establish a combined particle size of around 70 to 90nm, and individual spike proteins, mimicking the shape of a crown and hence the name "corona", which make contact to the infected host cell, each subunit weighing around 50kDa. The human angiotensin-converting enzyme 2 (ACE2) receptor which is known to bind with COVID-19 also weighs in the same order of magnitude at around 90kDa. As a result a scaled-down simulation for the potential mechanism of isolated spike protein and the ACE2 combination involves a system weighing at least around 100kDa, bar the solvent environment. A state-of-the-art supercomputer dedicated to molecular dynamics resolves for protease of COVID-19, one of the smallest scales in the drug design, in some 10 microseconds scale [10], while typical timescales in the laboratory operations are much longer as RNA tests for infection require tens of minutes, individual cell experiments takes hours, epidemiology report updates are based on days, and new samples from vaccine test stages reside in the scale of weeks. There is a huge gap between the computer simulations against what can be measured and manipulated in the real world.

A clear motif echoing from 70 years back is: molecular dynamics is very powerful, but it needs to be faster!

2. Introduction

Materialised are the conceptual understandings of the objective world when the duo of theory and experiment, in between the minds (hearts) and the hands, systematically constructs our knowledge. The two pillars support the well-established scientific paradigms have had been proved helpful in guiding the applications where new materials can be designed, synthesised, and put to good use, there is however the rising data science which leads to a revolutionary paradigm shift, as a third pillar which contributes with ever increasing significance to our cognition of the world, both from the perspectives of natural as well as social science [11, 12].

One of the ultimate goals of the data-based approach is to replicate the real world in the computer as a "digital twin". The accurate digital model with complete atomistic information allows for investigation of the material properties via fast computer simulations with unprecedented scalability as opposed to conventional experiments for the establishment of the qualitative structure-property relationship (QSPR). In molecular design and screening problems, the QSPR corresponds to the individual control elements, e.g. the singular atomistic defects, side groups, alloy mixing ratios, which can be modified upon the desired molecular structure whereas the overall change in the physical, chemical, or biological properties is varied respectively, e.g. the activity and selectivity of specific catalysts, before the data is validated from the experimental side. [13, 14] The establishment of the QSPR benefits all the 3 pillars in science of the theory, the experiment and the computer simulation. When computer simulation serves as a quick screening tool that propose experiments, from the novel discoveries in the experiment and the data analysis the theory is updated, and the break-through in theory provides improved models where new simulation methods are based on.

The computers, as a man-made tool, provides unprecedented power of abstraction in encoding arbitrary volume of information, as well as an overwhelmingly high volume of real-time data flow as an input, before yielding an output of wholistic information which exceeds human cognitive capacity. In particular for the digital replication of the world made out of atoms in the molecular dynamics simulations, the output is a wholistic movie about atoms presented to human eyes.

2.1. Molecular dynamics

As a brief retrospective, before the introduction of molecular dynamics (MD), two of the greatest directionss in which theoretical physics has been developing well are statistical mechanics, typically for the study of ideal gases, which offers great understanding of thermodynamics and phase transitions that in turn lead to invention of combustion engines, refrigerators, and etc., plus the enhanced understanding of the subatomic world from the perspective of quantum mechanics, which coupled well with the world of chemistry and led to promotions in fertilizer industry, modern electronics, etc. However, the gap between properties of individual atom groups and the overall statistical behaviours of an ensemble

of molecules cannot be closed merely through derivation of analytical formulae. From an alternative approach both in physics and mathematics on the study of the few body movements, the approach of dynamic system has been established in the early years of the 20th century. The evolution of a complex system is characterised by deterministic rules based on the instantaneous state, while the movement is solved via time-integration in revealing the whole trajectory. Defined as a dynamical system, a finite number of coupled ordinary differential equations can be utilized to describe the evolution of states that promoted the understanding of the stability of the likes of a 3-body system [15] or the population evolutions in an ecological system [16], however for many-body systems, the coupling between the ODEs are unintuitive to define and extremely expensive in the calculation. From the 2nd half of 20th century, molecular dynamics simulations started getting to be numerically achievable owing to the major progress in the performance of computers and the development of improved time-integrators, and immediately proved to be a powerful tool for studying atomistic motions for a variety of physical, chemical and biological systems.

Based on Newtonian equations of motion, molecular dynamics solves many-body differential equations for a many-body system numerically. For a Newtonian particle, its degrees of freedom are limited to coordinate and velocities, whose evolution can be solved via integration over pair-wise contributions for all interacting pairs iteratively following:

$$m\mathrm{d}^2/\mathrm{d}t^2 x = F(x) = -\nabla V(x) \tag{2.1}$$

where m the mass of individual atoms, and x the coordinates, while $F(x)$ and$V(x)$ are both coordinate-dependent, corresponding respectively to the forces that act upon the individual atoms and the interaction potential. For discrete timesteps, the individual atomic coordinates are updated subjected to the forces acting upon the atoms, and the force in turn is determined by the conformation that takes into account of relative arrangements of the atomic coordinates, forming an asynchronous, iterative, self-consistent protocol with discrete timesteps.

The determination of interaction potential V follows problem-specific force fields (FF). Generally speaking, a force field incorporates both the physical model and the corresponding parameterisation which includes non-bonded and bonded interaction terms in an atomistic system. Non-bonded contribution typically corresponds to van der Waals and Coulomb interactions, while bonded interactions refer to intra-molecular contributions of chemical nature. Force-field model and parameters are usually fitted from experimental measurements, or developed empirically fitting to established models, when some force-field are derived from more fundamental computer simulations, e.g. first-principle calculations for the atomistic force-field and all-atom MD simulations for the coarse-grained force-field.

Here, to serve as an example the optimized potentials for liquid simulations (OPLS) [17] is formulated as:

$$E_{\mathrm{OPLS}} = E_{\mathrm{nonbonded}} + E_{\mathrm{bonds}} + E_{\mathrm{angles}} + E_{\mathrm{dihedrals}} \tag{2.2}$$

where the individual terms are:

$$
\begin{aligned}
E_{\text{nonbonded}} &= \sum_{ij}\left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{C_{ij}}{r_{ij}^{6}} + \frac{q_i q_j e^2}{4\pi\epsilon_0 r_{ij}}\right) \\
E_{\text{bonds}} &= \sum_{\text{bonds}} K_r(r - r_0)^2 \\
E_{\text{angles}} &= \sum_{\text{angles}} k_\theta(\theta - \theta_0)^2 \\
E_{\text{dihedrals}} &= \sum_{\text{dihedrals}} \{V_1/2[1 + \cos(\phi - \phi_1)] + V_2/2[1 - \cos(2\phi - \phi_2)] \\
&\quad + V_3/2[1 + \cos(3\phi - \phi_3)] + V_4/2[1 - \cos(4\phi - \phi_4)]\}
\end{aligned}
$$

where the pair-wise distances r_{ij}, the bond lengths r, the angles between bonds θ, and the dihedral angles ϕ provide coordinate information of the current state.

The parameters for individual atoms, bonds, and torsional potentials are dependent on specific type designations encoded in the force-field model, e.g. in the OPLS-AA force field the tetrahedrally bonded carbon can be designated as C_{139}, while the aromatic carbon is designated as C_{145}. Between different atom types the mixing rules of the nonbonded part is defined as $A_{ij} = \sqrt{A_{ii}A_{jj}}$ and $C_{ij} = \sqrt{C_{ii}C_{jj}}$, i.e. the geometric mixing. [18, 19] Speical rules are applied in the force field models as in OPLS while the van der Waals interaction between different types are derived using the mixing rule, in a bonded conformation the 1-4 interaction for atoms on the farside of a bonded pair, the nonbonded interaction is scaled to half. In the nonbonded energy, the Coulomb interaction subjected to the point charges, usually assigned to the atoms in the form of partial charges as q_i and q_j, and the permittivity ϵ_0 of the environment. The bond and torsional potentials are harmonic, characterized by K_r and k_θ as the stiffness, when r_0 and θ_0 denote the equilibrium bond length and angle. The dihedral potential corresponds to the torsional stress between 4 atoms among which 2 is bonded in a shared connection. For the bonds that are not shared, the 4 atoms are divided into 2 groups between which the torsional potential is approximated via a series of 4 sinusoidal terms with incremental orders. In classical molecular dynamics, the topology of the molecules remain unchanged throughout the simulation, i.e. chemical reactions are not taken into account. For predefined bonds, the bonding potential dictates the scale of the strongest interaction, which supplemented with the information of atomic masses, determines the fastest oscillation modes in the atomistic system. Meanwhile, although the torsional potential is relatively soft, they are essential in the description in the chirality of the molecules, e.g. in the induction of alpha helices in the protein secondary structure.

Force fields are typically optimized with specific target properties, while for similar systems, e.g. for proteins, the force field formulations and parameterisation can be similar, as is the case between the assisted model building with energy refinement (AMBER) force field and the general AMBER force field (GAFF) [20,21], along with the OPLS force field [17]. There are notable efforts of designing a universal force field (UFF) [22] which can be extended with specific concerns for a narrower range of applications: universal force field for metal-organic frameworks (UFF4MOF). [23] In principle a hybrid parameterisation with mixed force fields definitions is empirically applicable, while for each force field there are limitations due to biased optimization target in the design stage. The performance of the MD simulation depends heavily on the accuracy of force field parametrisation and development of a force field model requires effort on the scale of decades. In special cases, the chemical elements involved in the MD simulations span beyond the typical elements in organic materials as C, H, N, and O, and reach across the periodic table where the current force field model may not be applicable. Recently in an adaptive approach, force field development utilizing data from *ab-initio* calculations are gaining popularity together with machine learning algorithms to bridge between the sub-atomic and atomic level knowledge. [24]

Among a selection of different force fields, for the nonbonded interaction potential, the Lennard-Jones (LJ) interaction is formulated as the first 2 terms in the $E_{\text{nonbonded}}$ from Equation (2.2). Commonly used in the forms of 12-6, 9-6, or the Buckingham variant, the LJ interaction is such widely adopted for molecular dynamics that the two terms almost always are mixed in the usage. Additionally, the LJ potential also applies for alternative simulation methods as Monte Carlo. [25, 26] The LJ potential is termed as a summation of a strong repulsive interaction and a wealy attractive interaction. For short distances, the pair-wise repulsion subjected to Pauli exclusion overwhelms in the potential profile so that the atom, although in Newton's law is considered as a 0-dimensional mass point, preserves a finite 3D size. For long distances, the mutual attraction of pairs which corresponds to the combination of dipole-dipole, and induced dipole interactions and van der Waals forces encompassing the London dispersion, asymptotically reduces to zero when the pair separation goes to infinity. In between the repulsion and the attraction is an equilibrium distance, near which balance distance the pair-wise vibrations are analogous to harmonic oscillations subjected to springs. To generalize from the combined strong but short-range repulsion and the weak but slow-decaying attraction, the Lennard-Jones-like interaction potentials are applicable to an extended scope of physical, chemical, and biological systems, ranging from the adsorption free energy profile of small proteins on metallic surfaces, the molecular order, dynamics, and bulk properties of shaped particles, e.g. Gay-Berne potential for liquid crystals [27], and may remotely be correlated to the Hartmann model which describes the mechanical adhesion of bacterial communities. [28]

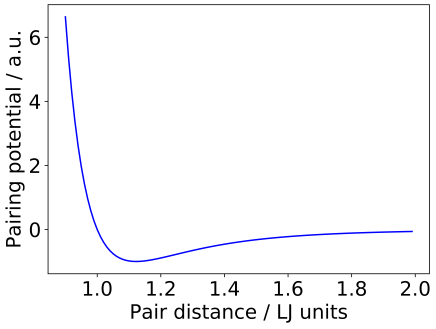


Figure 2.1.: The Lennard-Jones interaction. The widely-adopted Lennard-Jones interaction consitutes of a strong, short-range repulsion and a weak, longer-range attraction. Between the two an equilibrium separation is found at minimum potential.

Once the coordinate dependent pair-wise interaction potentials have been established, the forces acting on each individual particle changes their momentum via Newton's 2nd law of motion while the time integration of the Newtonian equations of motion of the whole dynamic system updates in turn the new coordinates hence a new coordinate dependent potentials. The time integration step in MD is often achieved in a velocity-verlet formulation:

$$\begin{aligned} \dot{x}_{t+\tau/2} &= \dot{x}_t + \frac{\tau}{2}\frac{F_t}{m} \\ x_{t+\tau} &= x_t + \tau \dot{x}_{t+\tau/2} \\ \dot{x}_{t+\tau} &= \dot{x}_{t+\tau/2} + \frac{\tau}{2}\frac{F_{t+\tau}}{m} \end{aligned}$$

Subjected to the time integration over discrete timesteps, the temporal resolution of MD is limited when the oscillation modes. The shortest timescale, which is determined by the most intense forces, must be separated from the integration timescale of τ so that the dynamics details is conserved. In the case of OPLS, one of the typical timescales of bonded

interaction is the C-H bond stretching at 1/(1E14 Hz) [29],which requires the integration timestep τ to be sub-femtosecond. In the trajectory of high frequency oscillations, each periodic cycle must be covered by multiple time-integration timesteps for the conservation of dynamics detail as well as the constraining the accumulation of numerical error.

In most cases, even though in MD it is possible to include atoms in the tens of thousands in the simulation, the system size is still not comparable to the lab conditions where the substance in investigation is usually counted in moles and weighted in grams. Length-scale wise, the dimensions of the simulation are also constrained to the order of nanometers. Specific to systems in investigation, the boundaries of the MD simulation can be a fixed wall with variable reflection potentials, typical for interfaces, or periodic, replicating the properties in bulk phase.

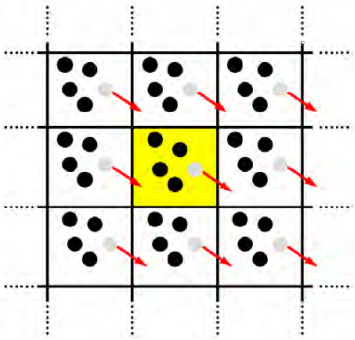


Figure 2.2.: Periodic boundary conditions in a 2D representation. The central colored box correspond to the current simulation domain whereas the mirrored simulation boxes expand infinitely. When an atom moves across the boundary, it reappears by making an entry from the mirrored position on the opposite edge of the simulation box, as in the game Pacman. [25]

Thermodynamics in MD is gauged by thermostats and barostats. In the most straight forward case the microcanonical ensemble (NVE) can be applied to the simulation, however compared to lab-controlled conditions, NVT and NPT ensembles are often adopted in simulations. When there is a selection of thermostat and barostats to choose from, the most widely used is the Nosé-Hoover thermostat and the Nosé-Hoover-Andersen barostat. The thermostat and barostat complies with the Boltzmann distribution by rescaling velocities of individual atoms to their thermo-equilibrium constantly which essentially for each atom the regularization is almost instantaneous. Such a strategy of applying thermostat enforces ergodicity for long time simulations while greatly reducing the degree of freedom by sampling. Alternative methods for temperature and pressure control allows involvement of dissipative forces, e.g. the Langevin thermostat and the dissipative particle dynamics.

The acquisition of information from the current simulation timestep to calculate properties and adjust for the desired ensemble average is the last step before the the MD simulation is propagated into the next timestep. The general protocol of MD for one timestep is to initialize, sample current state, time integrate, communicate across domains if necessary, apply temperature and pressure control, and propagate, illustrated in 2.3. As will be later shown, our customized algorithm for accelerated algorithm for molecular dynamics require modifications to multiple modules listed in the flowchart.

The state-of-the-art MD simulation packages, e.g. LAMMPS [30], GROMACS [32], and NAMD [33] are highly scalable via the implementation of the domain decomposition. For a whole molecular system, the spatial domain the molecular system occupies is divided into a grid of smaller subdomains. Within each subdomain the number of atoms is smaller when for pair-wise interactions, the neighbour list of each atom is greatly reduced. For

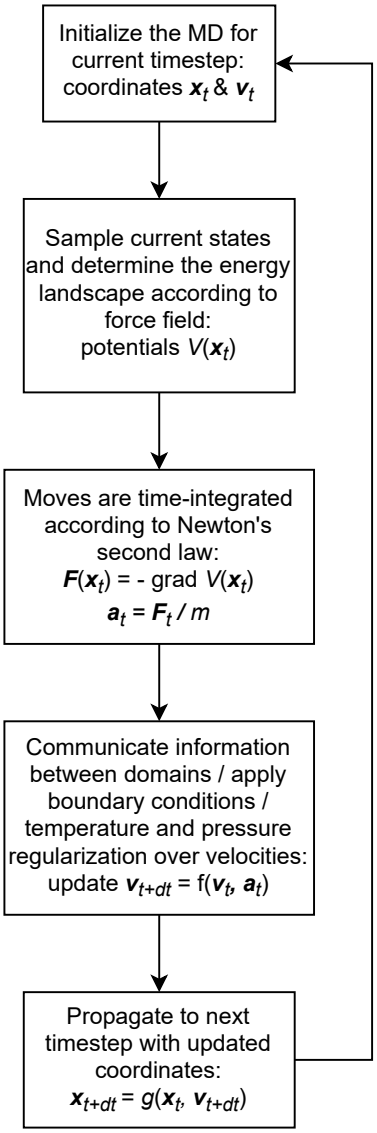


Figure 2.3.: A brief flowchart for molecular dynamics. The coordinates and velocities are updated for each timestep according to Newton's second law of motion when the force field, the boundary conditions, together with the additional image boundarys introduced from the domain decomposition, and the regulartization from thermostat and barostat, comprehensively define the updated coordinates and velocities.

each subdomain, the boundary includes a finitely extended region which the information of neighbouring subdomains are communicated across. Such a scalable architecture enjoys great advantages from the recent development of parallel computing units grid on GPU-based computing platforms when dedicated high performance computing machines, as was referred to in the preface, has been able to reach 1ms simulation timescale. [34]

2.2. Thesis outline

MD is an invaluable computer simulation tool in understanding how atoms interact with each other, when the generation of a motion picture for a collection of atoms facilitates a substantial foundation for the intriguing story-telling for the molecular modelling and dynamics. However, even with state-of-the-art, highly optimised simulation software running on dedicated hardware, the capability of molecular simulations is nevertheless restricted in the timescale far from lab timescale in routine chemical and biological studies.

Focused upon the story-telling via molecular dynamics, this thesis discusses several routes of advancing the performance of MD simulation from: improvement of the MD methodology, the automated, high-throughput performance of MD simulations, and the extension of MD timescale into laboratory relevant timescales via coarse-grained molecular modelling and MD simulations, in the meantime the thesis attempts to address upon the experiment / production guidelines for molecular design and synthetic control through the establishment of QSPR.

- **Chapter 2.** The MD methodology is improved via the implementation of the adaptive, locally softened pair-wise potentials for Lennard-Jones potential. The systematic optimization of for accelerated algorithm is achieved via a hybrid voting model as a combination of machine learning, genetic algorithms, and hand-designed rules. As a benchmark, the acceleration via locally softened potential is benchmarked in a grain boundary movement problem, where a global acceleration of 2x is achieved.
- **Chapter 3.** High-throughput MD simulation investigations are performed for the diffusion of ionic liquid embedded in metal-organic frameworks. As both ionic liquids and the metal-organic frameworks are both designer materials, the workflow for quick MD preparation, parametrisation, simulation, and information extraction in the *posteriori* analysis of both the separate components and the combined system are established. From the extracted data, the various diffusion mechanisms are revealed from different ionic liquid, dopant ion, and metal-organic framework combinations, when the simulation and experiment results show very good consistency. The QSPR of ionic liquid in nano-confinement opens pathways of designing selectively conducting / switching nano-supercapacitors and high-sensitivity probes.
- **Chapter 4.** Coarse-grained MD simulations applied to polymer system with empirical modelling and parametrisation for bond recombination reveals control strategy for light gated folding of single-chain polymeric nanoparticle and the mechanical properties of a 3D printed polymer network. Via coarse-grained molecular dynamics simulations, the timescale is extended to laboratory conditions when the quantitative predictions are shown to reproduce experiment results, when further predictions serve as guidelines in control of the polymer folding and polymeric printing.
- **Chapter 5.** Summary for accelerated MD methods and simulation protocols and the outlook of faster MD simulations.

3. Accelerating molecular dynamics via locally softened potentials

3.1. Introduction

Corresponding to how MD works and to which tasks MD has been applied to, a number of improvements to accelerate the simulations targets specific bottlenecks which limit the performance of MD. For the computational complexity from the shear number of pair-wise interaction calculations induced by the large number of atoms of the scale $O(N)$ in the atomistic system, the domain decomposition breaks them into smaller pieces of $O(n)$ to establish the $O(n^2)$ pair-wise interactions in the local vicinity. The assignment of atoms into local domains is further improved in calculation efficiency by slowly updating the neighbour list with curtailed the interaction potential via introducing a finite radius for the interaction sphere. The book-kept neighbour list facilitates a quick search mechanism, which reduces the computational efforts to approximately $O(N)$ for the look up effort when the neighbour list does not subject to frequent changes. For long-range interactions as Coulomb interaction, the Ewald summation based methods has been developed to integrate in the reciprocal space rather than the real space, taking advantage of the periodicity from image charges. [35, 36]

For the strong interaction potentials, one of the limitations resides in the small timesteps during the time integration due to the intrinsic time-scale, e.g. for the vibrational modes of the C-C bond. Nevertheless, the strong interaction potential establishes a high energetic barrier which is very difficult to overcome via the stochastic propagations regulated by the heat bath. As a result, the most of the computational capacity in the MD simulations is limited to local vibrations within the potential wells in the form of local vibrations and bond stretching. The limitation in the exploration in phase space due to the energetic barriers is known as kinetic trapping, whose prevention of collective movements, e.g. the formation of secondary structures in the protein, is one of the major obstacles that induces a gap between the MD simulation time-scale and the time-scale of interest.

Compared to alternative atomistic simulation methods, the kinetic trapping for individual atom movements is less of a problem in the Monte Carlo simulations. In Monte Carlo for each propagation only the potential difference between the start and finish states matters, the kinetic trapping in the transition states is promptly circumvented. However for collective movements the success rate of stochastic sampling drops exponentially with respect to the linear increase in the number of atoms. Meanwhile in the Monte Carlo simulations, when the exploration in the energy landscape is randomly probed between short-memoried attempts, the simulation time correspond to the stochastic sampling rather than the real timescale, hence the detailed dynamics is lost.

3.1.1. Accelerated MD simulations via enhanced sampling methods

To tackle the limitations induced by the kinetic trapping, a number of methods have been developed inspired by the Monte Carlo-like enhanced explorations in the energy

landscape. The coarse grained molecular dynamics method which treats groups of atoms as soft particles, allows overlapping of particles such that the pronouced energetic barrier of swapping of coordinates is lifted, promoting the particle movements. In both cases the energetic barrier is consistently down-scaled, such that the intrinsic time-scale is extended. For all-atom molecular dynamics where the details within atom groups cannot be disregarded, the accelerated MD (aMD) [37, 38] and metadynamics [39] have been established for fast exploration of the energetic landscape. In these methods, artificial propagations in the form of small energetic "kicks" are enforced upon the local potential wells. As the local minima gets filled, the molecule is released into a flattened energy surface as the original energy landscape gets incrementally stuffed by the artificial energetic profiles, before finding another energetic basin and repeat the local exploration. The additional energetic profiles are recorded such that for the calculation of physical observables, the states in the artificially biased energy landscape are reweighted to remove the artificial modifications of the energetic landscape. This process is termed as "back-mapping", after which the Boltzmann distribution is recovered from the enhanced sampling methods.

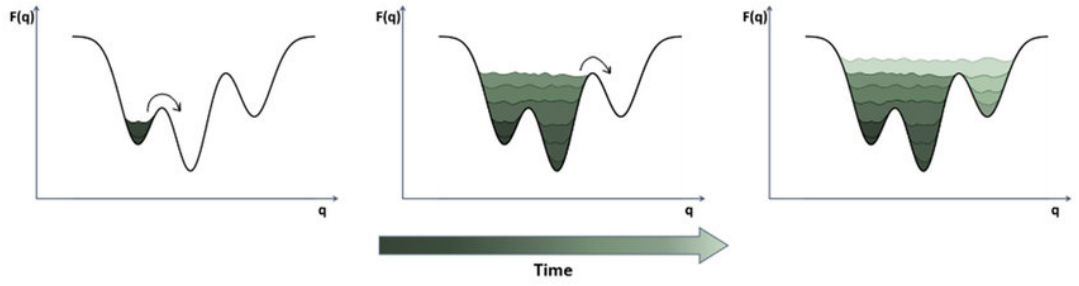


Figure 3.1.: Enhanced sampling scheme to prevent kinetic trapping in metadynamics-like algorithms. [40] In both methods with enhanced sampling, the acclerated molecular dynamics (aMD) [37, 38] and metadynamics [39] incorporate artificial modifications to the energy landscape for prevention of kinetic trapping in local minima, when the thermodynamics properties are recovered via the additional resampling through "back-mapping".

The enhanced sampling methods have been reported with great success in the conformational sampling for protein applications. From the literature, the aMD simulations routinely reaches the timescale of 1ms for small protein molecules, e.g. a bovine pancreatic trypsin inhibitor (BPTI) protein with 58 amino acid units. [41] In the applications of metadynamics simulation, a reaction coordinate have to be specified *a priori*, e.g. for the sampling of Fullerene-membrane interaction in terms of free energy, [42] making it generally faster than aMD simulations. However in both cases, due to the institutional manual input of information, either in helping all atoms from escaping the kinetic trapping, or in leading the molecule through a certain predefined trajectory along the reaction coordinate, the enhanced sampling methods under-perform in the conservation of detailed dynamics with a distorted trajectory. The simulation timescale achieved in the likes of aMD and metadynamics enforce overall modifications to the energy landscape such that the dynamics of individual atoms does not correspond to the real timescale, nevertheless, the thermodynamic properties cannot be monitored *in situ* but are retrieved only after the back-mapping, i.e. an additional artificial sampling process that deviates from the precise dynamics.

Arguably from the data analysis in the process of protein folding, both from experiment and simulation, the pivotal changes which facilitate the collective movement corresponding to the transition from the local energetic basin to the global minimum, e.g. the formation of the β sheet, does not necessarily involve prevention of kinetic-trapping for all atoms on the collective branch. On the contrary, a critical conformation change in the active binding site on the protein leads to the rearrangement of the whole macromolecule. If a

selection of smart and pinpoint modifications are applied to the pivotal segment, involving merely a fractional number of atoms across the whole system, the critical energetic "kick" to the kinetically trapped conformation encourages the collective movement, while for the majority of the atoms the thermodynamic properties are conserved throughout the simulation.

3.1.2. Machine learning

Machine learning has deep roots in the contemporary development of mathematics where the improved regularization allowed for precursor methods as compressed sensing. [43] One of the central ideas is to derive from a limited scope of knowledge the principle components that represents the whole structure of interest. For instance, consistent with cognitive models for object classification, in specific cases, many details in a higher resolution representation is not necessary to reconstruct the principle visual outline of the object, nor does a lack of resolution obscures too much. For the eye the pixelated image in Figure 3.2 does not include adequate information however the brain is not confused as much to identify Albert Einstein. Interestingly, from further the distance the reader gazes at it, i.e. with increased correlation between the pixels rather than the increased volume of raw information from the increased number of pixels, the clearer Einstein gets identified.

Figure 3.2.: A pixelated lower resolution image of A. Einstein. Even though the information is not complete, the subject can still be recognized from afar.

Take an image classification task as an example, information is encoded as a comprehensive, high dimensional matrix where discrete numeric values are measurements of the original object. A representation of lower resolution, i.e. in the representation matrix a large number of zero values, should be invariant once the representation matrix is fed to the cognitive model and yield a successful classification result. The huge matrix with a small number of non-zeros elements is considered as a sparse matrix, i.e. the high dimensionality being less relevant to approximate the solution in the conventional one-to-one mapping of $y^m = A^m_n x^n$ that all dimensions of m and n must fit. Fulfilling both the sparsity and incoherence conditions in the representational space of A^m_n the principle signal can be recovered from much less information required by the Nyquist sampling theorem. [44]

The abundance of observation records, dating as far back as the Galilean era when people start to systematically book-keep the trajectories of planetary movements, is available to the bottom-up analysis and model deduction based on data-driven methods, e.g. compressed sensing. Ever since the accessibility to the likes of barometers, weather data has been one the most well-maintained timeseries data for a genuinely complex dynamic system, in itself implying a causative model with only partial information. While the measured dimensions cannot provide a complete description as point-based temperature readings in the 17th century, improvemented theory and equipment gradually couples temperature reading with pressure data dating from 18th century, before the sensor array are deployed as a 3D mesh

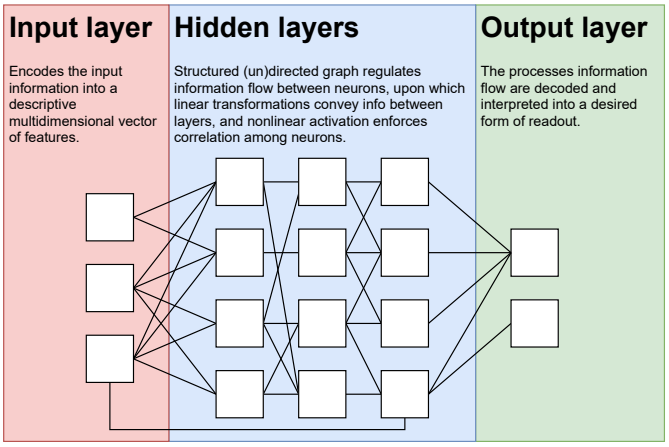


Figure 3.3.: A generic representation of a deep learning neuron network. A neuron network is "deep" when the hidden layers, i.e. neuron layers excluding the input and the output is more than a single layer. The topology of the neuron network directs the flow of information when between different layers the information is gathered and correlated between the neurons of the same layer via the activation function. Generic dense neuron networks with a finite depth of 3 layers is capable of solving complex tasks when the key feature in the design of the neuron network is to allow redundancy in the information flow, as illustrated. The limiting factor of application of deep learning often resides in the training efficiency.

across the global with much finer whilst finite resolution and accuracy as of nowadays. Since the digital era there has been a booming increase in both data availability and a booming rate of dataflow where data processing establishes itself to be a major role as the data science as a counter part to theory and experiment. On the one hand the abundance of data should cover the scope of application in deriving a causative model that understands the underlying mechanism, whilst on the other hand, the big data problem is unprecedently overwhelming to human cognitive power, when in principle tasks that should have been simpler than understanding the data are already extremely challenging as the visualization of the raw data. Machines learning methods takes advantage from both aspects as it takes raw data of great availability whilst on the other end produce automatically an abstract causative model for a specific goal functions where only a handful disciplinary components are of interest. In the realisation, the input of redundant information into the cognitive model, as well as the redundant construction with higher dimensionality than the causative model for which the compress sensing asks, takes advantage when the principle component in the signal is strengthened while the noise or irrelevant signal gets suppressed, before in the output, the cognitive model completes a relatively simple task as getting lower dimensional projected state vectors, in the extreme case binary answers to a highly-specific question as yes or no.

Concurrent machine learning algorithms utilises layer-structured neuron networks. [45] The graph-based network structure partially mimics a biological neuron network, whereas modern computing architectures, especially the readily available GPUs which are intended for fast vector operations subjected to computer vision, excel on these linear transformations and tensor operations extensively employed in graph operations. As the input representation goes through a series of neuron network layers, abstraction of the representation is achieved via resampling in the dimension changes through the forward mapping between layers of neurons, as the correlation among the neurons gets established via the activation functions. For the training of the network, the input to the neuron network goes through all the projection and is mapped to an output, where a loss function is defined by the difference between the projected output and the exact output. The loss function for each training

input-output pair is back-propagated in the form of corrected weights on the links in the neuron network. The training stops when the goal for the loss function is achieved, usually set to a point that the loss function is no longer improved with further training. For the layered neuron networks, it is not necessary that results in the intermediate layers are human-readable so that people understand which layer does what, however improvements of the network structure based on external information or design rules other than the training have been proved essential in enhancing machine learning capabilities for specific goal functions. For instance, in reduction of dimensionalities for image recognition, there is convolution neuron network based on 2D kernel-like structures fitting to the 2D nature of the graphics, while in timeseries analysis there is recursive neuron network where an invariance in timescale is hard-coded within the implementation.

3.1.3. Genetic algorithm

Inspired by the power of natural selection and gifted with the capability of fast iterative calculations in the computers, the genetic algorithm (GA) provides an automatic, generative tool of optimization. [46] The solution domain in the genetic algorithm is discretized and encoded as genetic sequences, mimicking the gene pool in an ecological system where the optimisation target is implied to be dependent on the sequence. To evaluate the solution domain, a fitness function is predefined, so that the individual genetic sequences are ranked with a bias inclining toward the optimization objective.

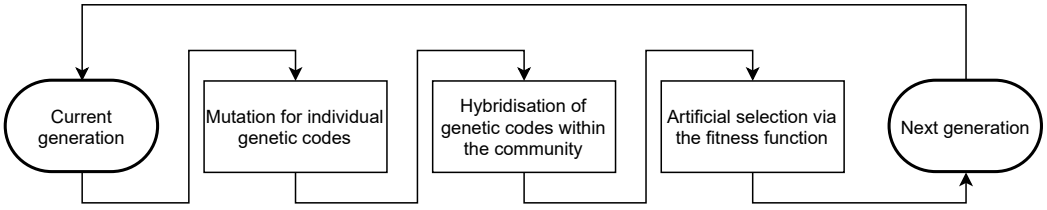


Figure 3.4.: Operations within an iteration in the genetic algorithm. Mimicking the natural selection, the population is iteratively forced to evolve subjected to the stochastic selection with an *a priori* fitness bias. The generations preserve a number of features which are deterministically encoded in an analogy to the gene, which through the operations as the mutation and hybridisation, pass on to the next generation, exploring the feature space based on genetic similarity.

The artificial evolution strategy modifies the genetic codes iteratively in a series of "generations". Mimicking the modification of genes in biological systems, the modification of genetic codes in the optimization process comes from mutation, which serves as a stochastic driving force to explore a larger parameter space, as well as hybridization, or crossover, which is an effort to improve the provisional "fitness" by sharing as well as preserving of the advantageous traits in the parental genetic codes. Both the mutation and hybridisation operations generate the next generation with a wider expanse exploring the parameter space based on readily available genetic codes whose "fitness" is ranked. Based on the assumptions of similarity in the genetic codes leads to similarity in genetic traits, the advantage in genetic codes thereafter translates to better fitness in a deterministic black-box mapping and *vice versa*. Such an exploration strategy in practice often proves more efficient than random search as the advantageous trait from a partial genetic code, comparable to building blocks in the gene sequences, gets incrementally revealed via local clustering across generations.

The inclusion of the fitness function mimics the natural selection process upon individual genetic codes. The genetic codes that yield better fitness evaluation preserves an advantageous sampling frequency within the gene pool, which can be interpreted as better

probability of survival and reproduction for individual biological entities. Comparable to the assignment of a goal function in machine learning, a low dimensional, quick-to-evaluate fitness function is crucial in defining the efficiency of genetic algorithm. In analogy to the back-propagation in machine learning, a poor fitness evaluation decreases the sampling frequency yet again as local information, while a good fitness evaluation increases the sampling frequency in the local vicinity, contributing to a finer local resolution in finding the local optima. This similarity-based search pattern, however, institutes bias to the local exploration in the parameter space.

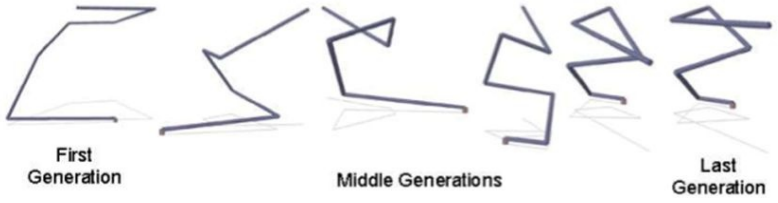


Figure 3.5.: Genetic algorithm is applied to antenna optimization for micro-satellites. Iteratively the transittng signal quality, gain, and gain consistency are simultaneously improved with modifications on rotation angles between antenna segments. [47]

Unlike machine learning, the intermediate generations in the genetic algorithm carry diverse traits that work as seeds in the iterative optimization and is human-conceivable as the "direction" of evolution is explicitly expressed, seen in Figure 3.5. The directional evolution implies a causative model as the later generations are better than the previous ones. For the completion of the artificial evolution, a stopping criterion is manually implemented usually when a certain number of generations are bred, or the improvement in fitness between generations becomes marginal.

3.1.4. Efficiency of parameter space exploration and hand designed rules

Both the machine learning and genetic algorithm suffer from efficiency penalties for complex problems. A multitude of competing factors may offset each other whilst the correlation among the factors render the signal-to-noise ratio poor in the determination of weight back-propagation and the fitness-dependent gradient in the sampling. For both methods, to compensate for the accuracy, the volume of training data / genetic population required inflates quickly as the dimension of parameter space grows.

If for machine learning there is preliminary knowledge encoded in the structure, a correspondent neuron network design reduces the training effort massively than the prototypical all-connected neuron network. While an all-connected multilayer neuron network with for each layer nodes is theoretically complete that can be recovered as any network structure for any problem, in practice the training cost would be infinitely large. Any preliminary knowledge in the network structure, i.e. a part of the network that needs not to be trained saves geometrically the training requirement. In the genetic algorithm, hand designed rules improve the efficiency as optimal seeds to start near the optima.

3.2. Methods

The adaptive acceleration algorithm is implemented as an extension of the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS), an extensively used MD simulation software package. [30] Two modular part constitutes the algorithm as:

1. an adaptive management module identifies from all pair-wise interactions the most vibrant pairs, performs online selection, and book-keeping of those pair-wise interactions whose potentials are softened, and performs the deselection which returns the softened potential to the vanilla Lennard-Jones potential;
2. potential transition module that modifies the pair-wise interaction between the Lennard-Jones potential and the softened potential.

Since LAMMPS is based on the message passing interface (MPI), we have implemented the acceleration algorithm parallel in accordance with the native architecture. Since the dynamic management in the book-keeping for pairs that are modified with softened potential requires constant sampling in the domain assigned to local MPI process, the acceleration algorithm is scalable only when the decision made by the the adaptive management module is a spatio-temporally local operation, meanwhile, the communication among the MPI processes must be kept minimal.

3.2.1. The adaptive management of the softened pair-wise potential

In the realisation of the dynamic and adaptive management of the pair-wise interactions with the softened potential, an Markovian approach is implemented to minimise the book-keeping effort and enforce locality and scalability of the modification decisions. For all the pair-wise interactions in the system, the acceleration algorithm book-keeps unique identifications for each pair. For the pair-wise interactions where the potential barrier for the two atoms in the pair to swap positions is very high due the strong repulsion, the vanilla Lennard-Jones interaction results in local vibrations between the pair. In such close vicinity where the Lennard-Jones can be approximated by a parabolic potential, the selection module samples pairs and calculates the probability of pair swapping based on the relative movement in between to replace the intense Lennard-Jones interaction potential with a parabolic one. The relative movement in the center-of-mass reference frame for the pair of atoms is filtered such that the pairs which moves towards each other is recorded. Also, within the center-of-mass reference frame, the momentum of the pair is projected along the line that connects the two particles, hence a projected kinetic energy V_{projKE} can be formulated as a function of velocity pair-wise difference $\mathrm{d}\boldsymbol{v}_{ij}$ and the pair-wise coordinate difference $\mathrm{d}\boldsymbol{x}_{ij}$ for the i-th and j-th atoms.

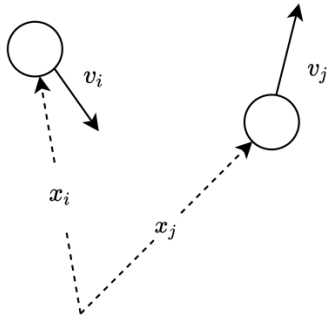


Figure 3.6.: Relative movements between pairs characterised by vectors of $\mathbf{x}_i$, $\mathbf{x}_j$, $\mathbf{v}_i$ and v_j in the MD simulation.

The relative movements for the atoms in a pair-wise interaction can be characterised with the vectors of of $\mathbf{x}_i$, $\mathbf{x}_j$, $\mathbf{v}_i$ and v_j in the MD simulation. When the differences of the coordinates $\mathrm{d}\boldsymbol{v}_{ij}^N$ and the difference in the velocities $\mathrm{d}\boldsymbol{x}_{ij}^N$ is defined as:

$$\mathrm{d}\boldsymbol{x}_{ij}^N = \boldsymbol{x}_i^N - \boldsymbol{x}_j^N$$

$$\mathrm{d}\boldsymbol{v}_{ij}^N = \boldsymbol{v}_i^N - \boldsymbol{v}_j^N$$

where in Cartesian coordinates the 3D velocities and coordinates correspond respectively to the N-th dimension. The sign functions which characterises whether the pair of atoms are approaching each other or departing from each other via the definition of:

$$\text{sgn} = \begin{cases} -1 & \mathrm{d}\boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x} \leq 0 \\ 1 & \mathrm{d}\boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x} > 0 \end{cases}$$

for individual atoms, when the projected velocity is established as:

$$\mathrm{d}v_{\text{proj}}^{N} = \mathrm{d}\boldsymbol{v}^{N} \cdot \frac{\mathrm{d}\boldsymbol{x}^{N}}{||\mathrm{d}\boldsymbol{x}^{N}||}$$

while the summation over all N dimensions yield the projected kinetic energy:

$$V_{\text{projKE}} = \text{sgn} \sum_{N} \frac{1}{2m} \cdot (\mathrm{d}v_{\text{proj}}^{N})^2 \tag{3.1}$$

The selection process monitors pairs within a predefined interval in terms of pair-wise distances as the accleration algorithm applies exclusively to the imminent collisions. In the example of the Lennard-Jones liquids, the typical radial distribution function (RDF) for an imperfect crystalline structure is shown in Figure 3.7. The RDF exhibits a significant concentration in the first peak ranging from approximately 1σ to 1.2σ where σ corresponds to the Lennard-Jones unit length.

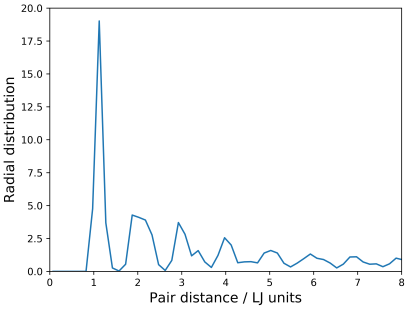


Figure 3.7.: Radial distribtuion fuction of an imperfect Lennard-Jones crystal. Most concentrated distribution of pair-wise distances resides in the interval ranging from approximately 1σ to 1.2σ where the σ is the Lennard-Jones unit length.

Empirically, the sampling within the limited scope for all pair-wise interaction in terms of the projected kinetic energy results in a Boltzmann-like profile that is comparable to the Maxwell-Boltzmann distribution of the vanilla kinetic energy, illustrated in Figure 3.8. This allows us to characterise the energy profile of the projected kinetic energy of all pairs in the local domain with few parameters and synchronise across the whole entire ensemble of domains in the parallelised simulation with minimal communication effort whilst maintaining the scalability.

Sampling from the projected kinetic energy a Boltzmann-like profile can be established in the form of:

$$P_{\text{projKE}} \propto e^{f(V_{\text{projKE}}; V_{\text{projKEmax}})/kT} \tag{3.2}$$

in which the reference maximum projected kinetic energy $V_{\text{projKEmax}}$ is summarised from all domains in the domain decomposition parallel scheme. In practice, among all of the pairs, the maximum projected kinetic energy is recorded every timestep, as the first descriptor of the distribution of the projected kinetic energy. The total number of pairs within the pair-wise distance dependent vicinity is recorded every timestep and serves part in the

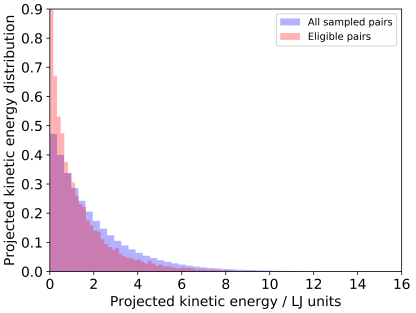


Figure 3.8.: Distribution profile of the projected kinetic energy in an imperfect Lennard-Jones crystal. The projected kinetic energy for all pairs, as well as the pairs sampled by the adaptive managment module, both follow a Boltzmann-like distribution in the potential, thus can be characterised with few parameters.

regulator term, as the second descriptor of the projected kinetic energy distribution. The two parameters are constantly updated and shared across all domains owned by individual processors so that the sampled profile of the projected kinetic energy from the previous timestep is a shared knowledge. For each pair-wise potential, the difference of individual projected kinetic energies is sampled according to:

$$P_{\text{selection}} = e^{(V_{\text{projKE}} - V_{\text{projKEmax}})/R} \tag{3.3}$$

where the regulator $R = \frac{V^2_{\text{projKEmax}}}{2V_{\text{projKEmax}} - 2}$ such that $P_{\text{selection}}$ is normalised over the interval $V_{\text{projKE}} \in (0, V_{\text{projKEmax}})$. The probability of finding individual pairs with a certain projected kinetic energy (Equation 3.2) is then proportional to the selection probability of the the pair being modified with the softened potential in Equation 3.3. As a result, the atoms in the pairs that approach each other with the most projected kinetic energy are most likely of being modified with a softened interation potential which lowers the energtic barrier to promote swapping of atom positions.

