



České vysoké učení technické v Praze  
Fakulta jaderná a fyzikálně inženýrská



## Diplomová práce

Počítačové simulace extrakčních roztoků

Computer simulations of extraction solutions

**Autor:** Lucie Celbová

**Vedoucí práce:** doc. Ing. Ladislav Kalvoda, CSc.

**Konzultanti:** Ing. Tomáš Koubský, Ph.D., Ing. Jaroslava Fojtíková

**Akademický rok:** 2018/2019



**Prohlášení:**

Prohlašuji, že jsem svou diplomovou práci vypracovala samostatně a použila jsem pouze podklady (literaturu, projekty, SW atd.) uvedené v příloženém seznamu.

Nemám závažný důvod proti použití tohoto školního díla ve smyslu § 60 Zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Praze dne

Lucie Celbová



## **Poděkování**

Srdečně děkuji doc. Ing. Ladislavovi Kalvodovi, CSc. za čas, který mi věnoval. Jsem za něj šťastná a vděčná.



*Název práce:*

**Počítačové simulace extrakčních roztoků**

*Autor:* Lucie Celbová

*Obor:* Inženýrství pevných látek

*Druh práce:* Diplomová práce

*Vedoucí práce:* doc. Ing. Ladislav Kalvoda, CSc., FJFI ČVUT v Praze

---

*Abstrakt:* Tato práce se zabývá především konformační analýzou dat získaných z počítačových simulací molekulární dynamiky. Analyzované molekuly představují zástupce extraktantů/*hold-back* agentů navržených pro chemické procesy přepracování jaderného vyhořelého paliva. Práce navrhuje nový přístup k pohledu na stabilitu těchto molekul v závislosti na jejich prostředí pomocí parametru  $L$ , definovaného jako počet změn v konformaci dané molekuly v čase. Dále je poukázáno na možnou souvislost zvýšené degenerace jednotlivých torzních úhlů a stability vazeb daných molekul.

*Klíčová slova:* Molekulární dynamika, Molekulární konformace

*Title:*

**Computer simulations of extraction solutions**

*Author:* Lucie Celbová

*Abstract:* This thesis primarily deals with conformational analysis of data obtained from molecular dynamics simulations. The analyzed molecules belong to families of extractants/*hold back* agents suggested for use in the field of reprocessing of spent nuclear fuel. New parameter  $L$  is defined as the number of changes in conformation of a molecule. Possible correlation between  $L$  and the stability of a molecule is discussed as well as relationship between higher degeneracy in values of specific torsion angles and the stability of given bonds of the molecule.

*Key words:* Molecular dynamics, Molecular conformations





# Obsah

<b>1</b>	<b>Úvod</b>	<b>1</b>
<b>2</b>	<b>Základy kapalinové extrakce</b>	<b>4</b>
<b>3</b>	<b>Molekulární mechanika</b>	<b>7</b>
3.1	Vazebné členy v předpisu silového pole . . . . .	7
3.2	Nevazebné členy v předpisu silového pole . . . . .	8
3.3	Minimalizace energie . . . . .	13
<b>4</b>	<b>Základy simulačních metod</b>	<b>16</b>
4.1	Periodické a neperiodické okrajové podmínky . . . . .	16
4.2	Výpočet nevazebných energií . . . . .	17
4.2.1	List sousedních atomů . . . . .	18
4.2.2	Ewaldova sumace . . . . .	18
4.2.3	<i>Particle-particle particle-mesh</i> metoda . . . . .	21
4.3	Molekulární dynamika (MD) . . . . .	23
4.3.1	NVT dynamika . . . . .	25
4.3.2	Potenciál středního pole a stochastická dynamika . . . . .	27
4.4	Monte Carlo simulace . . . . .	28
4.4.1	' <i>Biased</i> ' MC . . . . .	30
4.4.2	<i>J-walking</i> a Multikanonické MC . . . . .	30
4.4.3	Modul Amorphous Cell . . . . .	31
4.4.4	Modul Conformers . . . . .	32
4.5	Radiální distribuční funkce a tepelná kapacita . . . . .	33
<b>5</b>	<b>Nastavení počítačové simulace</b>	<b>34</b>
5.1	Volba silového pole . . . . .	35
<b>6</b>	<b>Konformační analýza</b>	<b>42</b>
6.1	Program pro konformační analýzu . . . . .	44
<b>7</b>	<b>Výsledky a jejich diskuze</b>	<b>51</b>
<b>8</b>	<b>Závěr</b>	<b>58</b>
<b>9</b>	<b>Přehled použité literatury</b>	<b>59</b>



# Kapitola 1

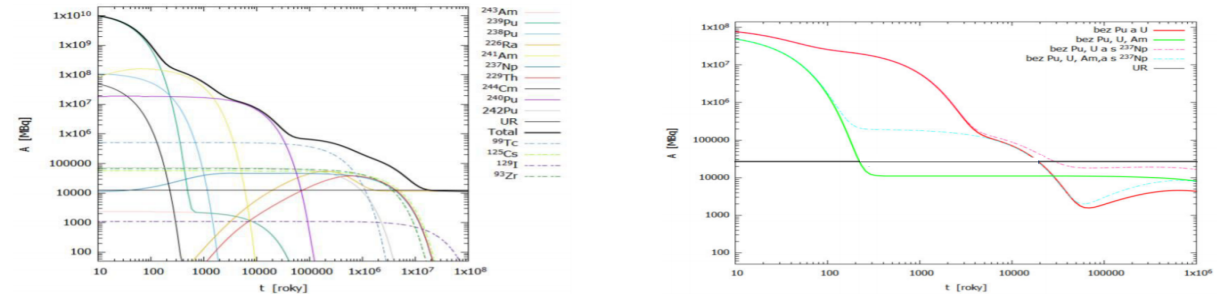
## Úvod

Jaderná energie v dnešní době představuje spolehlivý a zároveň šetrný způsob zisku elektrické energie. Zatímco uhlíková stopa jaderných reaktorů se pohybuje v hodnotách  $\sim 2,5-5,7$  g CO<sub>2</sub>/kWh, pro fosilní palivové cykly se jedná o hodnoty  $\sim 100-350$  g CO<sub>2</sub>/kWh a pro obnovitelné zdroje energie  $2,5-75$  g CO<sub>2</sub>/kWh [1]. Na úkor nízké uhlíkové stopy se ale objevuje jiný problém a tím je vyhořelé jaderné palivo, které je vysoce a především dlouhodobě radioaktivní.

Řešení soustavy diferenciálních rovnic popisujících postupný rozpad jader poskytl v roce 1910 H. Bateman [2]

$$A_n = N_0 \sum_{i=1}^n c_i e^{-\lambda_i t}, \quad c_m = \frac{\prod_{i=1}^n \lambda_i}{\prod_{i=1, i \neq m}^n (\lambda_i - \lambda_m)}, \quad (1.1)$$

kde  $A_n$  je aktivita  $n$ -tého produktu rozpadové řady,  $N_0$  je počet původně přítomných jader a  $\lambda_i$  je rozpadová konstanta  $i$ -tého jádra. Řešení je pro hlavní produkty obsažené v jedné tuně vyhořelého jaderného paliva z tlakovodního reaktoru při využití paliva 33 GWd/t (gigawatt-den/tuna) zobrazeno v levé části obrázku 1.1.



Obrázek 1.1: V levé části jsou zachyceny časové průběhy aktivity hlavních složek jedné tuny vyhořelého jaderného paliva s 33 GWd/t z tlakovodního reaktoru. UR značí aktivitu přírodního uranu. Hmotnosti jednotlivých složek byly převzaty z [3]. V pravé části je červeně zobrazena celková aktivita bez izotopů Pu a U a štěpných produktů (LLFP). Zeleně celková aktivita bez izotopů Pu, U, LLFP a aktinoidů (izotopy Am, Np a Cm). Černě je zobrazena aktivita přírodní uranové rudy.