3.2.2. The transition between the vanilla Lennard-Jones and the softened potential

The list that records all the pair-wise interacitons with a softened potential are updated constantly according to their pair-wise distances. If for a modified pair, their separation is larger than the upper threshold of the predefined sampling vicinity, they are gradually removed from the list of modified pairs via a delayed regularisation factor. This asymmetrical transition, i.e. an instantaneous switch from the vanilla Lennard-Jones to the softened potential versus a dampened recovery from the soft potential back to the Lennard-Jones, on the one hand prevents an abrupt switch between the modified potential and the vanilla Lennard-Jones potential, whist on the other hand this dampening mechanism prevents for artificial creation of a local hot-spot, i.e. a repeated selection-deselection cycle for pairs that preserve high frequency vibrational modes. In practice, a failure in preserving enough dampening not only leads to the introduction of artificial hot-spots which distorts the profile of the projected kinetic energy, and reduce the performance of the acceleration algorithm, while in extreme cases, due to the reservation of momentum via softened repulsions, a resonant amplification of momentum. The resonant accumulation of momentum between the softened pair in analogous to an acoustic feedback loop, which ultimately leads to a terminal divergence in the thermodynamics.

For the selected few pairs, the interaction potential is seamlessly switched from the 12-6

Lennard-Jones potential to a parabolic one, formulated as:

$$V_{\mathrm{LJ}} = 4\epsilon[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^{6}] \tag{3.4}$$

$$V_{\mathrm{SOFT}} = 4\epsilon^{*}(\frac{\sigma}{r + c_1})^{2} + c_2 \tag{3.5}$$

in which the paramter ϵ^{*} regulates the stiffness of the softened potential, as the combination of complimentary constants pair of (c_1, c_2) connects the parabolic V_{SOFT} smoothly to the 1st derivative, i.e. continuous force, to the vanilla Lennard-Jones potential at a predefined switching distance of $r = r_{\mathrm{switch}}$. By introducing V_{HYBRID}, the transition from V_{LJ} to V_{SOFT} is time-delayed the regulating factor $\lambda_{\mathrm{delayed}}$ gradually increases from 0 and 1 after the pair is excluded from the softened potential, before recovering to $V_{\mathrm{HYBRID}} = V_{\mathrm{LJ}}$.

$$V_{\mathrm{HYBRID}} = (1 - \lambda_{\mathrm{delayed}})V_{\mathrm{LJ}} + \lambda_{\mathrm{delayed}}V_{\mathrm{SOFT}} \tag{3.6}$$

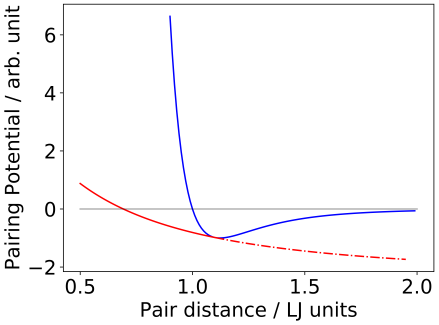


Figure 3.9.: Softened potential compared to 12-6 Lennard-Jones. The blue curve corresnpond to the vanilla 12-6 LJ potential V_{LJ} whereas the parabolic potential V_{SOFT} is connected to the vanilla LJ at the specific switching r_{switch}. For selected pairs, the pair-wise interaction potential dynamically switches between the vanilla LJ to the solid part of the softened potential.

The performance of the acceleration algorithm is directly regulated by the number of pairs whose pair-wise interactions are softened across the system, a large number of modifications intuitively provides better acceleration, when the ratio of modified pairs must be small enough that the overall thermodynamics properties do not deviate too much from vanilla dynamics. Meanwhile, the performance of the acceleration algorithm is also dependent on the softness of the modified potential, that is, a softer modified potential invites more deviations against the vanilla thermodynamics, while the softer potential enhances the acceleration performance. As a number of system-specific parameters are introduced, the optimal parameter setups vary according to the system in investigation, to increase acceleration performance whilst an ill-defined parameter setup may lead to numerical divergence.

The parameter space is compiled from of 5 physical parameters in the acceleration algorithm, when the additional parameters are either technical or auxiliary. For the reproducibility of the simulation data, the technical parameters are also recorded as metadata. For each parameter, the sampling interval is discretised either in linear or logarithmic scale, while the combinatorial number for paramter setups, encoded as 7 "bits" sequences, the total number of possibilities quickly reaches over 1 million. The list of modelling parameters are shown in Table 3.1, where D_{rate} is couple to the time-delayed regularitsation factor of $\lambda_{\mathrm{delayed}}$ as $D_{\mathrm{rate}} = 1/\lambda_{\mathrm{delayed}} \cdot \delta t$ as δt records how long the current pair has been released from the list of pairs with softened potential.

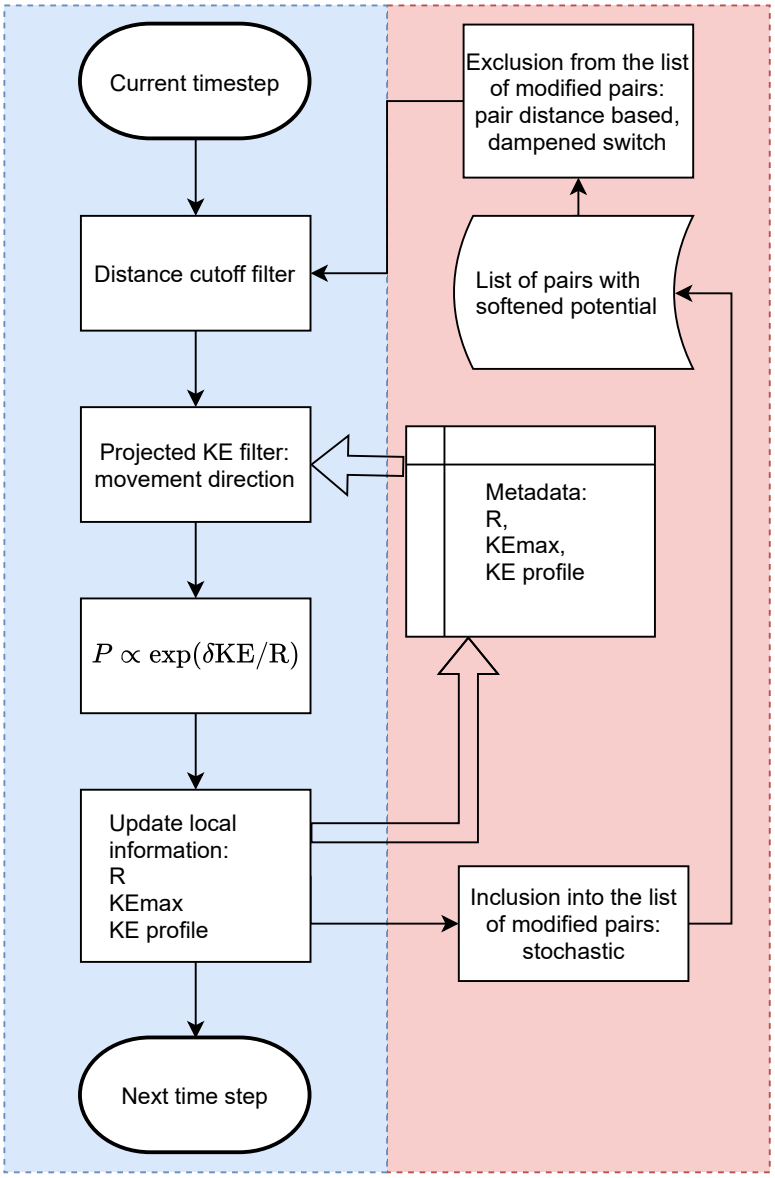


Figure 3.10.: Flowchart of the accelerated algorithm. The method is divided into 2 modules, distinguished by different colors as the a) local MD operation, in light blue, and b) the global book-keeping of modified pairs with softened potentials, in pink. The book-keeping part consists of 2 information loops. The inner loop with global communication, illustrated with thick, hollow arrows, records of the simulation parameters as the regularization factor, and the two parameters which characterises the KE profile, and are processed across all MPI processors as the KE profile is synchronised. In the outer loop, illustrated with thin, filled arrows, no global commuication is needed across all MPI processors, while the list of modified pairs book-keeps, and exchanges information in the local domain and in the neighbouring whereas the one-to-one mapping the local atoms and the image atoms within the domain boundary regions are checked against the global hash table so that all atoms are recorded with unique identities.

For a quick exploration in the parameter space to find an optimised parameter setup, a 2-stage, deep learning (DL), genetic algorithm (GA), and hand-designed rules (rules) hybrid approach has been employed. The hybrid optimization approach consists of two steps: an initial adaptive exploration of the parameter space based on the trio of DL, GA, and hand-designed rules (DL-GA-rules) that probes the boundary of a subdomain of the parameter space; a second stage optimisation utilises the Phoenics package [31] based

Physical parameters	Sampling range
Potential switching pair-wise distance r_{switch}	1.0σ to 1.25σ
Sampling cut-off distance for softened pairs r_{sample}	1.10σ to 1.5σ
Intensity of the softened potential ϵ	$(0.1, 2.0)$
Dampening rate of potential switching D_{rate}	$(10, 80)$
Upper limit for the number of modified pairs N_{total}	$(2^4, 2^{16})$

Technical parameters
Number of parallel processes
Random seed
Adaptive parameters
Global maximum projected kinetic energy $V_{\text{projKEmax}}$
Number of modification candidates in the sampling interval of $(r_{\text{switch}}, r_{\text{sample}})$
Auxiliary parameters for the 1st-order smooth potential c_1, c_2

Table 3.1.: List of parameters for accelerated molecular dynamics with softened potentials.

on machine learning to quickly sample the subdomain where the choice of parameters is interpolated with finer resolution that improves and thereafter validates the findings from the first stage, illustrated in Figure 3.11.

In practice, since the parameters a number of options alternate the thermodynamics both in conjunction with, or against one another, many of those parameter combinations breaks the overall physics that numerical divergence occur. Uniformly sampled from a predefined, reasonable range for each individual parameter, the combinatorial possibility of finding a "wrong" parameter setup in 1000 attempts is as high as 77%. In the hybrid optimization approach, a deep learning module aimed at quick estimating the validity for a given parameter setup is implemented. A dense deep neuron network classifies the simulation with a specific parameter setup as either convergent, marked as 1 or divergent, marked as 0. With sufficient training the deep learning module is able to predict whether or not the simulation run with a certain parameter setup is valid without the MD performance with a fairly good accuracy at over 75%. The genetic algorithm module explores the parameter space with the fitness function defined in a similar way with the performance measurement in the deep learning module. The fitness for each individual parameter setup is concluded *posteriori*. Based on the initialisation of the zero-th generation from either the evaluated data or from random sampling, new generations are iteratively populated via the hybridisation, the mutation, and the selection steps. In the exploration of the parameter space, for some parameters which are pair-wise dependent, hand designed rules are applied. For instance, for the softened pairs, their pair-wise distance should be larger than the switching distance whilst being smaller than the sampling cutoff distance. If the sampling cutoff distance is smaller than the switching distance, the acceleration algorithm is turned off, and the dynamics returns to vanilla Lennard-Jones.

The first opmitisation stage corresponds to the DL-GA-rules voting model whereas the second stage of the Phoenics machine-learning based quick sampling utilises the evaluation results from the first stage and perform an individual optimisation with finer sampling in the parameter space. The Phoenics algorithm is based on Bayesian optimisation which utilises the historical evaluation results for the sequential improvements in the global optimisation for expensive-to-evaluate functions. Iteratively, the bayesian kernels are placed on the discrete spot with prior evaluations, constructing a series of approxiamted local landscape for the objective function. The surrogate model is updated as new observation data is introduced from the current iteration. For the proposals of new observations, the

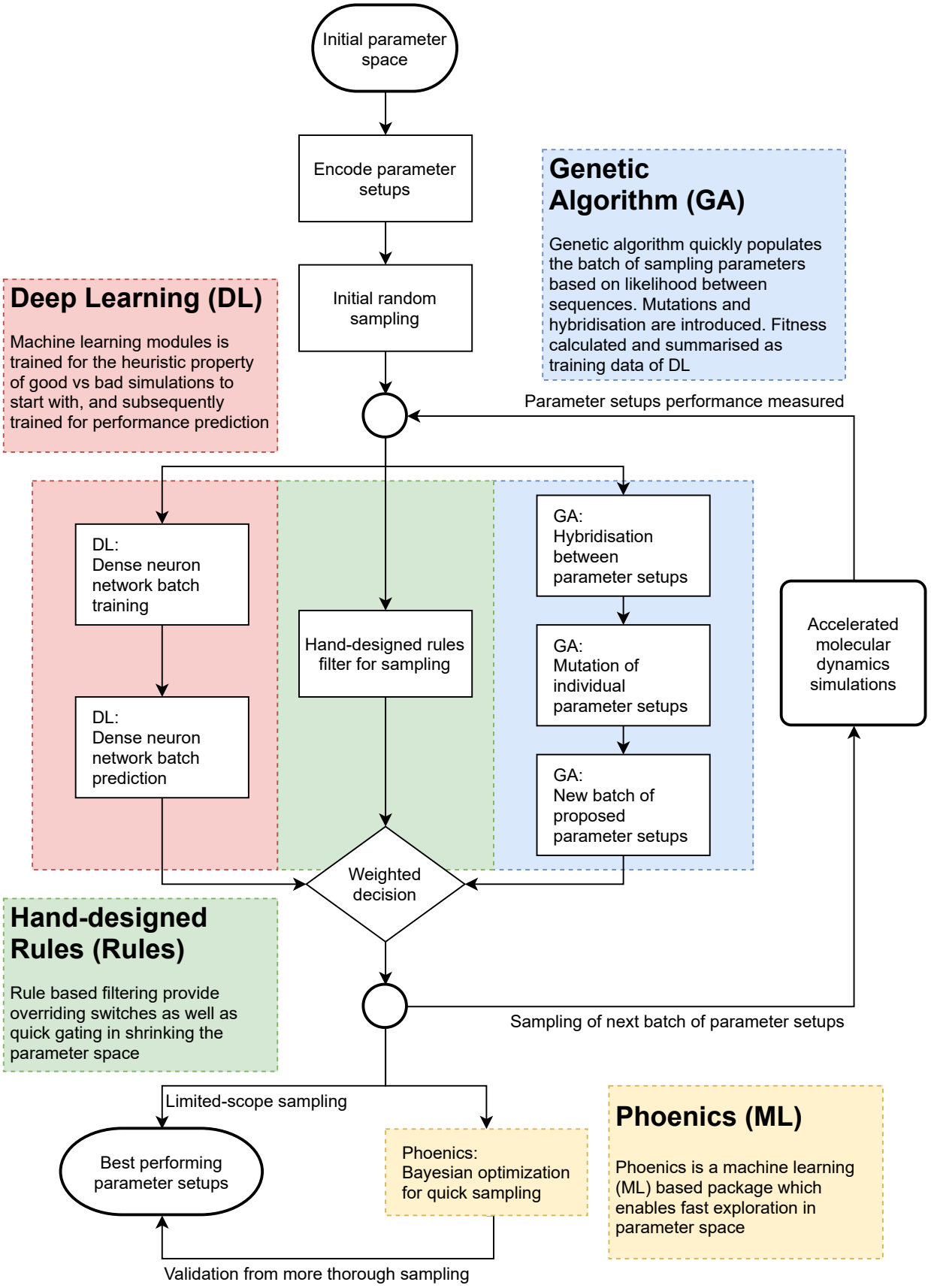


Figure 3.11.: Flowchart of the hybrid voting optimisation of parameter setups.

acquisition function in the Phoenics algorithm can be tuned to favour either exploration

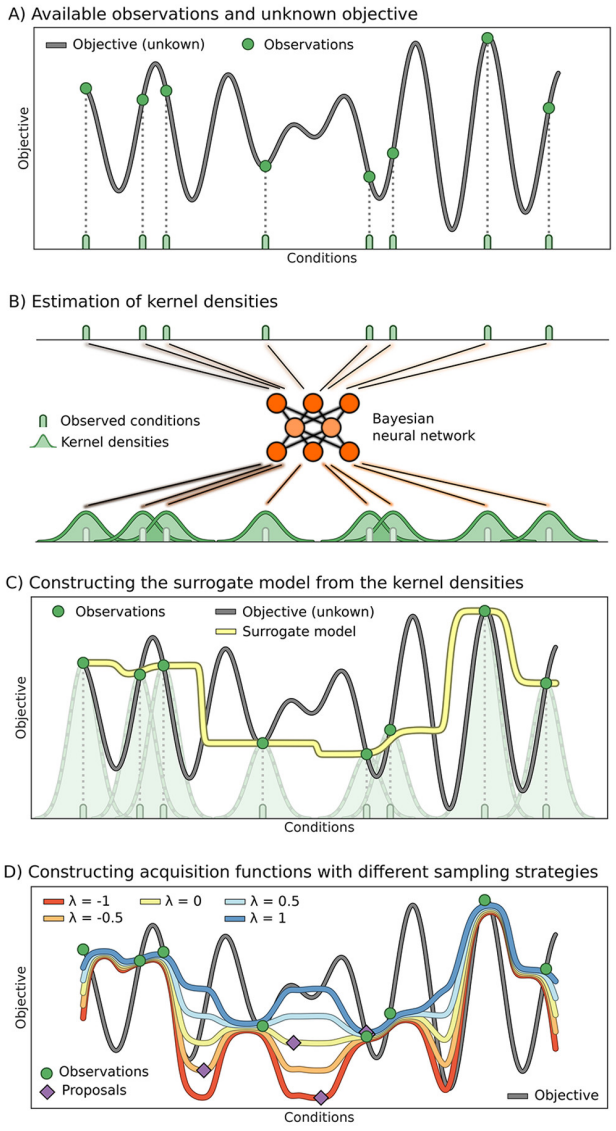


Figure 3.12.: Workflow of Phoenics algorithm. [31]

or exploitation of the search space. The Phoenics algorithm promises for a rapid global optimisation with optimal exploration efficiency in the exploration of the parameter space is balanced due to the advanced kernal density estimation methods, however in practice for the parameter setup optimisation of acceleration MD simulations, the initial randomly sampled observations lead to equally bad evaluations with numerical divergence which serverely undermines the accuracy and efficiency of the kernel density estimation near the sample points. As a result, the Phoenics algorithm, even though being a proven method which has claimed to be one of the quickest exploration techniques, cannot not work as a stand-alone optimisation tool in our case. However, as is typical with machine learning based methods, with a finite effort in the pre-training, i.e. set up the Phoenics to work as the 2nd stage optimisation when a number of observations are already available as prior knowledge, the Phoenics algorithm recovers consistent optimisation results faster than the DL-GA-rules voting by a factor of 3x.

3.3. Acceleration benchmark in a grain boundary movement problem

The performance of the acceleration algorithm is benchmarked on a 2D grain boundary system. In the preparation of the benchmarking system, a single sheet of Lennard-Jones

atoms are close-packed, filling the periodic simulation box, whereas a smaller grain, located in the center of the sheet, is cut away, tilted in the alignment, and thereafter overlaid back to the original sheet. In the relaxation of the 2D sheet, the tilted smaller grain will eventually shrink in domain area as the atoms on the boundary of the smaller grain realigns with the crystalline orientation of the base grain, when the excess of the energy cost induced by the grain boundary gets gradually dissipated.

The grain boundary movement is a slow process exclusively driven by thermodynamic fluctuations. On the one hand, the realignment of individual atoms with respect to its nearest neighbours is dictated by the short-range interaction and maintains no overall knowledge, as opposed to the reaction coordinates in metadynamics, which renders great difficulty for individual atoms to decide which in direction the grain boundary should move to minimise globally for the most stable crystalline structure. On the other hand, over time the grain boundary movement should cover the entire grain, i.e. the collective movement which results in a global realignment to the crystaliline structure is of orders of magnitude difference in the spatio-temporal scale. By monitoring the time-dependent size-shrinking of the smaller grain, the rate of shrinking serves as an indicator for the acceleration performance.

The benchmark systems are prepared with a variety of initial configurations. The 2D sheets are prepared with 1) a rectangular smaller grain tilted by 15 degrees anti-clockwise, 2) a diamond shaped grain tilted smaller grain tilted by 30 degrees anti-clockwise, and 3) a mixed grid of rectangular smaller grains which are tited by 25 degrees either anti-clockwise or clockwise. Due to the difference in the configurations, the dimensions of the prepared systems are changed accordingly. Meanwhile the parameter setups for the acceleration algorithm are shared across all benchmarking cases.

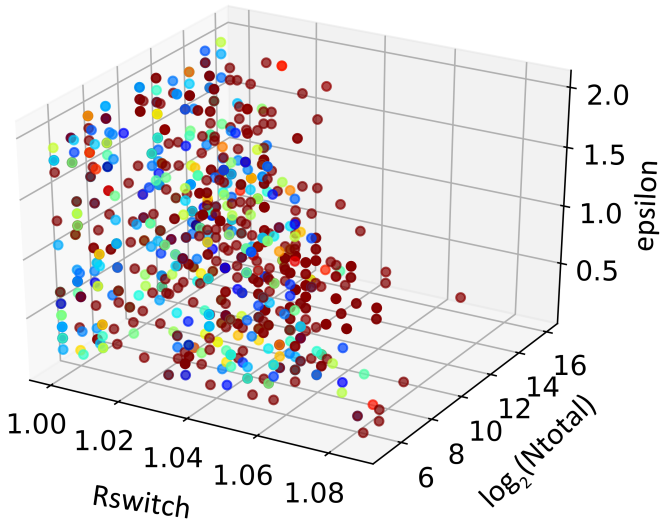


Figure 3.13.: The grid search in the parameter space visualised for 3 major components r_{switch}, r_{sample}, and N_{total}. Hot colors correspond to poor acceleration performance, cool colors correspond to good acceleration. Note the complex mixture of well performing vs poorly performing parameter setups and the barrier of red points suggesting for poor performance prevents further sampling in the DL-GA-rules optimisation.

For the 2D sheet of Lennard-Jones atoms, the sampling inteval be narrowed down for individual parameters, e.g. for the switching distance from the vanilla Lennard-Jones into

the softened potential, the upper limit of the sampling should be larger than 1.0σ and less than 1.25σ. In the individual sampling intervals, a finite number discrete sample points are selected for the DL-GA-rules optimisation in the first stage. This random, discrete sampling scheme is equivalent of the grid search in the parameter space however as has been discussed previously, on the one hand the grid is of high dimensions and the parameter space of great expanse, while the majority of the grid points will lead to in time unphysical behaviours in the form of divergent interaction potential. The deep learning module with 2 hidden layers, each with 32 nodes and sigmoid activation function and a dropout rate of 20% for the first hidden layer is implemented to tackle this stability problem. With more training epchs, shown in the learning curve of Figure 3.14, the prediction accuracy for a physically viable parameter setup increases to over 70%. The deep learning has proved successful for simple yes or no decisions, however the utilisation of deep learning for prediction of more complex predictions as acceleration performances, the requirement in the training becomes much more expensive computation-wise.

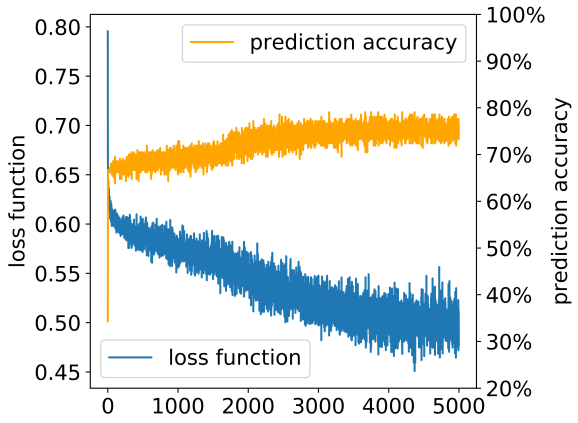


Figure 3.14.: Learning curve and accuracy of the deep learning (DL) module.

While the training of the deep learning module increases the sampling efficiency for physically viable simulations with enhanced stability, a starting ensemble of evaluated parameter setups serve as a initial population for the genetic algorithm module. The genetic algorithm module maintains a biased objective towards the improvement of acceleration performance, hence counter-balancing the stability-oriented decisions from the deep learning module. In the meantime, the incorporation of the genetic algorithm module circumvents the requirement of the extra amount of training data for the deep learning algorithms as genetic algorithm performs quick explorations in the parameter space based on similarity of the genetic codes than random sampling, i.e. with a moderate mutation probability, the mutation applied to the parameter setups effectively samples the neighbouring grid point in the high dimensional parameter space, when in the case of hybridisation between the pairs of parameter setups the grid points in the middle are evaluated. The fitness function is taken as $P_{\text{fit}} = 1 - t_{\text{shrink}}/t_{\text{end}}$ where t_{shrink} corresponds to the time when the smaller grain size first reaches below 50 atoms, as the t_{end} is the fixed total simulation time. After all evaluations of fitness functions corresponding to the current generation are ranked as the top performing parameter setups, plus a series of randomly selected sequences are chosen to establish a "gene pool". In practice, the number of individuals per generation can be only of a small fraction in the whole parameter space at 100, when for each generation the 20 top performing parameter setups plus a randomly selected 20 are selected to be the "breeders" in pairs. The hybridisation of the "genetic codes" are contributed 50-50 from the parent pair, and from each parent pair 10 child sequences are generated. The mutation probability is set to 5%, when an individual mutation in the genetic codes moves

the current sampling spot to an adjacent point in the grid of the parameter space.

For all newly proposed sequences, the deep learning modules predicts from the current neuron network whether or not the new parameter setup renders the MD simulation unstable, as from both genetic algorithm module, a rough estimation of the fitness based on the average ranking from the parent generation is passed on to the decision node, whereas the rules module checks certain criteria among the parameters, e.g. the r_{sample} should be larger than the r_{switch}, otherwise the sampling interval is null and the interaction potential from the acceleration algorithm is reduced to vanilla Lennard-Jones interation. In the decision node, the decision of rejecting the newly proposed individual parameter setups is voted from all 3 modules, where the deep learning module has a fixed weight of 0.6, the genetic algorithm prediction maintaining a limited filtering effect at 0.4 and from the rules module the most decisive influence at 0.9 in a 2-steps voting scheme, the 1st round vote is collected from the deep learning and the genetic algorithm module, once passed, the rules module votes again to filter out the ineffective sampling. In both rounds, the voting sum only applies to negative votes in filtering out the obviously unreasonable parameter setups.

3.3.1. Approaching the quickest shrinking path: the sandwich problem

The first benchmark is run on a sandwich-like Lennard-Jones 2D sheet where the smaller grain is of rectangular shape. The domain sizes are 30σx60σ for the smaller grain and 100σx70σ for the base grain. The base grain reaches across the periodic boundary whilst the smaller grain is of finite area in the middle. With time, when the whole Lennard-Jones system is relaxed, the extra potential resulted from the partial misalignment in the smaller grain moves the grain boundary such that the smaller domain shrinks. In time, the smaller grain totally vanishes into individual vacancy sites, as all the rest of the atoms are aligned to the crystalline grid. In particular for the rectangular smaller grain, on the one hand, the tilting angle for the smaller grain is 15 degrees counter-clockwise, i.e. a small misalignment with respect to the base so that the recovery to the perfect alignment is faster near the grain boundary, on the other hand, the area of the smaller boundary is nearly 1/3 of the overall system that the realingment of atoms and the momentum conserve from the local vibartions on the grain boudnary takes a significant amount of time to reach the inside of the grain.

With the adaptively softened interaction potential, subjected to different parameter setups, the smaller grain may experience global translational movements, and change in the aspect ratio as the grain boundary movement constantly reshapes the morphology when practically the selection of the pair-wise interactions are randomly distributed. In the extreme cases the benchmark will fail for both the vanilla Lennard-Jones potential or with the softened potential, the smaller grain reaches both the upper and the lower boundary of the simulation box, resulting in a aligned-tilted-aligned sandwich-like conformation that fundamentally changes the finite domain representation into a semi-infinite 2D plane, which in the post-mortem analysis is reflected by a constant domain size across the simulation time span. Meanwhile, for most parameter setups, the smaller grain eventually vanishes, which results in a single piece of domain as atoms are globally aligned. For all of the sampled parameter setups, the time series of the smaller grain size evolution is shown in Figure 3.16.

The sampling of the parameter setups is automated by the trio-voting DL-GA-rules model. For the parameter setup sampling, the fitness function and the training target of the neuron network as directly correlated with the optimisation target function of $P_{\text{fit}} = 1 - t_{\text{shrink}}/t_{\text{end}}$, where for the domains that do not shrink, the penalty is bounded such that $P_{\text{fit}} = 0$, i.e. equally poor accleration performance as physially unviable parameter setups. During the optimisation, the sampling efficiency is constantly improved as the best performer for the parameter setup continually moves the performance curve to the

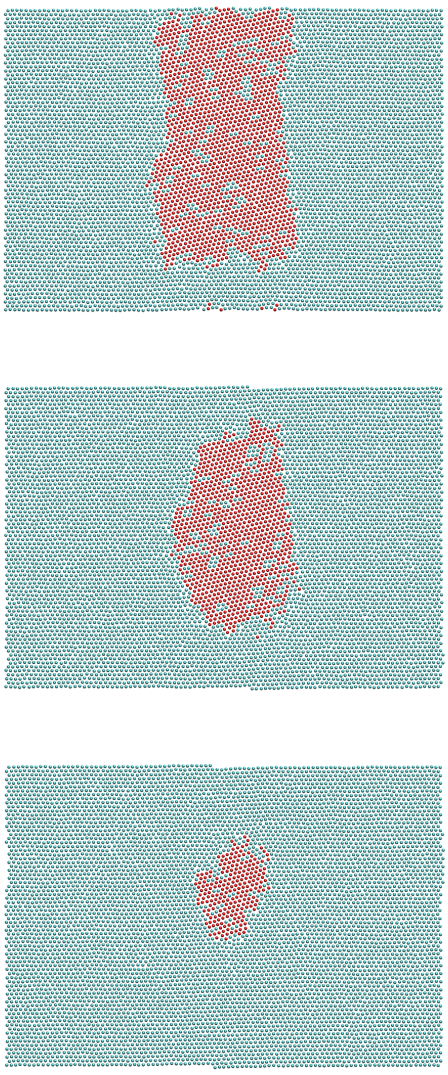


Figure 3.15.: The grain boundary movement of the rectangular smaller grain. The smaller domain, marked by red color, shrinks in size before eventually vanishes as the dislocation and misalginment of the atoms within the domain is realigned with the base grain.

left-hand side. For the estimation of the optimisation progress, the P_{fit} corresponding to each timeseries is summarised in the histogram, shown in Figure 3.17. After sampling 7300 unique parameter setups, an optimal parameter setup is found that corresponds to the fastest shrinking rate achievable by short range interactions, where the domain size decreases with a consistent slope, approximately 2x faster than the vanilla Lennard-Jones interaction potential. The best performer here corresponds to a genetic code of $r_{\text{switch}} = 1.0, r_{\text{sample}} = 1.205, N_{\text{total}} = 2048, \epsilon = 1.7$, and $D_{\text{rate}} = 20$, whereas the softened pair-wise interactions are divided among 8 CPU processes. Similarily, the top 10 performing parameter setups are traced back in Table 3.2.

Based on the data acquired from the DL-GA-rules parameter space sampling, the Phoenics optimisation samples the parameter space with a finer resolution from scratch, that is, all the new grid points in the Phoenics parameter space does not overlap with the DL-GA-rules sampling. Trained with the prior knowledge of 7300 evaluated entries, the Phoenic algorithm recovers comparable results in less than 2000 sampling attempts, the performance curve

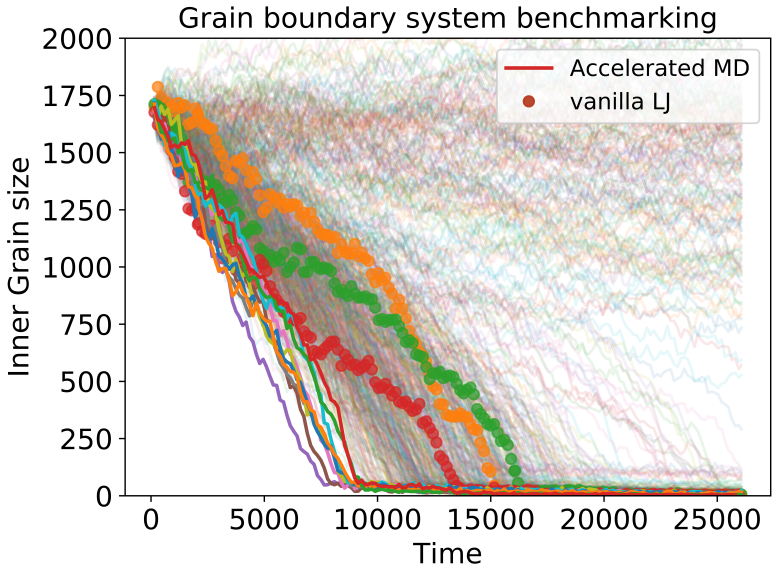


Figure 3.16.: Time series of the smaller domain size for 7300 unique parameter setups. The top performing parameter setups are marked with unique hash codes corresponding to their "genetic codes" are drawn in thick lines whereas 3 references of the vanilla Lennard-Jones trajectories are illustrated with filled circles. The faded lines correspond to the rest of the simulations with suboptimal parameter setups.

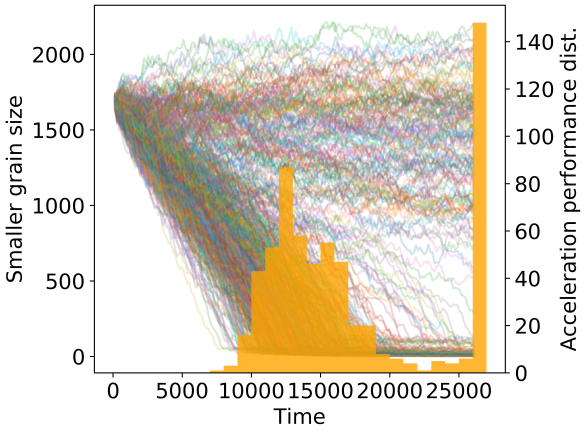


Figure 3.17.: The distribution of performances of the accleration algorithm. While for a number of the parameter setups the smaller grain does not shrink, indicated by the right-most peak in the distribution, the best performing parameter setups samples for the global optimum.

shown in Figure 3.18. The top performing parameter setups from the Phoenics optimisation are listed in Table 3.3.

The consistency between the DL-GA-rules and the Phoenics optimised parameter setups indicates several general rules for a good parameter combination. For the separate modules, i.e. the modification in the pair-wise interaction potential and the adaptive sampling, the ranked parameter setups are balanced internally within each module. More modified pairs, gauged by a larger number of softened pairs N_{total} is correlated to an improvment of acceleration performance however in specific cases the total number of modified pair cannot be fulfilled as the interval of $(r_{\text{switch}}, r_{\text{sample}})$ restricts the scope of sampling. A softer interaction potential leads to the improvement of the acceleration performance,

r_{switch}	r_{sample}	N_{total}	ϵ	D_{rate}
1.0	1.205	2048	1.7	20
1.043	1.311:	56	0.4	80
1.012	1.121	64	0.7	50
1.0	1.374	256	1.5	60
1.006	1.121	4096	0.7	70
1.037	1.184	32	0.5	50
1.049	1.479	64	1.1	70
1.043	1.5	32	0.8	40
1.043	1.5	64	0.7	40
1.031	1.205	64	0.1	80

Table 3.2.: Top 10 performing parameter setups in the rectangular grain boundary shrinking with DL-GA-rules sampling.

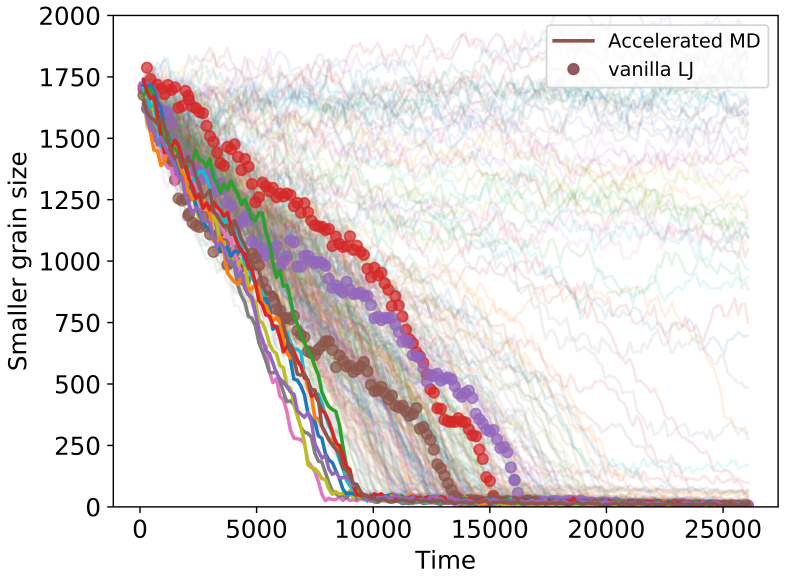


Figure 3.18.: GBshrinking Phoenics

while the switch between the softened potential and the vanilla Lennard-Jones for a softer parametrisation have to be balanced by an increase in the transition time via D_{rate}.

The balance between the two modules of our acceleration algorithm is also reflected in the top performing parameter setups as a softer interaction potential is counter-weighted by a more conservative sampling criterion for stability of the algorithm. Performance-wise, similar acceleration effects can be achieved among a range of different parameter combinations.

3.3.2. Limitations in the selection criterion: the domain shrinking a diamond shaped smaller grain and in a random grid of smaller grains

In practice, during the optimisation process for the rectangular domain shrinking problem, a major challenge in accelerating the grain boundary movement is to overcome the uncertainty in the movement direction. Since the Lennard-Jones interaction are strongly biased towards the short-range interaction, transmission of information is established via a chain of vibrations in an analogy to a series of impact events among a collection of billiards, such that the conserved momentum can be passed beyond the local vicinity and instantiate the collective movement.

r_{switch}	r_{sample}	N_{total}	ϵ	D_{rate}
1.038	1.401	289	1.89	30
1.038	1.208	279	0.894	20
1.025	1.312	55	0.277	44
1.026	1.178	260	0.538	58
1.028	1.166	93	1.464	15
1.015	1.408	434	1.327	34
1.022	1.476	99	1.616	55
1.037	1.316	46	0.866	43
1.005	1.404	1262	1.963	41
1.044	1.304	128	1.086	57

Table 3.3.: Top 10 performing parameter setups in the rectangular grain boundary shrinking with Phoenics sampling.

Ideally, the smart adaptive management of pairs should sample the pair-wise interactions on the grain boundary as the local misalignment allow for greater movements. We have also added a high-pass filtering mechansim with respect to the projected kinetic energy in the improved selection algorithm, and performed simulations on a a larger base grain of dimensions 100σx100σ in which a diamond-shaped smaller grain, tilted by 30 degrees counter-clockwise, is overlaid back to the base grid. Through the relaxation process, the diamond shaped smaller grain will vanish in a similar manner with the rectangular smaller grain however in the diamond pattern of the grain boundary we have the differentiation between the acute ends on the top and bottom of the diamond shaped domain, and obtuse angles on the left and right of the diamond shape. A snapshot of the diamond shaped domain shrinking is shown in Figure 3.19, where the acute-angled grain boundaries have not moved as rapidly as the grain boundaries with the obtuse angles. In the minimisation of the misalignment interface energy, the grain boundary in the acute-angled region incorporates more atoms with a mixed arrangement, i.e. an increased floppiness in the grain boundary movement when the dissipation takes away the extra momentum, however based only on local information provided by the short-range interaction potential, the accleration algorithm with the softened potential contributed more to the floppiness of the grain boundary and makes the decision of moving the grain boundary towards a certain direction even harder, thus reducing the efficiency of shrinking the smaller grain.

Espeically, the effect of additional filtering in the projected kinetic energy in the selection criterion is limited as the momentum passing through the grid of atoms results in a spatially dispersed distribution of higher pair-wise projected kinetic energy, or "hot-spots", while the acceleration algorithm generates such "hot-spots" when the vibrational momentum is better conserved as the the potential barrier from the pair-wise repulsion is artificially lowered. In Figure 3.19, when a number of pairs are identified by the selection criterion on the grain boundary, there still exists a number of pairs in the bulk, where the formation local clusters of trio-atoms correspond to the induction of "hot-spots" from the acceleration algorithm. The life-times of the softened pairs are generally short from the book-keeping and the "hot-spots" are constantly switched on and off across the whole simulation length-scale and time-scale.

A compromise is to test the accleration algorithm, whose selection of pair-wise interactions with the softened potential with the current selection criterion is inherently spatially dispersed, on a grain boundary system where the grain boundaries are also distributed across the whole simulation length-scale. A series of square shaped smaller grains are stacked together in a simulation box of 100σx100σ with periodic boundary conditions. The smaller grains are all sized 10σx10σ or 20σx20σ where the atoms within each smaller grain

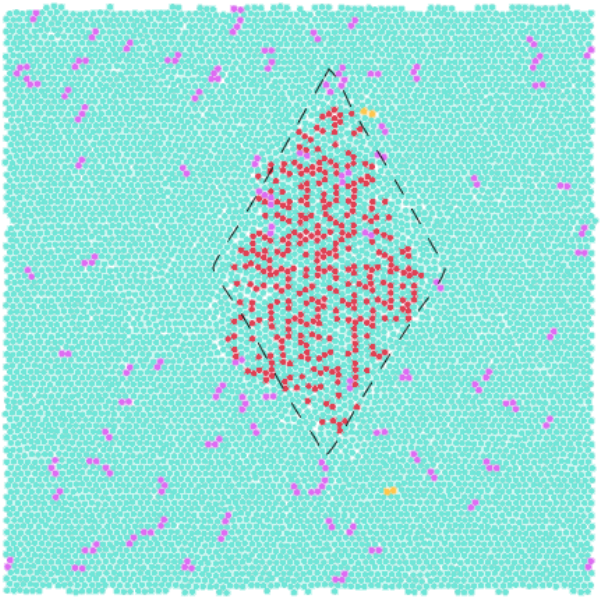


Figure 3.19.: Snapshot in the process of shrinking a diamond shaped smaller grain. The red colored area corresponds to the smaller grain with an alignment tilted 30 degrees counter-clockwise. Purple colored atoms are adaptively selected atoms with the softened interaction potential, and orange colored atoms the pair-wise interactions are in the transition from the softened potential back to vanilla Lennard-Jones. Dotted lines indicate the original grain boundary.

are tilted by 25 degrees either in the counter-clockwise direction or the clock-wise direction with probabilities of 20% and 20% respectively, while the base grain constitues 60% of the whole Lennard-Jones sheet.

Between the differently aligned domains, the grain boundary movements are highly competitive even if in the long time limit all atoms will be aligned with the base grain. In the relaxation process, local back-and-forth movements of the grain boundary becomes even more pronouced than in the case of the diamond shaped domain shrinking such that in the computationally viable timescale the grain boundary does not vanish to nil even when the misaligned area is comparable to the sandwich shaped domain shrinking problem, illustrated in Figure 3.20.

To circumvent the timescale limit, we benchmark the acceleration performance on the area of the grain boundary itself. The hexatic order parameter is used to detect hexagonal symmetry locally in a 2D system, formulated for the m-th atom as:

$$q_6(m) = \frac{1}{N_{\text{sample}}} \sum_{n=1}^{N_{\text{sample}}} \exp(6i\theta(\mathbf{r}_{mn})) \tag{3.7}$$

where the term $\theta(\mathbf{r}_{ij})$ corresponds to the individual relative orientations between the local atom and the sampled atom, when the parameter N_{sample} regulates the sampling range. The overall Q6 order parameter as an ensemble average $Q_6 =< q_6(m) >$ from the individual atoms reflect the misaligned portion of atoms across the entire domain.

In our implementation, the Q6 order parameter is calculated from the orientation from the next-nearest-neighbours with respect to the atom being investigated, representing the area covered by the grain boundary where the local alignment is mixed. In the time-series retrieved from the MD trajectory, the initial Q6 order parameter is very close to the perfect

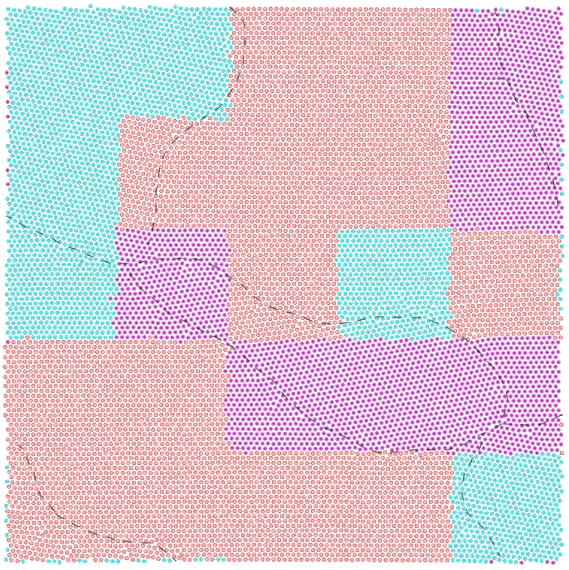


Figure 3.20.: Snapshot in the process of domain shrinking in a grid of smaller grains of dimensions 20σx20σ. The red colored area corresponds to the base grain constituting the 60% of the total domain area. Purple colored grains that are originally tiled 25 degrees clockwise, and cyan colored grains originally tilted 25 degrees counter-clockwise. Dotted lines indicate the in situ grain boundary.

alignment as the grain boundary is constituted from the bi-layer on the interfaces between the small grains, while during the relaxation process, the floppiness of the moving grain boundary disrupts crystalline-like the hexagonal alignment in the close-packed Lennard-Jones atoms such that the Q6 parameter is immediately reduced before the shrinkage in the grain boundary recovers the Q6 towards the constant of 1.0 for ideal packing, shown in Figure 3.21.

However due to the short-range effect of the Lennard-Jones interaction, the restructuring is limited to the close vicinity whereas the global passage of momentum for the collective movement is greatly hindered. For the 10σx10σ smaller grains, the realignment on the grain boundary spreads inwards in the analogy of a shock-wave however for the 20σx20σ smaller grains a single domain can be divided into multiple subdomains with different alignments subjected to the complex pattern in the grain boundary movement, shown in the division by *in situ* grain boundaries of the individually colored domains in Figure 3.20, and in the slow increase in the Q6 order parameter in Figure 3.22.

In both cases of the 10σx10σ grain and the 20σx20σ grain, the accleration via the softened potential with transferred parameter setups from the sandwich shaped domain shrinking optimisation is around 30% for the 10σx10σ grain and only marginal for the 20σx20σ grain compared to vanilla Lennard-Jones interaction simulations. The abrupt change in the grain boundary movment directions is clearly shown in the Q6 order parameter time-series as step-like changes for both the accelerated algorithm and the vanilla Lennard-Jones, while with the softened interaction potential the accelerated algorithm may induce worse orientation in the initial stage of the relaxation as the softened potential allowed for greater floppiness in the grain boundary movement.

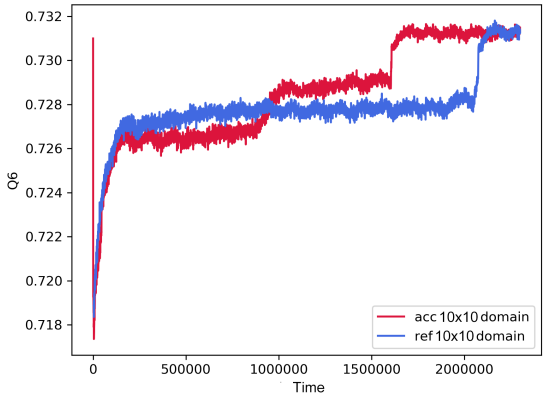


Figure 3.21.: Q6 orientation order parameter evolution with 10σx10σ smaller grains. The Q6 order parameter drop initially from a near-perfect crystal structure before the relaxation recovers a globally aligned base grain as the grain boundary shrinks to nil in the end-state. In the process the accleration algorithm maintains initially a worse performance due the increased floppiness in the grain boundary movement but performs superior overall.

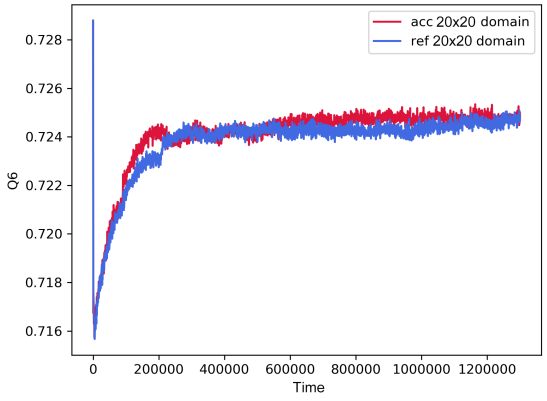


Figure 3.22.: Q6 orientation order parameter evolution with 20σx20σ smaller grains. The Q6 order parameter drop initially from a near-perfect crystal structure before the relaxation recovers gradually. In the process the accleration algorithm maintains marginally better performance than the vanilla Lennard-Jones.

3.3.3. Conservation of thermodynamic order parameters: overall potential energy and Q6

The acceleration algorithm is intended to work with minimal alternation to the vanilla thermodynamics, as the typical number of modified pair-wise interactions constitute less than 5% across the entire ensemble of pairs. On the one hand, through the implementation of the the adaptive selection process, a predefined upper limit of N_{total} is implemented such that once the modified pairs fills the list, the selection is switched off as the remaining pair-wise interactions follow the vanilla Lennard-Jones formulation. During the benchmark, for some parameter setups, the maximum allowed number of modifications N_{total} reaches around 40% of the total number, while practically through the adaptive sampling the number of modified pairs are always less than the threshold. Furthermore, with the additional implementation of the high-pass filtering in the projected kinetic energy, as the grain boundary vanishes, the number of modified pairs decreases accordingly.

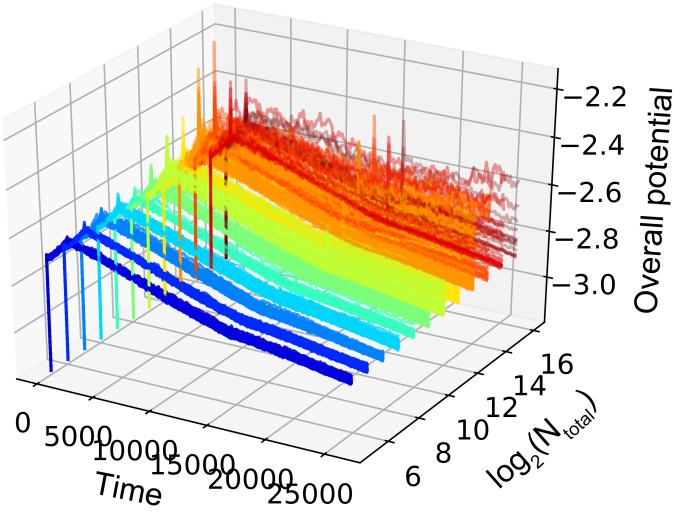


Figure 3.23.: Conservation of the overall potential energy in the smaller grain shrinking with respect to various control parameters. For all sampled parameter setups, the individual simulations are categorised and color-coded with respect to the maximum number of modified pair-wise interactions N_{total}, while the rest of the parameters are varied correspondingly for optimised performance.

On the other hand, the "softness" of the modified potential is dependent upon the total number of modified pairs through the optimisation process while a softer interaction allows for more conservation of momentum for individual atoms, eventually translating to a higher fluctuation in the overall interaction potential, in Figure 3.23. Among the simulation runs with the same setting of N_{total}, the time-series of the overall potential energy are overlaid, represented as a "potential band" with consistent trends which coincides with the reference potential energy of the vanilla Lennard-Jones simulation. For extremely large value of N_{total}, the excess number of modified pairs alternates too much the overall potential energy, seen as divergent time-series in the "hotter" colored section, whilst for the simulations with typical N_{total} values less than 2^{12}, the variation in the overall potential energy, visualised as the expanded width of the potential band, is less than 2%.

As has been discussed during the benchmark on the random grid of smaller grains, the Q6 orientation order parameter depicts the overall thermodynamic property of the atomistic system being investigated. For the shrinking of an individual smaller grain, an automated tracking algorithm calculates between the local atom and its nearest neighbours for the orientation differentitate atoms belonging to the smaller grain and the base grain. The Q6 parameter can be calculated for the separate grains, shown in Figure 3.24. A bifurcation behaviour is observed at around $T = 4\text{E}4$ when the base grain continues to grow, the Q6 measurement increasing, the atoms in the smaller grain are less well orientated and subject to large fluctuations with respect to the floppiness of the grain boundary movement and reduced size of the smaller grain. As is indicated in the convergence in the Q6 measurements for the whole system and the base grain between simulations with vanilla Lennard-Jones and the accelerated algorithm, plus the comparable Q6 measurements for the smaller grain, the overall Q6 order parameter is conserved while we artificially introduce the softened potential.

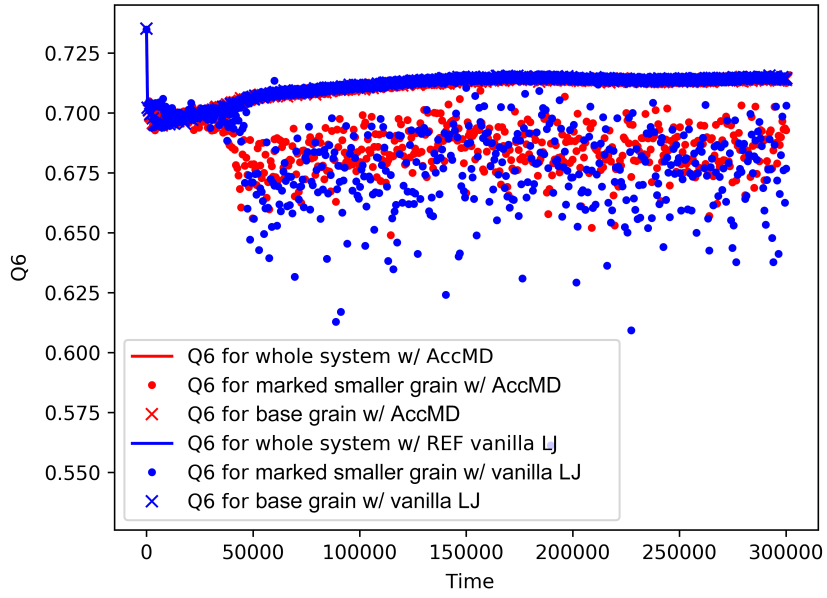


Figure 3.24.: Separate Q6 measurements for the whole system, the base grain, and the smaller grain. For the whole system and the base grain, the time-series from the vanilla Lennard-Jones and the accelerated algorithm overlaps consistently, while for the smaller grain, the dotted patterns are comparable.

3.3.4. Conclusion and outlook

The efficiency of molecular dynamics simulations is often limited by the kinetic trapping especially in the dense systems, where the strong short range repulsion renders high energetic barriers such that the movement of the atoms are mostly local vibrations, taking up most of the computational capacity whilst the collective movement beyond the local spatiotemporal scale, e.g. mixing and morphology formation with polymeric melts and the formation of secondary structures in the proteins, is crucially missing from the picture. Through the softening of the local interaction potentials for a limited small ratio of the pair-wise interactions adaptively, the kinetic trapping is alleviated which results for the benchmarking grain boundary movement system an acceleration factor approximately 200% the speed of the original system. For the system specific parametrisation of the acceleration algorithm, an systematic, artificial intelligence enhanced automation process is proposed in quickly exploring the high dimensional parameter space. During the accelerated MD simulation, due to the spatiotemporally local and rare modifications the true dynamic trajectory as well as the thermodynamics properties are conserved when no additional global resampling and back mapping is implied or required.

The acceleration algorithm is implemented compatible with the highly scalable architecture of LAMMPS software and is incorporated into the in-house developed Simstack workflow engine. [48] The modular development of both the acceleration algorithm and the automated optimisation method can be transferred to alternative atomistic systems beyond the Lennard-Jones like interaction potentials while the combined deep learning / genetic algorithm / hand designed rules / Phoenics parameter space exploration method is adopted for alternative in-house screening efforts of oligomeric folding predictions.

As of its current status, the acceleration algorithm is limited to the modification of the Lennard-Jones interaction, while the introduction of a series of parameter and hyperparameters due to technicalities restricts its scope of application. Moreover, the introduction

of a range of parameters complicates the system-specific optimisation problem and requires considerable training effort with the deep learning based optimisation process.

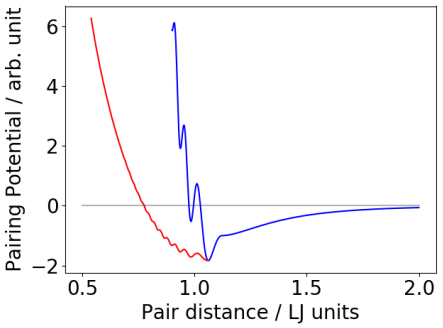


Figure 3.25.: Potential rescaling as the softening method for a generic form of pair-wise interaction

We are looking into improving the potential softening module with a potential rescaling method which applies to arbitrary pair-wise interactions while the parameter optimisation effort is significantly reduced as there are fewer control parameters. Meanwhile, the refinement in the adaptive management module aiming at reducing the floppiness and artificial generation of "hot-spots" may yield improved acceleration performance.

4. High-throughput molecular dynamics simulations of ionic liquid diffusion in metal-organic frameworks

4.1. Introduction

Limited to the case-specific application, e.g. the likes of ionic liquids and metal-organic frameworks, the all-atom MD simulations are performed with well-established, tailor-made force-fields. With an inclusion of an external electric field, the ionic liquids diffusion within the metal-organic frameworks are driven where the molecular movements can be accelerated via scaling up the external driving force, as the diffusion properties of the under the original external electric field can be retrieved via extrapolation [49, 50] and other analysis methods [51] during the the post-mortem data processing. Meanwhile, for similar materials with similar simulation setups, typical for the material screening and drug design cases, the implementation of workflows offers an automated protocol in quickly deploying a series of MD simulations with marginal additional effort for the preparation and the post-processing. In this chapter, for the diffusion of ionic liquids embedded within metal-organic frameworks, a range of MD simulations are performed corresponding to variable types of the guest ionic liquids, the hosting metal-organic framework, the external control parameters as the electric field intensities, the major diffusion channel alignments, and the ionic liquid loading factors. The comparison for the diffusion features characterised by the measurement in ion mobility of [BMIM][NTf_2] in HKUST-1 and UiO-66, 67, and 68 along with [$Li_{0.2}BMIM_{0.8}$][NTf_2] in HKUST-1 with respect to the ion loading factor is retrieved with only a few mouse clicks.

In addtition to the streamlined execution of MD simulations, the workflow engine also allows for modular structuring in the operations of MD, when the results from one workflow can be efforlessly transfered to another operation as an input, e.g. the all-atom metal-organic framework structure serve as a starting point for adsorption of other guest molecules as CH_4 [52], the morphologies of the ionic liquid molecules summarised as an reference state in *ab initio* calculations, etc.

The compiled all-atom MD workflow of ionic liquids diffusion in metal-organic frameworks as an independent package grants access to high-throughput simulations for fast screening in the combinatorial guest-host selection, and offers a systematic means in device design and control strategies for organic-electronics applications such as the supercapacitor and sensors via revealing the unique conduction mechanism of confined ionic liquids in the nanoporous host materials.

4.1.1. Ionic liquids

Room-temperature ionic liquids (ILs) are a class of molten salt constituted of designer molecular ions. Working as electrolytes they preserve advantageous properties of non-flammable, thermally stable, and negligible vapor pressure. [53, 54] In contrast to common

salts which tend to melt at higher temperatures as the ionic bond is stronger than van der Waals interactions, the IL constituted from molecular ions, exhibits a less pronouced ion pairing due to greater separation between ionic centers due to the molecular size, and in some cases delocalised distribtuion of net charges across the ionic molecule. The selective features in the "bulkiness" of the IL molecule and the charge distribtuion offers a great opportunity in molecular design for the organic electrolyte, whereas the combinatorial possibilities from the rich dictionaries separately from the cations' and anions' offers an unrivaled tunability. [55] Further more, mixing IL with common cations, e.g., Lithium ions (Li-ions), in the electrolyte-based applications, provides novel opportunities in the control strategy to gauge diffusion based on ion concentration where the vivid contrast between IL cations' bulkiness against the mercurial nature of Li-ions results in fundaementally different conduction mechanisms as well as distinguished patterns in the failure of transport.

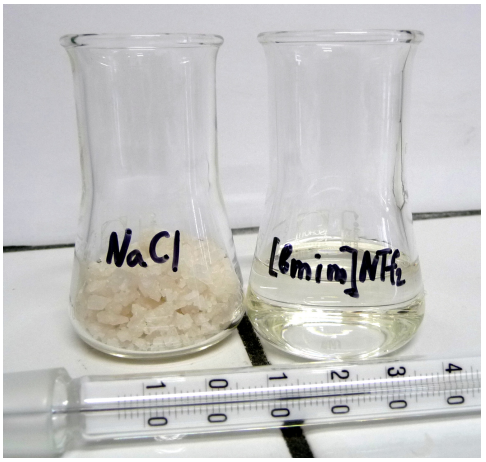


Figure 4.1.: Room temperature ionic liquid [BMIM][NTf_2] compared against a common salt of NaCl. Image license: public domian, original source: Zgonnik.

4.1.2. Metal-organic frameworks

It is conceptually interesting to debate whether the metal-organic framework (MOF) is polymer, while a sequentially and spatially ordered material with designer organic ligands grants poly- possibilities for material properties and functions in storing, filtering, and sensor applications. [56] Ever since the succesful synthesis of MOFs, more crystalline framework structures are inspired from MOF applications. Covalent organic frameworks (COF), replacing the metal-coordination bonds with colavent linking, can be derived from and synthesised with the help of MOFs, while a MOF/COF/polymer hybrid offer further functional and mechanical tuning capabilities. [57]

The study of man-made porous crystalline solid-state materials originated from the study of zeolites, which is an aluminosilicate found with multiple occurrence structures in nature. The fundamental building block is a silicate tetrahedron, with bonds from each silicon reaching out to four interdigitated oxygen atoms before linking to another silicon atom. In typical cases of zeolite, aluminum replaces silicon atoms with certain mixing ratios, introducing a net charge in the framework, thus providing active binding sites for adsorption applications, e.g. in water purification. Ordered multi-membered-ring structures formed from silicate work as diffusion channels for agent molecules with tunable nano-pore sizes. For instance, a 10-membered-ring of silicate grants more diffusion than a 6-membered-ring. The distinct "bricks" that build the zeolite network is termed as the secondary building units (SBU) hence with different components and arrangements of the SBU, the structure and porosity of the zeolite can be designed to fit certain sized guest molecules. While a few

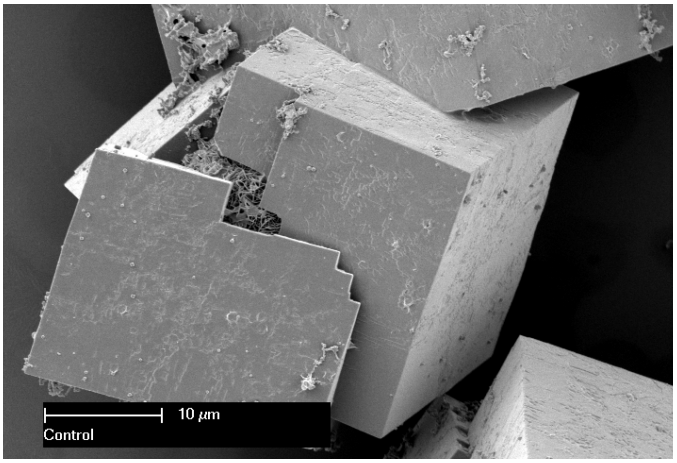


Figure 4.2.: Scanning electron microscope (SEM) image of a seed during MOF growth. Well-ordered surfaces of the artificially synthesised crystalline structure is clearly shown. Image licence: CC BY 3.0, original source: CSIRO.

tens of structures are discovered to exist in nature, some hundred occurrences of unique structures have been established from theory to compile a library for zeolites. [58]

Inspired by blocks building scheme in zeolites, MOFs are a class of designer porous materials constructed from metal ion clusters, termed as the SBU, bridged by organic linkers. Comparable to the SBU in zeolites, the SBU in MOF includes metal complexes, which also offers exposed binding sites. For the organic linker connecting the SBUs, the compoential combination can be substituted via introducing modifications to the linker backbone for control over linker length and thus the porosity of the MOF, as well as functional groups for even more complex linkages between individual linkers. Multiplying the design options both for the SBU in Cu, Zn, and Zr based metal-complexes and numerous complementing linkers that stabilizes the MOF, the total possibilities inflates way beyond that of zeolites. Since its first breakthroughs in the 1990s [59], tens of thousands of MOF structures have been fabricated, while around a million of MOF structures have been proposed. [60–62]

Generally speaking, the electric conductivity of MOF is negligible. In order to utilize MOFs for applications such as electronic devices, additional conduction mechanism have to be added to MOFs system either by modifications in the MOF linkers [63], or addition of alternative conduction pathways within the MOF pores such as introductions of Fullerene [64], etc. In the following discussions, we consider conductivity of the combined system of IL embedded inside of the MOF.

4.1.3. [BMIM], [NTf_2], and Lithium ions embedded in the HKUST-1 and UiO family MOFs

A common pair of IL ions are 1-Butyl-3-methylimidazolium hexafluorophosphate ([BMIM]) and Bis(trifluoromethane)sulfonimide ([NTf_2]). [BMIM] cations and [NTf_2] anions are weighted 139.22Da and 280.14Da respectively and are comparable in the calculated van der Waals sizes. Through the *ab initio* modelling of the individual ions, results suggest [BMIM] carries a net positive charge delocalized across the whole molecule, whereas [NTf_2] retains a concentrated negative charge in the nexus nitrogen. The fact that [BMIM] is also more flexible in terms of the tail section makes it hydrodynamically favorable in the transport compared to the hefty [NTf_2] when the wedge-shaped anion molecule possesses a more rigid form especially in the S and F components, resulting in effectively a larger molecular size versus the cation.

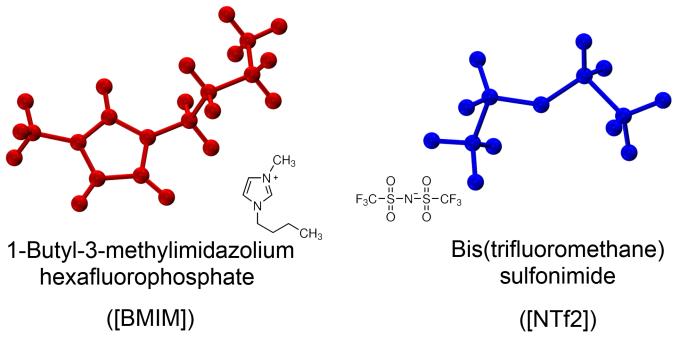


Figure 4.3.: The ionic liquid pair of [BMIM][NTf_2].

MOF	unitcell dim.	density	pore volume	1st pore size	2nd pore size
HKUST-1	2.63nm	$0.35 g{\cdot}cm^{-3}$	$0.57 cm^3{\cdot}g^{-1}*$	1.0nm	< 0.5nm
UiO-66	2.10nm	$1.20 g{\cdot}cm^{-3}$	$0.45 cm^3{\cdot}g^{-1}$	0.8nm & 1.2nm	
UiO-67	2.71nm	$0.71 g{\cdot}cm^{-3}$	$1.05 cm^3{\cdot}g^{-1}$	1.2nm & 1.6nm	
UiO-68	3.33nm	$0.46 g{\cdot}cm^{-3}$	$1.82 cm^3{\cdot}g^{-1}$	1.6nm & 2.1nm	

Table 4.1.: Pore specifications of the SURMOFs of HKUST-1 and UiO-66, 67, and 68. The HKUST-1 in its special case is composed of larger and smaller pores whereas the small pores are inacessible to the [BMIM] and the [NTf_2] molecules due to limitation in the pore aperture. * For HKUST-1 the routinely specified pore volume in commercial sources correspond to accessible probe center pore volume which excludes the volume in the secondary pores, whereas for the adsorption of gas molecules the reported pore volume of HKUST-1 is approximately 0.66. [65]

HKUST-1 is a typical copper-based MOF which has been widely synthesized and extensively investigated. The Cu-Cu SBUs in HKUST-1, also known as the paddle wheel complexes, are connected with the benzene-1,3,5-tricarboxylicacid (BTC) linker. There are 2 major types of pores in HKUST-1, the secondary pore with limited aperture blocks passage of [BMIM] or [NTf_2], while for the primary pores there are 2 variations with slightly different pore sizes due to distinguished linker-plane orientations aligned against the paddle wheel planes. UiO family MOFs on the other hand, is a well-studied MOF family with precisely controlled incremental pore sizes. The Zr-based SBUs are all identical across the UiO-66, 67, and 68 MOFs, whereas the Zr-complexes are connected with a linearly step-wise elongated series of linkers as phenyldicarboxylate, biphenyldicarboxylate, and triphenyldicarboxylate groups. Via systematically gauging the linker length, pore apertures in UiO family MOFs range from 0.8nm to $2.1nm$ where as the pore volumes range from 0.54 to 0.83. A comprehensive set of statistics registers pore information for HKUST-1 and UiO family MOFs in Table 4.1.

The molecular movement of IL in the MOFs results in a conducting system with conduction mechanism deviating from either the IL or the MOF. [66] From literature, the IL mobilities in the bulk phase be subjected to both charge-transfer and featured formation of ion pairs. [67]. However for the diffusion of IL For IL confined in the nanopores, frequently referred to as "ionogels", they behave vastly different from the bulk phase counterpart. [68] In the MOF, the IL density can be controlled by the specific loading factor, while, as will be later shown, the diffusion properties show strong IL concentration dependence. Comparable to the conventional fluids in the low concentration scheme, for low IL loading, the ion-ion interaction remains insignificant as the individual IL molecules are located in separate MOF pores; however in a stark contrast, in the high loading scenario the ion-ion interaction becomes the predominant factor, as the local properties, e.g. instantaneous ion density, pore aperture, and local free pore volume comprehensively lead to an inhomogeneous distribtuion

of ILs with spatio-temporal deviations from the overall homogeneity. In the confinement of MOF pores, the spontaneous local density fluctuations of IL distribution may lead to emergent transient blockage of IL transport as cations and anions bunch together. In the extreme cases the emergent blockage develops into a global terminal failure of transport at the ordered bottlenecks in the form of pore apertures limits within the relatively rigid HKUST-1. IL diffusion in UiO family MOFs shows a different mechanism by travelling along the diagonal transport channels as ILs manages to circumvent the blockage from the framework. In the case of IL doping with Li-ions, the dopants' ionic radius is so small that alternative routes can be taken in the secondary pores in HKUST-1, where the molecular ILs are denied access due to the limiting pore aperture.

4.2. Surface MOF synthesis: layer-by-layer growth, loading of IL inside the MOFs, and the conductivity measurement from experiments

In the experiment setups, the reference IL@MOF system is established on a prepared surface. Specifically, the MOF film is prepared upon the substrate in a layer-by-layer (LBL) fashion. [69] LBL utilizes robotic cycles of depositing exclusively SBU and linker parts on the substrate in separate iterations, at the end of which process the extra growth more than a single layer gets rinsed. Validated from X-ray diffractogram measurements the SURMOF growth is well controlled in the LBL assembly where the MOF grows in the (100) direction for HKUST-1 and (111) direction for UiO family MOFs such that the main diffusion channel in the MOF is parallel with respect to the substrate.

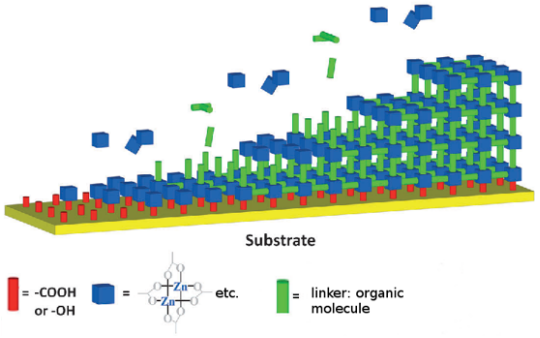


Figure 4.4.: Schematic representation of layer-by-layer assembly of SURMOF. The organic linkers and secondary building units are deposited on to the processes surface iteratively, forming a well-defined framework structure.

Distinguished from the pellets pressed from MOF powder, the precise pore channel alignment is crucial in the characterisation for diffusion of guest molecules, while in the MOF powder the diffusion also takes place on beyond the interfaces of the crystals where the MOF surfaces are "nanowetted" [70], rendering the diffusion mechanism of the guest a mixed pattern between the embedded within the MOF confinement and of bulk phase in the wetting region. For the characterisation. In the specific case of IL embedded in MOFs, from the literature a wide scope of investigations have utilised the MOF powder as the host framework and reported ever increasing conductivity as the IL loading is increased, when the LBL growth of the MOF allowed for more detailed investigations of the IL transport and its concentration-dependent failure of transport discovered for the embedded IL.

MOF used in the scope of this thesis are assembled on a quartz surface with interdigitated gold electrodes to measure electric conductance via the LBL technique. When an external AC current is applied to the interdigitated electrodes, the external electric field induced aligns with the HKUST-1 pore channel in the (100) direction and with the UiO pore channel in the (110) direction. Examined under the scanning electron microscope (SEM),

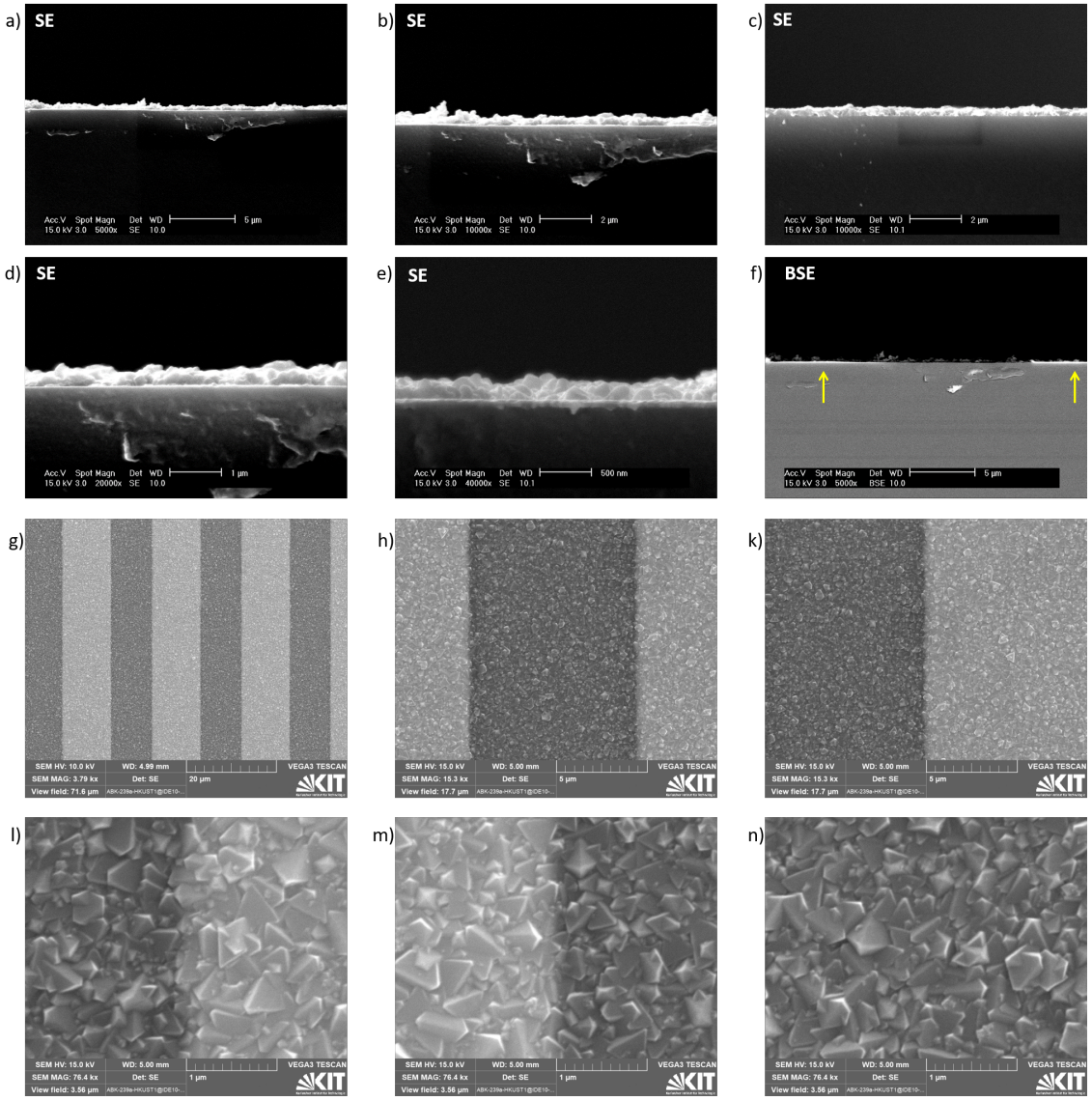


Figure 4.5.: Side-view and top-views of the scanning electron microsope (SEM) images of the SURMOF HKUST-1. Figures (a-f) corresponds to the side-views of the MOF-on-substrate system as figures (g-n) correspond to the top-view. Figure a) and c) are taken from different sections; figures b) and d) are enlarged images of a), whereas figure e) is a zoom-in of figure c). Figure f) is taken with identical view section as figure a) with backscattered electron detector (BSE), as the gold electrodes are seen as bright planar layers on the substrated, indicated by yellow arrows, when all the other images are taken with secondary electron (SE) detector. In the top-view images of the SURMOF, the interdigitated gold electrodes are visible as bright stripes in the background. The zoom-ins with variable magnification factors on the MOF thin film in different view sections yield homogeneous coating whereas the crystal facettes can be identified in high-resolution images of (l-n).

the side-views and top-views of the SURMOF in the case of HKUST-1 are summarised in 4.5.

Loading of IL into the MOF is achieved from IL-acetonitrile solutions and rinsing from the top side of the LBL-grown SURMOF such that no bulk phase IL resides outside of the MOF film. Often reported in the case of MOF powders, typically pressed in pellet shape, rather than SURMOF, Li-ion and IL conduction increase as loading of charge carriers increase. However, by using electrochemical impedance spectroscopy (EIS), it has been

discovered that an excess of charge carriers does not reside in MOF pores. While the excess of IL that is "nanowetted" on the external surface of the MOF improves contact of electrolyte to electrodes, the liquid IL outside distinguish themselves as "bulk phase" from the "gel like hybrids" IL embedded inside nanoconfinement. As will be later shown, for embedded IL, there is exceptional density dependence with regard to ion transport. For IL embedded inside HKUST-1, as [BMIM] and [NTf_2] measure in smallest cross-sectional diameters both around 0.5nm, the secondary pores enclosed by the BTC linkers, with pore aperture less than 0.5nm, generally do not allow ILs' access. While due to possible misalignment, component defects, and deformation of the linker, IL diffusion through the secondary pores is still possible, the mobility through those pores is significantly reduced. Other scattering effects can be routinely observed potentially due to other misalignment and defects in the MOF, while it is worth noting that in MOF synthesis defects are very hard, if at all possible, to be prevented in the assembly process, for diffusion of small ions, such as Li-ion, the deviation of simulation results on the perfect MOF away from that of the experiment may be comparable to that of MOF powders and SURMOF.

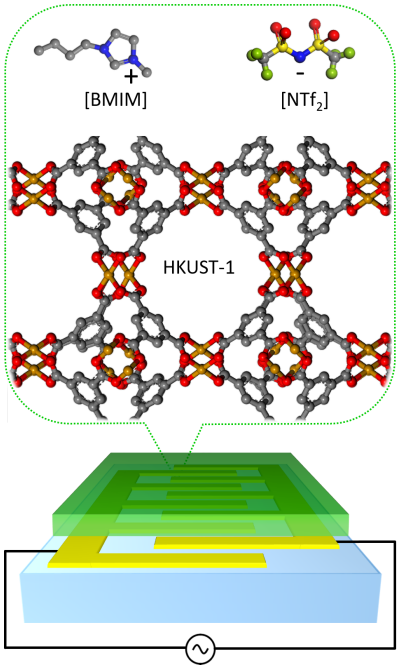


Figure 4.6.: Scheme of the [BMIM][NTf_2]@HKUST-1 system. The SURMOF is grown on the substrate with interdigitated electrodes and the IL is loaded into the MOF when extra IL from the MOF surface is removed.

The SURMOFs are loaded with the IL with various loading factors via immersion of the MOF sample in IL:acetonitrile solution with corresponding IL ratios at room temperature. Subsequently, as has been discussed, the essential step of rinsing with acetonitrile removes extra IL. For the MOF loaded with IL:acetonitrile, a drying step with pure nitrogen flow removes the acetonitrile, yielding an IL@MOF sample. Taking the pure IL@HKUST-1 system as an example, infrared reflection absorption spectra illustrates the vibration bands of [BMIM][NTf_2] after the IL:acetronitrile immersion between 2900 and 3200cm^{-1}. An insignificant vibration band at 2200cm^{-1} corresponds to the acetonitrile residual. The small intensity of the acetonitrile signal interprets into approximately 1/6 acetonitrile molecules per MOF unitcell, 2 orders of magnitude less than the embedded IL.

Loading of IL is measured both by quartz crystal microbalance (QCM) and energy-dispersive X-ray spectroscopy (EDX). For QCM, the frequency shift before and after the loading of IL is determined from the Sauerbrey equation with the SURMOF reference thickness at

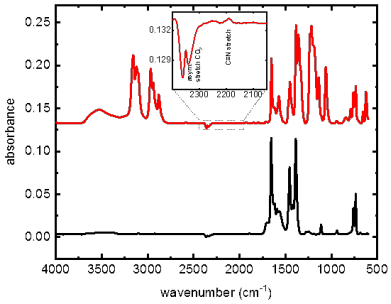


Figure 4.7.: Infrared reflection absorption spectra (IRAS) of loading the IL into HKUST-1. IRAS spectra before (black) and after (red) the immersion of MOF in the IL-acetronitrile solution show signals belonging to [BMIM][NTf_2] in the range of 2900cm^{-1} to 3200cm^{-1}. when the vibration signal at 2200cm^{-1} of acetonitrile, enlarged in the inset image, is minimal.

170nm, and the IL molar mass at 419.4g·mol^{-1}. A maximum loading of 17.4 ± 6.3 IL pairs per MOF unitcell is determined while the variation remains significant among different measurements. EDX improves the IL loading accuracy via measuring the ratio between Cu component, which is unique in the MOF, and F component, which is unique in the IL anion. EDX spectra are obtains at different locations in the sample and a maximum loading of 18.6 ± 0.7 IL pairs per MOF unicell is determined. From data from both measurements, the IL loading factor show a linear correspondence with the IL concentration in acetonitrile, see Figure 4.8.

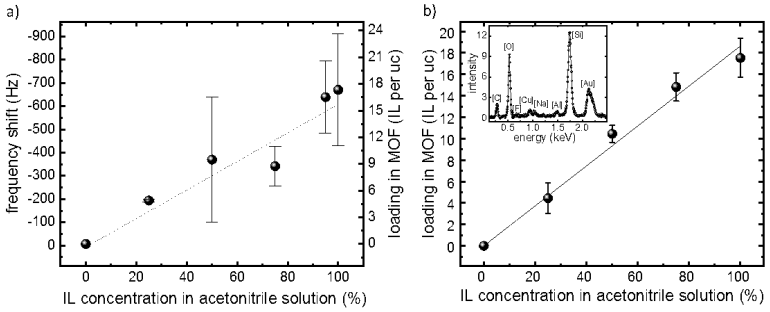


Figure 4.8.: a) Quartz crystal microbalance (QCM) and b) energy-dispersive X-ray spectroscopy (EDX) quantification of IL loading in the MOF.

In the case of Li-doped IL, the [BMIM][NTf_2] and Li[NTf_2] are mixed with a fixed ratio of 4:1, before the mixture is soluted in acetonitrile loaded into the MOF. In the case of HKUST-1 powder synthesis, the sample is prepared from dissolving $Cu(NO_3)_2 3H_2O$ (0.87g, 3.6mmol) and BTC (0.42g, 2.0mmol) in the 50:50 volumeric mixture of ethanol and de-ionised water. The solution is then heated at 100°C and annealed before the crystalline powder is recovered, centrifuged, and washed with the mixutre of 50:50 vol. % ethanol and water, and activated at 120°C. ILs are added to the MOF powder separately as different amounts of IL are mixed with separate MOF samples with constant heating at 120°C.

In the experiments the electrochemical impedance spectra are measured with the impedance analyzer with a frequency range of 0.5Hz to 5Hz. The AC current is applied to the interdigitated gold electrodes, upon which the IL@MOF system is placed, when the electric field between the adjacent electrodes approximately 0.03V·μm^{-1}. The conductivity as the reciprocal value can be inferred from the impedance spectra as the IL@MOF system can be reduced into the Randles circuit model, illustrated in Figure 4.9 except for the

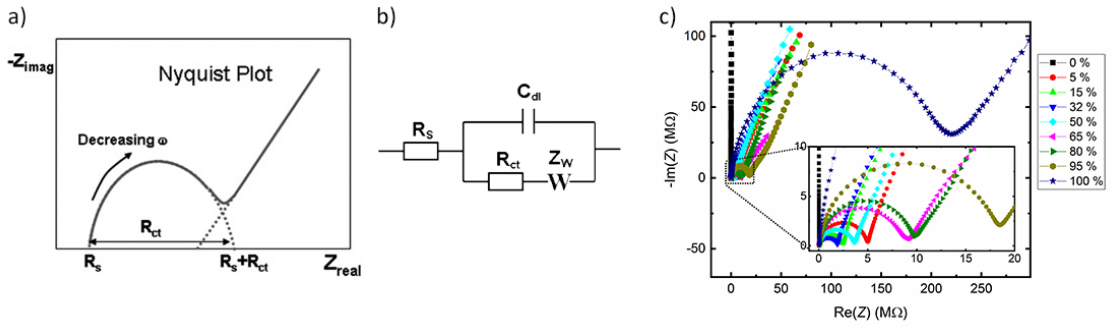


Figure 4.9.: Nyquist plots for the impedance measurements. The impedance spectra measured from the experiment corresponds to the a) Nyquist plot for the b) Randles circuit model. In the IL@MOF system, from the the experiment impedance spectra in c), resistance of the IL is recovered, whereas in the case of empty MOF, marked as black squares, only the capacitive resistance is measured.

MOF without IL (black data points). For the pure MOF, a capacitive resistance equals a minimal conductivity of $5E(-11)Sm^{-1}$, whereas the molar conductivity of the bulk phase [BMIM][NTf_2] is approximately $1E(-4)Sm^2mol^{-1}$ at room temperature, the embedded IL ranging from approximately $50nSm^2mol^{-1}$ to $0.07nSm^2mol^{-1}$.

4.3. Workflow design for high-throughput IL@MOF MD simulations

The dictionaries of commercially available ILs are greatly varied and the list of synthesizable MOF is constantly growing with time. While for the IL their chemical properties such as charge distributions are similar the molecular sizes of those ions are distinct. In the case of MOFs, the building blocks can be chosen separately from secondary building units, where for the metal ions, a number of choices as Cu, Zn, and Zr have been proven successful in holding a stable framework, and organic linkers, which is critically defined in the linker lengths, subsequently defining the pore volume of the MOF, can also be modified with various side groups that introduces additional chemical functions as proton-donor, light-switchable groups, etc. have been widely reported in the field. In the all-atom simulation where the electric conduction is majorly contributed from the travelling ILs embedded in the MOF, the diffusion subjected to guest-guest and guest-host interactions are well-described with the hybrid UFF/UFF4MOF for the MOF and OPLS-AA based forcefield dedicated to IL. A quick protocol to perform an all-atom simulation where for any cation and anion pair of IL, from typically PDB format atomistic descriptions, and an arbitrary MOF, routinely described with CIF crystal structures, can be established via an automated process by feeding information through a series of operations that builds for the MOF supercells, packs the IL within the MOF, and assign proper partial charges. Such a tool grants access for the scientific community a one-click preparation, execution, and analysis to the diffusion of IL@MOF problem.