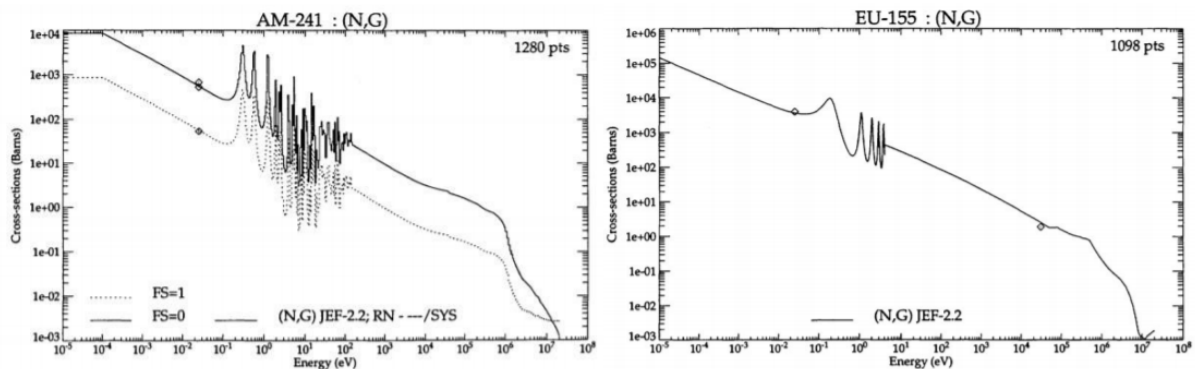
Průběhy aktivity vyhořelého paliva závisejí na jeho původním složení, typu reaktoru a jeho výkonu. Časová měřítka jsou ale obecně zachována a jak je vidět z obrázku 1.1, při odstranění izotopů uranu, plutonia a minoritních aktinoidů (Am, Cm, Np), poklesne čas potřebný pro redukci aktivity odpadu na/pod úroveň aktivity přírodní uranové rudy z  $\sim 10^7$  let do řádů stovek let. To je zásadní

informace vzhledem k tomu, že lidstvo nemá zkušenosti v horizontu  $\sim 10^7$  let a žádný z návrhů hlubinného úložiště jaderného odpadu nemůže v tomto měřítku adekvátně predikovat situaci.

Prvním krokem recyklace je odstranění izotopů uranu a plutonia. Tento proces se nazývá PUREX (Plutonium Uranium Redox EXtraction). Byl vyvíjen v rámci Manhattanského projektu ve skupině Glenna T. Seaborga a patentován byl v roce 1947 Herbertem H. Andersonem a Larnedem B. Aspreyem. Za motivaci tohoto výzkumu lze považovat především získání plutonia, které má relativně malou kritickou hmotnost a je vhodný k výrobě jaderných zbraní. Všechna pracoviště, která recyklují jaderné palivo jsou přísně kontrolována především díky obavám z ozbrojování. Dnes je z extrahovaného Pu a U vyráběno MOX palivo (Mixed Oxide), které může být opět využíváno jako palivo jaderných reaktorů.

V souvislosti s ostatními produkty jaderného palivového cyklu se nejčastěji hovoří o metodě P&T (Partition and Transmutation), tedy o extrakci jednotlivých složek z vyhořelého paliva a jejich následné transmutaci na stabilní izotopy. Mechanismem transmutace je v P&T metodě zmiňováno především ozáření neutronovým tokem. K tomuto ozáření se jako vhodné jeví ADS reaktory (Accelerator Driven System), které jako zdroj neutronů využívají spalační jaderné reakce vyvolané urychlenými protony (např. belgický reaktor MYRRHA - 600 MeV protony, průměrný proud 3,2 mA (ve vývoji)) v terci z těžkého prvku.

Zásadním problémem transmutace aktinoidů je fakt, že se chemicky chovají velmi podobně (podobná koordinanční čísla) jako lanthanoidy a při neselektivní kapalinové extrakci dochází k získání zástupců obou skupin. Některé z lanthanoidů mají vyšší účinné průřezy pro zachyt neutronu než aktinoidy. Příklad rozdílu je zachycen na obrázku 1.2.



Obrázek 1.2: Účinné průřezy pro zachyt neutronu v závislosti na jeho energii pro  $^{241}\text{Am}$  a  $^{155}\text{Eu}$ . G značí kvantum gama. Převzato z [5].

K účinné transmutaci aktinoidů neutronovým tokem je tak třeba proces, který bude selektivně extrahovat pouze zástupce této skupiny. Takové procesy se vyvíjejí a jednou z možností je, stejně jako v případě PUREX procesu, kapalinová extrakce z roztoku. Vzhledem k tomu, že molekuly navrhované jako extraktanty v těchto procesech jsou v kontaktu s radioaktivním prostředím, musí prokazovat odolnost především vůči produktům radiolýzy jejich rozpouštědla (extraktanty jsou rozpouštěny s nízkými koncentracemi a zanedbává se proto přímá interakce molekuly s produktem jaderného rozpadu).

Stabilita extraktantů se zkoumá nejen experimentálně, ale také počítačovými simulacemi a to především kvantově mechanickými výpočty metodou DFT (Density Functional Theory). Ty jsou schopné predikce nejnáchylnějších míst molekuly k rozpadu, tranzitivních stavů i produktů reakce.

Důležitou vstupní informací těchto výpočtů je konformace dané molekuly, která je závislá na prostředí a teplotě a je vhodnější pro její získání využít simulace molekulární dynamiky. Konformační analýza je obsahem této práce.

Využití neutronů není jediný zkoumaný způsob transmutace jader. Alternativou může být využití gigantické jaderné rezonance (GDR). Jedná se o fotojadernou reakci, při níž dochází ke kolektivní excitaci jádra a nejčastějším rozpadovým kanálem je emise neutronu. Nespornou výhodou GDR je, že energie potřebná k jejímu vyvolání je závislá na protonovém čísle a navíc ji prokazují všechna jádra. Eliminovat by se tak problém prvků s nízkým účinným průřezem pro záchyt neutronů (např.  $^{137}\text{Cs}$ ) a zároveň je prostřednictvím energie nalétávajících fotonů možná selektivní transmutace. Energie fotonů potřebné k vyvolání GDR se pohybují v řádech jednotek až desítek MeV a je nutné je získat například inverzním Comptonovým jevem, kdy se nízkoenergetické fotony  $\sim 1,5$  eV odrážejí od urychlených elektronů (řády stovek MeV). Urychlení elektronů v těchto schématech obvykle probíhá v brázdovém poli vyvolaném interakcí intenzivního laseru s plazmatem, což lze obecně považovat za budoucnost urychlování (nejen) elektronů. Tuto metodu zmiňuje i Gérard Mourou ve své nobelovské přednášce [6]. To, která z metod bude efektivnější, ukáže čas.

Práce je členěna do devíti kapitol včetně úvodu, závěru a seznamu použité literatury. Ve druhé kapitole jsou sepsány základy kapalinové extrakce z roztoku, ve třetí základy molekulární mechaniky z pohledu počítačové simulace, ve čtvrté pak základy simulačních metod molekulární dynamiky. V páté kapitole je uvedeno nastavení simulací provedených v této práci. V šesté kapitole je popsán postup provedené konformační analýzy, v sedmé jsou uvedeny a diskutovány její výsledky.

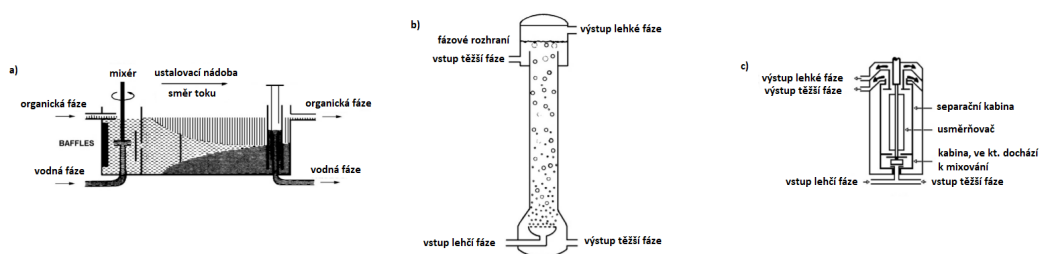
Všechny simulace obsažené v této práci byly provedeny pomocí softwarového balíku Materials Studio 8.0 (DASSAULT SYSTEMS BIOVIA) (MS) [7] a na výpočetním klastru Hyperion katedry Inženýrství pevných látek a katedry Matematiky [8]. Vlastní kódy byly napsány v Matlabu.

## Kapitola 2

### Základy kapalinové extrakce

I přesto, že se nejedná o jediný způsob chemického zpracování jaderného odpadu, představuje kapalinová extrakce z roztoku významného zástupce této skupiny procesů. Jedná se o postup, při kterém se smíchají dvě v rovnováze nemísitelné fáze (obvykle hydrofilní (vodná fáze) a hydrofobní (organická fáze)). Organická fáze obsahuje koordinační molekuly, které mají vysokou afinitu pro kov, který chceme extrahovat. Afinita musí být dostatečná k tomu, aby umožnila kovu překonat fázové rozhraní a aby mohlo dojít k vytvoření neutrálních organo-metalických komplexů rozpustných v organické fázi. Po dosažení rovnováhy celého systému se poté kov nachází v organické fázi, ve které se s ním dále operuje. Obecně tedy extrahovaná složka musí být rozpustná jak ve vodné, tak v organické fázi. Fáze, do které se extrahovalo se označuje jako extrakt (v této práci organická fáze), druhá fáze jako rafinát (v této práci vodná fáze).

Při procesu kapalinové extrakce z roztoku je vhodné vytvořit objekty organické fáze s co největším povrchem, obvykle velké množství malých kapiček. Dalšími požadavky je efektivní oddělení fází po smíchání a schopnost operace s velkými objemy. K dosažení těchto požadavků se využívá různých způsobů, některé jsou zobrazeny na obrázku 2.1.