One of the platforms of workflow engineering is Simstack [48] that allow modular design with an easy-to-access user interface. The workflow incorporates the modular parts of MOF parametrisation, IL parametrisation, IL packing inside of the MOF, the MD preparation, the MD execution and the post-processing. Among the modular parts of the IL@MOF workflow works are defined in operation nodes termed as WaNos whereas the information flow between WaNos generate a directed graph which guides the input and output from each WaNo to the connected nodes. The general purpose WaNos for MOF parametrisation and LAMMPS simulation executions are implemented in the Simstack platform, in particular, the MOF parametrisation Wano is being tested within the group for quick DFT screening

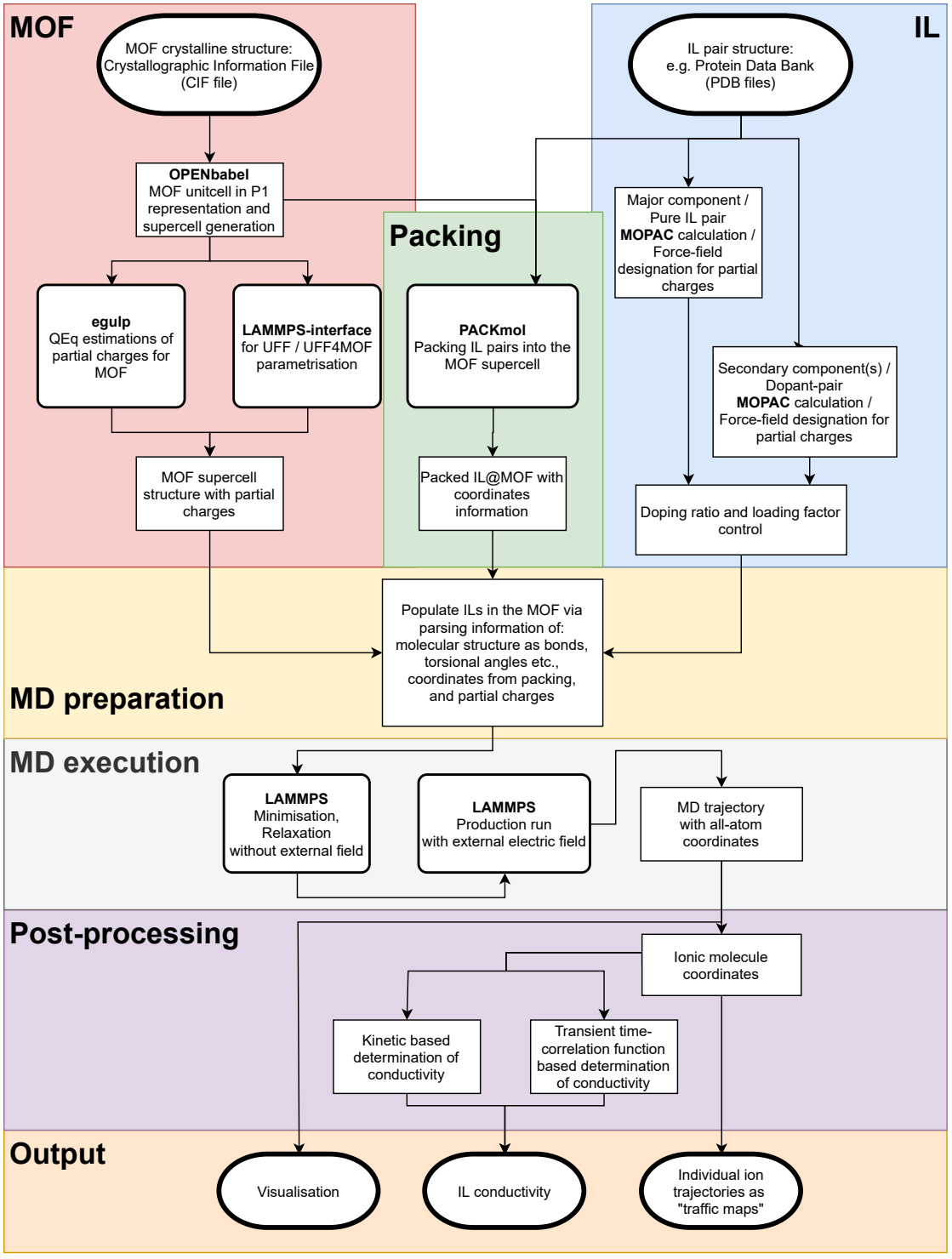


Figure 4.10.: Workflow of high-throughput IL@MOF simulations.

of MOF structures, while the implementation of the additional dedicated WaNos in support of the IL in MOF analysis is work in progress.

The parametrisation of the MOF and the IL can be separated when force-fields of UFF4MOF for the framework [23], the OPLS-AA based, dedicated force-field for IL [71], and the OPLS-AA force-field for Li-ion [17] are empirically compatible in the simulation of IL@MOF systems. With the increased interest on MOF studies and the development of simulation aided modelling, a great number of proposed MOF structures are available in online MOF

databases, e.g. the computation-ready, experimental MOF database (CoRE DB) [60], the cambridge structural database non-disordered MOF subset (CSDSS) [61] as crystallographic information files (CIF files). However, to generate the all-atom model from the CIF sources, validation have to be performed before the simulation unitcell of the MOF can be derived from the CIF info, e.g. in some CIF files the hydrogen atoms are omitted, or the definition of the MOF may not be unique as the crystalline structure can be either defined in a trilinear simulation box or a rectilinear simulation box. For conformity and compatibility with most features developed for the MD simulation, the P1 representation of the crystalline structure serves as a nexus for all of the conversions from CIF data to MD-ready data, when the CIF info of the MOF unicell are passed on as input to the egulp, an independent package derived from gulp for the designation of partial charges based on the charge equilibration (QEq) method, and to lammps-interface, a standalone MOF-oriented software package developed as a bridge between the MOF structure information and MD simulations. The QEq method requires a prescribed dictionary of ionisation potential, the electron affinity, and the atomic radius, which can be parametrised either from experiment measurements or from *ab initio* calculations based on the element. Parameter description file from OPENbabel, GULP, etc. include the 3 aforemented parameters for almost all elements in the periodic table, when the QEq method assumes the ionisation potential and the electron affinity for atoms within the certain molecular structure is well-represented by the properties of the individual atoms, the partial charge for those atoms can be derived as the electronegativity in the molecule is equalised. QEq benifits from superior performance in quick assignment of partial charges whilst the accuracy is generally acceptable compared to the ab initio calculations, e.g. the density-dervied electrostatic and chemical method (DDEC) in MOF modelling. [72] The UFF4MOF force-field provides more detailed atom type decriptions for the Zr-based UiO family MOFs, e.g. Zr with unconventional coordination numbers of 2, 4 and 8. While in principle the UFF and the UFF4MOF extension are designed for molecular modelling rather than molecular dynamics, empirically the UFF-based force-fields cover a great selection of chemical elements whilst the compatibility with the likes of OPLS for host-guest interactions proved advantageous over dedicated force-fields, e.g. DFT-derived force fields.

For the IL, the initial structure information is usually supplied from online sources, e.g. from commercial companies selling chemical products, in the format of protein data bank files (PDB files). Since the individual cations and anion as small molecules with common elements as C, H, N, O, the parametrisation via *ab initio* calculations is affordable whilst with the Hatree-Fock modelling tool of MOPAC, the electrostatic potential (ESP) level accuracy is usually sufficient. For the well-studied typical IL pairs of [BMIM][NTf_2], the DFT-derived, OPLS compatible force-fields works well together with the OPLS-AA Li-ion in the case of Li[BMIM][NTf_2] transport within HKUST-1. The [BMIM] is characterised with 7 unique atom types where as for the [NTf_2] has 5 atom types. Consistent with the delocalised charge distribution for both the cation and the anion, the net positive charge is mostly located on the nitrogen atoms within the aromatic ring while the partial charges between them rather comparable than focused on a single site. Meanwhile for the [NTf_2], the symmetric arms are relatively polar both on the oxygens and the fluorine while the nexus nitrogen atom is strongly negative, which results in a strong ion pairing in the Li[NTf_2].

The UFF4MOF extends the UFF force-field while adopting the UFF formulations for HKUST-1 as:

$$E_{\mathrm{UFF4MOF}} = E_{\mathrm{vdW}} + E_{\mathrm{Coulomb}} + E_{\mathrm{bonds}} + E_{\mathrm{angles}} + E_{\mathrm{dihedrals}} + E_{\mathrm{impropers}} \tag{4.1}$$

where the individual terms are:

$$
\begin{aligned}
E_{\text{vdW}} &= \sum_{ij} 4\epsilon[(\frac{\sigma}{r_{ij}})^{12} - (\frac{\sigma}{r_{ij}})^{6}] \\
E_{\text{Coulomb}} &= \sum_{ij} \frac{q_i q_j e^2}{4\pi\epsilon_0 r_{ij}} \\
E_{\text{bonds}} &= \sum_{\text{bonds}} K_r(r - r_0)^2 \\
E_{\text{angles}} &= \sum_{\text{angles}} C_\theta[1 - B_\theta(-1)^n \theta \cos(n_\theta \theta)] \\
E_{\text{dihedrals}} &= \sum_{\text{dihedrals}} K_\phi[1 + d\cos(n\phi)] \\
E_{\text{impropers}} &= \sum_{\text{impropers}} K_\omega[C_0 + C_1\cos(\omega) + C_2\cos(2\omega)]
\end{aligned}
$$

when the UFF4MOF introduces in the case of UiO family MOFs modified angle potentials involving the Zr modelled with 8 coordination bonds as:

$$
E_{\text{angles\#2}} = \sum_{\text{angles\#2}} K_\theta[C_0 + C_1\cos(\theta) + C_2\cos(2\theta)]
$$

where the $(r_{ij}, q_i, q_j, r, \theta, \phi, \omega)$ characterises the instantaneous state for the individual pair distances, partial charges, bond lengths, and torsional angles, whereas the reference (σ, r_0) corresponds to the equilibrium pair disntances, bond lengths, and torsional angles. The atom type designations are marked with colored spheres in the Figure 4.11.

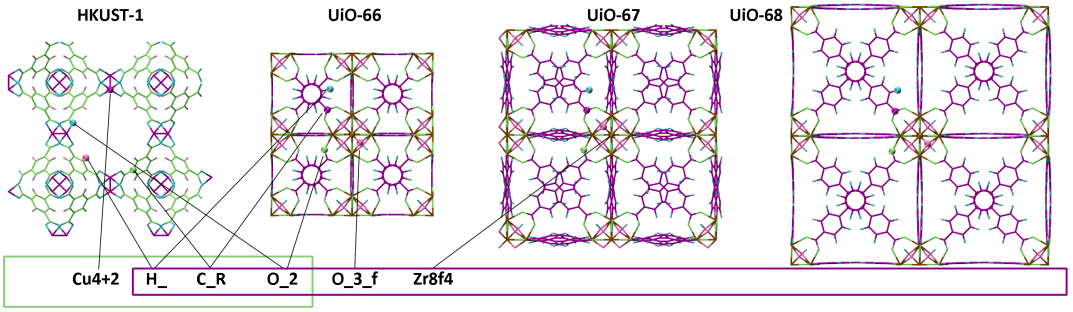


Figure 4.11.: HKUST-1 and UiO MOF unitcells and atom type designations. The Cu with a coordination number of 4 as well as the Zr with a coordination number of 8 are unconventional assignments in the UFF4MOF force-field.

The dedicated force-field developed for [BMIM][NTf_2] is OPLS-AA compatible while the Li-ion is parametrised in OPLS-AA are written in the conformal expressions of $E_{\text{OPLS-AA}}$ as:

$$
E_{\text{OPLS-AA}} = E_{\text{vdW}} + E_{\text{Coulomb}} + E_{\text{bonds}} + E_{\text{angles}} + E_{\text{dihedrals}} + E_{\text{impropers}} \tag{4.2}
$$

where the individual terms are:

$$
\begin{aligned}
E_{\text{vdW}} &= \sum_{ij} 4\epsilon[(\frac{\sigma}{r_{ij}})^{12} - (\frac{\sigma}{r_{ij}})^{6}] \\
E_{\text{Coulomb}} &= \sum_{ij} \frac{q_i q_j e^2}{4\pi\epsilon_0 r_{ij}} \\
E_{\text{bonds}} &= \sum_{\text{bonds}} K_r(r - r_0)^2 \\
E_{\text{angles}} &= \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 \\
E_{\text{dihedrals[BMIM]}} &= \sum_{\text{dihedrals[BMIM]}} K_\phi[1.0 + \cos(n_\phi\phi - d_\phi)] \\
E_{\text{dihedrals[NTf}_2\text{]}} &= \sum_{\text{dihedrals[NTf}_2\text{]}} \{K_1/2[1 + \cos(\phi)] + K_2/2[1 - \cos(2\phi)] \\
&\quad + K_3/2[1 + \cos(3\phi)] + K_4/2[1 - \cos(4\phi)]\} \\
E_{\text{impropers}} &= \sum_{\text{impropers}} K_\omega[1 + d\cos(n\omega)]
\end{aligned}
$$

where the $(r_{ij}, q_i, q_j, r, \theta, \phi, \omega)$ charaterises the current state whereas the the reference state for equilibrium is quantified by $(\sigma, r_0, \theta_0, d_\phi)$. The atom type designations are marked with colored spheres in the Figure 4.12

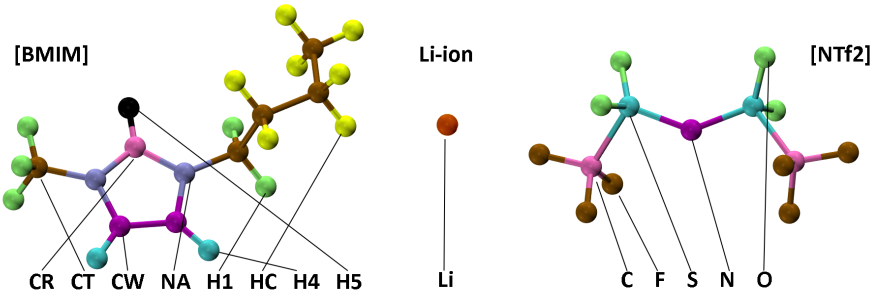


Figure 4.12.: Li-ion, [BMIM], and [NTf_2] in the IL and the corresponding atom type designations. The delocalised positive net charge of the [BMIM] cation is spared between the aromatic nitrogens designated as "NA", while on the [NTf_2] anion, the negative charge is mostly assigned to the nexus nitrogen atom, designated as "N".

The mixing between the UFF4MOF and the OPLS-AA derives for the intermolecular interaction potentials. The van der Waals interactions follow the geometrical mixing rule of $\sigma_{ij} = \sqrt{\sigma_i \sigma_j}$ as $\{\sigma_i\}$s are parameters determined by atom types. The Coulomb interactions are turned off for the intramolecular 1-2 and 1-3 atom pairs when for 1-4 interactions the Coulomb weights are scaled at 1.0.

The comprehensive force-field parameters of the HKUST-1 is summarised in Table A.1, parametrisation of UiO family MOFs in Table B.2, and parametrisation of IL in Table C.3 in the Appendix.

The loading of IL into the MOF from the simulation side is achieved through the pack algorithm from the software package of PACKmol [73] where the desired number of IL molecules are stuffed inside the simulation box containing MOF in force. The packing algorithm randomly places the molecules subjected to free translational and rotational

moves while the regularization minimises the overlaps between atoms. A certain threshold of 0.2nm is prescribed in the packing algorithm however in the end the forced configuration allows for marginal violations for the minimum separation criterion. To calculate what number of IL pairs corresponds to the full loading, a calculation based on van der Waals sizes are derived from the available pore sizes of the MOF and the van der Waals size estimations of the IL. For the individual cation and anion molecules, the estimation of the van der Waals sizes are quantified from both Hatree-Fock calculations on the [BMIM] and [NTf_2] via the MOPAC software [74] where the molecular size correspond to the conductor-like screening model (COSMO) [75] which scales approximately 120% of the van der Waals size. On the other hand, the van der Waals size of the IL molecules are calculated based on a fast-estimation algorithm from reference. [76] The consistency from both methods yield "fully loaded" MOFs with respect to IL pairs per MOF unitcell as:

IL	MOF	100% loading
[BMIM][NTf_2]	HKUST-1	19.2
[BMIM][NTf_2]	UiO-66	10.8
[BMIM][NTf_2]	UiO-67	32.2
[BMIM][NTf_2]	UiO-66	67.8
[$Li_{0.2}BMIM_{0.8}$][NTf_2]	HKUST-1	20.8

Table 4.2.: Statistics of number concentrations at 100% loading of IL@MOF.

Especially, the maximum IL loading measured from experiment for [BMIM][NTf_2]@HKUST-1 at 18.6 ± 0.7 IL pairs per MOF unitcell corresponds very well with Van der Waals sizes based theoretical predictions, whereas for [$Li_{0.2}BMIM_{0.8}$][NTf_2] the theoretical maximum loading is acceptable compared to the experiment measurement of 18 ± 1.5 IL pairs per MOF unicell. The term “maximum loading” or “fully loaded” should be taken with precaution as in an equilibrium, the IL density that fills MOF pores corresponds to this upper limit, whereas in non-equilibrium states when the ILs are driven by the external field, local density fluctuations lead to even denser packing of IL.

From the PACKmol package [73] which populates the MOF pores with variable amounts of IL molecules, the all-atom coordinates of both the MOF and the IL are updated when an in-house developed interface parses the IL pair information as molecular structure, partial charges, etc., along with the force-field parameters, into a single data file. The LAMMPS MD simulation software [30] reads from the data file with complete information and executes a templated recipe with a short relaxation, the application of the external electric field, and the coupling for the thermostat and barostat controls in a serial order. In the MD simulation, since the MOF is not conducting and in the classical all-atom MD the partial charges modelled as point charges, fixed on the center of mass of individual atoms, the conductivity of the overall IL@MOF system is solely contributed by the molecular diffusion of the IL. When the external electric field is applied, ions start to move driven by the field, as the electric conductivity results exclusively from ion diffusion. As ILs travel, the molar conductivity of the molecular charge carriers can be directly determined from drifting velocities as $\Lambda_m = E^{-1}F\ q_{cation} < v_{cation} > + q_{anion} < v_{anion} >$, where q is the charge of the ion, $< v >$ is the drift velocity, and F the Faraday constant. For the weaker external electric fields, the transient time-correlation function (TTCF) analysis [51] supplements the kinetics approach to improve accuracy via a time-integration from the field switching event on the response of ionic molecules. In the case of [BMIM][NTf_2]@HKUST-1, TTCF analysis showed consistency with the kinetics determination of conductivity for dopant-free IL@MOF systems, while in the case of [$Li_{0.2}BMIM_{0.8}$][NTf_2] the application of TTCF analysis is limited to high loading scenarios as the extraodinary fluctuation in the individual trajectories of Li-ion diffusion renders relatively poor signal-to-noise ratio. The

overall mobility summarised with respect to the ensemble of ions is directly compared with the conductivity measurement in the experiment.

Besides, advantages can be taken in the all-atom simulation approach as MD provides a complete motion-picture of the molecular movements with notably enhanced temporal resolution compared to the experiment measurement capabilities. Generally speaking, the individual tracer dynamics offers unprecedented details in revealing the diffusion pathways and diffusion mechanism in the equilibrium, as well as the transitioning behaviour towards the equilibrium as a response to the external driven force. In the IL@MOF applications, the diffusion pathways of the individual ions along with their instantaneous diffusion speeds are recovered as a "traffic map", in which the vehicular dynamic properties of [BMIM] and [NTf_2] are analogous to the traffic-jam phase transitions on the express road as the traffic map in the event of blockage by a bottleneck exhibit surprising similarity to the traffic motion on the Autobahn A5 recorded near Frankfurt. [77, 78]

The workflow of IL@MOF simulation offers an automatic, "one click" solution to start from simple input information as the MOF crystal structure and the selection of IL while the output compiles the full MD movie for visualisation, the comprehensive mobility analysis from both the kinetic based, and the TTCF methods, plus the condensed 2D mapping of vehicular ion movements onto the temporal dimension and the displacement along the electric field. All MD simulations for the IL@HKUST-1, IL@UiO, and Li-doped IL@HKUST-1 are performed utilising the established workflow when the modular parts can support or be integrated into alternative simulation tools. On small scale GPU-accelerated clusters from the input to the conductivity analysis of a 8x8x8nm^3 MOF supercell filled with IL molecules corresponding to a real timescale of 2ns, the whole workflow takes approximately 10hours to complete.

4.4. Bulk phase [BMIM][NTf_2] versus [BMIM][NTf_2]@HKUST-1

As AC current is applied in between the interdigitated electrodes, within the HKUST-1 an electric field drives guest IL to move along the MOF channel in the direction of (100). The [BMIM] cations and the [NTf_2] anions are singly charged such that they are driven by equal force albeit in opposite directions. The framework is charge-neutral and maintain a high mechanical stiffness as HKUST-1's Young's modulus measures at the level of 10GPa and strongly adhered to the substrate. The electric conductivity is exclusively contributed from the vehicular ion transport of the IL. Carrying opposite net charges, the cations and anions travel against each other when the movement of the molecular charge carriers results in electric conductiviity when an external eletric field is applied. In ambient conditions the IL maintains a fluid nature, with a inherent native ion density in the solution where the cations and anions experience strong ion pairing. Subjected to a homogeneous external electric field, the homogeneously mixed cations and anions travel in the form of homogeneous flows. MD simulations show that for the bulk phase IL, when an external electric field is applied, cations and anions dissociate due to the intense external field. For the transport of the IL a dissipative structure emerges as steady cation and anion flows along the direction of the external field, whereas in the cross sectional area, a fingerprint-like pattern is established as cations and anions travel within dedicated transport channels, when the hindrance between the cations and anions are minimised.

The diffusion pattern between embedded IL is fundamentally different from that of the bulk phase IL as IL loading concentration is variable in the MOF pores. The IL transport is subjected to the competing aspects of the IL-MOF and the IL-IL interactions in the MOF, while the crystalline structure of the MOF imposes geometrical limitations between the pores. Constrained by the pore apertures, as a contrast to bulk phase IL, the transport of ions through the formation of dedicated ion channels is no longer viable. In the case

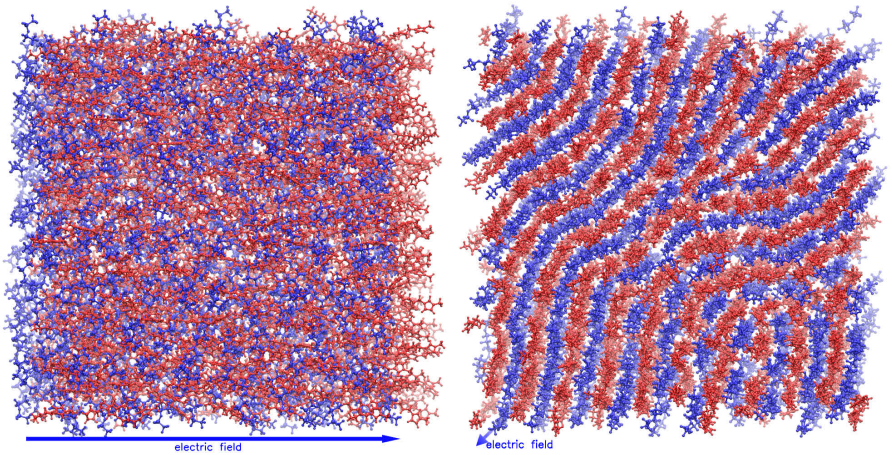


Figure 4.13.: Side-view and cross section patterns of the bulk phase IL transport under intense external electric field. [BMIM] cations are colored red, [NTf_2] anions in blue color. The cationic and anionic transport enjoyes dedicated ion channels as the travelling ions in opposite directions do not hinder each other. In the cross sectional area a fingerprint-like pattern for the ion channels emerge under intense external driving.

of HKUST-1, the pore aperture of the primary pore is limited to 0.9nm, allowing only a couple pairs of IL molecules to pass each other simultaneously. The IL collides with the framework as well as other ions which forces ILs and forms a labyrinth of an amorphous mixture of cations and anions, where individual ions are forced squeeze through and share the MOF pore apertures when no dedicated diffusion channels are present.

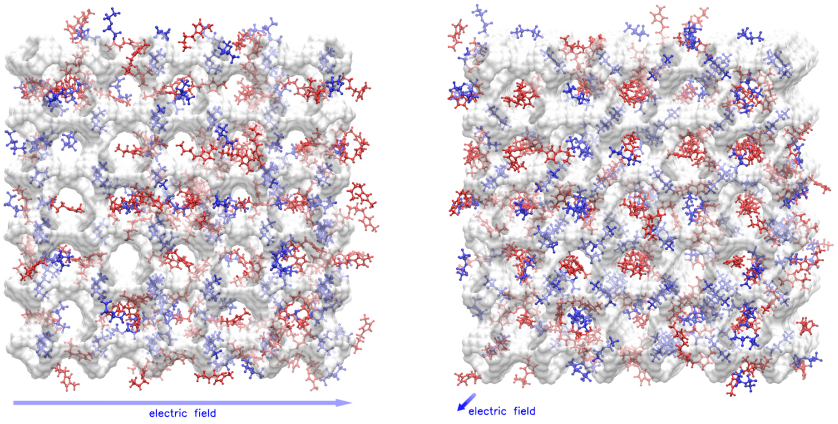


Figure 4.14.: Side-view and cross section patterns of the embedded IL transport under external electric field. [BMIM] cations are colored red, and the MOF in gray, [NTf_2] anions in blue color. The cationic and anionic transport are forced to share the MOF pores as the travelling ions in opposite directions mutually block the pore aperture, the blockages dependent upon local IL concentrations.

4.4.1. Density-dependent bunching and immobilization of confined IL

In the MD simulations for low IL loading, the IL flow through the MOF pores is statistically homogenous. In the extreme cases of very low loading, in the IL@MOF system there is essentially exclusively the guest-host interactions between ions and HKUST-1, when the ionic guest molecules generally do not interact with each other. In the interval of very low IL loading and low IL loading, the IL-MOF dispersion remains as the most prominent hinderance to IL diffusion, as in the experimental measurements there are no perceptible

changes in IL conductivity, while in the MD simulation as we have an defect-free MOF that is perfectly aligned with the external field, diffusion of IL decreases continually with respect to the increasing IL loading. Tracing individual ions, it is observed that their travelled distances scale linearly with time, as the travel distances of cations and anions exhibit a Gaussian form in the distribution profile.

However, as IL loading mounts up, the local cation-anion condensation appears intermittently. The stochastic nature of density fluctuation of IL in the MOF leads to temporal blockages of the transport most prominently on the MOF pore aperture, i.e. bottlenecks between MOF pores. This effect is exemplified as the temporally hindered diffusion turns into a grid-like mixture of cation and anion pairs stablised via ion pairing when extra IL gets deposited and incorporated into the grid in a snowballing manner. Individual trajectory of ions verified intermittent blockage with pronounced local residence which is shown in the tracer "traffic map" as a short plateau in Figure 4.17. However due to the amorphous nature of such IL condensation, the blockage dissolves in time, while the long-time statistics of IL diffusion remain spatio-temporally homogeneous.

In the case of very high loading, as HKUST-1 is almost fully loaded with the IL, a critical IL density around 90% maximum loading around 17 IL pairs per unit-cell is observed, where the transport experiences a terminal failure as the mobility drop drastically. Density fluctuations lead to blockage of ILs at MOF pore apertures when the elevated local density results in greater blockage in neighboring MOF pores, where the IL already filled the space and is highly constrained. The resultant global blockage of IL leads to the formation of a bunched layer of the IL, through which occasionally ions manage to squeeze through however the majority of ILs are immobile. For individual ions, due to the difference in hydrodynamic properties of [BMIM] and [NTf_2], the cation which is less rigid and maintains a smaller cross-sectional area, travels through the bunched layer faster than the wedge shaped anion. Any cation or anion which manages to travel through the bunched layer drifts freely in the remaining part in the MOF, before eventually hitting again into another bunched layer due the the periodic boundary condition, accumulating on the far, or "wrong" side of the bunched surface and result in a bias of charge distribution within the bunched layer, working against the external field and further slowing down the IL diffusion.

The result of the transition towards bunching for the IL transport translates into an astonishing drop of ion mobility and thence electric conductivity as a global bunched layer gets formed. Compared to the reference IL mobility in the low loading conditions, the molar conductivity of the IL@HKUST-1 system reduces gradually by 2-orders-of-magnitude before dropping an additional 3rd-order-of-magnitude in the transition towards bunching as the bunched layer traps most ions.

4.4.2. The traffic jam: molecular dyanmics insights from tracer trajectories

The emergent bunching from high charge carrier density in the event of a bottleneck in the passage, in the mean-time, is not an unknown phase-transition behaviour and quite often occurs in our daily experience. In the study of traffic flows for vehicles, it has been commonly conceived that a critical high vehicle density leads to a traffic jam in the statistical modellings of the traffic flow. [79] Specifically, in the study of traffic on the German Autobahn A5 [78] a three-phase traffic theory corresponds to the free flow in low density traffic, the synchronous flow, and the stop-and-go traffic jam, describes the emergence of the vehicular transport patterns depedent on different vehicle densities when subjected the presence of a single road narrowing site. Granted in the case of IL there are differences in the mechanism compared to the coordinated car traffic, as the passage channel is amorphous, whose opening degree is subjected to IL dynamics as cations and

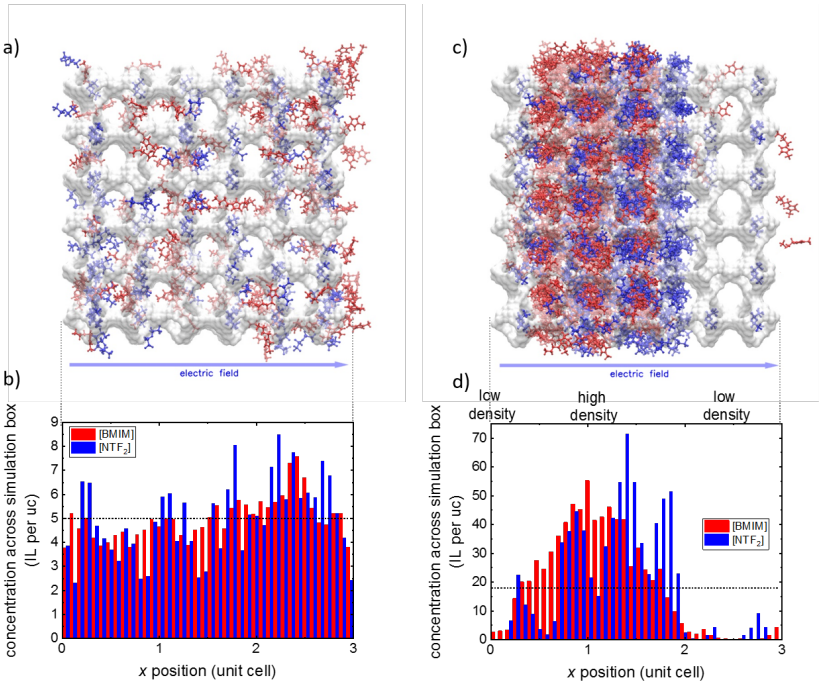


Figure 4.15.: Homogeneous flow with low IL loading and bunched layer with high IL loading. In low loading a) with 5 IL pairs per MOF unitcell, the IL driven by the externall field in the (100) direction forms a homogeneously flow, whose distribution b) over 200ps is uniform in the consistent density level of 5 pairs per unitcell, marked as a dotted horizontal line. In high loading c) the cations and anions in the IL attmepts to travel to the opposite sides whilst a bunched layer is formed in the MOF pores, resulting in a high density region where the maximum density is more than twice of the original loading factor (dotted line at 17 pairs per unitcell). Meanwhile, the distribution of [BMIM] (in red color) and [NTf_2] (in blue color) show an double-peaks formation as the [BMIM] distribtuion is biased towards the cathode when the [NTf_2] is biased towards the anode.

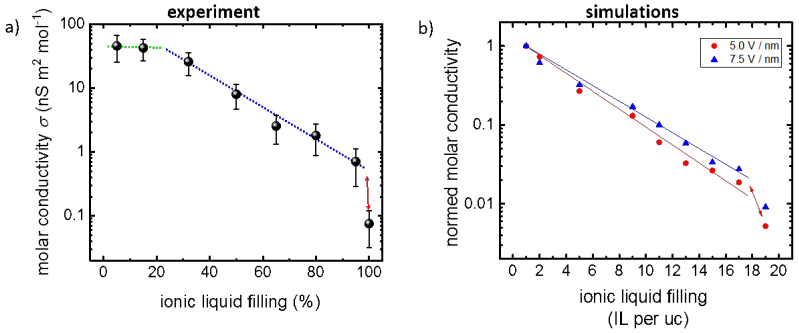


Figure 4.16.: Molar conductivity of [BMIM][NTf_2]@HKUST-1. a) From the experiments, the molar conductiivty exhibit a short plateau before an exponential decrease with respect to IL loading. Beyond 90% loading in MOF, the IL molar conductivity sees an additional order-of-magnitude drop. b) From the MD simulations, the expoential decrease of molar conductivity is consistently observed under external electric field strengths of $5V{\cdot}nm^{-1}$ and $7.5V{\cdot}nm^{-1}$, before the phase transition due to the formation of the bunching layer further immobilises the IL. In MD simulations the absolute conductivity estimations are not to scale with the experiments due to the adoption ofa much stronger external electric field, hence the molar conductivity from the MD simulations are normalised against the low loading conductivity.

anions intermittently accumulate near the MOF pore apertures, which themselves work as fixed bottlenecks in the first place, one shared critical identifier between the traffic jam

and the vehicular transport of IL in MOF is the residence time that one "vehicle" remains immobile.

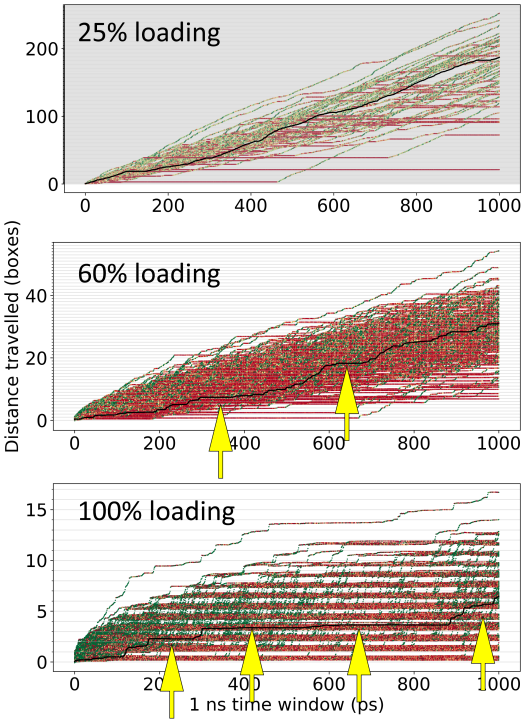


Figure 4.17.: Tracer trajectories of [BMIM] in [BMIM][NTf_2]@HKUST-1. Color coded are the instantaneous diffusion speeds with respect to the average mobility, where green color indicated fast diffusion and red color indicated slow diffusion, a typical tracer trajectory indicated in black. The local residence as ions are immobilised correspond to red plateaus, the most pronounced localisations marked by yellow arrows. Light gray reference grids correspond to simulation box lengths when in the case of formation of bunching layers, corresponding to each simulation box length, a bunching layer is formed, illustrated as a thick red layer in the "traffic map".

In the case of low IL loading factor around 25%, a typical trajectory of the tracer ion exhibits a linear transport with insignificant fluctuations along the distance travelled. In time, for some individual ions, a transient blockage interprets into the traffic map as a intermittent horizontal plateau, before the ion continues the with the linear transport as the temporal blackage dissolves. In the medium loading at approximately 60% IL loading, the intermittent blockages are more frequently observes while the local residence events are elongated in the form of long horizontal plateaus on the traffic map. Compared to the reference spped, the individual ions neverthesless travel with statistically same speeds when the instantaneous diffusion sees a step-like pattern. In the critically high loading, a dominant immobile plateau is observed for ions travelling in the bunched layer. Through this dense region, the ILs are forced to squeeze against the other ions when the local density s extraordinarily inhomogeneous across the simulation box as was discussed previously. Once released from the bunching layer, the ions enjoy a swift free diffusion, seen in the traffic map as green segments between the predominant red plateaus, before running into another bunched layer due to the periodic boundary conditions. The bunched layers exhibits a one-to-one correspondence with each simulation box however due to the distinguished hydrodynamic properties between the cations and anions and the amorphous nature of the bunched layer itself, the layer slowly moves withs respect to the MOF. For the pair of IL that travel in opposite directions, hinderance due to van der Waals sizes becomes the dominating factor in terms of interaction potentials, taking up over 80% of the overall

contribution, whilst the electro-static interactions contribute less than 15%.

4.5. IL transport in UiO family MOFs

When the mechansim in the concentration dependent bunching of IL is revealed for [BMIM][NTf_2] in HKUST-1, a natural question is whether in other MOFs the such bunching behaviour can be reproduced, also, how to tune the MOF pore properties such that the bunching becomes a controlled feature for specific guest-host combinations. In the case of [BMIM][NTf_2]@HKUST-1, the transport properties of molecular ions is determined both via the pore volume, the pore aperture, and the molecular size of the IL, while the ratio between IL size and pore aperture size can be interpreted as the granularity of the guest molecules. When the guest IL molecules are fixed to be the [BMIM][NTf_2] pair, the range of selection in the UiO family MOFs provide a systematic screening in the pore geometries with both incremental pore volume and pore aperture dimensions in the MOF series of UiO-66, 67, and 68.

4.5.1. Tunable pore geometries in UiO-66, 67, and 68 and tunable critical IL bunching

In the UiO family MOFs, there are 2 major pore channels in the direction of (111) and (110) where the IL diffusion rates are comparable. In the experiment as the SURMOF is deposited onto the substrate and grow along the (111) direction, the major diffusion axis is aligned with (110). The maximum [BMIM][NTf_2] loading factors for UiO-66, 67, and 68 are calculated to be 10.8, 32.2, and 67.8 IL pairs per unicell. At the same time, since the SBUs are all identical among the UiO-66, 67 and 68, the ratio of the pore aperture vs ion sizes are MOF-specific, e.g. UiO-68 offers the greatest opening in the pore apertures for ILs to diffuse through, whilst the UiO-66 is the most limiting. Especially, compared to HKUST-1 where the secondary pores deny passage of [BMIM][NTf_2], the UiO-66 preserves a comparable pore aperture while the pore volume is noticeably smaller. The weaker bottlenecking effect in the UiO MOFs, in an analogy to a smaller road block in the traffic flow, leads to a milder transition between the homogeneous flow and the bunching. During the IL transport, the MOF linkers in the UiO, in the form of a linear alignment of aromatic units, allows each individual aromatic ring to rotate with marginal penalties in the torsional potential, i.e. the individual rings are almost free to rotate. The flappy gates in the MOF induce constant fluctuations in the pore aperture as combined with the density fluctuations of the IL, as had been shown in the case of IL@HKUST-1, leads to transient jammings.

As the ionic affinity between the [BMIM][NTf_2] is relatively weak due to the delocalised charge distributions and the pronounced molecular sizes, the transient nature for the individual events of local jamming between the [BMIM] and [NTf_2] graudally gets dissovled with time. However for UiO-66 with the smallest pore volume, the MOF pore is so small that only a couple of IL molecules can be pack simultaneously within the local pore. In the event of a transient jamming, the traffic jam quickly spreads across the pore and reaches another pore aperture, thus spreading to the neighbouring pores and further into a global immobilisation of IL. As a result, although the UiO-66 has a similar pore aperture dimension compared to HKUST-1, i.e. the [BMIM][NTf_2] IL pair preserves a similar granularity with respect to the bottleneck, the bunching starts much earlier in the UiO-66 than in HKUST-1 with respect to the loading factor at around 40% to 60%. As the MOF get systematically lesser obstructions for UiO-67 and UiO-68, the density-dependent bunching gets delayed. From the kinetic-based analysis of the IL trajectory, the molar conductivity of [BMIM][NTf_2] generally decreases with respect to the increase in IL loading before the bunching steps in with a form of abrupt change in the scaling behaviour. Analogous to the formation of bunching layers in the case of [BMIM][NTf_2] in HKUST-1, as the MOF gets

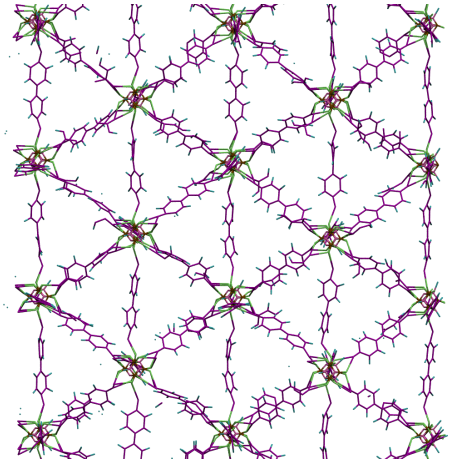


Figure 4.18.: View along the ion transport channel along the (110) direction in the UiO-67. The linkers are flexible in the rotation with minor torsional energy penalties and constantly reshapes the pore aperture.

filled with IL, the loading-depedent bunching and immobilisation of ions is also observed for all of the UiO-66, 67, and 68, whilst the severity in the reduction of conductivity and the critical phase transition exhibit distinguished features, illustrated in Figure 4.19.

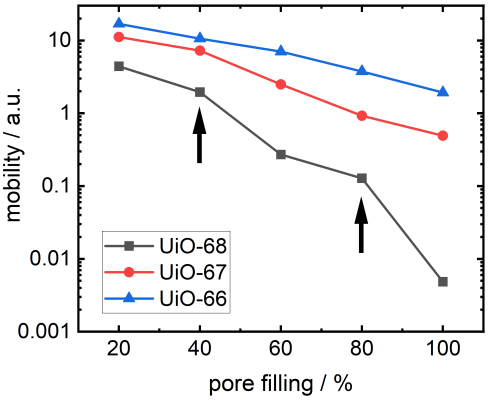


Figure 4.19.: Mobility of [BMIM][NTf_2] in UiO-66, 67, and 68 in the (110) direction. For UiO-66, a 2-step bunching is indicate by black arrows, where as for UiO-67 and UiO-68, partial bunching is observed with respective critical densities.

Single ion trajectories illustrate a clearer evolution in the formation of the bunching layers, marked by the extended local residence of the long plateaus in the traffic map. For UiO-66, the bunching layers forms as early as the IL loading increases to 40%, for UiO-67, the transition behaviiour is observable at 60% loading, while for the UiO-68, which maintains the largest pore volume, the largest pore aperture and the smallest relative blocking. Under the same external electric filed, the compaction of the bunching layer corresponds to the elevated local density while temporally the bunching layer is also amorphous as was the case with HKUST-1.

4.5.2. Two-stages traffic jam: spontaneous symmetry breaking in ions' path finding

Along the (110) direction in the UiO MOFs, the pore apertures form a ceratin angle with respect to the major diffusion channel. The IL has to zig-zag through the MOF pores

when for each cavity there are always two orthogonal diffusion routes leading into the neighbouring pores. For the homogeneous IL transport through the MOF, the diversion into either neighbouring pore is symmetrical with respect to the external field, the trajectory of single ions is analogous to the travelling on a unbiased binoimial labyrinth. However in the circumstances with spontaneous fluctuations in the ion density and the pore aperture, once the transient jamming emerge on one exiting pore aperture hinders the diffusion, the rest of the IL flow is diverted along the alternative route, which induces even greater density fluctuations as well as the spontaneous breaking in symmetry.

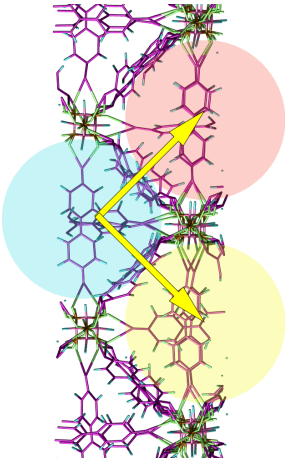


Figure 4.20.: Side-view of the ion transport channel along the (001) direction in the UiO-67. The diffusion channels indicated by yellow arrows, as individual pores are marked in light pink, cyan, and yellow colors.

Consequently, as the IL loading mounts up, while the external field is aligned with (110), the individual ions have to go through a series of bifurcating path finding problems subjected to the diffraction against the MOF and other ions. When all bifurcation branchs along a specific direction, e.g. the (100) route, is jammed, the alternative (010) direction becomes the singular preferred diffusion channel similary to the case of HKUST-1 where all MOF pores are aligned to the major diffusion axis. The bifurcation is relfected as difference in the ion mobilities projected along the separate diffusion channel directions, shown in Figure 4.21. For UiO-66, a consistent 2-step transition is observed with fluctuation differences between the transport speed of the separate channels when the critical loading levels are around 40% and 80%. Meanwhile, for UiO-67, a clear separation between the (100) direction diffusion and the (010) direction diffusion is marked at 60% loading, when the IL in UiO-68 exhibits a critical switching point at 80% loading. The partial bunching in the first transition hinders movement along a single diffusion channel, in the respective cases of UiO-67 and UiO-68 the channel along the (010) direction and the (100) direction, while the remaining diffusion channel does not subject to global jamming, i.e. a 1-step transition for the MOFs with greater porosities.

More details are revealed in the single ion trajectories. The loading-dependence of the IL transport behaviour is compared between UiO-66 and UiO-68 in Figure 4.22 where the transport in UiO-66 is already bogged down at 60% loading whilst for UiO-68 even at full loading there is no sign of global immobilisation. Analogous to previous traffic maps, the completely immobilised ions cannot diffuse along the direction with respect to the external field, reflected in the "traffic map" as a long horizontal plateau in the case of UiO-66, while the partial bunching where the pore channel is blocked only in one direction results in a slanted plateau as the ion is still able to diffuse along the alternative route albeit with a reduced speed.

In the case of UiO-67, the partial bunching results in a comparable maximum mobility

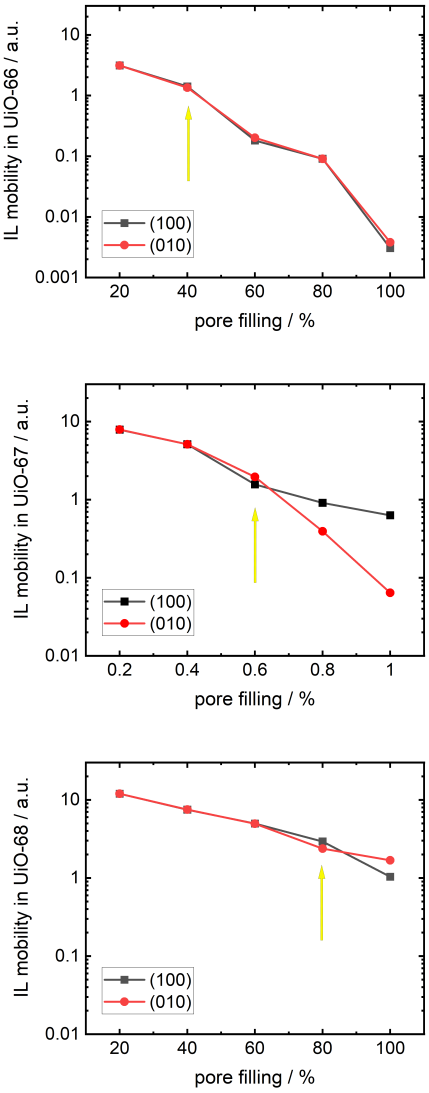


Figure 4.21.: Mobilities of the IL transport along the (100) and (010) directions in the UiO MOFs. For UiO-67, with IL loading less than 40%, the diffusion is symmetric as the mobility of IL along the (100) and (010) directions overlap perfectly. For high loading of IL beyond 60%, the transport in the (100) direction is significantly reduced whereas in the (010) direction the diffusion speed slightly decreases. In UiO-68 the (010) direction is the preferred channel with a less pronounced bias, while in UiO-66 marginal difference between the two channels is observed. The symmetry breaking points are marked by yellow arrows, coinciding with the critical densities in the transition of IL tranport into the formation of bunching layers.

between the 80% loading and 100% loading cases, however the statistical average diffusion speed of 100% loading is approximately half of 80% loading, their fundamental difference lying in the ratio of immobilised ions vs travelling ions. Indicated by the significant red patches in Figure 4.23, the 100% loaded UiO-67 sees a slower overall transport of IL with a series of transient jams occur in the second pore channel, whereas for the 80% loaded case, a long conformal slanted plateau correspond to homogeneous ion diffusion in the alternative pore channel as the other channel is globally blocked. In UiO-66, the bifurcation between the pore channels is less pronounced as the bunched morphology constantly rearranges within the small pore. The first transition stops the homogeneous flow of IL through the MOF whereas during the second transition the two pore channels are comprehensively blocked,

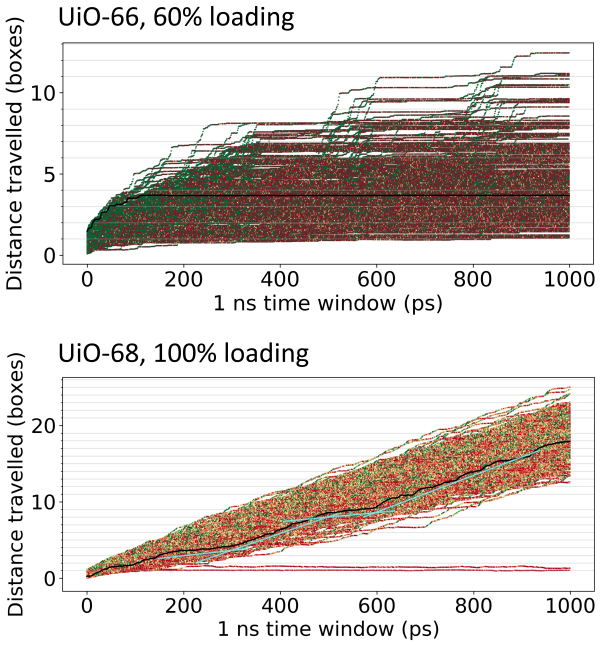


Figure 4.22.: Single ion trajectories of bunched IL in UiO-66 with 60% loading and partial bunching in UiO-68 at 100% loading. UiO maintains an advantageous porosity to UiO-66 such that the critical bunching density is vastly different between the two. For the IL diffusion in UiO-68, light cyan colored reference lines with specific slopes correspond to the switching of diffusion channels with paritial bunching.

the global bunched layers form in the similar manner as was the case with IL@HKUST-1, resulting in a 2 orders of magnitude drop in mobility.

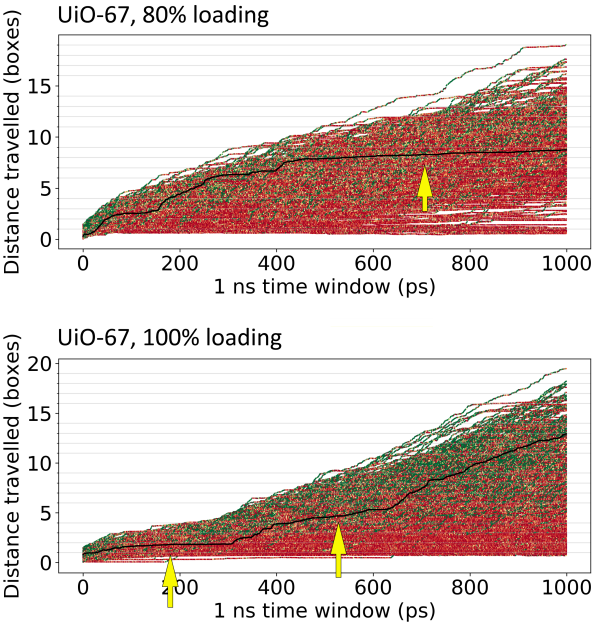


Figure 4.23.: Single ion trajectories of partial bunching of IL in UiO-68 at 80% loading and 100% loading. The 1-step partial bunching is indicated by the long slanted plateau marked with yellow arrows.

4.6. Li-doped-IL@HKUST-1

The concurrent Lithium-ion batteries serve a paramount role in our daily lives in the first decades of 21st century. The introduction of Li-ion batteries into consumer electronics is essential in reducing product weight, size, and design complexity, whereas Li-ion batteries are just starting to get introduced in automobile industry which combined with better communication technologies and distributed sensor/AI capabilities will fundamentally reshape our ways of life in the coming decades. The performance of Li-ion based batteries have been continuously improved while the quest for better Li-ion based batteries aiming at re-designing the electrolytes, such as Li-sulfur [80] and Li-air [81] prototypes, which eventually could replace petroleum for common vehicles with a comparable energy density, attracts unprecedented interest in the research, while the profound influence of novel energy storage applications extends beyond the science community.

Meanwhile, since ILs maintain certain advantageous properties over conventional Li-ion batteries, one of the most essential traits being its nonflammability, the application of IL as electrolyte opens new routes to safe and environment-friendly energy storage. Mixing the two types of Li-ion based and IL based electrolytes offers further tunability upon their distinguished conduction mechanisms. In the hope of combining the best traits from both routes, the systematic design and improvement of the Li-ion conduction and IL conduction in the mixture requires well-established understanding from both worlds.

4.6.1. Bi-modal conduction of Li-ion in IL: small size and alternative diffusion routes versus strong ion pairing

The MOF serving as a solid state support, a electrolyte separator, and a nanoporous containment medium makes it an optimal designer framework in the electrolyte-based applications. From the literature, the embedding of the mixed Li-ion and IL electrolytes has been reported successful when insights from the theoretical investigation is much desired. As has been discussed for the pure [BMIM][NTf_2] in HKUST-1, the bunching of molecular ions emerges through the density fluctuations of IL in the MOF pores. When molecular ions drift in opposite directions subjected to the external electric field, they always try to squeeze past each other whereas in the high IL loading scheme, the steric repulsion characterized by the LJ interaction accounted for approximately 85% of the total potential while the ion pairing, which is a crucial feature characterising the IL in its bulk phase, offers limited contribution to the overall interaction potential.

The addition of Li-ion introduces charge carriers with new traits since they are marginally hindered due to the small ionic size as well as the highly localised charge. Prepared with a fixed doping ratio, from both the perspective of the experiment and simulation the conductivity of [$Li_{0.2}BMIM_{0.8}$][NTf_2] in HKUST-1 is investigated. The experiment measurement of conductivity on both the pure [BMIM][NTf_2] versus the [$Li_{0.2}BMIM_{0.8}$][NTf_2] is shown in Figure 4.24, where the their respective conductivity is measured from EIS and featured a clear disparity between the conducting mechanisms of pure IL and Li-doped IL. For the Li-doped [BMIM][NTf_2]@HKUST-1, the loading dependence in the IL conductivity is much less pronounced than in the case of Li-free [BMIM][NTf_2] while the phase-transitional behaviour at full IL loading with serveral orders-of-magnitude conductivity reduction is absent in the experiment. From the MD simulations the conductivity with respect to IL loadings is summarised in Figure 4.25, in good agreement with the experiment measurements.

The bi-modal conduction of Li-ions, subjected to both the small size and strong ion pairing, distinguishes the Li-ion conduction from the vehicular transport of the molecular charge carriers. From the calculation of conductivity contributions from the Li, [BMIM], and [NTf_2] components, Figure 4.25 illustrates a 2-region behaviour in the conductivity of Li-doped IL.

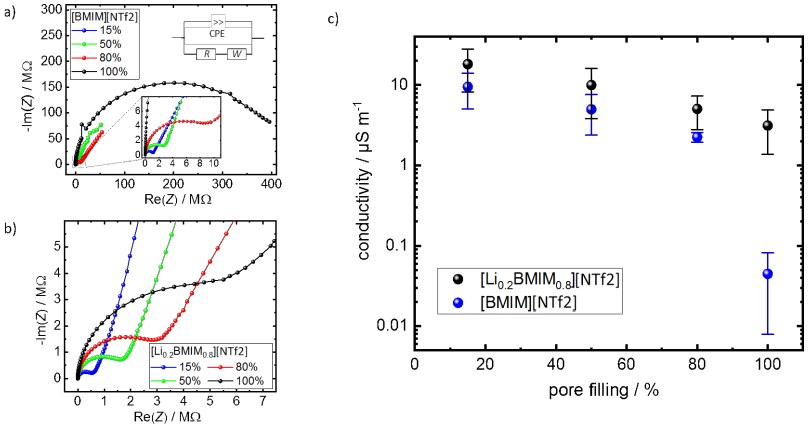


Figure 4.24.: Conductivity of [BMIM][NTf_2] versus [$Li_{0.2}BMIM_{0.8}$][NTf_2] embedded in SURMOF HKUST-1. The conductivity of pure IL pair as [BMIM][NTf_2] and the Li-ion doped IL are measured respectively from EIS (a-b), and are compared against each other (c). When for the pure [BMIM][NTf_2]@HKUST-1 the bunching effect leads to a drastic drop in ionic conductivity the Li-doped IL@HKUST-1 exhibits minimal concentration dependence reaching the full loading.

In the region of low IL loading upto around 60% IL loading, Li-ion conductivity increases accompanied by an increase in Li-ion contribution in the overall conductivity. Despite Li-ion only constitutes 10% of the number of all charge carriers, the high mobility of Li-ion makes up from around 26% overall conductivity in low IL loading cases to around 37% in high IL loading. In this region, the conductivity contribution from Li-ion compensates the concentration-dependent mobility decrease for the [BMIM][NTf_2] pair. On the one hand, the ionic radius of Li-ion is much smaller than the [BMIM][NTf_2], with the Li-ion dopant the estimated van der Waals size of the IL pair of [$Li_{0.2}BMIM_{0.8}$][NTf_2] is around 267 cubic angstroms while for the Li-free IL, the [BMIM][NTf_2] pairs has a van der Waals size is around 303 cubic angtroms. For the full loading of IL in HKUST-1, the smaller size of Li-doped IL facilitates an increased IL loading. The higher loading factor translates into a more sensitive loading-dependence of molecular ion, especially for the bulkier and more rigid [NTf_2] anion, while Li-ion is marginally hindered due to the steric repulsion. On the other hand, the ion pairing of Li[NTf_2] is more intensive than the [BMIM][NTf_2] due to the delocalisation of partial charges on the molecular ions. For low loading, the Li-ions are trapped near the [NTf_2] anion complexes with very long local residence time however as IL loading is increased, a more homogeneous charge distribution lowers the energetic barrier of dislocating the Li-ion from the ion pairing, which contributes to the increase in conductivity. For the region of 60% upto full IL loading, bunching layers form between [BMIM] and [NTf_2], when the Li-ion transport also see a sharp drop, as Li-ion contribution in the overall conductivity decreases to less than 30%. For some Li-ions, they are able to find alternative passage through the secondary MOF pores to circumvent the bunching region, while the majority of the Li-ions are able to navigate through the bunched layer via a distinct Grotthuss-like hopping mechanism. Compared to the intense compression during the bunching of Li-free IL in HKUST-1, the steric effect of Li-ion against the framework and the molecular ions quantified by the Lennard-Jones interaction, comprises only 2% in the whole system.

4.6.2. The Grotthuss-like conduction of Li-ion

The ion pairing of light charge carriers in close proximity of counterion complexes, followed by a series of discretised, short hopping in the transport is famously known as the Grotthuss mechanism of conduction. [82] The conduction model had been proposed by Grotthuss

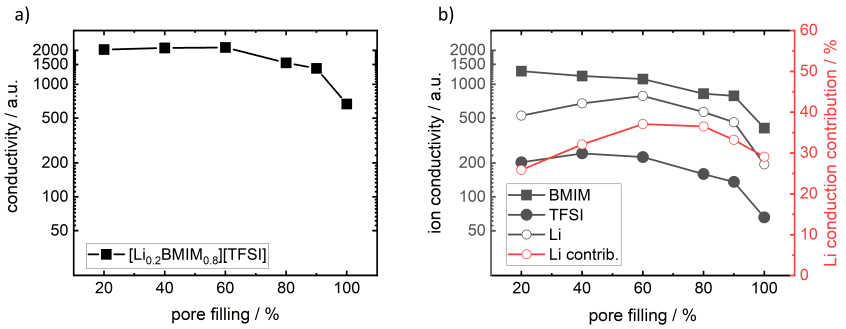


Figure 4.25.: Conductivity of $[Li_{0.2}BMIM_{0.8}][NTf_2]$@HKSUT-1 from MD simulations. a) In high IL loading the conductivity exhibits a slight loading-dependent extra drop in conductivity, whereas from the contribution from separate components (b), the conductivity of the $[BMIM][NTf_2]$ pair decreases exponentially with respect to loading while the concentration-dependence of the Li-ions is marginal below 80% loading. For low loading the increase in Li-ion conductivity compensates the mobility drop in [BMIM] as Li-ion contribution rises from 26% to around 37%.

in the early years' discussion of proton conduction in water, while experiments provided proofs that spontaneous formation of ionic complexes, e.g. a blob of water molecules or transformations of iodine complexes between I_3- and I_6- [83], continuously ever since. New discoveries of Grotthuss-like mechanisms in more materials and the analysis is still a developing field of research. One quantitative characterization method for the Grotthuss-like conduction is featured by the pronouced localization of ions. Here in the case of Li-ion doped IL, the formation of $Li[NTf_2]$ ion pairs results in a highly localized distribution of Li. In a sharp contrast against the [BMIM] cation, whose charge is highly delocalized, the $[NTf_2]$ anions preserve a strongly negatively charged central nitrogen atom and showed strong ionic affinity in to Li-ion in the pure $Li[NTf_2]$ solution such that any mixing of $Li[NTf_2]$ into the $[BMIM][NTf_2]$ reduces the bulk conductivity. [84]

Meanwhile, the Li-ion holds an advantageous charge-to-mass ratio where the positive charge is highly localized in its small ionic radius, featuring vastly different dynamics compared to the bulky, slow-to-react ionic molecules. The $[NTf_2]$ anion complexes emerge within the MOF due to the confinement of the framework as well as the anions' significant molecular size and rigidity restricts the relocation of individual anions, making the formed anion complexes maintain its close vicinity for longer time, which promotes the trapping of Li-ion. In the low loading scheme, the formatino of the $Li[NTf_2]$ complexes are sparsely located within the framework such that the high energy barrier for the Li-ion to dislocate from the current anion complex and hop into the next renders a slow Grotthuss-like hopping rate. In the higher loading scheme as the the intermittent jamming occurs, the molecular ions form an amorphous mixture where the local density fluctuation constantly changes the energy barrier of ion pairing. The Li-ion trapped in the anion complexes gets intermittently released once the ion pairing barrier is overcome by the driving from the external electric field, thence the Li-ion diffuses relatively freely before getting captured in the next anion complex which in turn reshape the local ion pairing configuration. This effect is attenuated in the high IL loading in the bunched layer where the preferential pockets near the anion complexes are closely located such that a chain of Grotthuss-like hopping event allows for a continuous Li-ion conduction in the analogy to the Newton's cradle.

The radial distribution function indicates a strong ion pairing effect for Li-ion as a significant peaking is centered around 0.25nm of Li-ion near the $[NTf_2]$ in Figure 4.26. With the increase of IL loading, the localisation of $Li[NTf_2]$ sees a drop in amplitude, when the mixing cations and anions reduce the overall inhomogeneity of charge distribution.

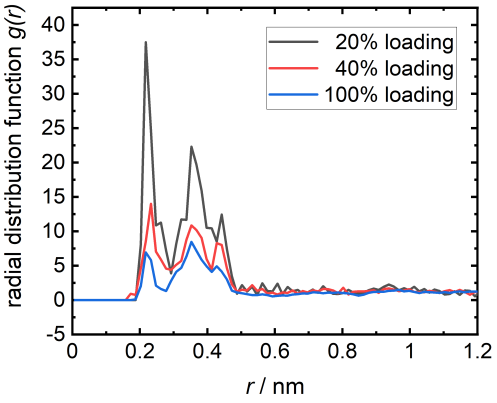


Figure 4.26.: The localization of Li-ion near anion complexes. For small loading of IL, Li-ion exhibits a highly-concentrated peak at approximately 0.25nm near the [NTf_2] anion, whereas for increased loading of IL, the peak diminishes when increased loading of IL results in a more homogeneous mixture of the charged environment. In the inset image, the Li-ions are represented as orange spheres.

Consecutive snapshots for individual Li-ions offer high resolution analysis on the individual Li-ion dynamics in Figure 4.27. In low loading, the [BMIM][NTf_2] pairs travel through the MOF pores in free, homogeneous flows however due the diffusion timescale difference between Li-ion and the molecular IL, the sparse distribution of slow moving [BMIM] and [NTf_2] renders the local electrostatic landscape spatially inhomogeneous. When the tracer Li-ion meets an anion complex, the reordering of the local ion pairing structures dislocates the previously captured Li-ions and in time releases them into the sparsely charged environment. When the probability of another incident Li-ion against the newly formed ion paired blob remains low due to the small overall ion density, the local ion paired structure remains stable for a long time. The Li[NTf_2] complex is relatively large compared to the MOF pore aperture, hence the IL pair is often kept in a local residence in conjuction with the anion complex, reducing the Li-ion mobility greatly. In the high IL loading, even when the [BMIM] and [NTf_2] forms a bunching layer, due to the dense deposition of cationic and anionic molecules the electrostatic landscape within the mixture is nevertheless more homogeneous compared to low loading. Meanwhile, the abundance of local preferential ion pairing sites within the amorphous mixutre of [BMIM] and [NTf_2] provides a number of energentically equally favourable sitting spots with a significantly lower barrier for the Li-ion to hop from one spot onto another. In the tracer trajectory, the incident Li-ion replaces consecutively the previously bound Li-ion in a sereis of anion complexes, when the newly released Li-ion relocate into alternative anion complexes and form new ion-paired structures within the 2ps time window, in vivid contrast against the 20ps-stable Li[NTf_2] complex formed in 20% loading.

4.6.3. Fast and well-tempered: the traffic map for Li-ion diffusion

Analysis of individual ion traffic maps provides greater insights into the distinguished diffusion behaviors between the molecular ILs and Li-ions. Compared to the pure IL@HKUST-1 case, the molecular components in the Li-doped IL exhibit similar loading-dependence the transport of [BMIM] and [NTf_2] in the left panel in Figure 4.29. In the case of doped IL, the maximum loading in HKUST-1 rises to 20.8 pairs per unitcell while the number density of [NTf_2], the rigid molecular anion whose size is pronounced further contributes to the concentration dependency. In the traffic flow through the MOF, the 3 components of Li-ion, [BMIM] and [NTf_2] are forced to share passage through the MOF pore apertures, when

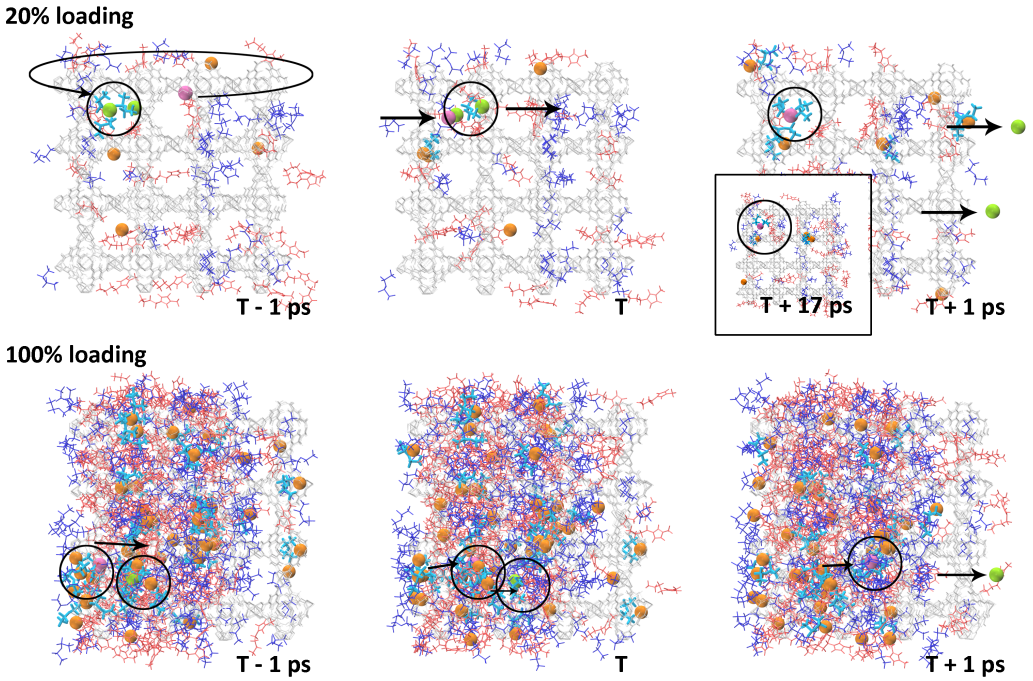


Figure 4.27.: The Grotthuss-like conduction of Li-ion. Consecutive snapshots are taken from the MD trjectory from both the 20%, and the 100% IL loading, when between the snapshots a time difference corresponds to 1ps in the MD simulation. The incidnet Li-ion, shown as pink colored sphere, is captured by the anion complex, illustrated as thick cyan frames. When the incident Li-ion gets captured by the pocket of anions, marked by black circles, the previously localised Li-ion, illustrated as the green spheres, are dislocated from the anion complex. In low loading due to the more pronounced ion pairing between the Li-ion and the [NTf_2], when the incident Li-ion replaces the previous Li, the residence time span reaches upto $>$ 17ps, shown in the inset image. In a stark constrast, in the high IL loading when the [BMIM] and [NTf_2] form a bunched layer, Li-ions nevertheless finds passage through. In the bucned layer within the time span of 2ps, the incident Li-ion in a series of hopping events in between multiple anion complexes, dislocating the previously captured Li. In the snapshots, the [BMIM] is illustrated in red frames, the [NTf_2] the blue frames, Li-ion the orange spheres, and the MOF light gray frames. Black arrows represent the Grotthuss-like hopping events of Li-ions.

the bulkier molecular IL cations and anions subsequently see through the traffic phases of free and homogeneous flow, transient bunching and immobilization with the increase of IL loading, Li-ion mobility is subjected to the Grotthuss-like hopping with less pronouced local residence with respect to the loading increase, as the zoom-in trajectories of Li-ion always show a step-wise pattern, in Figure 4.29.

Due to the emergence of the traffic jams when the MOF is fully loaded, the travelling speed of individual ions are reduced by a factor of 1.6 from 90% loading to 100% loading. Meanwhile, for the Li-ion whose ionic radius is much smaller and charge-to-mass ratio greatly advantageous, the mobility from 90% loading to 100% decreases with a greater concentration dependence by a factor of 2.1. In the transport driven by the external field, the Li-ion quickly finds passage much less obstructed in the porous medium of amorphous [BMIM][NTf_2] mixture in the MOF, however the inhomogeneous accumulation of cations and anion works against the Li-ion transport, as the moleulcar ions accumulate on the "wrong" side of the bunched layer. The higher concentration of cations near the anode repels Li-ion due to same charges, whilst the concentrated anions near the cathode forms a labryinth with greater charge negativity and promotes localisation of Li, both factors contribute to the pronounced drop of Li-ion conductivity and slight decrease of the Li-ion

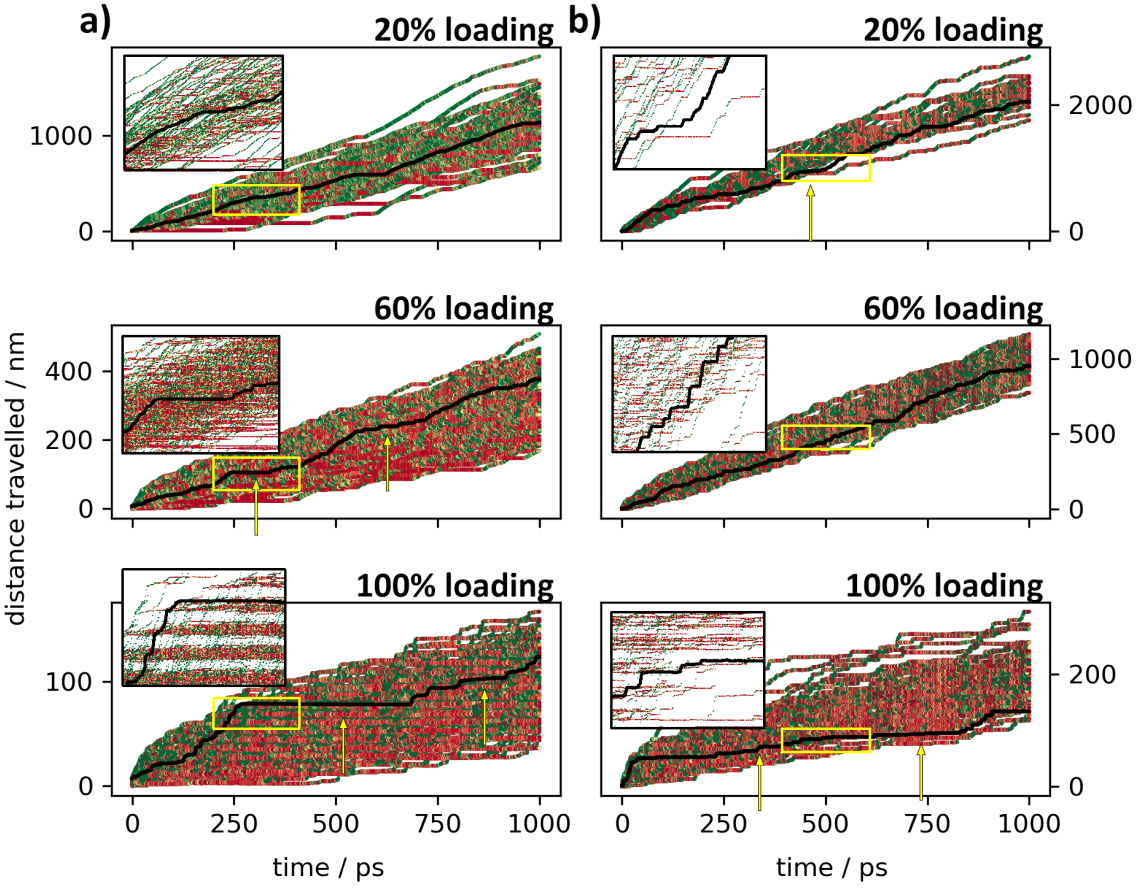


Figure 4.28.: The trajectories of individual [BMIM] (a) versus Li-ions (b) in low, medium, and high IL loadings. Typical tracer ion trajectories are makred as black lines, red segements correspond to slow diffusion with respect to overall mobility. Long local residence periods in both [BMIM] and Li-ion diffusion are marked with yellow arrows. Insets show zoom-ins of the single ion trajectories with equal time spans of 200ps, corresponding to the yellow boxes shown in the main trajectory.

contribution to the overall conductivity at full IL loading.

The Li-ion hardly accumulate on the surface of the bunching layer and the diffusion timescale of Li-ion is in another scale compared to the molecular ions. In very high loading the travelling speed of Li-ion is also reduced as it zigzags through the labyrinth of the bunched IL, however, there is no drastic drop in mobility for the Li-ion component as had been observed in the case of pure IL, while the relative travelling speed of Li-ion versus [BMIM] for Li-ion in high loading gets even better. Utilising the [BMIM] diffusion speed as a reference, in low loading and full loading the Li-ion travels approximately 2 times faster than the [BMIM], whereas for medium loading the Li-ion travels over 3 times faster than the [BMIM]. The [NTf_2] maintains a scaled down diffusion pattern compared to [BMIM] with marginal contribution to the overall conductivity. As a straight-forward impression, the Li-ion single ion traffic map resembles the jagged [BMIM] transport in high loading, when the mechanisms of the hindered diffusion maintain foundamental differences. As the loading is increased, the muted step-like Grotthuss-like hopping renders a trajectory more in resemblance to the homogeneous flow of [BMIM] subjected to low IL loading.

4.7. The IL transport in SURMOF vs in MOF pellets

From the perspective of experiments, in the literature a wide range of MOF applications are based on MOF pellet samples pressed from MOF powder. As for the IL@MOF mea-

surements, the contradicting behaviours of IL@MOF-pellets vs IL@SURMOF have to be taken with caution. For the evaluation of IL@MOF-pellets vs IL@SURMOF difference, the conductivity of $[Li_{0.2}BMIM_{0.8}][NTf_2]$@HKUST-1 is also measured for in-house synthesised pellets pressed from HKUST-1 powder. Besides, even when the current-generation SURMOF synthesis is of great precision, as the typical defect density of HKUST-1 SURMOF is only 4%, it nevertheless does not compare to the defect-free, ideally aligned MOF structure in the MD simulations whereas the SURMOF consists of different crystalline domains with dimensions ranging from 100nm to 1μm, when the grain boundaries in between such length scaled domains unviable for the implementation in all-atom MD simulations.

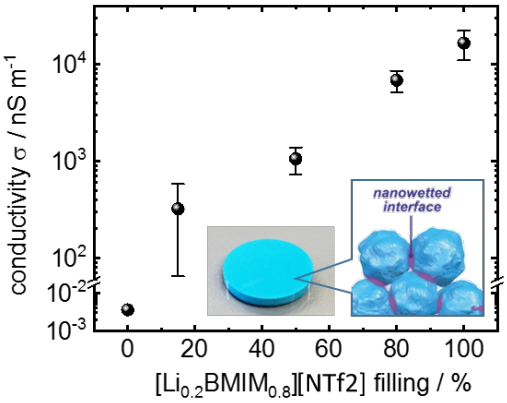


Figure 4.29.: Conductivity of $[Li_{0.2}BMIM_{0.8}][NTf_2]$ in HKUST-1 pellets. The IL shows ever increaing conductivity with respect to the increase in IL loading. The inset image shows a picutre of the HKUST-1 pellet and the sketch of a nanowetted interface area with bulk phase IL out side of the MOF.

For the MOF pellets, it has been reported that when the MOF gets immersed in IL, a significant amount of IL resides on the external MOF surface, which behaviour is referred to as the "nanowetting". In between the nanowetted areas, the IL arguably maintains its bulk phase, which as has been discussed, contrasts greatly in the conduction behaviours from the confined IL. Additionally, as has been shown in the Li-doped IL conduction, when the transport of $[BMIM][NTf_2]$ is sensitive to the local IL density in the confined IL, the transport of Li-ion relies on the morphology of the $[BMIM][NTf_2]$ mixture subjected to the local density and confinement. In the bulk phase IL where the morphology of the molecular ion pairs is not regulated by the framework, the quantitive investigation from the simulations' side may be ill-founded, when similar problems are expected for the hybrid of bulk phase IL and confined IL mixtures.

4.8. Conclusion

The investigation of diffusion of IL embedded in SURMOFs with well-defined structure reveals conterintuitively that their ionic conductivity is increased when the amount of charge carriers is decreased. The discovery that guest molecules the mutually hinder each other in the diffuion owning predominantly to their van der Waals sizes with respect to the dimensions, especially the bottlenecking pore apertures in the nanoconfinement, opens new possibilities for guest-host screening, pore design, and threshold densitiy gauging for the control of ionic conductivity. All-atom MD simulations on $[BMIM][NTf_2]$ in HKUST-1 and UiO-66, 67, and 68 reveal the vehicular transport and the bunching mechanism with unique transitional behaviours and the corresponding tunabilities with respect to IL loading in unprecedented details. In a stark contrast to the homogeneously flowing bulk phase $[BMIM][NTf_2]$, as IL loading increase in HKUST-1 where the primary MOF pore apertures

constituent a series of bottlenecks hindering the ion transport, a global bunching emerges in the form of an inhomogeneous layered structure as the cations and anions travel in opposite directions but cannot make passage through. While in the UiO family MOF, due to the unique alignment of the framework with respect to the pore channel, a binary switching in the ion diffusion routes leads to partial bunching first along a single direction followed by another. Particularly with respect to different MOF configurations with tunable pore volume, pore opening and pore alignment, the quantitative structure-property relationship is direct established between the MOF and the threshold IL density during the phase transition in IL transport.

For Li-doped IL in HKUST-1, an additional characteristic diffusion timescale is added to the previous IL@MOF system, along with a distinct Grotthuss-like conduction mechansim. While the molecular components in the IL still subject to the vehicular transport, the ion pairing in the mixed electrolyte is tunable via the loading of different IL components and different loading factors. In the case of $[Li_{0.2}BMIM_{0.8}][NTf_2]$, with 10% number density in the charge carriers, the Li transport contributes upto more than 1/3 in the overall conductivity in medium IL loading and compensates the density-induced reduction in the conductivity of molecular IL. Compared to Li-free $[BMIM][NTf_2]$, the conductivity of Li-doped IL is tremendously improved.

For the similar diffusion of guest molecules in the MOF simulation, the workflow engine allows for a quick setup-and-run simulation based protocol in material screening evaluation. In practice, disregarding the delay in the experiment due to the coronavirus outbreak, once the $[BMIM][NTf_2]$ digital model is established, the follow up data collection, processing, and analysis is accomplished even before the experiment started. The very good consistency between the experiment and simulation results proved the capability of QSPR enabled one step beyond the current empirical research. Compared to the methodologies from the data science where the training a conceptual model on available data grants capabilities of the very model to predict novelty, the high throughput approach in the all-atom MD simulation of guest in MOF diffusion takes the experiment-simulation consensus as a starting point, accomplishes the fortune-telling in the digital model, before the logic cycle is closed as the result is verified in the lab.

In addition to the understanding of the model system, for applications the on/off switching behavior in the bunching transition may prove desirable for iontronics as sensors etc., while the prevention of bunching serves as a design rule in the IL based supercapacitors. Further more, the insights of the conduction mechansim of the Li dopant offers impoved tunability of Li-ion battery applications involving nanoporous separator frameworks and IL electrolytes.

5. Coarse-grained molecular dynamics of polymerisation and polymer folding

5.1. Introduction

Coarse-grained molecular dynamics (CGMD) takes groups of atoms as an individual entity and assign the proper parameters to the entity whereas the detailed dynamics within the group are disgarded. [85] The coarse-graining operation as an abstraction of the original all-atom system introduces many benefits. On the one hand, for the groups of atoms, the group-group interactions are usually softer than the all-atomic interaction potentials, leading to an up-shift in the characteristic timescale. For the coarse-grained groups, represented typically as an isotropic bead with an increased interaction sphere whose internal morphology is statistically parametrised, the degrees of freedom within the coarse-grained group is compressed into a handful order parameters; on the other hand, the coarse-grained groups preserve a greater mass as a sum of all atoms within the bead, when the particle numbers in the simulation gets reduced, which interprets as slower responses to external potentials as the collisions, the thermostat control etc. are all shifted to a larger timescale; the combination of the two folded extension in timescales allows CGMD simulations to run typically more than 3 orders of magnitude faster than the conventional all-atom MD. [86]

Depedent on the coarse-graining schemes, the CGMD requires dedicated force-fields to perform the simulation. Notable force-fields developed for CGMD are Martini [87], SIRAH [88], and dissipative particle dynamics (DPD) [89] when a number of applications of CGMD simulations aim at fast calculations for macromolecule dynamic properties, e.g. protein structure predictions [90], rheology [91], etc. In the typical applications the correct characterisation of the solvent environment is crucial in the force-field development thus the CGMD simulation accuracy is heavily dependent on the problem-specific force-field development. In some recent efforts, machine learning supported force-field parametrisation based on *ab intio* and conventional MD simulations have achieved success in separate cases. [92]

On the application side, for polymerisation and polymer folding processes the light gated chemical reactions are popular control strategies typically orthogonal to other control routes as temperature gauging etc., as photo-chemical reactions can be spatio-temporally switched on and off in the polymeric 2D chain or 3D network on request. [93, 94] In this chapter, the focus will be placed upon CGMD simulations of polymer systems when chemical reactions reorder the molecular structure to print from the monomer resin a solid 3D object via polymerisation, or to form intramolecular crosslinks and fold a polymer coil into a 3D compact, dynamic photostationary state.

5.1.1. Direct laser writing

Modern electronics are mostly based on integrated circuits where individual chips are mass-produced with photolithography due to requirements in precision in the manufacturing

process and very high resolution in the design. An everlasting race toward a finer feature size in individual device length scales surpasses the diffraction limit of extreme ultraviolet light source via speical techniques, e.g. immersion of the operational interface in liquid with a high refractional index. A smaller feature size for individual devices greatly enhances the operating efficiency, reducing the waste heat, whilst the as the footprint of an individual device is decreased, more sophisticated design in system architecture is possible in the design, which in turn further improves the overall performance, better manages waste heat, and reduce size of the final product. Based on semiconductor transistors, the development of feature sizes on the circuit board follow what was famously known as the Moore's Law, becoming halved every two years, doubling the performance of the processor. The current generation smartphones preserve more computing power than desktop PCs only one decade old, whilst the palm size economic computer models as the Raspberry Pi can be compared to Tower PCs in late 1990s based on passive cooling. However, as the minimum feature size become smaller in the range of several nanometers, complexity in circuit design, high requirement of light source and complexity of the optical system and photomasks have become the limiting factor in further extending of the Moore's Law. The current generation photolithography for feature sizes in the scale of 10nm requires extreme ultraviolet (EUV) light sources and operates submerged in high refractive index photoresist, a protocol optimised and highly specific for the mass production of general purpose devices. On the very opposite end in the demand spectrum, the versatile, small scale production of highly customisable electronics with 3D features with an arbitrary template with variable feature sizes is nonviable via the standard photolithography techniques.

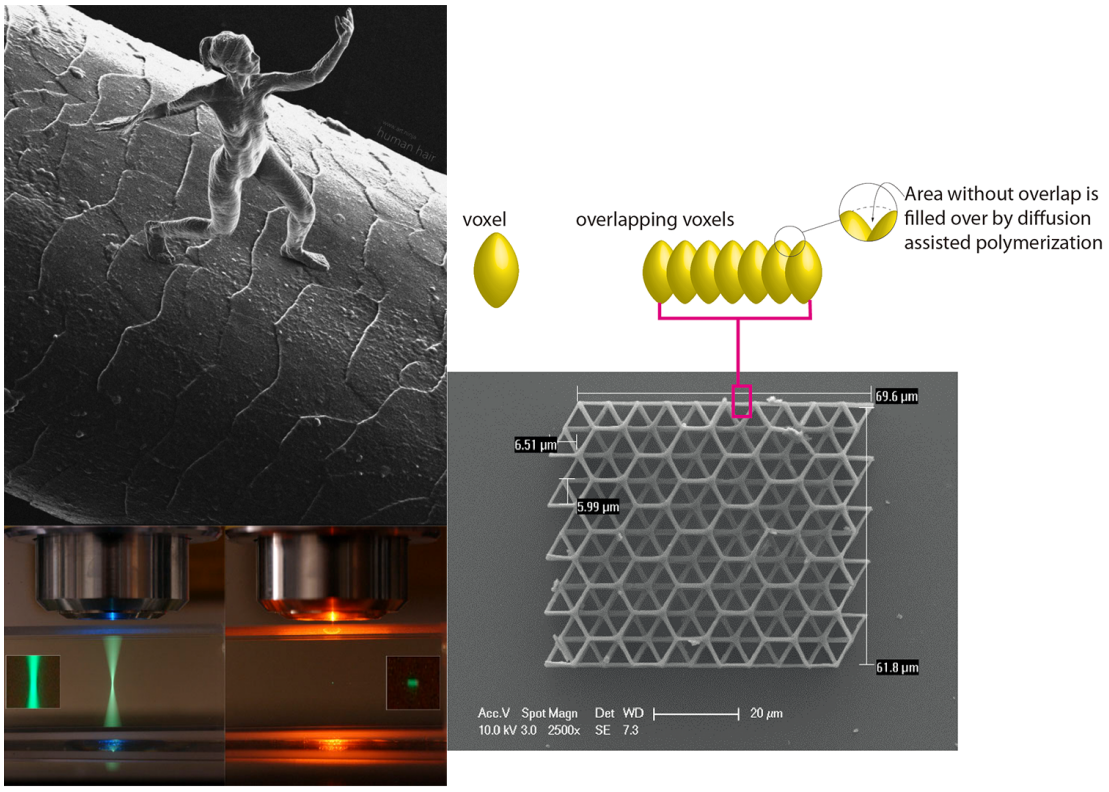


Figure 5.1.: The 3D printed sculpture on a human hair, the two-photon absorption, and the printing voxels. Image sources: 3D printed sculpture by Jonty Hurwitz, Karlsruhe Institute of Technology, and the two-photon absorption image by Steve Ruzin and Holly Aaron, UC Berkeley.