Obrázek 2.1: a) Míchačka; b) Extrakční věž; c) Odstředivý kontaktor. Převzato z [9].

Ze tří zobrazených zařízení pro kapalinovou extrakci z roztoku si vzhledem k požadavkům nejlépe vede odstředivý kontaktor, který umožňuje efektivní promíchání fází a jejich nejefektivnější oddělení. Zároveň toto zařízení umožňuje nejkratší kontaktní dobu organické a vodné fáze. Z hlediska největší spolehlivosti je nejlepší volbou extrakční věž, protože neobsahuje pohyblivé komponenty. Všechny tři zmíněné typy kontaktorů jsou součástí např. závodu La Hague ve Francii. Vzhledem k tomu, že extrakční účinnost jednoho cyklu smíchání není stoprocentní, probíhá extrakce obvykle v několika po sobě následujících extrakčních kontaktorech. Přitom se může přidávat čerstvá organická fáze nebo se organická fáze po určitém počtu extrakčních kroků podrobí tzv. *stripping* fázi, ve které se organická fáze kontaktuje s novou vodnou fází, tentokrát za účelem extrakce kovu

z organické fáze. Tento krok se provádí i na konci celého extrakčního cyklu. Konečná vodná fáze obsahující extrahované kovy se dále podrobuje purifikaci (*washing stage*) a konečný produkt je získán vypařením vodného roztoku.

Jedním z hlavních parametrů využívaných v kapalinové extrakci je rozdělovací poměr  $D$ , definovaný jako poměr koncentrací všech složek obsahujících  $M$  v organické a vodné fázi [9]

$$D = \frac{[M]_{org}}{[M]_{aq}} \quad (2.1)$$

a je využíván pro kvantifikaci extrakčních schopností daných ligandů. Pokud je koncentrace v organické a vodné fázi definována pro jednu složku obsahující  $M$ , pak se hovoří o rozdělovací konstantě  $K_D$ . Dalším parametrem je separační faktor [9]

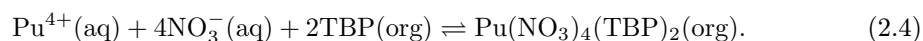
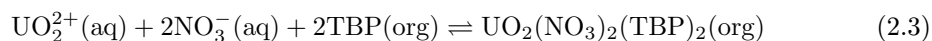
$$SF_{M1/M2} = \frac{D_{M1}}{D_{M2}}, \quad (2.2)$$

kde  $D_{M1}$  a  $D_{M2}$  jsou rozdělovací poměry kovů  $M1$  a  $M2$  v soustavě.

V radiační fyzice se také obvykle používá  $G$ -value, kterou lze považovat za míru stability zkoumané molekuly. Vyjadřuje počet entit vzniklých, pokud látka absorbovala energii rovnou 100 eV. Hodnota  $G=1$  tedy znamená, že po absorpci 100 eV vznikl (například) jeden volný radikál daného typu. V radiační chemii se obvykle objevují hodnoty  $G$  v jednotkách mol/J.

Navrhovaných extrakčních procesů je několik (už jen v rámci přepracování radioaktivního odpadu). Zde bude popsán nejrozšířenější PUREX proces. Ten má tři základní kroky [11]

- **Rozpuštění vyhořelého paliva v koncentrované kyselině dusičné a extrakce U(VI) a Pu(IV)** tributylfosfátem (TBP) rozpuštěným v koncentraci 30% ve směsi alifatických uhlodivků z 3-4 M  $\text{HNO}_3$  vodného roztoku. V tomto kroku dojde k extrakci až 99,9% U(VI) a Pu(IV) vytvořením tributylfosfátových komplexů



Zároveň se jedná o vysoce selektivní krok. Štěpné produkty a ostatní transurany zůstávají ve vodné fázi. Výjimkou je neptunium, kterého je extrahováno  $\sim 90\%$ .

- **Separace Pu, Np od U (*partitioning*)**. V tomto kroku se zredukuje Pu(IV) na Pu(III) pomocí N,N-diethyl-hydroxylaminu a využije se toho, že trivalentní plutonium interaguje s TBP jen minimálně a dochází k přechodu  $\text{Pu}(\text{NO}_3)_3$  zpět do vodné fáze. To platí i pro Np(VI).
- **Oddělení uranu z organické fáze (*stripping*)** do roztoku s nízkou koncentrací  $\text{HNO}_3$  při teplotě 45-50°C (nízká koncentrace  $\text{HNO}_3$  výrazně snižuje rozdělovací poměry všech kovů).

V dalších krocích probíhá několik kroků purifikace uranu. Separace Pu od Np probíhá opět selektivní redukcí. Obnovený uran a plutonium jsou konvertovány na MOX palivo.

Všechny procesy kapalinové extrakce pracují na stejném principu. Liší se především typem extraktantů a použitými rozpouštědly, které výrazně ovlivňují jak radiolytickou stabilitu ligandů, tak například rychlost kinetiky daných procesů. Jako příklad jmenujme proces nazvaný TRUEX (TransUranic EXtraction), který rozšiřuje PUREX proces přidáním koordinační molekuly CMPO. Ta umožňuje současnou ko-extrakci americia a curia (selektivita vůči lanthanoidům).

Dalším způsobem vytvoření vhodných podmínek pro selektivní chování extraktantů je využití *hold-back* molekul. Jedná se například o hydrofilní diglykolamidy, které jsou rozpuštěné ve vodné fázi a

selektivně ovlivňují chování určitých kovových iontů.

V souvislosti s extrakčními činidly sloužícími k separaci aktinoidů od lanthanoidů se také často zmiňuje CHON princip. Ten říká, že by tato činidla měla být složena pouze z uhlíku, vodíku, kyslíku a dusíku. Toto složení zaručuje možnost úplného spálení daných sloučenin a extraktanty, které tento princip splňují jsou upřednostňovány (například molekula CMPO tento princip nesplňuje, obsahuje fosfor).



## Kapitola 3

# Molekulární mechanika

Základem fyzikální simulace je funkční předpis popisující potenciální energii daného systému, tzv. silové pole (SP, z angl. *force-field*). Díky uplatnění Born-Oppenheimerovy aproximace, je taková funkce v případě molekulární mechaniky závislá pouze na polohách jader a pohyb elektronů se zcela zanedbává.

Silová pole se obecně mohou lišit jak svým funkčním předpisem, tak hodnotami parametrů, které se v nich vyskytují. Každý z nich může být vhodný k popisu různých systémů. Parametry se v dnešní době získávají především z kvantově mechanických výpočtů, které jsou sice schopny simulovat pouze systémy s relativně malým počtem částic, nicméně v některých případech jsou tyto výpočty schopny podat přesnější informace než je možné získat experimentálním měřením (například koeficienty pro odpudivé síly na základě překryvu atomových orbitalů). Jiné parametry můžeme získat porovnáním s experimenty (lepší např. pro zachycení interakcí mezi molekulami v roztoku).

Obecný předpis pro SP je  $E = E_{vazebná} + E_{nevazebná}$ , kde  $E_{vazebná}$  představuje energii vázaných atomů, která je závislá na vnitřních souřadnicích dané molekulární struktury, a  $E_{nevazebná}$  popisuje elektrostatické, van der Waalsovy a další interakce mezi nevázanými atomy. V této kapitole budou popsány některé z těchto členů.

### 3.1 Vazebné členy v předpisu silového pole

Zde uvedeme pouze významy vazebných členů, které se vyskytují v silovém poli COMPASS [12].

- **Bond-Stretch** člen je základním členem popisujícím změnu energie v závislosti na změně délky vazby. V nejjednodušším případě se aproximuje Hookovým zákonem  $E = \frac{k}{2}(l - l_0)^2$ , kde  $k$  je silová konstanta a  $l_0$  je referenční délka vazby. Jedná se o délku, kterou bude vazba mít, položíme-li všechny ostatní členy v předpisu pro potenciální energii rovny nule. Silové konstanty se pohybují v řádech stovek  $\text{kcal}\cdot\text{mol}^{-1}\text{Å}^{-2}$ . Například pro  $\text{sp}^3$  uhlíky je  $k \sim 300 \text{ kcal}\cdot\text{mol}^{-1}\text{Å}^{-2}$ , pro vazbu C-N je  $k \sim 700 \text{ kcal}\cdot\text{mol}^{-1}\text{Å}^{-2}$  [13]. Hookeův zákon dobře popisuje vazby v blízkosti základního stavu molekul, kdy dochází k dobré aproximaci Morseova potenciálu [13]. Pro lepší shodu ve větších vzdálenostech se započítávají vyšší členy Taylorova rozvoje Hookova zákona. V předpisu COMPASSu je použit rozvoj do čtvrtého řádu včetně.
- **Angle Bending** člen popisuje změnu energie v závislosti na úhlu mezi třemi atomy a v nejjednodušším případě ho lze opět aproximovat harmonickým potenciálem. Pro molekulu je jednodušší změnit úhel než délku vazby, což se odráží na hodnotách silových konstant. Např. pro tři  $\text{sp}^3$  hybridizované uhlíky je  $k \sim 0,01 \text{ kcal}\cdot\text{mol}^{-1}\text{Å}^{-2}$  [13]. Silové pole COMPASS aproximuje tento člen opět do čtvrtého řádu včetně.