Direct laser writing, on the other hand, has been known as a versatile maskless method in printing sub-micronmeter scale polymers with 3D features. Two-photon absorption in

the concurrent direct laser writing allows for finer control in the polymerisation process as volume pixels (voxel) via interference patterns from an external laser source allows for printing with feature lengths in the tens of nanometers. [95] Upon a functionalised substrate, polymerisation is initiated within the printing voxel, and a polymer network is fabricated where properties as conversion ratio for monomer conversion, mechanical rigidity of the polymer network, and printing resolution are dependent on laser power, exposure time for the voxel, and voxel size. Feature size of the printing process is determined both by the wavelength of the laser source, as had been the case in photolithography, together with a termination of polymerisation, typically with oxygen quenching. The printed polymer network gets initialised and stabilised within the monomer solution via photo-induced crosslinking and thus exhibit a dramatic change in the solubility as the polymerisation product gets larger. Finally, dependent on the properties of the photoresist, when the solution of free monomers gets removed, the left-over part can be induced via positive-tone or negative-tone exposure as was the case with conventional photolithography. Such 3D structuring and fabricating granted tuning capabilities for mechanical strengths for microstructures and biocompatible electronics. Current challenges lies in the better understanding and the precise control of the printing process as there is a known threshold of laser power below which the polymer does not get printed while the underlying mechansim is still sought after.

5.1.2. Single-chain nanoparticles

In a three-dimensional world we live, a structure is 3D becomes one of the first impressions when a new object needs to be described. While in most cases, compatible with human cognition, a specific 3D form indicates a specific function. Proteins, as has been found in all known lifeforms, preserve specific functionalities when a linearly connected sequence of amino acids get folded into the unique 3D conformers where they serve with unique functionalities. The mechanism of structure-property relationship goes as natural macromolecules, such as enzymes, preserve their highly selective reaction control via exposing active reaction sites whilst intricate intramolecular crosslinking stabilises the molecular conformation. On the other hand, improperly folded proteins lead to malfunctions of the macromolecule and thence the biological system, for instance Alzheimer's [5], whereas given the circumstances in the current coronavirus crisis, the precise understanding of how virus, or virus-like-particles behave in vivo may lead to a much-needed break-through in design and production of vaccines and targeted therapy drugs.

Through the decades, the mechanism of protein folding has been one of the holy grail questions of modern chemistry and biology while much effort has been put to molecular dynamics investigation as computational software packages along with dedicated hardware platforms get developed to investigate the folding process. However due to great complexity of amino acids (typical molar weight 110 Da), great size of the protein (typically hundreds up to tens of thousands of amino acids), and the intrinsically complex interactions, the simulation timescale is limited to typically nanoseconds, in stark contrast to the typical timescales measured in experiments and biological systems.

Recent progress in polymer chemistry allows for precise control in synthesis single chain nanoparticles (SCNP) [96], an ideal candidate in mimicking natural macromolecules as nanoreactors. SCNPs can be synthesised with industry standard chemical building blocks in the polymer design, hence circumventing the complexity and limitation in the amino acid monomers in the protein problem. For short polymers, typically an oligomer segment that consist of tens of units, the very order of the linear segment can be synthesised in a sequence-controlled way. [97] Further more, various crosslinking interactions can be chosen with hydrogen bonds [98], metal-complexation [99], and as will be shown in this chapter,

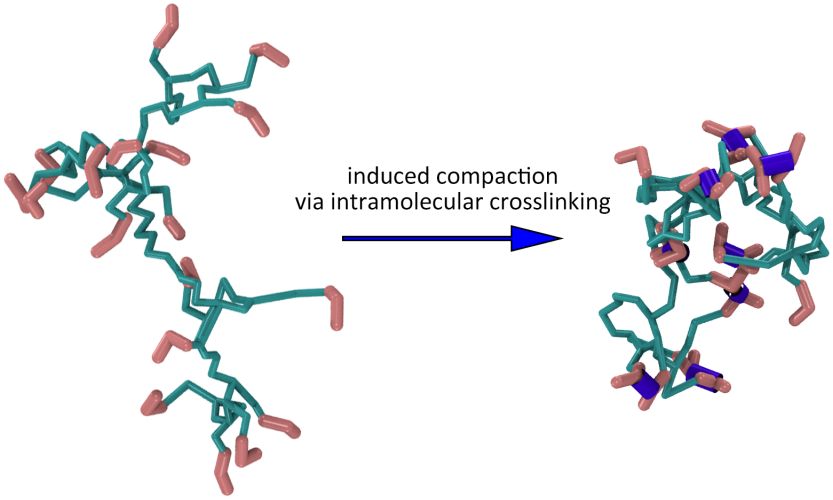


Figure 5.2.: Schematic representation of folding from a polymer coil into the single-chain nanoparticle via intramolecular crosslinking. Crosslinking candidate sites are in pink color, as crosslinks are indicated with thick blue symbols.

with covalent bonds. Thus, the combinatorial possibilities of polymer design, along with their respective folding mechanisms, multiples to an unprecedented level of freedom.

In experiments, from a top-down point of view in designing a polymer so that it folds with a controlled manner, light gated dimerisation is a crosslinking mechanism which mostly does not interfere other controlling means such as temperature modulation, as photodimerisation offer great spatio-temporal resolution in SCNP folding. [100] Amongst choice of photoreactive monomers, styrylpyrene is a versatile component in reversible photocycloaddition reactions for introduction of reversible covalent crosslinking into SCNP. Light of longer wavelengths (blue light, $\lambda > 400$nm) promote photocycloaddition of styrylpyrene, while shorter wavelengths (UV light, $\lambda = 330$nm) induce photocycloreversion in solution. However, confined upon a polymer backbone, the diffusion of styrylpyrene groups is strongly limited, resulting in a stark contrast of photocycloreversion behaviour against monomolecular styrylpyrene. Held in close proximity, photocycloaddition of styrylpyrene remains the predominant reaction over photocycloreversion, thus the folding and unfolding processes cannot be tuned separately via selection of wavelengths in the light source, rather leading to a photostationary state for UV.

Development of a holistic CGMD model for the investigation of the folding mechanism grants detailed understanding of selective control over the wavelengths-gated photoreversible cycloaddtition reactions and the precision control of polymer folding with selective polymer chain length, and the disparity between monomer based units, and the steric properties on side groups.

5.2. Methods

For the CGMD simulations, the degrees of freedom within the macromolecule is reduced depedent on the predefined coarse-graining scale. For a relatively homogeneous system as the pentaerythritol triacrylate (PETA) resin in the direct laser writing model, the dynamic details for each monomer is not so important. The 3 arms where the polymerisation reactions take place can be approximated by an isotropic reaction sphere. Meanwhile in the case of the SCNP folding, the polymer chain consists of large side groups with styrylpyrene which in terms of size and mass are comparable to the backbone units, thus requiring in the case of more complex structure reordering dynamics as the reversible

folding process more details in the modelling in the coarse-graining scheme. For the SCNP folding where the formation of intramolecular crosslinks transforms the linear chain into a compact conformation, the steric repulsion between the side groups gauges the crosslinking reation such that only when a pair of side groups remain close to each other while a preferential alignment is achieved the crosslining reaction should occur. Through a process of trail and error in tuning the scale of coarse-graining the explicit representation in the molecular structure and the enforcement of an implicit means of modelling where the finer details are smeared asks for a balance in the spatio-temporal resolution. The anisotropic nature of the interaction potential reflected from the qualitative instantaneous structural arrangement can hardly be recovered if the coarse-graining scheme does not conserve enough detailed information regardless of the data fitting or machine learning training effort in the quantitive force-field parametrisation. System specificly, the granularity have to be decided such that the desired level of detail preserves dynamics properties of interest while all the other parametrisation such as how many water molecules should be included in one coarse-grained atom group must be tuned for consistency with the rest of the system. In short, there is no silver-bullet choice of scales in the coarse-graining.

5.2.1. One point for one monomer: Langevin dynamics

In the simulation of direct laser writing, the system is prepared with all-identical monomers in the resin. The states of the monomers are differentiated via the definition of monomer types, whereas the physical properties as the mass, reaction spheres characterisation remain unchanged among the various types. The modelling of a dense pool of monomers, the Lennard-Jones fluid model works well in regulating the diffusion properties for the neutrally-charged individual as the fast diffusion timescale is well-separated from the slow chemical reaction timescale. The monomers are categorised into various coarse-grained particle types (CG types) when the differenct chemical reactions apply to the specific types in Table 5.1.

The dynamics of the point-like monomers subject to the Langevin dynamics as:

$$F_{\text{Langevin}} = F_{\text{conservative}} + F_{\text{dissipative}} + F_{\text{random}} \tag{5.1}$$

in which the conservative force $F_{\text{conservative}} = F_{\text{LJ repulsive}}$, contributing exclusivlely the repulsion as a characterisation of monomer sizes. The friction term $F_{\text{dissipative}} = -\gamma v$ which depicts the friction in a linear dependence with respect to the diffusion velocity v. The random force $F_{\text{random}} = \sqrt{2\gamma k_B T}\hat{r}$ is correlated to the friction as $\hat{r}$ is a stochastic process.

CG types	Descriptions
1	Common free monomers
2	Activated free monomers
3	Common monomers in the polymer network with 1-2 bond(s) formed
4	Activated monomers in the polymer network with 1 bond formed
5	Common monomer in the polymer network with 3 bonds formed
6	Crosslinked monomer in the polymer network with 3 bonds formed
7	Deactivated free monomers
8	Deactivated monomers in the polymer network

Table 5.1.: Coarse-grained particle types assignment in the simulation of direct laser writing.

In the implementation of the one point per monomer model, the individual monomers correspond to a feature-less spherical representation, when the monomer-monomer Lennard-Jones interaction characterises the size of the sphere. All the bonds formed on the monomer does not interact, e.g. no steric effects, due to the flexibility of the PETA arms, when the

polymer network is established as a wireframe without torsional potentials. Such a simplistic model captures the essential features of the rubbery elastomer [101] as the stablisation of the conformation and the mechanical properties of the printed polymer network depends on the spring-like interactions between the bonded monomers. The diffusion results in limited "hairiness" on the boudaries of the printing area, when the hairs quickly grow and maintain great flexibility due to the reaction scheme later discussed in the direct laser writing case study. Meanwhile the formation of crosslinks keeps the printed structure in place so that the desired patterns can be transcribed into polymer network conformations, illustrated in Figure 5.3.

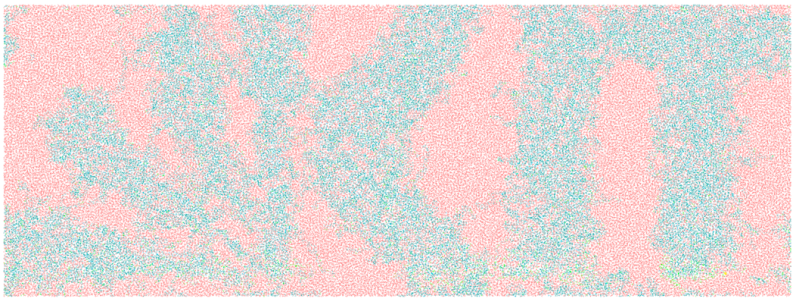


Figure 5.3.: Showcase of 3D printing the KIT logo in a thin 3D layer. Free monomers are colored in pink, the polymer network in cyan, and the active printing sites in light green.

The point-like monomer model can be extended for the simulation of the SCNP folding as during the irreversible compaction of the polymer chain, the crosslinking candidates also subject to the similar reaction scheme as with the direct laser writing such that the additional constraints gradually stablises the conformation of the SCNP. However treating individual monomers as a single coarse-grained (CG) point-like atom group smears the diffusion timescale of the side groups in the SCNP problem when the intermittent partial unfolding of the SCNP cannot be observed with the simpler models. The details in the dynamics can be only recovered not only be introducing the side group as addtional one CG atom group, but a selection of CG atom groups.

5.2.2. Dissipative particle dynamics in SCNP folding

Dissipative particle dynamics (DPD) is a coarse-grained MD simulation protocol widely adopted in simulating diffusive properties of macromolecules due to its advantageous efficiency compared to the Lennard-Jones-based model. [89] DPD is self-consistent in the definition of dissipative and random terms while the design of a soft conservative force which replaces the strong repulsion in the Lennard-Jones grants additional acceleration on top of the reduction of interal degrees-of-freedom from the coarse-graining scheme. In the DPD model for SCNP, the macromolecule is divided into several subgroups based on their mass and chemical functions, where each subgroup of atoms is represented as a single DPD bead. Unbonded interaction between DPD beads is characterised in a pair-wise fashion with 3 additive force components.

$$F_{\mathrm{DPD}} = F_{\mathrm{conservative}} + F_{\mathrm{dissipative}} + F_{\mathrm{random}} \tag{5.2}$$

Among the 3 components, the conservative force depicts short-range repulsive forces induced by mutual solubility for the polymer mix and the solvent. By adopting a soft repulsive

formulation with a short cutoff threshold:

$$F_{\text{conservative}} = a_{ij}\omega\left(r_{ij}\right)\omega\left(r_{ij}\right) = 1 - r_{ij}/r_{\text{cutoff}} \tag{5.3}$$

where r_{ij} characterises the distance between DPD beads and a_{ij} correspond to the repulsion amplitude. The conservative force is scaled via the scaling function $\omega(r_{ij})$, whose form is routinely chosen as linearly vanishing towards cut-off distance r_{cutoff}. [102] The dissipative force and the random force work together in characterising the thermal equilibrium the DPD bead resides within, operating effectively as a thermostat conventional MD simulations.

$$F_{\text{dissipative}} = -\gamma\omega(r_{ij})(\hat{r_{ij}} \cdot \vec{v_{ij}}) \tag{5.4}$$

$$F_{\text{random}} = \sigma\omega(r_{ij})\hat{\theta}/\sqrt{\delta t} \tag{5.5}$$

as r_{ij} characterises the pair-wise distances for the DPD beads, $\hat{r}_{ij}$ the unit vector along the direction of r_{ij}, and $\hat{\theta}$ a Gaussian stochastic process.

For the force-field parameters, the repulsive conservative force, gauged by the a_{ij}, reflects the DPD bead sizes between the same DPD bead type, whereas the inter-species interaction reflects the mutual mixing affinity between the atom groups. Correlation of the dissipative force and the random force is obliged to the fluctuation-dissipation theorem, as the friction coefficient γ and the random force term σ are coupled as $\sigma = \sqrt{2\gamma k_B T}$, where k_B is the Boltzmann constant and T is the temperature. In such formulation DPD disregards finer details below the fluctuation-dissipation intrinsic timescale as high-frequency vibrational modes are smoothened out, whereas the low-frequency vibrational modes, e.g. the collective motion in conformational changes, is reserved. Comparable to the enhance sampling method of aMD and the acceleration algorithm described in Chapter 2, through the definition of a soft conservative force, DPD allows for a much faster simulation compared to conventional all-atom MD. The spatio-temporal renormalisation introduces an intrinisic DPD length L_{DPD}, intrinsic mass unit m_{DPD}, and corresponding intrinsic timescale DPD $\tau_{\text{DPD}} = L_{\text{DPD}}\sqrt{m_{\text{DPD}}/\varepsilon_{\text{DPD}}}$.

In practice, the thermostat from the dissipative and random terms in DPD works well if the intrinsic timescales for all DPD types are comparable. The coarse-graining scheme from the all-atom model to the DPD is then dependent on the masses of individual DPD subgroups. In DPD the solvent is usually considered explicitly, as a result, for different solvents, the DPD schemes may be vastly different as light solvent molecules, such as water, are usually considered in combined groups, while heavier solvent molecules, in the case of SCNP folding, the tetrahydrofuran (THF) is considered as a singular DPD bead for each molecule. The conservative and dissipative parameters, as is reported, correlates with the mixing properties between various DPD types. As experimental data is not available for an arbitrary combination of DPD subgroup types, in the case of SCNP, where explicit solvent population is dominant, the conservative parameters corresponds proportionally to the subgroup size, while for abundantly large number density of DPD beads, the coarse-graining scheme is fine enough that a generic fluid-like behavior is described in invariance to a selected range of dissipative parameters. [103] Failure in the equal partition of molecular masses into various DPD types, or an inappropriate designation of dissipative parameters results in large timescale deviations in the diffusion modes, and can be identified in terms of simulation temperature extraordinarily large fluctuations, i.e. the failure of the DPD thermostat.

5.2.3. Finitely extensible nonlinear elastic model for bonds for bond length constraint, and segmental repulsive potential to prevent bond crossing

Between DPD beads, finitely extensible nonlinear elastic bonds (FENE) conserve the geometry of the SCNP. [104] Formulation of the FENE bond goes as:

$$E_{\text{FENE}} = -K{r_{\max}}^2 \ln\left[1 - \frac{(r_{mn} - \delta)^2}{{r_{\max}}^2}\right] \tag{5.6}$$

When the FENE bond is stretched with small displacement, the bond potential reduces to harmonic, yet for large displacement the FENE bonding potential diverges so that the maximum deformation is bounded to preserve molecular geometry. The vibrational modes introduced by the FENE bond have to be compatible with the intrinsic timescale in DPD. Desynchronised FENE oscillation modes from the diffusion timescales in DPD will introduce artifacts in the momentum of DPD beads and in extreme cases lead to divergence in FENE bond extension and FENE energies. With explicit bonds between DPD beads, segmental repulsive potential (SRP) is adopted to prevent bond crossing in the dynamics. [105] SRP introduces pseudo interaction sites in the mid-point of the FENE bonds, which repels each other in the form of:

$$F_{\text{SRP}} = C(1 - r_{uv}/r_{\text{cutoff-SRP}}) \tag{5.7}$$

The cut-off scaling of SRP is formulated in a comparable form with respect to DPD cut-off, and the parameterisation is consistent with that of DPD. Practically, a number of pseudo-particles on the mid-points of the bonds are introduced to the original system, as the forces of F_{SRP} are calculated to be the force between the pairs of pseudo-particles. An additional mapping step projects the forces on the pseudo-particles to the DPD beads in analogy to a lever mechanism.

5.2.4. Stochastic process of dynamic bond recombination

In both case studies of the direct laser writing and the SCNP folding where the chemical reactions of polymerisation and both intra- and inter-molecular crosslinking are present, the bonded structure is dynamically updated throughout the CGMD simulation. Conventional classic MD simulations cannot deal with chemical reactions where the bond dynamics requires redistribution of the partial charges upon the atoms before the bond formation and breaking takes place as a prerequisite and the rearrangement of the molecular structure after the reaction. Some *ab initio* MD protocols, e.g. ReaxFF, allows implementation of chemical reactions for a limited scope of elements during the all-atom MD simulation at the expense of greatly reduced simulation time coverage to 10ps. For CGMD the details the local rearrangement in the molecular structure are smeared such that the local energy barriers cannot be retrieved, the bond recombination is implemented in a heuristic and phenomenological, distance-based approach where the individual event of bond recombination is regulated by a stochastic process.

From the perspective of an atom group as the reaction center, a reaction sphere with a predefined, fixed radius is sampled around the the host atom group for possible bond formation reactions. When another reactive atom group is present in the reaction sphere, the formation of bonds corresponds to a stochastic event with a fixed probability. The current number of bonds differentiate the state of the reaction center and the transformation from between the states, e.g. a single-bonded atom group into an atom group with three bonds, is indicated by the transitioin between atom group types. This is a comparable classification method to the atom-types designation in the all-atom simulations when a

tetrahedral carbon is differentiated from an aromatic carbon via checking the number of bonds formed on the central atom and to whom the central atom is connected, while here the chemical reaction rules explicitly depends on the atom group types as the atom group types also represents the reactivity, e.g. in the direct laser writing problem when a monomer forms 3 links the atom group designation is changed into a deactived monomer so that it can no longer form additional bonds. The newly formed bonds are appended to the comprehensive list of bonds as the overall interaction potential is updated. In the reversible folding of the SCNP, the bond formation and bond breaking reactions are implement with a similar algorithm, while the added complexity requires the molecular structure to be recalculated every time for the whole polymer.

5.3. 3D printing a polymer network

The fabrication of arbitrary 3D micro- and nanoscale objects is achieved with high precision with the 3D direct laser writing technique however in the applications, a number of limitation remain in the feature size [106], the controllable mechanical stiffness [107], and fabrication speed [108], etc. In specific application to 3D print the functional devices, e.g. tissue engineering with biocompatible materials [109], and the conducting polymer devices [110], the requirements in the mechanmical strengths are case dependent, whilst the control over mechanical properties via tuning the printing parameters in the fabrication protocol, e.g. the laser power, the exposure time, the degrees of monomer conversion yet has to be understood. With the development of a MD simulation method as the correspodant digital test bed of the fabrication process, we seek to understand and improve on the QSPR for the 3D printed polymeric network with a focus on mechanical stability.

5.3.1. Coarse-grained building blocks, the reaction scheme, and printing 3D features via voxel stacking

The CGMD simulation of the direct laser writing process, through which a 3D polymer network gets printed initiated via a series of 3D voxels, starts from a homogeneous pool of all-identical monomers of pertaerythritol triacrylate (PETA). In the polymer network, each PETA monomer may serve as a 3-degree node as the monomer maintains 3 identical arms which contributes to the polymerisation reaction.

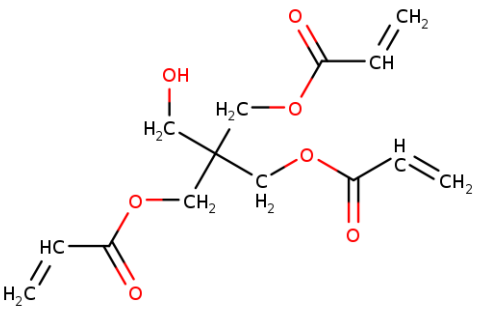


Figure 5.4.: The pentaerythritol triacrylate (PETA) monomer used in the 3D polymer printing.

The two-photon polymerisation involves radical dynamics [111], however, In the CGMD model, the photo-initiators that provide radicals, the radicals themselves, and photo-inhibitors which terminates the polymerisation, are all considered implicitly. Diffusion of the monomers and the crosslinking reactions between pairs of monomers are explicitly represented within the monomer resin, when the polymerisation is modelled in a step-wise series chemical reaction where the 3 steps are: 1) the initiation reaction that starts the

polymerisation and generate monomer bound radicals, 2) the propagation reaction which elongates the polymer and pass the radical along the polymer chain, and 3) the termination reaction where radicals are quenched thus the polymerisation stopped. Radicals dynamics on the polymer chain are represented via multitude conversions between monomer types. One PETA molecule is considered as a common monomer, denoted by M without the radical while an active monomer, denoted by A, is "initiated" within the printing voxel with a certain activation rate. Monomer type T corresponds to PETA within the printed polymer network. In addition to the 3 step reaction setup for the polymer network, the propagation also extends the network, termed as the branching reaction, while for the active monomer in the polymer network, a crosslink between polymer segments links combines the two parts into one. A comprehensive chart illustrating all reaction types of the initiation, the propagation and braching, the crosslinking, and the termination reaction by oxygen quenching is shown in Figure 5.5.

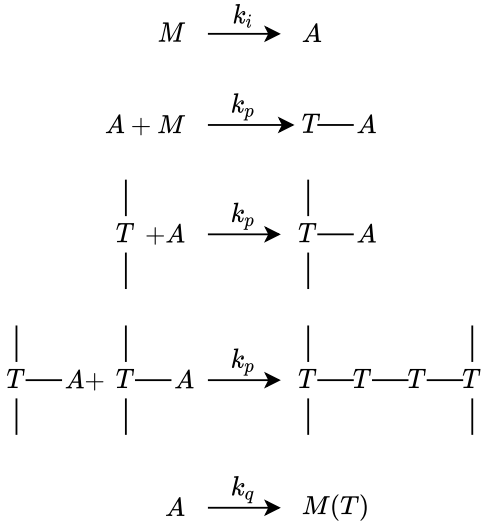


Figure 5.5.: Reaction schemes in the 3D polymer printing.

Corresponding to each type of chemical reaction, the reaction rate is determined via reaction rate constants of k_i, k_p, and k_q. In the initiation reaction, a monomer within the printing voxel is transformed into an active center, which initiate polymerisation reactions, via a stochasic process corresponding to the initiation rate k_i. When free monomers react with the active center, the translation of the radical is accomplished as the activate center move along the newly formed polymer in the propagation reaction to the new adduct in the form of a chain reaction. This results in a spontaneous elongation as without the requirement of supplementary radicals. Since the monomer of PETA retains multiple polymerisation sites, the branching reaction works in the similar form by extending the polymer network on the side branches and results in an active "loose end" site in the polymer network, subjected to further linear elongation. When two active ends meet, a crosslink is formed and the active centers are terminated, whereas for all active centers a constant rate of termination as stochastic process with the rate k_q, which characterises the oxygen quenching. Since the propagation, the branching, and the crosslinking reaction take place on monomers with 1, 2, or 3 polymerised arms, every time involving the formation of one additional polymerised arm, the reactions rates are identical as k_p.

For the determination of the reaction rate, k_i, the activation rate in the two-photon absorption activation, depends on laser power with a quadratic relation [111]. An explicit, moving ellipsoidal region in the simulation replicates the primary printing voxel as the focus from the two-photon absorption setup in the experiment. To print a 3D structure, the

priniting voxels are moved in discrete steps, where throughout the stationary time for each voxel, the laser exposure allows continuous initiation reactions within the printing voxel region. For the oxygen quenching, k_q is controlled via the local oxygen concentration, when the termination is modelled as a homogeneously distributed field across the simulation box throughout the simulation time span. The propagation, the branching, and crosslinking requires the involvement of active monomers, occuring across the whole simulation system when the as the active monomers is dynamically populated via the activation with a spatio-temporal pattern and reduced by the termination reactions. Analogous to the experimental setup, the 3D laser printing is based on a "substrate" of inactive bottom layer of the resin, while the printing is accomplished along the z-axis layer by layer. In the simulation however, the exposure protocol does not limit the printing process to continuous writing or substrate based printing as the voxel can be placed arbitrarily, e.g. within a enclosed shell, while the realisation of the pixel-by-pixel printing follows a scanning path as each newly printed layer corresponds to a 2D pattern projected from the 3D target object.

The concentration of the PETA is taken from experiment as 1.33mol/L and in the CGMD model for the PETA, the characteristic length scale of the monomer is $\sigma = 1$nm. In the bead representation of the monomer, the number density of the monomer is set to 0.8, each monomer bead has a diameter of 1σ, and in the bead-spring representation of PETA-PETA links the maximum bond lengths are 1.22σ. Bond formation is limited within a spherical region for pairs whose distance is below the reaction radius of 1.22σ. The CGMD characteristic timestep corresponds to the Lennard-Jones timestep τ when the reaction sampling frequency is set to 1 event per 10τ. The exposure time for each voxel ranges typically from 400τ to 2000τ when the total simulation time extends beyond $1E6\tau$. The reaction rates are calculated correspondent to experimental measurements in separate cases, gauged by laser power, as the comprehensive summary is listed in Table 5.2.

Control parameters	Value	Description
k	$k_p = 4.3\text{e}4(\text{L/mol·s})$ $k_q = 2.6\text{e}6(\text{L/mol·s})$	reaction rate constants
$[C]$	$[M] = 1.33\text{mol/L}$	concentration of monomers
P	$k_i = \alpha P^2$ $\alpha = 0.007\text{L}/(\text{s·m·W}^2)$	laser power in the printing
E	$N\tau$	exposure time
V	$(24\sigma, 24\sigma, 48\sigma)$	voxel dimensions

Table 5.2.: Summary of control paramters in the 3D polymer printing.

In the simulation, a desired geometry is pixelated into individual printing voxels with a finite resolution, as series of ellipsoidal printing voxels, whose size and aspect ratio is selective, covers all printing blocks while the voxels overlaps with a fixed ratio of 20% in all dimensions. A series of snapshots of fabricating a 3D polymeric network in the pyramidal shape via a moving voxel is shown in Figure 5.6. For the pyramids printing, the pyramidal shape is approximated by multiple layers of cubic blocks when each cubic block is filled by 2x2x1 overlapping voxels. Within the printing voxel, the activation reaction populates a number of reaction centers, which work as seeds as the growth of the linear polymer segments from those seeds quickly extends beyond the coverage of the printing voxel. Qualitatively speaking, this results in the "hairy" features on the surface of the voxel when the conformation of the hairs are subjected to diffusion properties regulated by temperature and local monomer density. Meanwhile, the diffusion, albeit limited to small travel distances, of the active monomers also result in polymerisation outside of the voxel. These two factors combined result in the diffusion regulated filling between the adjacent printing voxels, i.e. a "fluffy" area where the polymer network is indeed

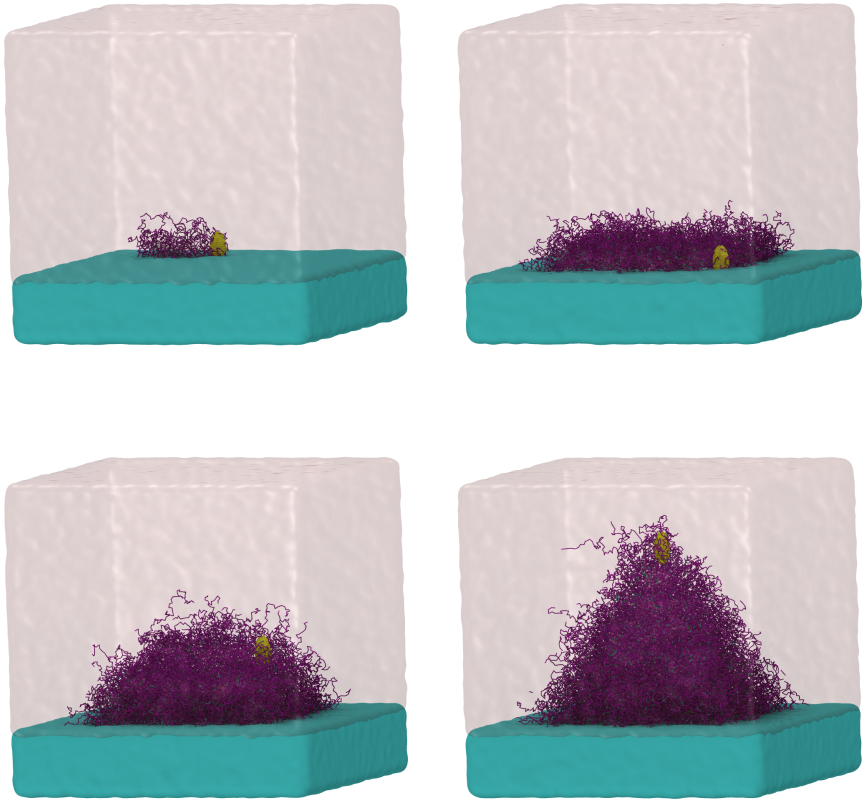


Figure 5.6.: Snapshots during the 3D priniting of a pyramidal structure. The light green layer corresponds to the inactive substrate, the light pink region as the homogeneous monomer resin, the yellow ellipsoid the printing instantaneous voxel, and the bright green colors within the pyramidal structure the crosslinking sites.

expanded compared to the printing template, while the crosslinking degree is substantially limited due to hindered diffusion in the dense monomer resin, its mechanical contribution to the polymer network being marginal. Throughout the parameter scans indicate that due to the existence of the "fluffy" region, the resolution of printing 3D features cannot be further improved via the adoptation of smaller printing voxels, while an elevated oxygen concentration in the quenching should help with the printing resolution by reducing the "hairiness", the mechanical strength of the printed polymer network suffers as crosslinking is suppressed.

5.3.2. Monomer conversion kinetics and efficiency tuning

To investigate how the interplay of various controlling parameters quantitively affect the properties of the printed polymer network, we draw a comparison between experimental data reported in a series of publications and the result from our CGMD model. monomer conversion degree is by definition the fraction of functional groups which is bounded versus the total number of functional groups on the monomers in the whole monomer pool. In the field of polymer chemistry, a pronounced threshold behaviour has been reported [112, 113] as the saturation is dependent on exposure time, suggesting a drastic difference in the conversion ratio between short time exposure and long exposure per voxel in the printing process. In the case of a brief exposure, the “life-time” of the reaction centers may extend beyond the exposure interval such that even after the laser has been switched off, the polymerisation reactions continue. During the non-irradiated period, the oxygen quenching

gradually reduces the number of remaining reaction centers before a stationary state with constant monomer conversion degree is reached. Whereas in the case of long exposure, the free monomers are depleted long before the laser is switched off, i.e. all monomers within the resin is either incorporated into the polymer network, or in rare cases remain deactiavted after going through the activation and quenching processes. This results in a very high monomer conversion degree, illustrated by the vanishing monomer concentration in Figure 5.7.

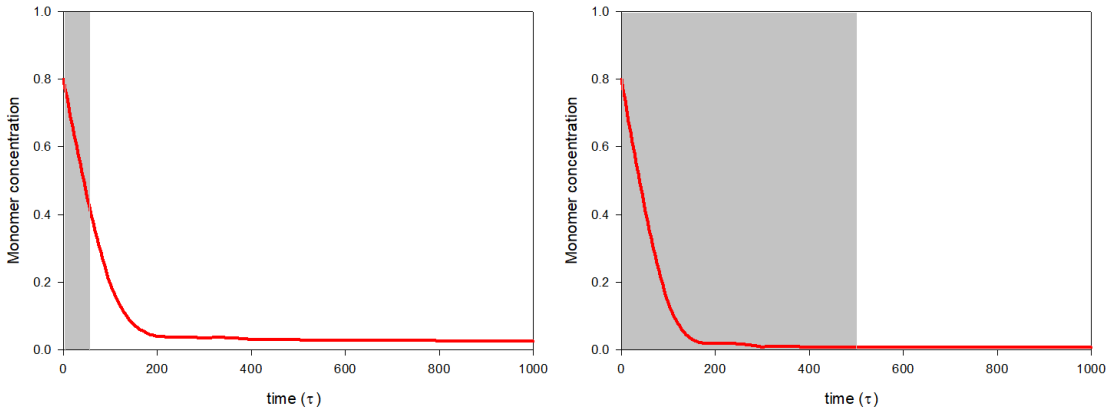


Figure 5.7.: Short vs long exposure on the monomer conversion degree.

Reported from multiple sources in the literature, the monomer conversion data is available for a range of direct laser writing applications with both commercial and non-commercial photo resist. Dependent on their specific printing conditions, the monomer conversion degrees of 60% to 75% has been achieved for non-commercial photo resists based on the mixture of two different tri-acrylates [114–116], while for a range of hybrid organic-inorganic photo resists, the conversion ratio ranges between 35% to 75%. [117, 118] That is, the polymer network that results from the direct laser writing withholds consistently a relatively low conversion ratio well below 100%.

In the simulation we comprehensively investigates the different aspects in improving the ratio of monomer conversion. Here in a cubic simulation box of size 40x40x40σ, the photoresist fills the space with periodic boundary conditions which corresponds to the bulk phase monomer solution when the distribution of the PETA monomers is homogeneous. The activation reaction initiated by the laser exposure is considered as always-on across the whole simulation box, while the laser power ranges correspondingly from 23 to 45mW. The monomer conversion degrees are sampled various exposure times and our simulation results is in very good agreement with the experimental results reported from tri-acrylate based photoresists. [119] Shown in Figure 5.8, the plateau with conversion ratio at its height near 75% to 77% suggests a moderate laser power of 40 to 45mW in the printing process is sufficient as further gains with greater laser power becomes marginally beneficial.

During the exposure period, the monomer conversion degree is regulated by the 3 competing processes as the diffusion, activation, and the termiation, when the oxygen concentration is considered as being constant, the termination rate is fixed. The propagation reaction incorporates the monomers into the polymer network while the radical site is moved to the end of the linear chain. For a low laser power printing which corresponds to a slow activation rate in relation to the constant termination rate, the growth of linear polymer chains is achieved with only finite length and the crosslinking degrees between polymer chains remains at a low level as the active ends are quickly quenched. In the simulation this polymer networks with limited crosslinking degree and finite size nevertheless exhibits

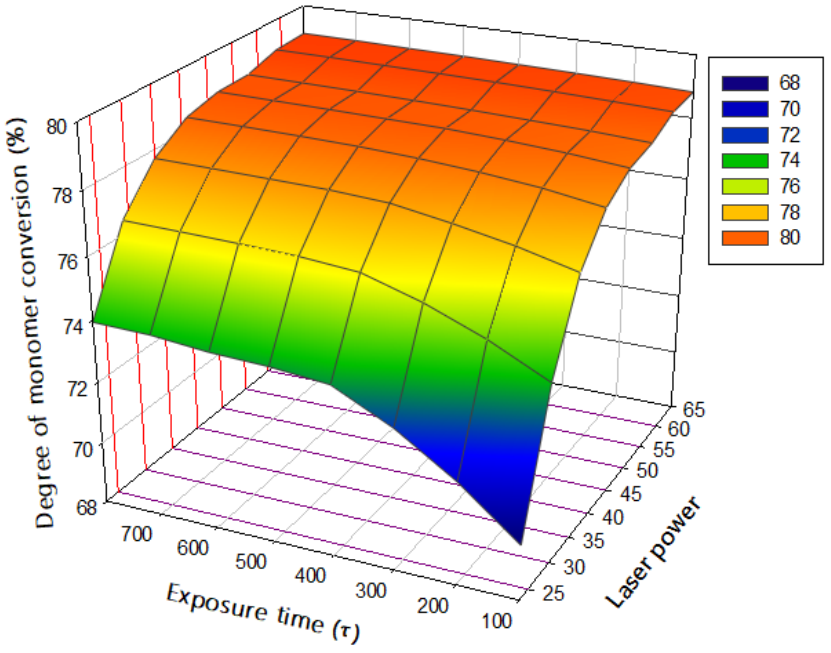


Figure 5.8.: The comprehensive effect of laser power and exposure time on the monomer conversion degree.

desirable monomer conversion degrees however in the experiments, the detection and measurement for such small polymer networks after rinsing is difficult if at all viable.

On the other hand, a high laser power corresponds to fast activation compared to termination however the polymerisation process as a swift construction of the polymer network, i.e. a framework comparable to the MOF in Chapter 3, limits the supply of fresh building blocks. That is, incrementally each new bond formed in the polymerisation reaction or each new crosslink formed adds a constraint to the morphology of the polymer network, while the increase in size of the polymer network renders greater steric repulsion which makes it more difficult for the free monomers from the resin to get access to the polymer network. For the intra-network crosslinking, the active sites on the end of the network branches have to find a partner for the inter-chain crosslinking reaction, which also gets hindered due to the loss of interal degrees of freedom within the polymer network. All factors combined, for time difference between occurences of bond formation events take exponentially long. Meanwhile, the termination reaction reduces the active sites on the ends of the polymer chains to quenched sites irreversibly, at a rate unaffected by the growth of the polymer network. In the long time limit, all monomers within the resin are either incorporated into the polymer network or become inactive, suspended monomers in the solution. The saturation level of conversion is reached faster with greater activation rates, however the terminal monomer conversion degree is merely slightly improved.

5.3.3. Mechanical strengths characterisation of the printed network

Apart from the conversion ratio, the printed polymer network also exhibits different mechanical properties mostly due to the difference in crosslinking degrees. In the experiments, the polymer network is only obtained after the resin gets rinsed, for which process the printed product must withstand with enough structural integrity. In the simulation our printed network does not subject to bond recombination under stress while the final product bar the suspended monomers is deformed to measure the corresponding Young's modulus. To compare the polymer printed in the simulation with that of the experiment, we have

incorporated the printing voxels as a serial sequence, the polymer network is fabricated voxel by voxel between which a 20% overlap is applied for the inter-voxel crosslinks while the writing velocity, the voxel size, and the laser power are control variables. For the bulk phase tests, a cubic simulation box with periodic boundary conditions is prepared with homogenous monomer resin. A series of polymer blocks are printed with 4x4x2 voxels with the aspect ratio of 1:1:2, when the printing voxels reach beyond the boundaries of the simulation box, i.e. the overlap between voxels connect the printed areas on all of the x, y and z directions across the simulation box boundaries. The laser power ranges from 8.8mW to 63mW in the printing process, which corresponds to the monomer conversion degree ranging from 65% to 77%.

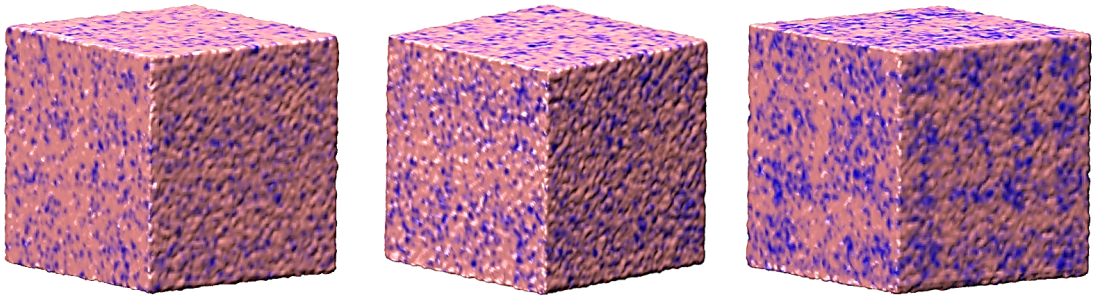


Figure 5.9.: Polymeric cubes printed from voxels with aspect ratio of 1:1:2 with variable laser powers. Blue areas correspond to crosslinks within the polymer network, indicating an enhanced crosslinking with the increase in printing laser power.

The Young's modulus of the printed polymer network is calculated via the stress-strain simulations. The simulation box is expanded at a constant engineering strain rate, the internal stress of the polymer network as a response to the expansion is monitored. The snapshots of the printed polymer networks, illustrated in Figure 5.9, indicate that more intense laser exposures result in enhanced crosslinking degrees, which under various external strain intensities, exhibit a two-region behavior consisting of a linear elastic behavior for small deformation which then exhibits a transition to hyper-elasticity for large deformation with nonlinear response. A detailed inspection of the CGMD trajectory shows that the linear response for small deformations correspond to the elasticity in the stretching of the polymer coils, while the nonlinear response is predominantly contributed from the elongation of the FENE bonds at extreme extension. The regression from the linear response segment is then compared among stress-strain analysis of the printing results from different laser powers. A transitional behavior is also observed in the relationship between Young's modulus and printing laser power, illustrated in Figure 5.10, suggesting the sufficient stiffness in the printed polymer network can be achieve via printing at moderate laser power, typically at 35mW to 40mW with the monomer conversion at 70%. The Young's modulus gradually saturates as the laser power is increased further, while the monomer conversion degree increases from 61% to 73%.

Another way of modulating for the desired Young's modulus of the printed polymer network is to change the writing velocity, gauged practically with the exposure timing. (ref 44 Elahe) In the simulation, we have again adopted the cube grid printing via stacking voxels with aspect ratio of 1:1:2 which fills the simulation box as the adjacent voxels are overlapped by 20% along all axial dimensions. After the printing is accomplished, the resulting polymer network is expanded in z-direction for the stress-strain relationship benchmarking, and once again the transitional behaviour in the elastic response is observed as the linear response corresponds to small deformation whereas a hyperelastic response corresponds to the FENE extension for extreme stretching. In Figure 5.11.

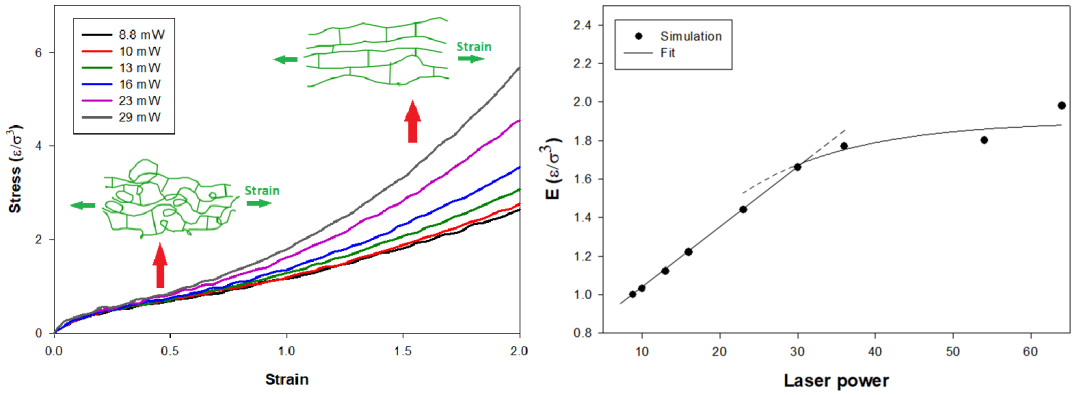


Figure 5.10.: Stress-strain relationships of the printed cubes with respect to laser power. Insets in the stress-strain curve represent the elastic regions of linear response, corresponding to the stretching of polymer coilds, vs the polynomial, corresponding to the FENE potential. The Young's modulus exhibit a transitional dependency with respect to the laser power, marked by the linear fitting and the exponential fitting segments, indicating a level of saturation in the mechanical strength when the laser power is large enough.

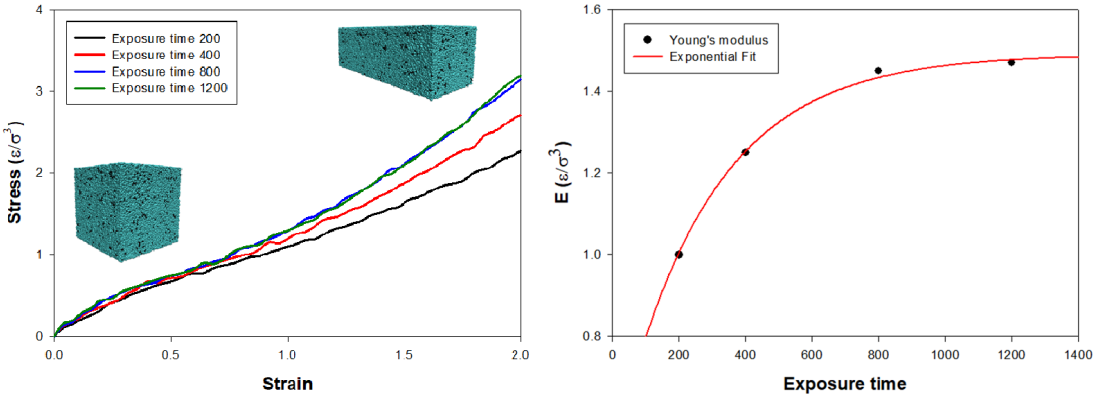


Figure 5.11.: Stress-strain relationships of the printed cubes with respect to exposure time. Insets in the stress-strain curve illustrates the deformed polymer network when the elastic response is measured. The Young's modulus exhibit an exponential trend with respect to the exposure time, marked by the fitting curve, when a level of saturation in the mechanical strength is reached when the exposure is long enough.

The Young's modulus increases with respect to the exposure time while the monomer conversion degree also see increases from 59% for $t_{\text{exposure}} = 200\tau$, 63% for $t_{\text{exposure}} = 400\tau$, to the saturation platform of 65% monomer conversion at $t_{\text{exposure}} = 800\tau$ and 66% conversion at $t_{\text{exposure}} = 1200\tau$. The crosslinking degree and the mechanical strengths show a consistent increase, which also saturates quickly for long exposure times. Meanwhile, the Young's modulus show a tunability range from $1.0\epsilon/\sigma^3$ to over $1.5\epsilon/\sigma^3$.

Additionally, the mechanical strength of the printed polymer network depends on the geometrical properties, where in the most simple cases, the aspect ratio for a cylindrical rod. For the bending response benchmark among cylindrical rods of the same height as the cross-sectional area is varied, the cylindrical rods are fixed upon the substrate, whereas the top of the rods are deflected along the x-direction, snapshots shown in Figure 5.12 as the "hairy" shells are manually removed. Due to the finite scale of the sample

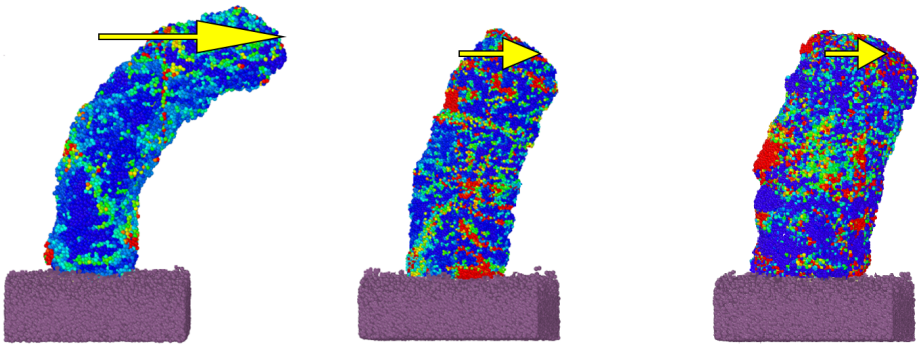


Figure 5.12.: Projected strains during the bending of printed cylindrical rods with variable aspect ratios. Yellow arrows represent the deflection under stress, whereas the color map correspond to local strain intensities, warm color for high strain and cold color for low strain.

and the stochastic bond formation, the strain measurement exhibit strong local patterns, whereas the overall bending deflection curve is linear with respect to the cross sectional area, positively proportional with respect to the printing laser power.

5.4. Folding a single-chain nanoparticle

Mimicking the folding process of the polypeptide from a random polymer coil into a unique, functional conformation, the artificially synthesised single-chain polymers can be programmed with a selection of units subject to folding mechanisms as hydrogen bonds [98], metal-complexation [99] , and covalent bonds, or a combination of those intramolecular crosslinks. At the same time, the triggering condition of the folding can be also selected from the polymer toolbox where control over concentration of the crosslinking units [120], sequence synthesis with a predefined order among the monomers [97], and temperature gating [98, 121] are all viable options while the light-gated reaction maintains the advantageous orthogonality against other control factors and can be spatiotemporally applied with high precision.

5.4.1. Photodimerisation of styrylpyrene containing monomers: in the solution vs on the polymer

Here we consider a fixed length polymeric single chain populated with photoreactive groups of styrylpyrene. The polymer chain of approximately 50 monomer units is constructed from two types of monomers: the styrylpyrene containing photoreactive monomer and the non-photoreactive foundation of methyl methacrylate, plus the explicit molecular solvent of THF. In the solution of monomers with the styrylpyrene groups, the photocycloaddtion and reversion reactions can be triggered by 455nm laser and 330nm laser respectively. In the solution, dimers are formed when the crosslinking candidate monomers diffuse into a very close proximity with the longer wavelength laser exposure, and with time all monomers manage to find partners. In a similar scheme, for the reverse reaction induced by the shorter wavelength exposure, the parted monomers quickly diffuse away from each other until all dimerised pairs are reduced to individual monomers.

In the previous study [120], with the irreversible dimerisation induced by longer wavelength exposure, the high degree of dimerisation can also be achieved for the monomers confined to a polymer chain. However, due to the geometrical constraints induced by the polymer

backbone, the styrylpyrene groups on the chain are limited in diffusion as results in a very high effective local concentration, the photocycloaddtion between the styrylpyrene groups on the polymer chain is much more efficient than the hydroxy-styrylpyrene. The vivid contrast between the confined photo-reactive groups on the polymer vs its free-roaming hydroxy-styrylpyrene counterparts presses a question that further extends the application of styrylpyrene on polymers, i.e. how does the photocycloreversion perform for confined styrylpyrene groups, and are we able to open and close the covalent bonds and reshape the SCNP however we wish?

The polymer chain is synthesised from a mixture of MMA and **M1** monomers. 2-cyanopropan-2-yl benzodithioate, methyl methacrylate (MMA), 2-(methacryloyloxy)ethyl(E)-5-(4-(2-(pyren-1-yl)vinyl)phenoxy)pentanoate (**M1**) and AIBN are dissolved in toluene as the reaction mixture. The mixture is thence purified via size-exclusion chromatography (SEC) to remove unreacted monomers before being dissolved in an excess of THF. This results in an extremely low concentration of the polymer in the solution where the intermolecular reactions are very rare. Characterisation of the end product is performed with ^{1}H-NMR spectroscopy and SEC where the MMA to **M1** ratio, and the average polymer chain length are derived. The synthesis process and experiment characterisation is illustrated in Figure 5.16, and the summary of the photoreactive monomer ratios is shown in Table 5.3.

	M_n	Đ	Styrylpyrene content	x/y
P1	13600	1.22	28%	17/44
P1'	19300	1.26	30%	25/58

Table 5.3.: Statistic of polymer systems **P1** and **P1'**. x, and y correspond to the number of styrylpyrene containing **M1** units, and the base MMA units.

From the experiment side, the exposure with a constant photon count at varied wavelegnths is achieved with an tunable laser source. The sample is loaded in a glass vial which is supported by a metal sample holder. The monochromatic laser pulse is regulated by an

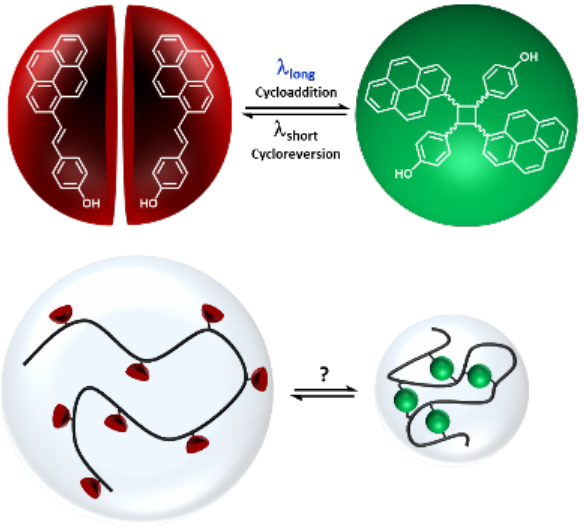


Figure 5.13.: Photocycloaddition and reversion in the covalent bond recombinations of styrylpyrene. The covalent bonding between styrylpyrene groups can be gated via the exposure under different ligh wavelengths for hydroxy-styrylpyrene, the motivation of this study is to investigate whether the wavelength-gated reversible crosslinking can be achieved with the same tunability.

polarising attenuator integrated in the laser source, before being reflected by a prism and then shines through the whole cross-sectional area of the sample vial before the number of photons reaching the energy meter is recorded, the laser path illustrated in Figure 5.14. For the estimation of number of photons absorbed by the sample in the vial, the measurement with an empty sample vial is used to calibrate the photon count.

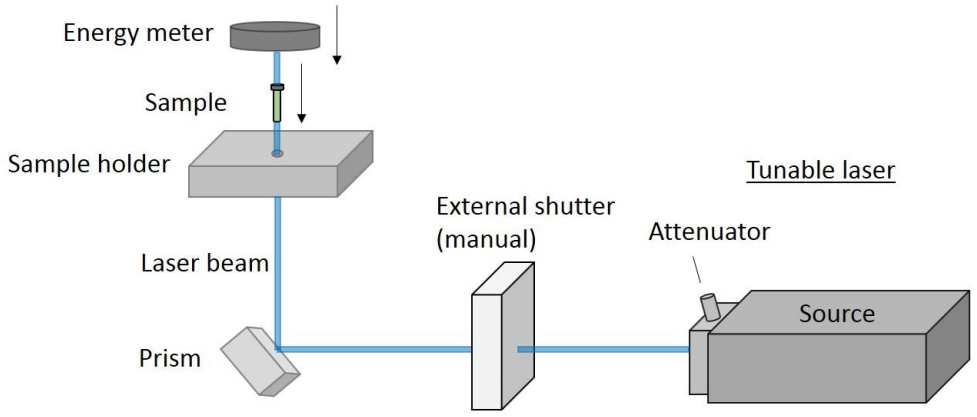


Figure 5.14.: Experimental setup of the SCNP folding.

For the CGMD modelling of the polymer, where the diffusion properties are vastly different between the non-photoreactive MMA monomers and the styrylpyrene containing **M1**, the **M1** monomer unit being 5 times heavier than the MMA, an DPD coarse-graining scheme of dividing **M1** into smaller subgroups is required. In the explicit representation of the **M1** monomer, for the consistency of intrinsic timescales for all DPD types, the DPD bead weights among all types need to be comparable. The consistently weighted DPD beads are coupled to the "DPD thermostat" which includes the mutual friction between the DPD beads, and the heat bath, implemented as Gaussian random forces. Additionally, the parametrisation of the DPD beads reflect the steric repulsion of the subgroups with various sizes, which then work together with inter subgroup linkages with various strengths in controlling the overall geometry of the polymer chain.

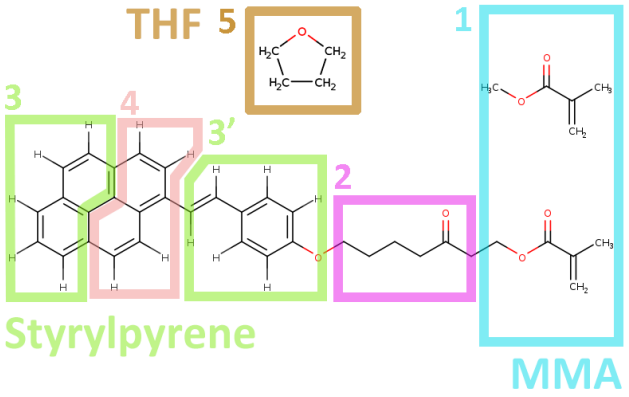


Figure 5.15.: DPD coarse-graining scheme of the monomers and the THF solvent.

For all the interaction parameters, the characteristic timescales have to be consistent with the DPD intrinsic timescale in design of the coarse-graining scheme. In an example where the coarse-graining scheme leads to poor performance, an ill-defined mass ratio between the DPD beads leads to difference in diffusion timescales, and thence results in poor

temperature control of the DPD thermostat as the vibrational mode are time separated, while the assignment of improper sizes, in the form of vastly different repulsive potentials in the DPD conservative term, undermines the characterization of the steric effect. In the implementation, a 5-types grouping scheme assigns 1 DPD type for the THF solvent, 1 shared DPD type between the MMA and the styrylpyrene containing **M1** monomer, 1 DPD type for the connection of the MMA and the styrylpyrene in the **M1**, and 2 types for the photoreactive styrylpyrene groups. Across all DPD types, the molecular weights reside in the scale of 100Da, where the styrylpyrene group is noticeably divided into 2 DPD types in 3 parts as the atoms in the interlocked aromatic complex is shared between the DPD beads of different types. Such implementation on the one hand improves the thermodynamics regularisation as the subgroup size and mass are comparable, on the other hand, the C=C double bond in the stytylpyrene is then modelled as being on the middle compoenent in the photoreactive complex, such that when crosslinked, the head-to-head or head-to-tail alignment of the styrylpyrene groups are addressed. The comprehensive parametrisation of DPD, FENE, and SRP modelling of the SCNP according to the DPD coarse-graining scheme, illustrated in 5.15, is summarised in Table D.5 in the Appendix. The characteristic length scale of the DPD model is $L_{DPD} = 1.0nm$, the characteristic mass at $m_{DPD} = 100Da$ and energy unit $\varepsilon_{DPD} = k_BT = 4.1416\text{E}{-}21\text{J}$ at 300K. The CGMD simulation timestep δt is chosen to be $0.01\tau_{DPD}$, when the intrinsic time scale of DPD is $\tau_{DPD} = L_{DPD}\sqrt{m_{DPD}/\varepsilon_{DPD}} = 7\text{ps}$, around 4 orders of magnitude longer the typical all-atom simulations.

The polymer chains with fixed lengths of 61 and 83 monomer units are designated as **P1**, and **P1'**, corresponding to the experimental polymer chain length. In the synthesis, the polymer chains follow a Gaussian distribution in chain length, whereas the photoreactive units are randomly distributed across the polymer chain. In the simulation, while the chain length is fixed, the number of photoreactive units are Gaussian distributed, with the expectation value centered at 28% for **P1**, and 30% for **P1'**. An ensemble of 500 distinct starting conformations are constructed in the linear form, while THF solvent molecules are later added to the simulation environment before the polymer-THF combination is relaxed.

5.4.2. Intermittent folding features and the photostationary state in the reversible folding

The folding dynamics are investigated in 3 processes as the: a) irreversible folding of the polymer coil into SCNP subjected to blue laser with the wavelength of 430nm, b) reversible folding of the polymer coil into a dynamic photostationary state via UV laser whose wavelength is 330nm, and c) the partial unfolding of compact SCNP, which is the end product of blue laser folding, into an UV-equilibrium state under the exposure of UV laser.

During the exposure with different wavelengths, the measured UV-vis spectra show a time-evolution bias in the absorption, from which the degree of crosslinking between the styrylpyrene groups are calculated. At the styrylpyrene adsorption peak of $\lambda = 383\text{nm}$ the blue laser exposure folds **P1** from the polymer coil into the **SCNP1**, while a monotonical decrease in the ratio of open bonds is observed with respect to time, illustrated in Figure 5.17, with an end-state at less than 5% open bonds. With exclusive UV laser exposure, the ratio of open bonds in the polymer coil also sees a monotonical decrease however the terminal state shows an equilibrium at 23% open bonds, shown in Figure 5.18f. However, as we change the starting conformation from the polymer coil into the pre-folded SCNP subjected to blue laser exposure, an intermittent increase in the number of bonds with an overshoot beyond the UV-laser equilibrium is observed. Shown in the case of the parital unfolding from **SCNP1** to the UV-equilibrium state at 35% openbonds in a short time

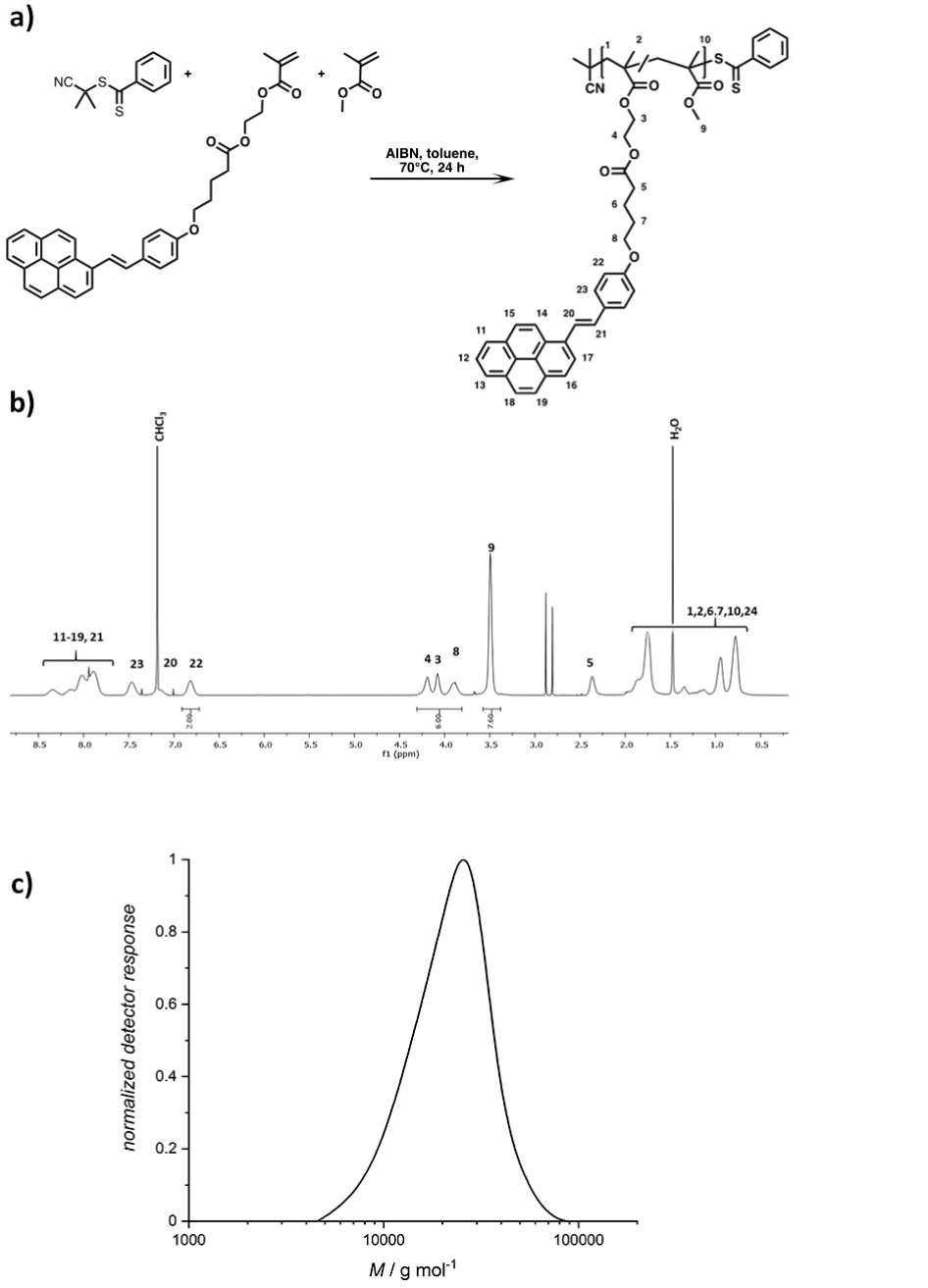


Figure 5.16.: The synthesis and characterisation of the polymer with mixed MMA and **M1** units. a) The polymerisation reaction of MMA and **M1** units and unique atom designations corresponding to the NMR characterisation. b) The ^{1}H-NMR spectrum of **P1'** in $CDCl_3$ solution show distinct singal peaks corresponding to the numbering scheme in a). c) Size-exclusion chromatography trace of the **P1'** polymer in THF calibrated with polymethyl methacrylate (PMMA).

interval, before the ratio of open bonds gradually recovers to the equilibrium state consistent with exclusive UV exporsure at around 25%.

From the perspective of simulation, the dynamic bond formation and fission for the styrylpyrene complex is modelled as a local stochastic process in the CGMD simulation when the bond recombination rates are empirically fitted to the experiment measurements. An individual bond can be formed with a finite probability when a pair of non-crosslinked

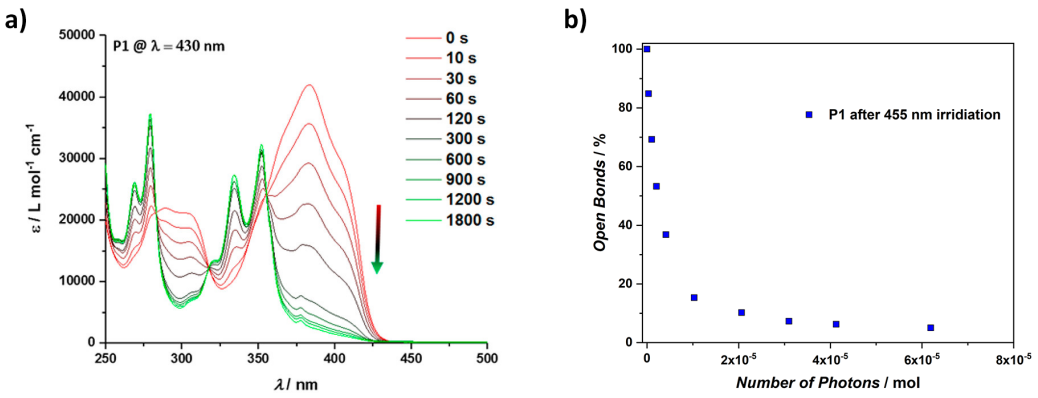


Figure 5.17.: The UV/vis spectra and the number of open bonds of folding **P1** into the photostationary state (d-f).

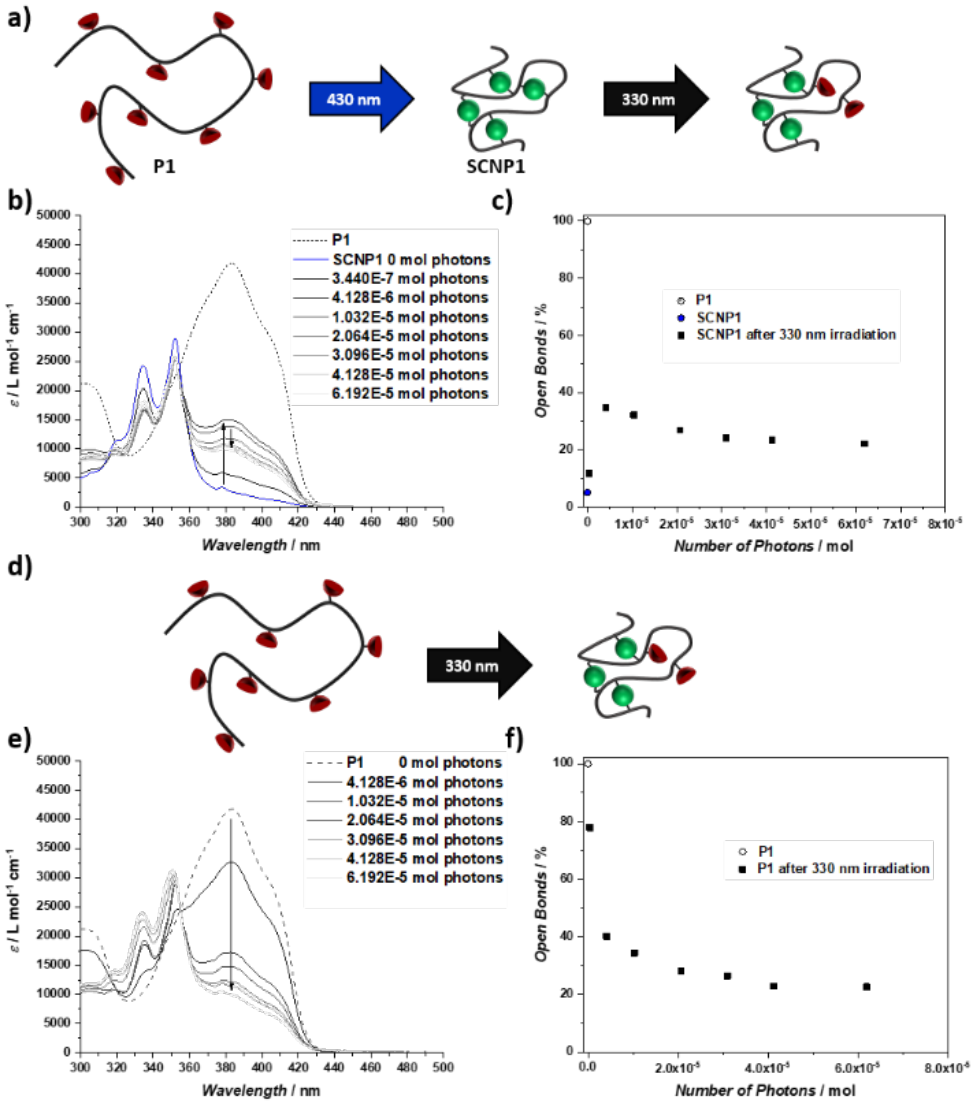


Figure 5.18.: The schematic representation, UV/vis spectra, and the number of open bonds of partially unfolding **SCNP1** into the photostationary state (a-c), and folding of **P1** into the photostationary state (d-f).

photoreactive sites get closer than $1.5L_{DPD}$. For photo-crosslinks, there is a certain probability that the bond is broken regardless the bond length. This induces in the dynamic bond formation and fission an asymmetric pattern, whose effect is exemplified by the bias in reaction rates. The reactions rates of bond formation and fission cannot be directly acquired from experiments, however for the processes considered in the SCNP folding under the exclusive blue laser exposure, the exclusive UV laser exposure, and the partial unfolding of blue laser folded SCNP subjected to UV laser, the pair-wise ratios of bond formation and fission rates can be determined with respect to the reference of diffusion rates. The interplay between bond formation and fission timescales and the diffusion timescales, results in a stark contrast between dimerisation behaviours of photocycloaddtion and photocycloreversion of styrylpyrene in the solution and on the polymer. For hydroxyl-styrylpyrene, free monomers containing styrylperene complexes diffusion away once the bond fission releases the geometrical constraint, however, for styrylpyrene confined to the polymer chain, they are always held in close proximity.

For the modelling of bond recombination, the timescale separation between bond recombination and diffusion comes into play. Modelled as a stochastic process, the bond recombination rate is dependent not only upon the local bond recombination probability, but also upon the sampling frequency when a bond recombination is attempted. Through some experimentation, it is found out that for exclusive blue laser exposure or UV light exposure, the bond recombination rate is scalable as the bond recombination sampling rate and local reaction probability offsets each other. For example, in the same time span, between a more frequently attempted chemical reaction whose reaction probability is low and a sparsely sampled chemical reaction rate whose event preserved a high reaction probability, the final results are equivalent. However, as will be later discussed, this scalability does not hold for the dynamic overly unfolding when the compact SCNP gets unfolded subjected to UV laser exposure. In the case of high sampling rate, e.g. 1 chemical reaction attempted per timestep, in the same order as the thermodynamic transport event, the DPD thermostat fails to take into account any information below the DPD intrinsic timescale whilst keeping for each timestep a thermodynamic equilibrium. The very high sampling rate is thence equivalent to the thermodynamic limit, as intermittent dynamic modes are not explicitly shown. This observation indicates that a sufficiently large timescale separation between chemical reactions and diffusion is crucial to depict the intermittent state. The parameterisation of bond recombination follows Table 5.4.

Photochemical reaction	Bond formation rate	Bond fission rate
Irreversible	$0.2057/1e2\tau_{\mathrm{DPD}}$	0
Reversible	$0.6127/1e2\tau_{\mathrm{DPD}}$	$0.00035/1e2\tau_{\mathrm{DPD}}$

Table 5.4.: Bond recombination parameters of the SCNP folding

From the CGMD simulation, an individual trajectory of a **P1** conformer quickly folds as each newly formed intramolecular crosslink confines the polymer conformation under blue laser exposure, while the geometrical constraint on the one hand increases the effective concentration of photoreactive units locally, on the other hand the bond formation locks the conformation step-by-step, which limits the diffusion of monomers in the local vicinity, rendering it harder for the remaining unsaturated crosslinking sites to find each other. In the end, a compact conformation is achieved, as the quantum yield of the photocycloaddtion reaches around 95%. However, for some starting **P1** conformers, along the timeseries towards the correspondent **SCNP1**, it is extremely hard for the very last remaining individual unsaturated crosslinking site, or the last pair of unsaturated crosslinking sites to find a proper partner when the rest of the polymer is locked. As a result, in the timeseries of number of open bonds, a plateau is gradually reached, when the lowest number of

open bonds always retains a finite distance from 0%. In the radius of gyration timeseries, considerable fluctuations are observed in contrast to the monotonically decreasing ratio of open bonds, when the conformer, although constrained by incrementally added crosslinks, still constantly explores the conformation space.

Interestingly, in the case of reversible folding under UV laser exposure, an individual trajectory reaches intermittently a completed saturated number of crosslinks, as well as an almost completely open conformation as a polymer coil as the constant photocyclo-addition and -reversion recombines the intramolecular bonds, when diffusion of monomers help the polymer chain to explore a greater extent in the conformation space. That is, the UV-equilibrium state does not correspond to a locked conformation but in its nature a photostationary state. Such photostationary states are observed as a plateau in the ratio of open bonds as well as in the radius of gyration.

For the partial unfolding at 330nm UV exposure upon the compact **SCNP1**, a transition from the plateau level of 430nm exposure to the plateau level consistent with exclusive 330 nm exposure is observed, when in the early stages of the unfolding, an intermittent overshoot both in the case of ratio of open bonds and the fluctuation in the radius of gyration is highly pronounced in the dynamics, shown in Figure 5.19.

The statistical average summarised from the ensemble of 500 simulations from distinct initial conformations, the transitional **P1**-**SCNP1**-photostationary state dynamics is reproduced very well with the CGMD simulation. From the analysis of MD data, the greater amplitude in the radius of gyration calculations versus the ratio of open bonds data indicate a pronounced timescale separation between the bond recombination process and the diffusion behavior. This explains why when in the simulation if the chemical reaction is attempted at a higher frequency, even though the photostationary state depends solely on the competition between bond formation and bond fission probabilities in the bond dynamics model, the intermittent overshoot cannot be retrieved. In the specific case of styrylpyrene group on the monomer, the aforementioned steric hindrance and strict alignment conditions of the photocycloaddtion reaction renders the bond recombination a much slower process than the diffusion. The comprehensive geometrical constraint of the photoreactive monomer on the polymer chain, the relatively slow bond formation, and the rapid morphology changes of the polymer chain eliminates the orthogonality between the wavelength-gated photocycloaddtion and photocycloreversion reactions for the hydroxy-stylrylpyrene, as both the blue light and UV light folds the polymer coil into the SCNP or a dynamic photostationary state.

Additionally, the parameterisation based on fitting the experimental data for **P1** is transferred to **P1'**. In the longer polymer of **P1'** with a comparable ratio of photoreactive monomers, the statistically longer segmental length between the crosslinking sites on the polymer corresponds to a less constrained conformer shape, as is discussed in the literature [98]. As a result, the dynamic overshoot from unfolding **SCNP1'** to its corresponding photostationary state is of weaker amplitude, illustrated in the Figures E.1 and E.2 in the Appendix.

5.5. Conclusion

Coarse-grained molecular dynamics allows for the fast simulations of larger scale assemblies in tens to hundreds of nanometers where the chemical reactions are implemented empricially as a stochastic process. Compared to the minutes to hours of experiments in the laboratory, the simulation of the 3D printing and the polymer folding are accomplished in a comparable timescale in minutes for individual simulations and hours for a statistic ensemble, extraordinarily faster than all-atoms simulations discussed in the previous chapters.

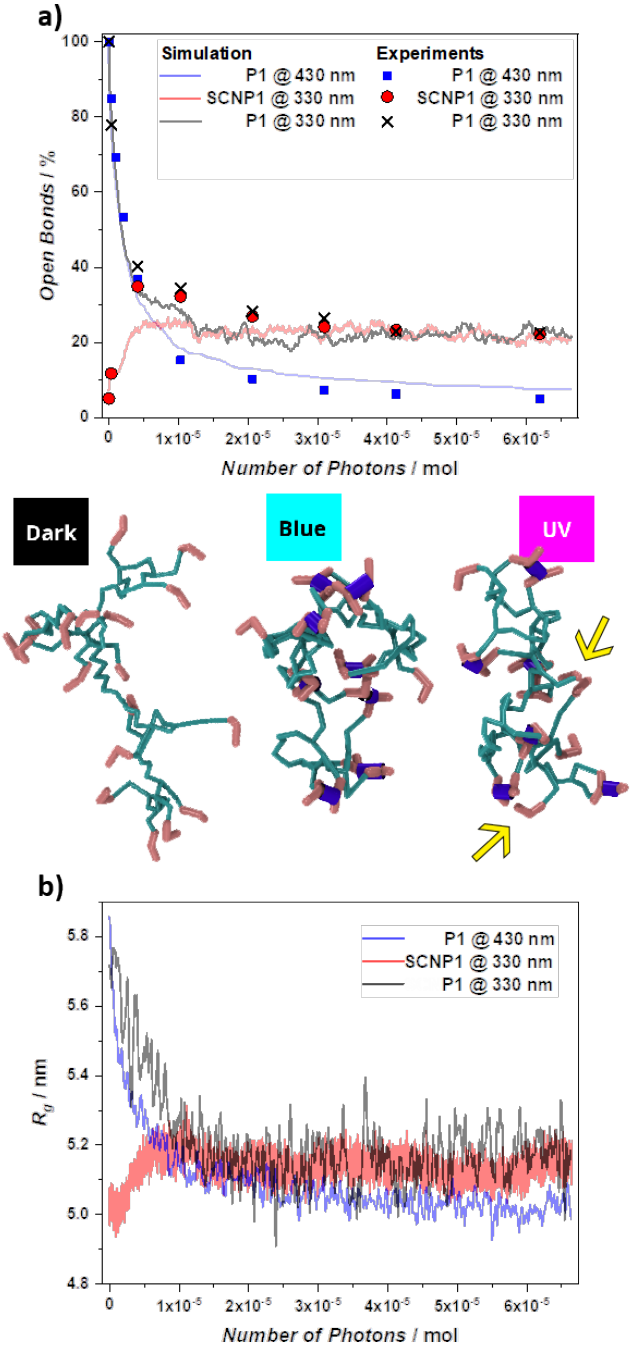


Figure 5.19.: The ratio of open bonds and the radius of gyration during folding **P1** into **SCNP1**.

For the 3D printing of a polymer network, the monomers are represented as individual spheres when the polymerisation reaction scheme is gauged by distinguished monomer types and the corresponding reaction rates. Through the printing process, a voxel by voxel printing protocal activates monomers and starts polymerisation, followed by propagation, crosslinking, and termination reactions, before the monomers are incorporated into the printed polymer networks with defined structural and mechanical properties. From the coarse-grained model, the kinetics in the polymerisation reaction with respect to the monomer conversion degree, as well as the optimisation of monomer conversion in relation to the printing control parameters as laser power and exposure time, reproduce well the measurements in experiments, establishing the quantitative structure-property relationship.

Based on the topologies assembled in the CGMD model, the predictions of simulated deformation yield mechanical benchmarks of the 3D printed network in bulk phase as well as with a 3D template. We illustrated the mechanical strength the printed polymer network with respect of the printing laser power and printing velocity with saturating behaviours and propose optimised printing laser powers between 35mW and 40mW for acrylate-based monomers, reaching the monomer conversion degree of around 60% to 70%.

In the folding of a single-chain nanoparticle, the wavelength-gated reversible photo-crosslinking between styrylpyrene groups simulated as stochastic bond recombinations captures the folding process from the polymer coil into a photostationary state. Experiment and simulation data in both number of open bonds and molecular sizes agree well in the time-series, where an intermittent state transition from the compact SCNP to the corresponding photostationary state is consistently recovered. Simulation data reveals a incrementally constrained conformational compaction from the polymer coil to the SCNP, whereas the photocycloreversion, instead of breaking the crosslinks and completely consumes the dimers as was the case for hydroxy-styrylpyrene, nevertheless folds the polymer into a stationary state as the styrylpyrene group are kept in close proximity by the backbone. MD trajectories reveals the constant rearrangement in the polymer conformation and bond recombination between open and close states for individual bonds in the dynamic balance during the transition from the compact state to the photostationary state as the competition in the diffusion, bond formation and bond breaking induces a temporal overly unfolding beyond the photostationary unit before the equilibrium is gradually recovered.

In both applications, the CGMD simulation provides highly detailed information throughout the dynamic assembly of the polymer network and the SCNP and bridges the gap between atomistic modelling and the statistical census available only in the macroscopic scale in the form of order parameters. In resemblance to the protein folding, the most interesting 3D assembly of features and functional components for polymers happen between the individual atoms level and the macromolecule level when the collective movements facilitates the unique structure-property definitions.

6. Summary

Throughout the several months' writing of this thesis, the coronavirus outbreak has fundamentally changed the way of life for the entire population on this planet. Facing a challenge of such scale, countless efforts have been devoted in the analysis of the RNA virus as every day updates of the information on the sequence, structure, functionality and its interaction mechanism with human receptor proteins is revealed little by little. On the positive side, the development of vaccine has already covered a number of milestones. As of today the large scale testing has already been deployed and proved quite promising. However to confront the RNA virus which is typically subjected to high mutation rates, some of the current vaccines based on inactivated virus may not be as effective as expected in the long run. To comprehensively stop the virus from spreading, the targeted molecular design based on a good understanding of the viral functionality is indispensible, where fast molecular simulations serve as the corner stone.

As of its current state, the applicable range of molecular dynamics is limited to microseconds in the timescale and nanometers in the length-scale, placing the method in the middle as a bridge between the *ab initio* calculations and the modelling in the coarse-grained and continuum level, illustrated in Figure 6.1. In specific cases, e.g. with highly optimised software package running on dedicated hardware, and purpose-designed force-field parametrisation in the field of protein folding simulations, molecular dynamics extends the applicable block shown in Figure 6.1 in the timescale to around 1ms for individual simulations, covering a relatively small range in the multi-scale diagram and hardly overlaps with experiment methods.

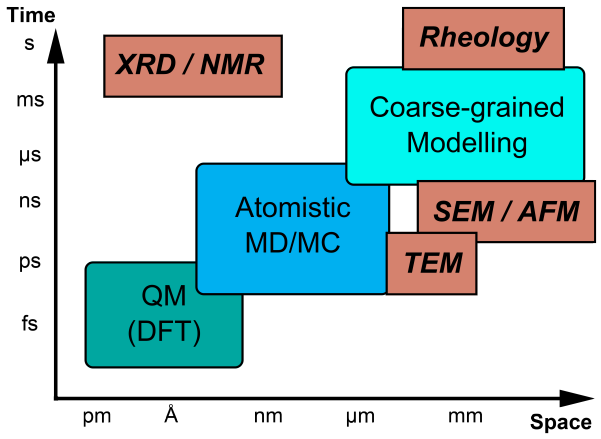


Figure 6.1.: A schematic representation of simulation and experiment methods in a multi-scale hierarchy.

In the thesis, the acceleration of molecular dynamics is discussed in 3 separate directions that extends both in the timescale and length-scale coverage of the molecular dynamics

applicability. From the method development perspective, the modification in the interaction potential, elaborated in Chapter 2, is inspired by the enhanced sampling methods as the accelerated molecular dynamics (aMD) etc. The accleration is achieved from both the adaptive sampling which switches the potential modification on and off in preferential spots prevents kinetic trapping which leads to an accelerated collective movement, and the conservation of the macroscopic order parameters as the overall energy and grain orientation in the grain boundary movement problem. The system-specific parametrisation for the acceleration algorithm is optimised iteratively with an artificial intelligence enhanced parameter space sampling, where the top performing parameter setups achieves an acceleration speed of approximately 200% of the original molecular dynamics method.

From the applications' perspective, respective routes of accelerated simulation and analysis are taken in seeking a faster experiment-theory-modelling-validation loop as well as the establishment of the quantitative structure-property relationship. In separate cases, fast molecular dynamics simulations serve for the better understanding of molecular functionalities and the molecular design in the respective polymer systems.

For the precise modelling of guest molecule diffusion in the nanoporous host, in Chapter 3, the all-atom simulations of ionic liquids drifting inside the metal-organic framework is compiled into an automated workflow. By the input of molecular ion types for the ionic liquids and the description of the metal-organic framework, the parametrisation, the preparation, diffusion simulation and the post-processing is accomplished without additional interference. The one-click solution from electrolyte and separator selection to conductivity estimations serves as a digital counterpart to the experiment in the lab, when a series of unqiue mechanisms in the ionic liquids conduction in various frameworks are identitied, validated, and even predicted before the experiment measurements become available. In the nanoporous host of the MOF, the guest IL molecules, due to their pronounced size and delocalised charge distribution, experiences a transition from the homogenoues flow to immobilisation and formation of a bunching layer subjected to IL density fluctuations in HKUST-1. The bunching effect is also observable for IL in UiO family MOFs with variable critical IL densities and a 2-step transition due to unique MOF geometries. While in the case of Li-ion doped IL, a bi-modal swift diffusion of the small Li-ion through the cavities through the amorphous bunching layer of IL and the diffusion through alternative pores, as well as a distinct Grotthuss-like hopping mechanism greatly reduces the bunching effect of the molecular IL. Proved successful, the ionic liquid in metal-organic frameworks simulation workflow enables high-throughput materials screening and envisions simulation guided iontronics design for both the electrolyte and separator.

For the modelling of coarse-grained polymers, in Chapter 4, the molecular dynamics simulations are performed with different coarse-grained levels of detail in the polymerisation and assembly of a polymer network, along with the reversible folding of a single polymer chain into a compact single-chain nanoparticle or the photostationary state. In both applications, the bond recombination is implemented as a local stochastic process when the diffusion constantly competes with the dynamic formation and breaking in the geometrical constaints. In separate cases, the modelling of 3D polymer printing follows a top-down reaction scheme where individual reaction rates are derived *a priori*, the folding of the single-chain nanoparticle takes a bottom-up route when the interplay between the bond formation rate, bond breaking rate, and the diffusion rate is fitted to data provided by separate experimental protocols. The fitted parameters in single-chain nanoparticle folding is transferred for the prediction of a separate model system before being validated. The acceleration provided by the coarse-grained model allows for a much faster individual simulations when a larger ensemble of simulations have to be performed before the statistical properties are compared with the experiments. The resulting molecular dynamics structures offer printed polymer networks comparable to the life-sized printing voxels in the order of

1E6 cubic nanometers, while the polyer folding trajectories are consistent with real world timeseries with unprecedented intermittent partial unfolding dynamics through the 45min photocycloreversion process, which extensively expands the length-scale and timescale of molecular dynamics simulations.

Acknowledgements

It is a great pleasure to express my gratitude to my supervisors Prof. Dr. Wolfgang Wenzel for his friendly and deeply inspiring discussions and guidance throughout my PhD period. He has had provided countless valuable ideas not only on how the scientific project at hand could be approached as well as precious advices in how to pursue a scientific career. Also I would like to express many thanks to Prof. Dr. Jörg Schmalian for his inspiring inputs especially regarding the linear response theory.

I owe deep gratitudes towards the dear faculty and staff in KIT for their support, my dear colleagues and friends in the research group for the professional discussions along with the happy times and cherished memories. Additionally, I would also thank deeply my collaborators mostly from the experiment side as chemists who granted unique perspectives and approaches outside of the square of a physist's.

Most sincere appreciation resides to my wife, Yuan, without whose support my research period would be unimaginable. Great thanks should also be attributed to my cats for the happy accompanying.