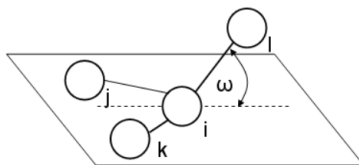
*Bond-stretch* a *angle-bending* členy se někdy nazývají také jako tvrdé stupně volnosti, protože je k jejich změně třeba dodat relativně velké množství energie.

- **Torzní člen** se snaží o popis vztahu mezi energií a torzními úhly v molekule. Torzní úhel je úhel mezi dvěma rovinami, které jsou tvořeny třemi atomy, přičemž dva jsou každé rovině stejné. Je to tedy úhel mezi čtyřmi atomy. Jinak řečeno se jedná o způsob měření rotace kolem vazby. Obvykle bývá vyjádřen jako Fourierův rozvoj kosinu. COMPASS není výjimkou, funkční předpis je následující

$$E(\phi) = \frac{1}{2} \sum_j [B_j (1 - d_j \cos(n_j \phi))], \quad (3.1)$$

kde  $\phi$  je torzní úhel,  $B_j$  je odhad hloubky potenciálu,  $n_j$  je periodičita funkce, tj. udává počet minim, kterých funkce nabude během rotace vazby o  $360^\circ$  a  $d_j$  se nazývá fázový faktor a udává úhel, který těmto minimům odpovídá.

- **Out-of-plane** člen má za úkol se vypořádat se skutečností, že použijeme-li k popisu potenciální nadplochy pouze předchozí členy, dojde ke generaci struktur, které ač by měly být planární, planárními nejsou. Způsobů, jak tento fakt napravit je několik. V případě silového pole COMPASS dochází k přeměření úhlu mezi rovinou tvořenou atomy  $j, i, k$  a atomem  $l$ , jak je znázorněno na obrázku 3.1, a následnému přesunutí atomu do roviny molekuly. Závislost energie na úhlu  $\omega$  potom popisuje harmonickým potenciálem  $E = \frac{k}{2} \omega^2$ .



Obrázek 3.1: Ilustrace k popisu ou-of-plane členu.  $i, j, k, l$  jsou atomy,  $\omega$  je úhel, který svítá rovina  $j, i, k$  a atomem  $l$  [7].

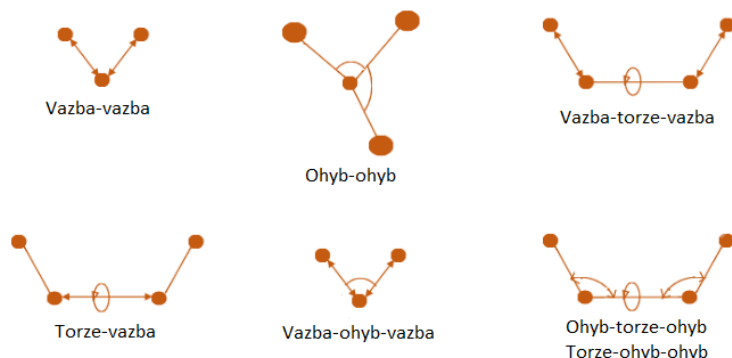
- **Cross terms** popisují změny jedné sady prostorových proměnných v závislosti na změně jiných. Jako příklad můžeme uvést změny délek dvou vazeb mezi třemi atomy, které jsou kompenzovány změnou vazebného úhlu mezi nimi. Takové členy jsou významné především pro určování vibračních spekter molekul. Všechny sedm kombinací, které jsou použity v silovém poli COMPASS, je zobrazeno na obrázku 3.2.

### 3.2 Nevazebné členy v předpisu silového pole

Nevazebné členy můžeme rozdělit do dvou skupin. Ty, které popisují elektrostatickou interakci a ty, které popisují van der Waalovu interakci. Zmíňme ještě, jak je to s vodíkovými můstky. Názor na ně není úplně jednotný. Jsou autoři, kteří vodíkové můstky řadí do van der Waalových sil (vdW) a jiní, kteří tyto dvě interakce považují za různý typ intermolekulárních sil. Ti pak klasifikují vdW interakci jako takovou, které se účastní alespoň jeden nepermanentní dipól. Jiní je řadí do dvou skupin, protože vodíkové vazby jsou silnější než vdW síly [18].

- **Elektrostatické interakce**

Jednou z možností, jak se vypořádat s interakcí např. dvou molekul, je taková, kdy ke každé



Obrázek 3.2: Ilustrace *cross-term* členů, které jsou obsaženy v COMPASSu.

přístupujeme jako k jediné entitě. Tyto entity potom reprezentujeme přiřazeným členem (nebo členy) *multipólového rozvoje*. Je-li přiřazen jeden člen, pak je to nejnižší nenulový moment. Tak třeba molekulu  $N_2$  bychom reprezentovali jako kvadrupól, náboj bychom rozložili do čtyř bodů. Může se ale stát, že některý z vyšších členů rozvoje má pro molekulu vyšší váhu a i proto by se mělo započítávat více členů. Tento přístup je efektivní zejména proto, že kdybychom měli třeba jen dipólové rozložení náboje, nemuseli bychom se zabývat členy typu  $1/r$ .

Multipólové momenty se získávají buď z QM výpočtů nebo z experimentu.

Tento přístup je ale možné použít pouze pro molekuly, které jsou od sebe vzdáleny více, než je součet vzdáleností od středu molekuly po nejvzdálenější bod jejího nábojového rozložení [13]. Populárnějším přístupem je přidělování parciálních nábojů na různá místa v molekule s tím, že elektrostatická interakce je potom počítána Coulombovým zákonem

$$V = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}, \quad (3.2)$$

kde  $N_A$  a  $N_B$  jsou počty parciálních nábojů na molekulách  $i$  a  $j$ ,  $r_{ij}$  je jejich vzdálenost,  $q_i$  a  $q_j$  jsou přiřazené náboje a  $\epsilon_0$  je permitivita vakua.

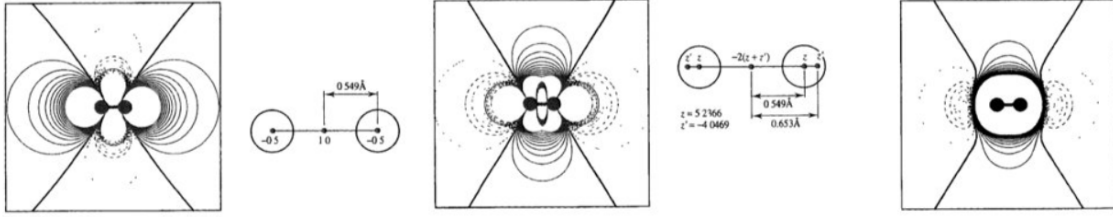
To, na jaká místa v molekule se parciální náboje přiřadí je podstatné. Na obrázku 3.3 jsou zobrazeny elektrostatické potenciály pro  $N_2$ , které se vygenerují různým zadefinováním těchto nábojů. V prvním případě jsou přiděleny tři parciální náboje. Dva na každý dusíkový atom a jeden do středu vazby. V druhém případě je přiřazeno celkem pět parciálních nábojů, což lépe reprodukuje QM výpočet zobrazený na témže obrázku.

Důležitou úlohou tedy představuje získání hodnot a poloh těchto parciálních nábojů. K tomu se využívají především QM výpočty. Parciální náboje nelze z vlnových funkcí systému získat jednoznačně [13].

Jednou z možností je odvodit náboje z *elektrostatického potenciálu molekuly* [13]

$$\phi(\vec{r}) = \phi_{nucl}(\vec{r}) + \phi_{el}(\vec{r}) = \sum_{A=1}^M \frac{Z_A}{|\vec{r} - \vec{R}_A|} - \int \frac{d\vec{r}' \rho(\vec{r}')}{|\vec{r} - \vec{r}'|}, \quad (3.3)$$

kde  $M$  je počet jader v systému,  $Z_A$  je protonové číslo jádra  $A$  s polohovým vektorem  $\vec{R}_A$  a  $\rho(\vec{r})$  je nábojová hustota.



Obrázek 3.3: Elektrostatické potenciály vygenerovány různým přiřazením parciálních nábojů. Poslední obrázek zachycuje výsledek QM výpočtu. Přerušovaně jsou záporné hodnoty. Převzato z [13].

Vybere se určitý počet bodů, ve kterém se tato funkce vypočítá a metodou nejmenších čtverců se odhadne soubor parciálních atomových nábojů, který nejlépe reprodukuje tento elektrostatický potenciál s podmínkou, že se součet těchto nábojů musí rovnat náboji na molekule.

Body k fitování by se měly volit tak, abychom zachytili místa ‚nejvýznamnější‘ pro intermolekulární interakce. V nejjednodušším případě se jedná o oblasti v přímé návaznosti na van der Waalsův poloměr daného atomu.

Výsledné hodnoty parciálních nábojů získané fitováním elektrostatického potenciálu jsou závislé na volbě bazických funkcí, které se využívají k odvození vlnové funkce [13] a na konformaci molekuly, která byla pro QM výpočty použita. Jedním z řešení tohoto problému je provedení výpočtů pro různé konformace a následné přidělování vážených hodnot nábojů na základě pravděpodobnosti vygenerování dané konformace z Boltzmannova rozdělení.

Jiným přístupem k výpočtu nábojů je takový, kdy se vychází pouze z obecných vlastností atomů jako je jejich elektronegativita a informace o vazbě, ve které se nacházejí. Výhodou takových metod je jejich rychlost.

Elektronegativita  $\chi_A$  atomu  $A$  je v tomto případě definována jako průměr součtu jeho ionizačního potenciálu  $I_A$  a elektronové afinity  $E_A$  [13]

$$\chi_A = \frac{1}{2} (I_A + E_A). \quad (3.4)$$

Ionizační potenciál i afinita závisí na okolnostech, kterým atom podléhá, jako je jeho orbitální stav. Například  $sp$  orbital má větší elektronegativitu než  $sp^3$ . Tato orbitální elektronegativita závisí na počtu elektronů v orbitalu i na nábojích v orbitalech ostatních.

Gasteiger a Marsili předpokládají vztah mezi orbitální elektronegativitou  $\chi_{\mu A}$  orbitalu  $\phi_\mu$  v atomu  $A$  a nábojem  $Q_A$  v parabolickém tvaru [27]

$$\chi_{\mu A} = aQ_A^2 + b_{\mu A}Q_A + c_\mu. \quad (3.5)$$

Použitím tabulovaných afinit a ionizačních potenciálů můžeme parametry  $a$ ,  $b$  a  $c$  vypočítat. Přiřazené orbitální afinity potom bereme jako referenční hodnoty pro simulaci ‚přemísťování‘ elektronů. Kdybychom přiřazovali elektrony tak, aby došlo k úplnému vyrovnání elektronegativit, dospěli bychom k výsledkům, které neodpovídají experimentům. Místo toho se využívá iteračního procesu, při němž je v každém kroku hodnota přesunutého náboje zmenšena [13]

$$Q^{(k)} = \frac{\chi_B^{(k)} - \chi_A^{(k)}}{\chi_A^+} \alpha^k, \quad (3.6)$$

kde  $k$  značí iteraci,  $Q$  je počet přesunutých elektronů,  $\chi_A$  a  $\chi_B$  jsou elektronegativity atomů  $A$  a  $B$  (v tomto případě  $\chi_B > \chi_A$ ),  $\chi_A^+$  je elektronegativita kationtu atomu  $A$  a  $\alpha$  je tlumící

faktor. Gasteiger s Marsilim volili  $\alpha = 1/2$ .

V prvním kroku iterace jsou atomům přiřazeny formální náboje. Následně jsou vypočítány elektronegativity podle 3.5, čímž se získá nový náboj.

Jiný přístup představuje metoda Rappého a Goddarda [28]. Ta oproti předchozí vnáší do problematiky vliv geometrie molekuly, která se může v průběhu simulace dynamiky systému měnit. Začínáme s Taylorovým rozvojem energie podle náboje pro izolovaný atom, přičemž se omezíme na členy do druhého řádu a na hodnoty energie v nábojích 0,-1 a 1

$$v_A(0) = v_{A0} \quad (3.7)$$

$$v_A(+1) = v_{A0} + q_A \left( \frac{\partial v}{\partial q} \right)_{A0} + \frac{1}{2} q_A^2 \left( \frac{\partial^2 v}{\partial q^2} \right)_{A0} \quad (3.8)$$

$$v_A(-1) = v_{A0} - q_A \left( \frac{\partial v}{\partial q} \right)_{A0} + \frac{1}{2} q_A^2 \left( \frac{\partial^2 v}{\partial q^2} \right)_{A0}. \quad (3.9)$$

$v_A(+1)$  představuje ionizační potenciál  $IP$  a  $-v_A(-1)$  elektronovou afinitu  $AF$ , tedy

$$\left( \frac{\partial v}{\partial q} \right)_{A0} = \frac{1}{2} (IP + AF) = \chi_A^0 \quad (3.10)$$

$$\left( \frac{\partial^2 v}{\partial q^2} \right)_{A0} = IP - AF. \quad (3.11)$$

Ionizační energie je energie, kterou musíme neutrálnímu atomu dodat, aby ztratil jeden elektron. Elektronová afinita je energie, která se uvolní, stane-li se z něj aniont. Rappé a Goddard považují rozdíl v těchto energiích (pro orbital s jedním elektronem) za důsledek Coulombovské interakce mezi elektrony a značí tento rozdíl jako  $J_{AA}^0 = \frac{\partial^2 v}{\partial q^2}$  [28].

Elektrostatická energie systému je nyní [28]

$$V(q_1 \dots q_N) = \sum_{A=1}^N (v_{A0} + \chi_A^0 q_A) + \frac{1}{2} \sum_{A=1}^N \sum_{B=1}^N q_A q_B J_{AB} \quad (3.12)$$

kde  $J_{AB}$  představuje Coulombovu energii mezi náboji  $q_A$  a  $q_B$ . Derivací 3.12 podle náboje  $q_A$  získáme elektrochemický potenciál atomu  $A$

$$\frac{\partial V}{\partial q_A} = \chi_A^0 + \sum_{A=1}^N q_B J_{AB} = \chi_A^0 + J_{AA}^0 q_A + \sum_{B=1, B \neq A}^N q_B J_{AB}. \quad (3.13)$$

Elektrony přechází z místa vysokého elektrochemického potenciálu (nižší elektronegativita) do míst s potenciálem nižším. V rovnováze by se měly elektrochemické potenciály atomů  $A$  a  $B$  vyrovnat. Spolu s podmínkou, že se součet nábojů musí rovnat náboji na molekule jsme získali řešitelný soubor rovnic. Pro výpočet náboje přitom potřebujeme pro každý atom tři vstupní informace. Elektronegativitu,  $J_{AA}$  a kovalentní poloměr. Zdůrazněme ještě, že díky členu  $q_A q_B$  je v tomto modelu zabudována závislost na geometrii molekuly.

V konsistentních silových polích (tedy i v COMPASSu) se atomové parciální náboje odvozují z *ab-initio* elektrostatických potenciálů.

- **van der Waalsovy interakce**

Přitažlivé síly jsou dlouhodosahové, odpudivé jsou důležité zejména na krátkých vzdálenostech.

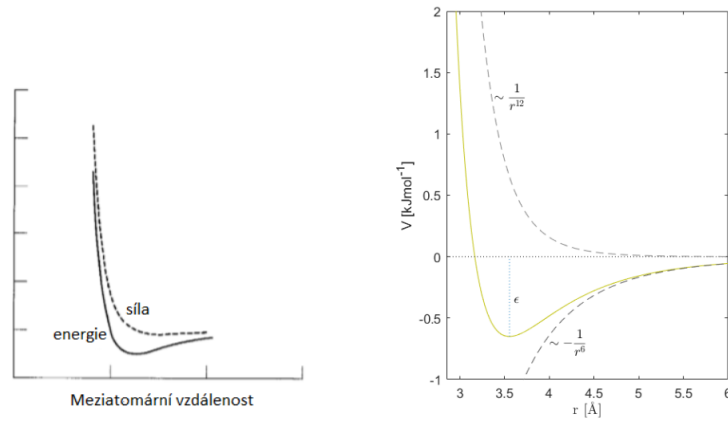
Přitažlivé síly jsou důsledkem tzv. disperzních sil, také značeny jako Londonovy disperzní síly. Jsou důsledkem fluktuací elektronových oblaků, tedy momentálních dipólů. Drude ukázal na modelu dvou molekul, kdy každá je charakterizována dvěma opačnými náboji v určité vzdálenosti, že je disperzní interakce  $\sim 1/r^6$ . Pokud bychom uvažovali i vyšší členy než dipól-dipólovou interakci, pak by platilo [13]

$$v(r) = \frac{C_6}{r^6} + \frac{C_8}{r^8} + \frac{C_{10}}{r^{10}} + \dots, \quad (3.14)$$

kde konstanty v čitatelích jsou záporné.