Appendix

A. UFF4MOF parametrisation for HKUST-1

UFF atom types	Masses	ϵ	σ
O_2	15.9994	0.060000	3.118146
Cu4+2	63.5460	0.005000	3.113691
C_R	12.0107	0.105000	3.430851
H_	1.0079	0.044000	2.571134

Bonds	K_r	r_0
Cu4+2-O_2	221.999373	2.029549
O_2-C_R	714.557442	1.269010
Cu4+2-Cu4+2	70.296579	3.032715
C_R-C_R	462.655054	1.379256
C_R-C_R	391.669513	1.458000
C_R-H_	357.440381	1.081418

Angles	C_θ	B_θ	n_θ
Cu4+2-O_2-C_R	59.956996	-1	3
O_2-Cu4+2-O_2	54.106605	1	4
O_2-Cu4+2-Cu4+2	17.208069	1	4
C_R-C_R-C_R	101.626578	-1	3
C_R-C_R-C_R	117.010657	-1	3
O_2-C_R-O_2	256.805977	-1	3
C_R-C_R-O_2	148.862838	-1	3
C_R-C_R-H_	76.817955	-1	3

Dihedrals	K_ϕ	d	n
Cu4+2-O_2-C_R-C_R	2.245703	-1	2
C_R-C_R-C_R-C_R	1.497136	-1	2
C_R-C_R-C_R-O_2	0.555556	-1	2

Impropers	K_ω	C_0	C_1	C_2
C_R-C_R-C_R-C_R	2.000000	1.000000	-1.000000	0.000000

Table A.1.: UFF4MOF parametrisation of HKUST-1.

B. UFF4MOF parametrisation for UiO-66, 67, and 68

UFF atom types	Masses	ϵ	σ
H_	1.0079	0.044000	2.571134
C_R	12.0107	0.105000	3.430851
O_2	15.9994	0.060000	3.118146
O_3_f	15.9994	0.060000	3.118146
Zr8f4	91.2440	0.069000	2.783168

Bonds	K_r	r_0
H_-C_R	357.440381	1.081418
C_R-C_R	462.655054	1.379256
O_2-C_R	714.557442	1.269010
O_2-Zr8f4	199.959824	2.410484
H_-O_3_f	602.435128	0.966432
C_R-C_R	391.669513	1.458000

Angles	C_θ	B_θ	n_θ
C_R-C_R-H_	114.578032	-1	3
C_R-C_R-C_R	222.595017	-1	3
C_R-O_2-Zr8f4	158.766650	-1	3
C_R-C_R-O_2	275.428492	-1	3
O_2-C_R-O_2	413.556852	-1	3
C_R-C_R-C_R #2	204.366560	-1	3

Angles#2	K_θ	C_0	C_1	C_2
Zr8f4-O_3_f-H_	78.498356	0.343737	0.374972	0.281246
Zr8f4-O_3_f-Zr8f4	195.233589	0.343737	0.374972	0.281246
O_3_f-Zr8f4-O_3_f	76.804805	0.343737	0.374972	0.281246

Dihedrals	K_ϕ	d	n
H_-C_R-C_R-C_R	1.497136	-1	2
Zr8f4-O_2-C_R-O_2	2.245703	-1	2
O_2-C_R-C_R-C_R	0.555556	-1	2

Impropers	K_ω	C_0	C_1	C_2
C_R-C_R-H_C-C_R	2.000000	1.000000	-1.000000	0.000000

Table B.2.: UFF4MOF parametrisation of UiO family MOFs.

C. OPLS-AA compatible parameters for Li-ion, [BMIM], and [NTf2]

Atom types	Masses	ϵ	σ
CR	12.0110	0.086	3.4
CT	12.0110	0.086	3.4
CW	12.0110	0.086	3.4
H1	1.0080	0.0157	2.471
H4	1.0080	0.015	2.511
H5	1.0080	0.015	2.511
HC	1.0080	0.0157	2.650
NA	14.0067	0.17	3.25
C	12.0107	0.066	3.5
F	18.9984	0.053	2.9
N	14.0067	0.17	3.25
O	15.9994	0.21	2.96
S	32.0650	0.25	3.55
Li	6.9410	0.01826677	2.12645

Bonds	K_r	r_0
Cu4+2-O_2	221.999373	2.029549
O_2-C_R	714.557442	1.269010
Cu4+2-Cu4+2	70.296579	3.032715
C_R-C_R	462.655054	1.379256
C_R-C_R	391.669513	1.458000
C_R-H_	357.440381	1.081418
CR-H5	367	1.080
NA-CR	477	1.343
CT-CT	310	1.526
CT-H1	340	1.090
CT-HC	340	1.090
NA-CT	280.11	1.472
CW-CW	409.89	1.343
CW-H4	367	1.080
NA-CW	427	1.3816
C-F	441.8	1.323
C-S	235.42	1.818
N-S	372.01	1.570
S-O	637.07	1.442

Dihedrals [NTf2]	K_1	K_2	K_3	K_4
F-C-S-N	0.0	0.0	0.64	0.0
F-C-S-O	0.0	0.0	0.70	0.0
S-N-S-C	15.66	-4.98	-1.52	0.0

Impropers	K_ω	d	n
CR-CW-NA-CT	2.00	-1	2
H4-CW-NA-CW	1.1	-1	2
H5-CR-NA-NA	1.1	-1	2

Table C.3.: OPLS-AA compatible parametrisation of Li, [BMIM], and [NTf2] I.

Angles	K_θ	θ_0
CR-NA-CT	49.95	126.30
CR-NA-CW	70.00	120.00
CT-CT-CT	40.00	109.50
CT-CT-H1	50.00	109.50
CT-CT-HC	50.00	109.50
NA-CT-CT	70.03	112.20
CW-NA-CT	49.95	125.70
CW-CW-H4	30.11	130.70
NA-CW-CW	119.98	107.10
H1-CT-H1	35.00	109.50
NA-CT-H1	54.97	109.50
NA-CW-H4	50.00	120.00
NA-CR-H5	50.00	120.00
HC-CT-HC	35.00	109.50
NA-CR-NA	70.00	120.00
C-S-N	97.51	100.20
C-S-O	103.97	102.60
F-C-F	93.33	107.10
S-C-F	82.93	111.80
O-S-N	94.29	113.60
O-S-O	115.80	118.50
S-N-S	80.19	125.60

Dihedrals [BMIM]	K_ϕ	n_ϕ	d_ϕ
CT-CT-CT-CT	0.18	3	0
CT-CT-CT-H1	0.16	3	0
CT-CT-CT-HC	0.16	3	0
CT-CT-CT-NA	0.25	3	0
CT-CT-NA-CR	-0.24	1	0
CT-CT-NA-CW	-0.18	1	0
CW-CW-NA-CR	12.00	2	180
CW-CW-NA-CT	2.00	2	180
HC-CT-CT-H1	0.16	3	0
H1-CT-NA-CR	0.16	3	0
H1-CT-NA-CW	0.24	3	0
NA-CW-CW-H4	1.50	2	180
H4-CW-NA-CR	2.00	2	180
H4-CW-NA-CT	1.50	2	180
H5-CR-NA-CT	1.50	2	180
H5-CR-NA-CW	1.50	2	180
HC-CT-CT-HC	0.15	3	0
NA-CT-CT-HC	0.16	3	0
NA-CR-NA-CT	2.00	2	180
NA-CR-NA-CW	12.00	2	180
NA-CW-CW-NA	12.00	2	180

Table C.4.: OPLS-AA compatible parametrisation of Li, [BMIM], and [NTf2] II.

D. DPD, FENE, and SRP parametrisation for SCNP folding

Mass species	Mass
1 base	1.001170
2 monolink towards styrylpyrene	1.150000
3(3') styrylpyrene ring compound	1.020000
4 styrylpyrene ring compound, crosslink site	1.000000
5 THF	0.721070

Pair interactions	DPD conservative A	DPD dissipative γ	DPD cut-off R_{cutoff}
1-1	36.0	4.5	1.0
1-2	25.0	4.5	1.0
1-3	60.0	4.5	1.0
1-4	60.0	4.5	1.0
1-5	15.0	4.5	1.0
2-2	25.0	4.5	1.0
2-3	25.0	4.5	1.0
2-4	25.0	4.5	1.0
2-5	15.0	4.5	1.0
3-3	60.0	4.5	1.0
3-4	60.0	4.5	1.0
3-5	15.0	4.5	1.0
4-4	60.0	4.5	2.0
4-5	15.0	4.5	1.0
5-5	32.0	4.5	1.0

FENE bonds	K	R_0	δ
1-1 backbone	24.0	1.5	0.6
1-2	5.0	2.5	0.8
2-3	5.0	2.5	0.8
3-4	40.0	1.5	0.6
4-4 crosslink	30.0	1.5	0.6

Pseudo atom	C	$R_{\text{cutoff-SRP}}$
SRP	8.0	0.25

Angles	harmonic K	θ_0
1-1-1 base	1.4	109.2
1-1-2	1.4	109.2
1-2-3	1.4	109.2
2-3-4	1.0	120
3-4-3	1.0	120
3-4-4 crosslink	2.0	90

Table D.5.: DPD, FENE, and SRP parameterization of the SCNP folding

E. Additional figure for reversible folding of P1' into SCNP1'

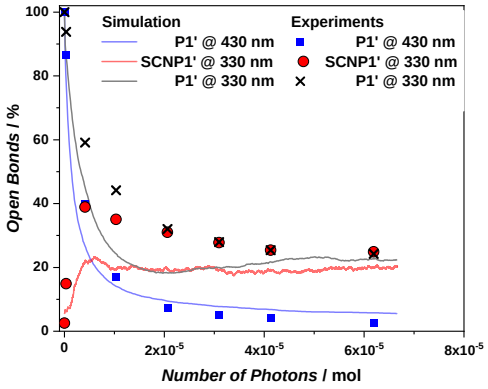


Figure E.1.: The ratio of open bonds during folding **P1'** into **SCNP1'**.

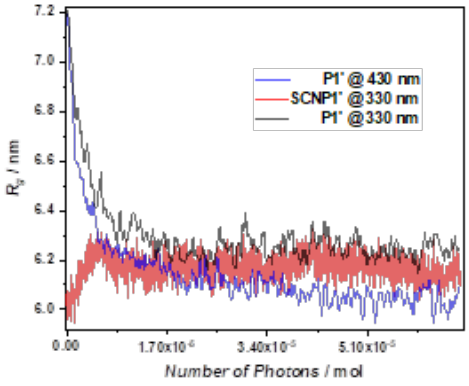


Figure E.2.: The ratio of radius of gyration during folding **P1'** into **SCNP1'**.

Bibliography

[1] Wei-jie Guan, Zheng-yi Ni, Yu Hu, Wen-hua Liang, Chun-quan Ou, Jian-xing He, Lei Liu, Hong Shan, Chun-liang Lei, David S.C. Hui, Bin Du, Lan-juan Li, Guang Zeng, Kwok-Yung Yuen, Ru-chong Chen, Chun-li Tang, Tao Wang, Ping-yan Chen, Jie Xiang, Shi-yue Li, Jin-lin Wang, Zi-jing Liang, Yi-xiang Peng, Li Wei, Yong Liu, Ya-hua Hu, Peng Peng, Jian-ming Wang, Ji-yang Liu, Zhong Chen, Gang Li, Zhi-jian Zheng, Shao-qin Qiu, Jie Luo, Chang-jiang Ye, Shao-yong Zhu, and Nan-shan Zhong. Clinical Characteristics of Coronavirus Disease 2019 in China. *New England Journal of Medicine*, 382(18):1708–1720, April 2020. Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMoa2002032.

[2] Qihui Wang, Yanfang Zhang, Lili Wu, Sheng Niu, Chunli Song, Zengyuan Zhang, Guangwen Lu, Chengpeng Qiao, Yu Hu, Kwok-Yung Yuen, Qisheng Wang, Huan Zhou, Jinghua Yan, and Jianxun Qi. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell*, 181(4):894–904.e9, May 2020.

[3] Linlin Zhang, Daizong Lin, Xinyuanyuan Sun, Ute Curth, Christian Drosten, Lucie Sauerhering, Stephan Becker, Katharina Rox, and Rolf Hilgenfeld. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science*, 368(6489):409–412, April 2020. Publisher: American Association for the Advancement of Science Section: Report.

[4] PDB101: Molecule of the Month: Coronavirus Proteases.

[5] Justin A. Lemkul and David R. Bevan. The Role of Molecular Simulations in the Development of Inhibitors of Amyloid β-Peptide Aggregation for the Treatment of Alzheimer's Disease. *ACS Chem Neurosci*, 3(11):845–856, August 2012.

[6] Fabio Finocchi, Giulia Galli, Michele Parrinello, and Carlo M. Bertoni. Microscopic struture of amorphous covalent alloys probed by ab initio molecular dynamics: SiC. *Phys. Rev. Lett.*, 68(20):3044–3047, May 1992. Publisher: American Physical Society.

[7] Jiuning Hu, Xiulin Ruan, and Yong P. Chen. Thermal Conductivity and Thermal Rectification in Graphene Nanoribbons: A Molecular Dynamics Study. *Nano Lett.*, 9(7):2730–2735, July 2009. Publisher: American Chemical Society.

[8] Shenshen Wang and Peter G. Wolynes. Communication: Effective temperature and glassy dynamics of active matter. *J. Chem. Phys.*, 135(5):051101, August 2011. Publisher: American Institute of Physics.

[9] Zhihong You, Daniel J. G. Pearce, Anupam Sengupta, and Luca Giomi. Geometry and Mechanics of Microdomains in Growing Bacterial Colonies. *Phys. Rev. X*, 8(3):031065, September 2018. Publisher: American Physical Society.

[10] D. E. Shaw Research. Molecular Dynamics Simulations Related to SARS-CoV-2, 2020.

[11] Dror Y. Kenett, Fred Morstatter, H. Eugene Stanley, and Huan Liu. Discovering Social Events through Online Attention. *PLOS ONE*, 9(7):e102001, July 2014. Publisher: Public Library of Science.

[12] An Zeng, Zhesi Shen, Jianlin Zhou, Jinshan Wu, Ying Fan, Yougui Wang, and H. Eugene Stanley. The science of science: From the perspective of complex systems. *Physics Reports*, 714-715:1–73, November 2017.

[13] Tu Le, V. Chandana Epa, Frank R. Burden, and David A. Winkler. Quantitative Structure–Property Relationship Modeling of Diverse Materials Properties. *Chem. Rev.*, 112(5):2889–2919, May 2012. Publisher: American Chemical Society.

[14] Felix Studt, Irek Sharafutdinov, Frank Abild-Pedersen, Christian F. Elkjær, Jens S. Hummelshøj, Søren Dahl, Ib Chorkendorff, and Jens K. Nørskov. Discovery of a Ni-Ga catalyst for carbon dioxide reduction to methanol. *Nature Chemistry*, 6(4):320–324, April 2014. Number: 4 Publisher: Nature Publishing Group.

[15] H. Poincaré. Mémoire sur les courbes définies par une équation différentielle (I). *Journal de Mathématiques Pures et Appliquées*, 7:375–422, 1881.

[16] Vito Volterra. Fluctuations in the Abundance of a Species considered Mathematically 1. *Nature*, 118(2972):558–560, October 1926. Number: 2972 Publisher: Nature Publishing Group.

[17] William L. Jorgensen, David S. Maxwell, and Julian Tirado-Rives. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.*, 118(45):11225–11236, November 1996. Publisher: American Chemical Society.

[18] Michael J. Robertson, Julian Tirado-Rives, and William L. Jorgensen. Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. *J. Chem. Theory Comput.*, 11(7):3499–3509, July 2015. Publisher: American Chemical Society.

[19] Leela S. Dodda, Israel Cabeza de Vaca, Julian Tirado-Rives, and William L. Jorgensen. LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Res.*, 45(W1):W331–W336, 2017.

[20] Junmei Wang, Romain M. Wolf, James W. Caldwell, Peter A. Kollman, and David A. Case. Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25(9):1157–1174, 2004. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.20035.

[21] Junmei Wang, Wei Wang, Peter A. Kollman, and David A. Case. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.*, 25(2):247–260, October 2006.

[22] A K Rappe, C J Casewit, K S Colwell, W A Goddard, and W M Skiff. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.*, 114(25):10024, 1992.

[23] Damien E. Coupry, Matthew A. Addicoat, and Thomas Heine. Extension of the Universal Force Field for Metal–Organic Frameworks. *J. Chem. Theory Comput.*, 12(10):5215–5225, October 2016. Publisher: American Chemical Society.

[24] Pascal Friederich, Artem Fediai, Simon Kaiser, Manuel Konrad, Nicole Jung, and Wolfgang Wenzel. Toward Design of Novel Materials for Organic Electronics. *Advanced Materials*, 31(26):1808256, 2019. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.201808256.

[25] Martin Oliver Steinhauser. Introduction to Molecular Dynamics Simulations: Applications in Hard and Soft Condensed Matter Physics. *Molecular Dynamics - Studies of Synthetic and Biological Macromolecules*, April 2012. Publisher: IntechOpen.

[26] Xipeng Wang, Simón Ramírez-Hinestrosa, Jure Dobnikar, and Daan Frenkel. The Lennard-Jones potential: when (not) to use it. *Phys. Chem. Chem. Phys.*, 22(19):10624–10633, May 2020. Publisher: The Royal Society of Chemistry.

[27] Douglas J. Cleaver, Christopher M. Care, Michael P. Allen, and Maureen P. Neal. Extension and generalization of the Gay-Berne potential. *Phys. Rev. E*, 54(1):559–567, July 1996. Publisher: American Physical Society.

[28] Raimo Hartmann, Praveen K. Singh, Philip Pearce, Rachel Mok, Boya Song, Francisco Díaz-Pascual, Jörn Dunkel, and Knut Drescher. Emergence of three-dimensional order and structure in growing biofilms. *Nature Physics*, 15(3):251–256, March 2019. Number: 3 Publisher: Nature Publishing Group.

[29] J. A. McCammon, B. M. Pettitt, and L. R. Scott. Ordinary differential equations of molecular dynamics. *Computers & Mathematics with Applications*, 28(10):319–326, November 1994.

[30] Steve Plimpton. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *Journal of Computational Physics*, 117(1):1–19, March 1995.

[31] Florian Häse, Loïc M. Roch, Christoph Kreisbeck, and Alán Aspuru-Guzik. Phoenics: A Bayesian Optimizer for Chemistry. *ACS Cent. Sci.*, 4(9):1134–1145, September 2018. Publisher: American Chemical Society.

[32] H. J. C. Berendsen, D. van der Spoel, and R. van Drunen. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, 91(1):43–56, September 1995.

[33] James C. Phillips, David J. Hardy, Julio D. C. Maia, John E. Stone, João V. Ribeiro, Rafael C. Bernardi, Ronak Buch, Giacomo Fiorin, Jérôme Hénin, Wei Jiang, Ryan McGreevy, Marcelo C. R. Melo, Brian K. Radak, Robert D. Skeel, Abhishek Singharoy, Yi Wang, Benoît Roux, Aleksei Aksimentiev, Zaida Luthey-Schulten, Laxmikant V. Kalé, Klaus Schulten, Christophe Chipot, and Emad Tajkhorshid. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.*, 153(4):044130, July 2020. Publisher: American Institute of Physics.

[34] David E. Shaw, Ron O. Dror, John K. Salmon, J. P. Grossman, Kenneth M. Mackenzie, Joseph A. Bank, Cliff Young, Martin M. Deneroff, Brannon Batson, Kevin J. Bowers, Edmond Chow, Michael P. Eastwood, Douglas J. Ierardi, John L. Klepeis, Jeffrey S. Kuskin, Richard H. Larson, Kresten Lindorff-Larsen, Paul Maragakis, Mark A. Moraes, Stefano Piana, Yibing Shan, and Brian Towles. Millisecond-scale molecular dynamics simulations on Anton. In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*, SC '09, pages 1–11, New York, NY, USA, November 2009. Association for Computing Machinery.

[35] P. P. Ewald. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Annalen der Physik*, 369(3):253–287, 1921. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19213690304.

[36] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.*, 98(12):10089–10092, June 1993. Publisher: American Institute of Physics.

[37] Donald Hamelberg, John Mongan, and J. Andrew McCammon. Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. *J. Chem. Phys.*, 120(24):11919–11929, June 2004. Publisher: American Institute of Physics.

[38] Yi Wang, Christopher B. Harrison, Klaus Schulten, and J. Andrew McCammon. Implementation of Accelerated Molecular Dynamics in NAMD. *Comput Sci Discov*, 4(1), 2011.

[39] Alessandro Barducci, Massimiliano Bonomi, and Michele Parrinello. Metadynamics. *WIREs Computational Molecular Science*, 1(5):826–843, 2011. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.31.

[40] Dario Gioia, Martina Bertazzo, Maurizio Recanatini, Matteo Masetti, and Andrea Cavalli. Dynamic Docking: A Paradigm Shift in Computational Drug Discovery. *Molecules*, 22(11):2029, November 2017. Number: 11 Publisher: Multidisciplinary Digital Publishing Institute.

[41] Levi C.T. Pierce, Romelia Salomon-Ferrer, Cesar Augusto F. de Oliveira, J. Andrew McCammon, and Ross C. Walker. Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics. *J Chem Theory Comput*, 8(9):2997–3002, September 2012.

[42] Marine E. Bozdaganyan, Philipp S. Orekhov, Alexey K. Shaytan, and Konstantin V. Shaitan. Comparative Computational Study of Interaction of C60-Fullerene and Tris-Malonyl-C60-Fullerene Isomers with Lipid Bilayer: Relation to Their Antioxidant Effect. *PLoS One*, 9(7), July 2014.

[43] David L. Donoho. For most large underdetermined systems of linear equations the minimal L1-norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics*, 59(6):797–829, 2006. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpa.20132.

[44] E.J. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, February 2006. Conference Name: IEEE Transactions on Information Theory.

[45] Frank Emmert-Streib, Zhen Yang, Han Feng, Shailesh Tripathi, and Matthias Dehmer. An Introductory Review of Deep Learning for Prediction Models With Big Data. *Front. Artif. Intell.*, 3, 2020. Publisher: Frontiers.

[46] D.S. Weile and E. Michielssen. Genetic algorithm optimization applied to electromagnetics: a review. *IEEE Transactions on Antennas and Propagation*, 45(3):343–353, March 1997. Conference Name: IEEE Transactions on Antennas and Propagation.

[47] Gregory Hornby, Al Globus, Derek Linden, and Jason Lohn. Automated Antenna Design with Evolutionary Algorithms. In *Space 2006*, AIAA SPACE Forum. American Institute of Aeronautics and Astronautics, September 2006.

[48] Nanomatch GmbH. SimStack – Computer-Aided Molecule Design, 2020.

[49] Jiayuan Miao, Ram S. Bhatta, Darrell H. Reneker, Mesfin Tsige, and Philip L. Taylor. Molecular dynamics simulations of relaxation in stretched PVDF nanofibers. *Polymer*, 56:482–489, January 2015.

[50] Wei-Jian Xu, Chun-Ting He, Cheng-Min Ji, Shao-Li Chen, Rui-Kang Huang, Rui-Biao Lin, Wei Xue, Jun-Hua Luo, Wei-Xiong Zhang, and Xiao-Ming Chen. Molecular Dynamics of Flexible Polar Cations in a Variable Confined Space: Toward Exceptional Two-Step Nonlinear Optical Switches. *Advanced Materials*, 28(28):5886–5890, 2016. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.201600895.

[51] Jerome Delhommelle, Peter T. Cummings, and Janka Petravic. Conductivity of molten sodium chloride in an arbitrarily weak dc electric field. *J. Chem. Phys.*, 123(11):114505, September 2005. Publisher: American Institute of Physics.

[52] Thaer M. Al-Jadir and Flor R. Siperstein. The influence of the pore size in Metal-Organic Frameworks in adsorption and separation of hydrogen sulphide: A molecular simulation study. *Microporous and Mesoporous Materials*, 271:160–168, November 2018.

[53] Michel Armand, Frank Endres, Douglas R. MacFarlane, Hiroyuki Ohno, and Bruno Scrosati. Ionic-liquid materials for the electrochemical challenges of the future. *Nature Materials*, 8(8):621–629, August 2009. Number: 8 Publisher: Nature Publishing Group.

[54] Douglas R. MacFarlane, Naoki Tachikawa, Maria Forsyth, Jennifer M. Pringle, Patrick C. Howlett, Gloria D. Elliott, James H. Davis, Masayoshi Watanabe, Patrice Simon, and C. Austen Angell. Energy applications of ionic liquids. *Energy Environ. Sci.*, 7(1):232–250, December 2013. Publisher: The Royal Society of Chemistry.

[55] Natalia V. Plechkova and Kenneth R. Seddon. Applications of ionic liquids in the chemical industry. *Chem. Soc. Rev.*, 37(1):123–150, December 2007. Publisher: The Royal Society of Chemistry.

[56] Hong-Cai Zhou, Jeffrey R. Long, and Omar M. Yaghi. Introduction to Metal–Organic Frameworks. *Chem. Rev.*, 112(2):673–674, February 2012. Publisher: American Chemical Society.

[57] Qi An, Isabelle D. Wessely, Yannick Matt, Zahid Hassan, Stefan Bräse, and Manuel Tsotsalas. Recycling and self-healing of dynamic covalent polymer networks with a precisely tuneable crosslinking degree. *Polym. Chem.*, 10(6):672–678, February 2019. Publisher: The Royal Society of Chemistry.

[58] Anthony K. Cheetham, Gérard Férey, and Thierry Loiseau. Open-Framework Inorganic Materials. *Angewandte Chemie International Edition*, 38(22):3268–3292, 1999. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291521-3773%2819991115%2938%3A22%3C3268%3A%3AAID-ANIE3268%3E3.0.CO%3B2-U.

[59] Stephen S.-Y. Chui, Samuel M.-F. Lo, Jonathan P. H. Charmant, A. Guy Orpen, and Ian D. Williams. A Chemically Functionalizable Nanoporous Material [Cu3(TMA)2(H2O)3]n. *Science*, 283(5405):1148–1150, February 1999. Publisher: American Association for the Advancement of Science Section: Report.

[60] Yongchul G. Chung, Jeffrey Camp, Maciej Haranczyk, Benjamin J. Sikora, Wojciech Bury, Vaiva Krungleviciute, Taner Yildirim, Omar K. Farha, David S. Sholl, and Randall Q. Snurr. Computation-Ready, Experimental Metal–Organic Frameworks: A Tool To Enable High-Throughput Screening of Nanoporous Crystals. *Chem. Mater.*, 26(21):6185–6192, November 2014. Publisher: American Chemical Society.

[61] Peyman Z. Moghadam, Aurelia Li, Seth B. Wiggin, Andi Tao, Andrew G. P. Maloney, Peter A. Wood, Suzanna C. Ward, and David Fairen-Jimenez. Development of a Cambridge Structural Database Subset: A Collection of Metal–Organic Frameworks for Past, Present, and Future. *Chem. Mater.*, 29(7):2618–2625, April 2017. Publisher: American Chemical Society.

[62] Yongchul G. Chung, Emmanuel Haldoupis, Benjamin J. Bucior, Maciej Haranczyk, Seulchan Lee, Hongda Zhang, Konstantinos D. Vogiatzis, Marija Milisavljevic, Sanliang Ling, Jeffrey S. Camp, Ben Slater, J. Ilja Siepmann, David S. Sholl, and

Randall Q. Snurr. Advances, Updates, and Analytics for the Computation-Ready, Experimental Metal–Organic Framework Database: CoRE MOF 2019. *J. Chem. Eng. Data*, 64(12):5985–5998, December 2019. Publisher: American Chemical Society.

[63] Kai Müller, Julian Helfferich, Fangli Zhao, Rupal Verma, Anemar Bruno Kanj, Velimir Meded, David Bléger, Wolfgang Wenzel, and Lars Heinke. Switching the Proton Conduction in Nanoporous, Crystalline Materials by Light. *Advanced Materials*, 30(8):1706551, 2018. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.201706551.

[64] Xiaojing Liu, Mariana Kozlowska, Timur Okkali, Danny Wagner, Tomohiro Higashino, Gerald Brenner-Weiß, Stefan M. Marschner, Zhihua Fu, Qiang Zhang, Hiroshi Imahori, Stefan Bräse, Wolfgang Wenzel, Christof Wöll, and Lars Heinke. Photoconductivity in Metal–Organic Framework (MOF) Thin Films. *Angewandte Chemie International Edition*, 58(28):9590–9595, 2019. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.201904475.

[65] Daniele Ongari, Peter G. Boyd, Senja Barthel, Matthew Witman, Maciej Haranczyk, and Berend Smit. Accurate Characterization of the Pore Volume in Microporous Crystalline Materials. *Langmuir*, 33(51):14529–14538, December 2017. Publisher: American Chemical Society.

[66] Anemar Bruno Kanj, Rupal Verma, Modan Liu, Julian Helfferich, Wolfgang Wenzel, and Lars Heinke. Bunching and Immobilization of Ionic Liquids in Nanoporous Metal–Organic Framework. *Nano Lett.*, 19(3):2114–2120, March 2019. Publisher: American Chemical Society.

[67] Barbara Kirchner, Friedrich Malberg, Dzmitry S. Firaha, and Oldamur Hollóczki. Ion pairing in ionic liquids. *J. Phys.: Condens. Matter*, 27(46):463002, October 2015. Publisher: IOP Publishing.

[68] Kazuyuki Fujie, Teppei Yamada, Ryuichi Ikeda, and Hiroshi Kitagawa. Introduction of an Ionic Liquid into the Micropores of a Metal–Organic Framework and Its Anomalous Phase Behavior. *Angewandte Chemie International Edition*, 53(42):11302–11305, 2014. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.201406011.

[69] O. Shekhah, H. Wang, T. Strunskus, P. Cyganik, D. Zacher, R. Fischer, and C. Wöll. Layer-by-Layer Growth of Oriented Metal Organic Polymers on a Functionalized Organic Surface. *Langmuir*, 23(14):7440–7442, July 2007. Publisher: American Chemical Society.

[70] Ziqi Wang, Rui Tan, Hongbin Wang, Luyi Yang, Jiangtao Hu, Haibiao Chen, and Feng Pan. A Metal–Organic-Framework-Based Electrolyte with Nanowetted Interfaces for High-Energy-Density Solid-State Lithium Battery. *Advanced Materials*, 30(2):1704436, 2018. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.201704436.

[71] Zhengjie Li, Yuanlong Xiao, Wenjuan Xue, Qingyuan Yang, and Chongli Zhong. Ionic Liquid/Metal–Organic Framework Composites for H2S Removal from Natural Gas: A Computational Exploration. *J. Phys. Chem. C*, 119(7):3674–3683, February 2015. Publisher: American Chemical Society.

[72] Daniele Ongari, Peter G. Boyd, Ozge Kadioglu, Amber K. Mace, Seda Keskin, and Berend Smit. Evaluating Charge Equilibration Methods To Generate Electrostatic Fields in Nanoporous Materials. *J. Chem. Theory Comput.*, 15(1):382–401, January 2019. Publisher: American Chemical Society.

[73] L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry*, 30(13):2157–2164, 2009. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.21224.

[74] James J. P. Stewart. MOPAC: A semiempirical molecular orbital program. *J Computer-Aided Mol Des*, 4(1):1–103, March 1990.

[75] A. Klamt and G. Schüürmann. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2*, (5):799–805, January 1993. Publisher: The Royal Society of Chemistry.

[76] Yuan H. Zhao, Michael H. Abraham, and Andreas M. Zissimos. Fast Calculation of van der Waals Volume as a Sum of Atomic and Bond Contributions and Its Application to Drug Compounds. *J. Org. Chem.*, 68(19):7368–7373, September 2003. Publisher: American Chemical Society.

[77] B. S. Kerner. Experimental Features of Self-Organization in Traffic Flow. *Phys. Rev. Lett.*, 81(17):3797–3800, October 1998. Publisher: American Physical Society.

[78] Boris S. Kerner, Hubert Rehborn, Ralf-Peter Schäfer, Sergey L. Klenov, Jochen Palmer, Stefan Lorkowski, and Nikolaus Witte. Traffic dynamics in empirical probe vehicle data studied with three-phase theory: Spatiotemporal reconstruction of traffic phases and generation of jam warning messages. *Physica A: Statistical Mechanics and its Applications*, 392(1):221–251, January 2013.

[79] Shin-ichi Tadaki, Macoto Kikuchi, Akihiro Nakayama, Akihiro Shibata, Yuki Sugiyama, and Satoshi Yukawa. Characterizing and distinguishing free and jammed traffic flows from the distribution and correlation of experimental speed data. *New J. Phys.*, 18(8):083022, August 2016. Publisher: IOP Publishing.

[80] Ariel Rosenman, Elena Markevich, Gregory Salitra, Doron Aurbach, Arnd Garsuch, and Frederick Francois Chesneau. Review on Li-Sulfur Battery Systems: an Integral Perspective. *Advanced Energy Materials*, 5(16):1500212, 2015. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aenm.201500212.

[81] Jake Christensen, Paul Albertus, Roel S. Sanchez-Carrera, Timm Lohmann, Boris Kozinsky, Ralf Liedtke, Jasim Ahmed, and Aleksandar Kojic. A Critical Review of Li/Air Batteries. *J. Electrochem. Soc.*, 159(2):R1, December 2011. Publisher: IOP Publishing.

[82] Guillermo A. Ludueña, Thomas D. Kühne, and Daniel Sebastiani. Mixed Grotthuss and Vehicle Transport Mechanism in Proton Conducting Polymers from Ab initio Molecular Dynamics Simulations. *Chem. Mater.*, 23(6):1424–1429, March 2011. Publisher: American Chemical Society.

[83] Jesse G. McDaniel and Arun Yethiraj. Grotthuss Transport of Iodide in EMIM/I3 Ionic Crystal. *J. Phys. Chem. B*, 122(1):250–257, January 2018. Publisher: American Chemical Society.

[84] Angelo Andriola, Kulwinder Singh, James Lewis, and Lei Yu. Conductivity, Viscosity, and Dissolution Enthalpy of LiNTF2 in Ionic Liquid BMINTF2. *J. Phys. Chem. B*, 114(36):11709–11714, September 2010. Publisher: American Chemical Society.

[85] Sebastian Kmiecik, Dominik Gront, Michal Kolinski, Lukasz Wieteska, Aleksandra Elzbieta Dawid, and Andrzej Kolinski. Coarse-Grained Protein Models and Their Applications. *Chem. Rev.*, 116(14):7898–7936, July 2016. Publisher: American Chemical Society.

[86] Djurre H. de Jong, Svetlana Baoukina, Helgi I. Ingólfsson, and Siewert J. Marrink. Martini straight: Boosting performance using a shorter cutoff and GPUs. *Computer Physics Communications*, 199:1–7, February 2016.

[87] Siewert J. Marrink, H. Jelger Risselada, Serge Yefimov, D. Peter Tieleman, and Alex H. de Vries. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B*, 111(27):7812–7824, July 2007. Publisher: American Chemical Society.

[88] Matías R. Machado, Exequiel E. Barrera, Florencia Klein, Martín Sóñora, Steffano Silva, and Sergio Pantano. The SIRAH 2.0 Force Field: Altius, Fortius, Citius. *J. Chem. Theory Comput.*, 15(4):2719–2733, April 2019. Publisher: American Chemical Society.

[89] Robert D. Groot and Patrick B. Warren. Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. *J. Chem. Phys.*, 107(11):4423–4435, September 1997. Publisher: American Institute of Physics.

[90] Luca Monticelli, Senthil K. Kandasamy, Xavier Periole, Ronald G. Larson, D. Peter Tieleman, and Siewert-Jan Marrink. The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.*, 4(5):819–834, May 2008. Publisher: American Chemical Society.

[91] J. T. Padding and W. J. Briels. Systematic coarse-graining of the dynamics of entangled polymer melts: the road from chemistry to rheology. *J. Phys.: Condens. Matter*, 23(23):233101, May 2011. Publisher: IOP Publishing.

[92] Jiang Wang, Simon Olsson, Christoph Wehmeyer, Adrià Pérez, Nicholas E. Charron, Gianni de Fabritiis, Frank Noé, and Cecilia Clementi. Machine Learning of Coarse-Grained Molecular Dynamics Force Fields. *ACS Cent. Sci.*, 5(5):755–767, May 2019. Publisher: American Chemical Society.

[93] Hendrik Frisch, Fabian R. Bloesser, and Christopher Barner-Kowollik. Controlling Chain Coupling and Single-Chain Ligation by Two Colours of Visible Light. *Angew. Chem. Int. Ed. Engl.*, 58(11):3604–3609, March 2019.

[94] Pascal Kiefer, Vincent Hahn, Martina Nardi, Liang Yang, Eva Blasco, Christopher Barner-Kowollik, and Martin Wegener. Sensitive Photoresists for Rapid Multiphoton 3D Laser Micro- and Nanoprinting. *Advanced Optical Materials*, n/a(n/a):20200895, August 2020. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adom.202000895.

[95] Joachim Fischer and Martin Wegener. Three-dimensional optical laser lithography beyond the diffraction limit. *Laser & Photonics Reviews*, 7(1):22–44, 2013. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/lpor.201100046.

[96] E. Johan Foster, Erik B. Berda, and E. W. Meijer. Metastable Supramolecular Polymer Nanoparticles via Intramolecular Collapse of Single Polymer Chains. *J. Am. Chem. Soc.*, 131(20):6964–6966, May 2009. Publisher: American Chemical Society.

[97] Michael A. R. Meier and Christopher Barner-Kowollik. A New Class of Materials: Sequence-Defined Macromolecules and Their Emerging Applications. *Advanced Materials*, 31(26):1806027, 2019. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.201806027.

[98] Denis Danilov, Christopher Barner-Kowollik, and Wolfgang Wenzel. Modelling of reversible single chain polymer self-assembly: from the polymer towards the protein limit. *Chem. Commun.*, 51(27):6002–6005, March 2015. Publisher: The Royal Society of Chemistry.

[99] Nicolai D. Knöfel, Hannah Rothfuss, Pavleta Tzvetkova, Bragavie Kulendran, Christopher Barner-Kowollik, and Peter W. Roesky. Heterobimetallic Eu(III)/Pt(II) single-chain nanoparticles: a path to enlighten catalytic reactions. *Chem. Sci.*, August 2020. Publisher: The Royal Society of Chemistry.

[100] David E. Marschner, Hendrik Frisch, Janin T. Offenloch, Bryan T. Tuten, C. Remzi Becer, Andreas Walther, Anja S. Goldmann, Pavleta Tzvetkova, and Christopher Barner-Kowollik. Visible Light [2 + 2] Cycloadditions for Reversible Polymer Ligation. *Macromolecules*, 51(10):3802–3807, May 2018. Publisher: American Chemical Society.

[101] Ralf Everaers, Kurt Kremer, and Gary S. Grest. Entanglement effects in model polymer networks. *Macromolecular Symposia*, 93(1):53–67, 1995. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/masy.19950930110.

[102] P. J. Hoogerbrugge and J. M. V. A. Koelman. Simulating Microscopic Hydrodynamic Phenomena with Dissipative Particle Dynamics. *EPL*, 19(3):155–160, June 1992. Publisher: IOP Publishing.

[103] J. Smiatek, M. P. Allen, and F. Schmid. Tunable-slip boundaries for coarse-grained simulations of fluid flow. *Eur. Phys. J. E*, 26(1):115–122, May 2008.

[104] Kurt Kremer and Gary S. Grest. Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *J. Chem. Phys.*, 92(8):5057–5086, April 1990. Publisher: American Institute of Physics.

[105] Timothy W. Sirk, Yelena R. Slizoberg, John K. Brennan, Martin Lisal, and Jan W. Andzelm. An enhanced entangled polymer model for dissipative particle dynamics. *J Chem Phys*, 136(13):134903, April 2012.

[106] Linas Jonušauskas, Tomas Baravykas, Dovilė Andrijec, Tomas Gadišauskas, and Vytautas Purlys. Stitchless support-free 3D printing of free-form micromechanical structures with feature size on-demand. *Scientific Reports*, 9(1):17533, November 2019. Number: 1 Publisher: Nature Publishing Group.

[107] Shi-Jie Zhang, Yan Li, Yang-Kai Wang, Li-Pu Liu, Hong-Da Wang, Yun-Feng Xiao, Hong Yang, and Qihuang Gong. Controlling Young's modulus of polymerized structures fabricated by direct laser writing. *Appl. Phys. A*, 118(2):437–441, February 2015.

[108] Zhen Bai, Jingsong Wei, Xin Liang, Kui Zhang, Tao Wei, and Rui Wang. High-speed laser writing of arbitrary patterns in polar coordinate system. *Review of Scientific Instruments*, 87(12):125118, December 2016. Publisher: American Institute of Physics.

[109] Angelo Accardo, Marie-Charline Blatché, Rémi Courson, Isabelle Loubinoux, Christophe Thibault, Laurent Malaquin, and Christophe Vieu. Multiphoton Direct Laser Writing and 3D Imaging of Polymeric Freestanding Architectures for Cell Colonization. *Small*, 13(27):1700621, 2017. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/smll.201700621.

[110] Eva Blasco, Jonathan Müller, Patrick Müller, Vanessa Trouillet, Markus Schön, Torsten Scherer, Christopher Barner-Kowollik, and Martin Wegener. Fabrication of Conductive 3D Gold-Containing Microstructures via Direct Laser Writing. *Advanced Materials*, 28(18):3592–3595, 2016. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.201506126.

[111] Catherine A. Leatherdale and Robert J. DeVoe. Two-photon microfabrication using two-component photoinitiation systems: effect of photosensitizer and acceptor

concentrations. In *Nonlinear Optical Transmission and Multiphoton Processes in Organics*, volume 5211, pages 112–123. International Society for Optics and Photonics, November 2003.

[112] A. Pikulin and N. Bityurin. Spatial resolution in polymerization of sample features at nanoscale. *Phys. Rev. B*, 75(19):195430, May 2007. Publisher: American Physical Society.

[113] Benjamin Harke, Paolo Bianchini, Fernando Brandi, and Alberto Diaspro. Photopolymerization Inhibition Dynamics for Sub-Diffraction Direct Laser Writing Lithography. *Chemphyschem*, 13(6):1429–1434, April 2012.

[114] Uzi Landman and W. D. Luedtke. Small is different: energetic, structural, thermal, and mechanical properties of passivated nanocluster assemblies. *Faraday Discuss.*, 125(0):1–22, November 2004. Publisher: The Royal Society of Chemistry.

[115] Nitin Uppal and Panos S. Shiakolas. Modeling of temperature-dependent diffusion and polymerization kinetics and their effects on two-photon polymerization dynamics. *JM3*, 7(4):043002, October 2008. Publisher: International Society for Optics and Photonics.

[116] Jonathan B. Mueller, Joachim Fischer, Frederik Mayer, Muamer Kadic, and Martin Wegener. Polymerization Kinetics in Three-Dimensional Direct Laser Writing. *Advanced Materials*, 26(38):6566–6571, 2014. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.201402366.

[117] Anna Y. Matveeva, Sergey V. Pyrlin, Marta M. D. Ramos, Helmut J. Böhm, and Ferrie W. J. van Hattum. Influence of waviness and curliness of fibres on mechanical properties of composites. *Computational Materials Science*, 87:1–11, May 2014.

[118] Jun Liu, Haixiao Wan, Huanhuan Zhou, Yancong Feng, Liqun Zhang, and Alexey V. Lyulin. Formation mechanism of bound rubber in elastomer nanocomposites: a molecular dynamics simulation study. *RSC Adv.*, 8(23):13008–13017, April 2018. Publisher: The Royal Society of Chemistry.

[119] Klaus Cicha, Zhiquan Li, Klaus Stadlmann, Aleksandr Ovsianikov, Ruth Markut-Kohl, Robert Liska, and Jürgen Stampfl. Evaluation of 3D structures fabricated with two-photon-photopolymerization by using FTIR spectroscopy. *Journal of Applied Physics*, 110(6):064911, September 2011. Publisher: American Institute of Physics.

[120] Hendrik Frisch, Jan P. Menzel, Fabian R. Bloesser, David E. Marschner, Kai Mundsinger, and Christopher Barner-Kowollik. Photochemistry in Confined Environments for Single-Chain Nanoparticle Design. *J. Am. Chem. Soc.*, 140(30):9551–9557, August 2018. Publisher: American Chemical Society.

[121] Yiliu Liu, Petri Turunen, Bas F. M. de Waal, Kerstin G. Blank, Alan E. Rowan, Anja R. A. Palmans, and E. W. Meijer. Catalytic single-chain polymeric nanoparticles at work: from ensemble towards single-particle kinetics. *Mol. Syst. Des. Eng.*, 3(4):609–618, August 2018. Publisher: The Royal Society of Chemistry.