Při zmenšujících se vzdálenostech atomů se začínají projevovat odpuzivé síly. Odpovědnost za ně na kvantově mechanické úrovni můžeme připsat Pauliho vylučovacímu principu. Ve velmi blízkých vzdálenostech atomů, kdy dochází k poklesu stínícího efektu elektronových oblaků, ji můžeme připsat repulzi kladných jader. V této oblasti tedy bude energie se vzdáleností klesat jako  $1/r$ , ve větších vzdálenostech potom exponenciálně jako  $\exp(-2r/a_0)$ , kde  $a_0$  je Bohrov poloměr [13].

Tyto síly mohou být vypočítány prostřednictvím kvantové mechaniky. V případě předpisu silového pole nám jde především o věrné napodobení funkční závislosti síly na meziatomární vzdálenosti, která je zobrazena na obrázku 2.4 a o to, aby předpis nebyl složitý na výpočet.



Obrázek 3.4: Nalevo je zobrazena závislost energie a síly na meziatomární vzdálenosti [13]. Napravo Lennard-Jonesův potenciál 12-6 pro parametry  $r_m \sim 3,55 \text{ \AA}$  a  $\epsilon = 0,65 \text{ kJ}\cdot\text{mol}^{-1}$ , které jsou použity v modelu SPC, speciálně navrženém k popisu vody. Přerušované je vyznačen přitažlivý člen  $1/r^6$  a odpuzivý člen  $1/r^{12}$ .

Nejznámější formou popisující tyto vazby je Lennard-Jonesův potenciál 12-6, který je zobrazen na obrázku 2.4 a pro dva atomy ho popisuje vztah [13]

$$V = \epsilon \left[ \left( \frac{r_m}{r} \right)^{12} - 2 \left( \frac{r_m}{r} \right)^6 \right], \quad (3.15)$$

kde  $\epsilon$  je hloubka potenciální jámy,  $r_m$  je vzdálenost, kde jsou částice vázány nejsilněji, tj. potenciál nabývá hodnoty  $-\epsilon$  a  $r$  je vzdálenost částic. Místo parametru  $r_m$  se často používá také kolizní parametr  $\sigma$ . Platí  $r_m = 2^{1/6}\sigma$ .

V předpisu 3.15 je pro přitažlivé síly volena závislost  $1/r^6$ , se kterou jsme se setkali u Drudeho

molekulárního modelu. Závislost  $1/r^{12}$  ale nemá teoretické opodstatnění. Kvantová mechanika, jak bylo zmíněno, diktuje exponenciální tvar.

Výsledky jsou takové, že zahrnutí právě této mocniny vede k velmi dobrému popisu vzácných plynů, pro jiné systémy ale představuje příliš strmý nárůst odpuzivých sil. I přesto se Lennard-Jonesův potenciál 12-6 hojně využívá a to zejména proto, že je výpočetně velmi výhodný. Pro výpočet dvanácté mocniny stačí umocnit přitažlivý člen.

Lennard-Jonesův potenciál se nicméně používá v různých SP s různými mocninami. Například v COMPASSu je volen Lennard-Jonesův potenciál 9-6.

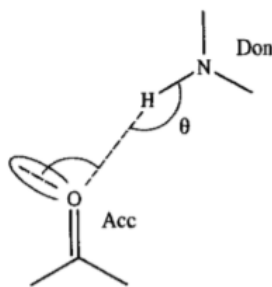
Popis vdW sil je zde ještě doplněn Morseho potenciálem, jenž má obecný tvar

$$V(r) = D_e \left(1 - e^{-a(r-r_e)}\right)^2, \quad (3.16)$$

kde  $a$  je empirický parametr udávající tloušťku potenciálové jámy,  $D_e$  je jeho hloubka a  $r_e$  je rovnovážná vzdálenost dvou atomů [29]. V COMPASSu je využit složitý předpis a čtenář se na něj může podívat v [7].

- **Vodíkové vazby**

Některé předpisy SP popisují vodíkové vazby Lennard-Jonesovým potenciálem 12-6. Jiné, jako například silové pole Dreiding, Lennard-Jonesovým potenciálem 12-10, přičemž se je snaží popsat i v závislosti na jejich geometrii prostřednictvím úhlu  $\theta$ , který je zobrazen na obrázku 3.5. Silové pole COMPASS v sobě člen explicitně popisující vodíkové vazby neobsahuje.



Obrázek 3.5: K úhlu  $\theta$ . Přerušovanou čarou je zvýrazněna vodíková vazba [13].

### 3.3 Minimalizace energie

Obečně přijímaného faktu, že se příroda snaží šetřit energií, se využívá i v počítačových simulacích. Jak jsme si ukázali, potenciální energie je závislá na velkém počtu parametrů. Zde si uvedeme některé minimalizační algoritmy, které mají za úkol hledat body minimální energie na potenciálové nadploše.

- **Metoda nejstrmějšího sestupu** je metodou využívající prvních derivací energie vzhledem ke svým proměnným. Nejdříve se zvolí směr, kterým se budeme po nadploše pohybovat. Směrový vektor  $\vec{s}_k$  se vypočítá jako

$$\vec{s}_k = -\frac{\vec{g}_k}{|\vec{g}_k|}, \quad (3.17)$$

kde  $\vec{g}_k$  značí gradient. V dalším kroku je nutné se rozhodnout, jak dlouhý krok podél tohoto směrového vektoru učinit. K tomu můžeme použít následující dvě metody

- **Line Search** se snaží najít minimum energetické plochy ve směru definovaném směrovým vektorem  $\vec{s}_k$ . Nejdříve hledá tři body v tomto směru tak, aby prostřední z nich měl minimální hodnotu. Iterativně potom zmenšuje vzdálenost dvou ‚vnějších‘ bodů a hledá tak v této oblasti minimum. Tento přístup může být výpočetně velmi náročný, protože může vyžadovat velký počet výpočtů potenciální energie. Jiným postupem může být, že nafitujeme tři nalezené body analytickou funkcí, například kvadratickou. Pro tuto funkci nalezneme minimum. To potom můžeme vyměnit za předchozí minimum a fit provést znovu.
- **Arbitrary Step Approach** spočívá v tom, že se jednoduše posuneme ve směru  $\vec{s}_k$  o určitou vzdálenost. Nové souřadnice po  $k$ -té iteraci tedy jsou

$$\vec{x}_{k+1} = \vec{x}_k + \lambda_k \vec{s}_k, \quad (3.18)$$

kde  $\lambda_k$  značí velikost kroku. Pokud první iterace vede ke snížení energie, délka kroku je zvýšena. Tento postup se opakuje tak dlouho, dokud dochází ke snižování energie. Pokud v dalším kroku dojde ke zvýšení energie, rozhodneme, že jsme prošli minimem, zjermníme krok a s větší opatrností projdeme oblast posledního a předposledního kroku.

V místě nalezeného minima má gradient kolmý směr ke směru předchozímu ( $\vec{s}_k \perp \vec{s}_{k+1}$ ) a v něm tedy probíhá další iterace. Díky této ortogonalitě po sobě jdoucích kroků má algoritmus problém v úzkých jamách a tedy i v blízkosti minim.

- **Metoda sdružených gradientů** se od předchozí metody liší přiřazováním směru jednotlivým krokům. V předchozí měli směr gradientů. V tomto případě se přiřazují tzv. sdružené směry (Polak-Ribiere) [13]

$$\vec{v}_k = -\vec{g}_k + \frac{(\vec{g}_k - \vec{g}_{k-1}) \cdot \vec{g}_k}{\vec{g}_{k-1} \cdot \vec{g}_{k-1}} \vec{v}_{k-1}, \quad (3.19)$$

kde  $\vec{g}_k$  je gradient v bodě  $\vec{x}_k$ , přičemž první krok se realizuje ve směru gradientu. Všechny směry a gradienty musí splňovat podmínky [13]

$$\vec{g}_i \cdot \vec{g}_j = 0 \quad \vec{v}_i \cdot V''_{ij} \cdot \vec{v}_j = 0 \quad \vec{g}_i \cdot \vec{v}_j = 0, \quad (3.20)$$

kde  $V''_{ij}$  je matice druhých derivací potenciálu, která zde slouží pro ověření podmínek.

V dalších metodách se využívá přímo druhých derivací. Nejjednodušší z nich je Newton-Raphsonova metoda, kdy je poloha minima  $\vec{x}^*$  [13]

$$\vec{x}^* = \vec{x}_k - V'(\vec{x}_k) (V'')^{-1}(\vec{x}_k). \quad (3.21)$$

Právě výpočet Hessiánu  $V''$ , který navíc *musíme zinvertovat*, je výpočetně náročný. Různé varianty Newtonovy-Raphsonovy metody se proto snaží jeho výpočet zkrátit.

- **Quasi-Newtonovy metody** budují inverzní Hessián postupně během iteračního procesu. Vygeneruje tedy  $k$  matic  $H_k$  tak, že  $\lim_{k \rightarrow \infty} H_k = (V'')^{-1}$ . Během každé iterace se vypočítají nové pozice jako

$$\vec{x}_{k+1} = \vec{x}_k - H_k \vec{g}_k, \quad (3.22)$$



což je přesný výraz pro kvadratickou funkci. V jiných případech se může využít *line search* podél vektoru  $\vec{x}_{k+1} - \vec{x}_k$  [13]. V této nové poloze je vypočítán nový Hessián podle předpisů, které závisí na blíže zvolené metodě a čtenář si může některé z nich prohlédnout v [13]. Důležité je, že všechny využívají pouze  $k$ -tou polohu, Hessián a polohu bodu v  $k + 1$  iteraci.

- ***Adopted basis Newton-Raphson (ABNR)*** metoda začíná aplikací několika kroků metody nejstrmějšího sestupu a Hessián počítá pouze pro určitý podprostor s dimenzí  $s$ , na kterém v předchozích krocích došlo k největším změnám. To vede k tomu, že přidáváme informaci o zakřivení potenciálové plochy pouze ve směrech, které jsou nejvíce relevantní. Systém se potom pohybuje pouze po tomto podprostoru. Dimenze  $s$  je většinou brána mezi 4-10 [30], což výrazně snižuje výpočetní náročnost algoritmu.

## Kapitola 4

### Základy simulačních metod

V reálném experimentu obvykle měříme časovou střední hodnotu nějaké funkce souřadnic a hybností  $A(\vec{r}^N, \vec{p}^N)$

$$\bar{A} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t A(\vec{r}^N(t'), \vec{p}^N(t')) dt'. \quad (4.1)$$

Statistická fyzika nám umožňuje nahradit jeden takový, v čase se vyvíjející systém, mnoha různými systémy. Takovými, u nichž  $\vec{p}^N$  a  $\vec{r}^N$  mohou nabývat různých hodnot. Máme-li k dispozici dostatečný počet těchto stavů, můžeme získat střední souborovou hodnotu veličiny  $A$

$$\langle A \rangle = \int \int d\vec{p}^N d\vec{r}^N A(\vec{r}^N, \vec{p}^N) \rho(\vec{r}^N, \vec{p}^N), \quad (4.2)$$

kde  $\rho(\vec{r}^N, \vec{p}^N)$  značí hustotu pravděpodobnosti nalezení dané konfigurace  $\vec{r}^N$  a  $\vec{p}^N$  v závislosti na energii tohoto stavu.

Podle tzv. *ergodické hypotézy* se tyto dvě střední hodnoty rovnají [31]

$$\bar{A} = \langle A \rangle. \quad (4.3)$$

Rovnice 4.3 tedy v principu umožňuje dva přístupy k simulaci.

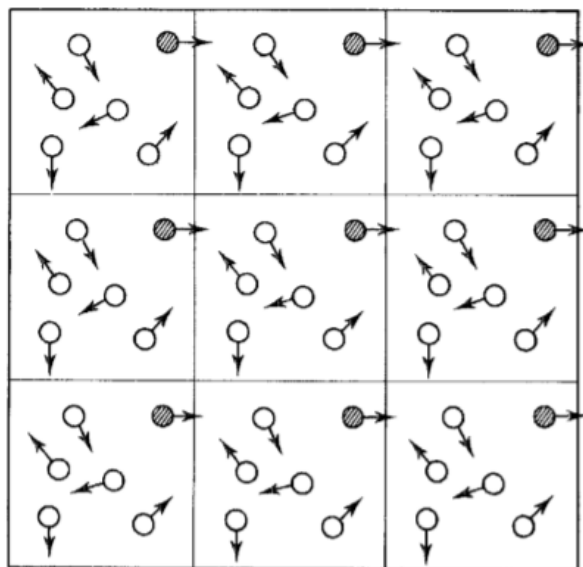
Molekulární dynamika počítá střední časovou hodnotu. Může poskytnout přehled o časovém vývoji dané veličiny. Metoda Monte Carlo počítá střední hodnotu souborovou.

V této kapitole budou nejprve uvedeny základy týkající se obou druhů simulací, jako jsou okrajové podmínky a podmínky pro výpočet ne vazebných členů. Následně jsou shrnuty principy obou výše zmíněných přístupů.

#### 4.1 Periodické a neperiodické okrajové podmínky

Jedním z problémů počítačových simulací je to, že si nemůžeme dovolit výpočty s velkým počtem částic. To pro simulace dějů probíhajících uvnitř reálné látky představuje problém, protože budeme-li mít malý box částic, potom bude většina z nich buď přímo na povrchu nebo budou alespoň efektům povrchu podléhat. Tento problém se nejčastěji řeší zavedením **periodických okrajových podmínek**. Původní box s původními částicemi se zreplikujeme a těmito replikami se obloží ve všech směrech. Takto vytvoříme nekonečnou makromřížku. Pokud některá z částic během simulace opustí na jedné straně centrální box, pak se tatáž objeví na straně druhé a počet částic v centrálním boxu tak zůstává konstantní. Ilustrace je pro 2D případ na obrázku 4.1.

Box může nabývat různých geometrických tvarů, s podmínkou, že se musí operací translace vyplnit celý prostor. Je ale vhodné volit takový box, který odpovídá geometrii systému. Například



Obrázek 4.1: 2D periodické okrajové podmínky [13].

pro sférickou molekulu by nemusel být pravoúhlý box neoptimálnější a lepší by bylo volit například rombový dodekahedron, protože bychom tak snížili počet částic v boxu. Nevýhodou periodických okrajových podmínek je například to, že nemůžeme dosáhnout fluktuací systému, které jsou větší než délka boxu. Další problém nastává s dalekodosaňovými silami a často je nutné přijmout to, že uvedeme do systému určitou uspořádanost na dlouhou vzdálenost i přesto, že ji systém ve své podstatě nemá.

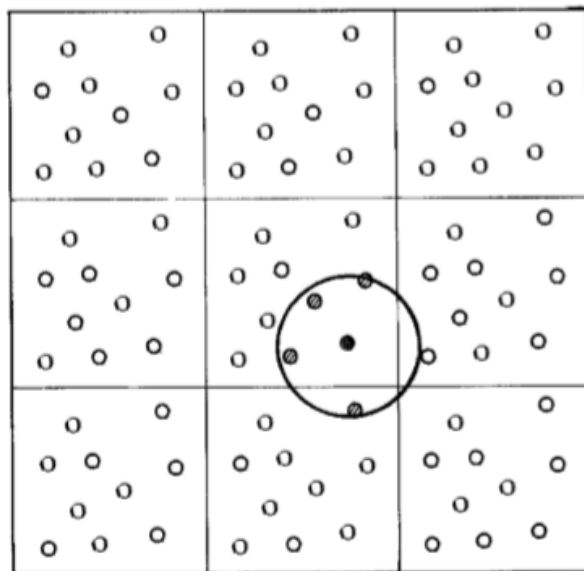
Jinou možností je zavedení **neperiodických okrajových podmínek**. Ty se většinou používají při simulaci nehomogenního prostředí nebo nerovnovážných systémů. V ostatních případech je vhodné je použít, pokud musíme snížit počet simulovaných částic. V této metodě se namísto tvorby boxu molekula obklopí vrstvou příslušného rozpouštědla. Pokud si představíme, že vložíme molekulu do boxu tvaru krychle tak, aby byla od každé stěny vzdálena o  $a$ , pak bude box obsahovat více molekul rozpouštědla než kdybychom molekulu obložili vrstvou o tloušťce  $a$ .

Při všech simulacích v této práci byly využity periodické okrajové podmínky.

## 4.2 Výpočet ne vazebných energií

Nejvíce času během simulace potřebuje obecně výpočet ne vazebných interakcí v systému. Ty se většinou počítají mezi dvojicemi atomů, tzn. že počet nutných výpočtů roste jako  $N^2$ . Nicméně pro mnoho interakčních potenciálů není třeba započítávat všechny tyto dvojice. Například Lennard-Jonesův potenciál klesá se vzdáleností velmi rychle a ve vzdálenosti  $2,5\sigma$  má hodnotu 1% své původní velikosti. Proto se v případě ne vazebných interakcí často používá *cutoff* a konvence minimálního obrazu (z angl. *minimum image convention*). Konvence minimálního obrazu znamená, že každý atom vnímá ostatní atomy pouze jednou a to vždy ty nejbližší, jak je zobrazeno na obrázku 4.3.

Použitím *cutoff*u bereme v úvahu pouze atomy, které se vyskytují ve sféře s jeho poloměrem. Pokud používáme periodické okrajové podmínky, *cutoff* bychom měli volit tak, aby atom „neviděl“ svůj obraz. Neměl by tedy být větší než je polovina nejkratší hrany boxu.



Obrázek 4.2: Konvence minimálního obrazu [13].

#### 4.2.1 List sousedních atomů

K tomu, abychom mohli zavést *cutoff*, musíme vypočítat vzdálenosti mezi všemi dvojicemi atomů v systému, což znamená  $N(N-1)/2$  výpočtů. Oproti výše zmíněným  $N^2$  se nejedná o výrazné zlepšení.

Řešením je zavedení listu sousedních atomů, kdy se vychází z toho, že se během 10–20 kroků molekulární dynamiky nebo iterací Monte Carla polohy sousedů jednotlivých atomů příliš nemění [13]. Potom si můžeme zapamatovat zkrátka všechny atomy, které leží v *cutoff* oblasti (a jeho blízkém okolí) bez toho, abychom museli počítat jednotlivé vzdálenosti.

Zavádíme dvě cely, seznam sousedů  $L$  a ukazatele  $P$ . Ukazatel  $P$  si pamatuje index prvních sousedů jednotlivých atomů, všechny sousední atomy všech atomů mohou tedy být obsaženy v jedné cele  $L$ . Důležité je správně zvolit frekvenci, s jakou se bude seznam sousedních atomů obnovovat. Pokud bude vysoká, přicházíme o efektivitu. Pokud bude příliš malá, mohli bychom chybně vypočítat energie a síly vzhledem k pohybu atomů. Jednou z možností je seznam aktualizovat po výše zmíněných 10–20 krocích simulace. Jinou možností je automaticky ukládat změny v pozici atomů během jednotlivých kroků simulace. Pokud bude součet maximálních posunů jakýchkoli dvou atomů větší než rozdíl mezi *cutoff*em a jejich vzdáleností v seznamu sousedů, aktualizujeme celý seznam.

Každá aktualizace stále vyžaduje výpočet všech vzdáleností. Pomohli jsme si ale snížením počtu těchto nutných výpočtů. Pokud je simulovaný systém mnohem větší než je *cutoff*, můžeme použít metodu indexování cel. Ta spočívá v rozdělení simulačního boxu na menší cely, přičemž délka každé z nich je o něco větší než náš *cutoff*. Všechny sousedy tedy nalezneme ve stejné nebo sousední cele. Pokud box rozdělíme na  $M^3$  cel, snížíme tím počet atomů, které musíme uvažovat o  $\sim M^3$  [13].

#### 4.2.2 Ewaldova sumace

Jako dalekodosažové interakce klasifikujeme takové, které se vzdáleností neklesají rychleji než  $r^{-d}$ ,  $d$  je dimenze systému. Pro třídímní systém tedy jde o náboj-nábojovou interakci  $\sim r^{-1}$  až

po dipól-dipólovou a náboj-kvadrupólovou interakci  $\sim r^{-3}$ . Problém je především v tom, že jejich dosah přesahuje délku poloviny hrany simulačního boxu a to i v případech, kdy pracujeme s velkými systémy.

Dalekodosahovost tedy znamená, že si nevystačíme s konvencí minimálního obrazu, který intrinsicky zavádí *cutoff* a započítává pouze atomy, které do něj spadají. Oblast vymezená tímto *cutoffem* je elektricky neutrální. Podobně nabitě ionty se budou snažit nabývat pozic v opačných částech boxu, čímž do systému, který by měl být izotropní, zavádíme nechtěnou periodicitu.

Tzv. mřížkové metody započítávají dalekodosahové interakce částice v centrálním boxu a všech ostatních částic v něm i ve všech jeho obrazech. Nejznámějším zástupcem těchto metod je Ewaldova sumace. Potenciální energie náboj-nábojové interakce je

$$V^{nn} = \frac{1}{2} \frac{1}{4\pi\epsilon_0} \sum_{\vec{m} \in \mathbb{Z}^3} \left( \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{|\vec{r}_{ij} + \vec{m}L|} \right), \quad (4.4)$$

kde  $g_i$  a  $q_j$  jsou náboje,  $N$  je počet částic v centrálním kubickém boxu o délce hrany  $L$ ,  $\vec{m} = (m_x, m_y, m_z)$  je vektor celých čísel a  $\vec{m}L$  reprezentuje polohu středu daného boxu. Zanedbáváme člen  $i = j$  pro  $\vec{m} = 0$ . Konvergence výrazu 4.4 je pomalá a navíc závisí na pořadí, ve kterém členy sčítáme. Obvykle se sčítání provádí postupně pro zvětšující se velikost  $|\vec{m}|$ . Tj. nejdříve přidáme šest boxů se středy v  $(\pm L, 0, 0)$ ,  $(0, \pm L, 0)$ ,  $(0, 0, \pm L)$  atd. Tímto způsobem postupně budujeme přibližně sférický systém. Nakonec definujeme prostředí, které tuto sféru obklopuje. Ewaldova sumace efektivně počítá  $V^{nn}$  pro sféru obklopenou dobrým vodičem (kovem), definovaným dielektrickou konstantou.

Ideou Ewaldovy sumace je rozdělení výrazu 4.4 na dva s tím, že každý bude konvergovat mnohem rychleji než suma původní. Jeden v reálném, druhý v reciprokém prostoru. V Ewaldově metodě se každému náboji přiřadí neutralizující nábojové rozdělení  $\rho_i^n$  s opačným znaménkem a stejnou velikostí, obvykle tvaru Gaussovy křivky

$$\rho_i^n(\vec{r}) = g_i \alpha^3 \exp(-\alpha^2 r^2) / \pi^{3/2}, \quad (4.5)$$

kde  $\alpha$  udává její šířku. Tyto distribuce mají za úkol stínit interakce mezi sousedními náboji tak, aby se staly krátkodosahovými. Celkový stíněný potenciál se vypočítá jako suma přes všechny náboje v centrálním boxu i v jeho replikách v reálném prostoru.

V dalším kroku se snažíme napravit umělé zavedení stínících distribucí a přidáváme distribuce se znaménkem původních nábojů a tvarem 4.5, čímž adekvátně snižujeme velikost původních nábojů. Tyto distribuce sčítáme v reciprokém prostoru. Nakonec musíme zavést dvě korekce. Jednu na fakt, že jsme do sumy započítali i interakce distribucí (přidaných v druhém kroku) se sebou samými a musíme je odečíst. Druhou korekci zavádíme vzhledem k faktu zmíněnému výše, a sice, že Ewaldova sumace odpovídá situaci, kdy kolem kvazi-sféry simulačních boxů zavedeme vrstvu kovu. Potenciál, který vznikne s touto vrstvou je s potenciálem, který vznikne s vrstvou vakua vztažen jako [16]

$$V_{kov}^{nn} = V_{vakuum}^{nn} - \frac{2\pi}{3L^3} \left| \sum_i q_i \vec{r}_i \right|^2 \quad (4.6)$$

a to díky tomu, že ve vakuu se na povrchu sféry vytvoří vrstva dipólů. Je to tedy druhý člen této rovnice, který musíme k potenciálu přičíst.

Výsledný výraz má tvar [13]

$$V_{vakuum}^{nn} = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N q_i q_j \left( \sum_{|\vec{n}|=0}^{\infty} \frac{\operatorname{erfc}(\alpha |\vec{r}_{ij} + L\vec{n}|)}{|\vec{r}_{ij} + L\vec{n}|} + \frac{1}{\pi L^3} \sum_{\vec{k} \neq 0} \frac{4\pi^2}{k^2} \exp\left(\frac{-k^2}{4\alpha^2}\right) \exp(-i\vec{k} \cdot \vec{r}_{ij}) \right) - \frac{\alpha}{\pi^{1/2}} \sum_{i=1}^N q_i^2 + \frac{2\pi}{3L^3} \left| \sum_{i=1}^N q_i \vec{r}_i \right|^2, \quad (4.7)$$

kde jsme položili  $1/4\pi\epsilon_0 = 1$  a  $\operatorname{erfc}(x)$  značí chybovou funkci

$$\operatorname{erfc}(x) = \frac{2}{\pi^{1/2}} \int_x^{\infty} \exp(-t^2) dt. \quad (4.8)$$

Rigorózní odvození Ewaldovy sumace předpokládá celkovou neutralitu systému. Pokud má systém nenulový celkový náboj, potom do něj musíme zavést neutralizující rozdělení, a do vztahu 4.7 přidat další člen

$$V_{nab} = -\frac{\pi}{2L^3\alpha^2} \left| \sum_{i=1}^N q_i \right|^2. \quad (4.9)$$

Chybová funkce 4.8 klesá k nule s rostoucím  $x$ . Z tohoto hlediska by pro nás bylo nejlepší volit velké  $\alpha$  tak, aby jediným členem přispívajícím k reálné sumě byl člen  $s = 0$ , což by odpovídalo konvenci minimálního obrazu. Nesmíme ale zapomenout na druhý člen rovnice 4.7, probíhající v reciprokném prostoru přes vektory  $\vec{k} = 2\pi\vec{n}/L$ , kde  $\vec{n} = (n_x, n_y, n_z)$  je vektor celých čísel. Velká hodnota  $\alpha$  v reciprokném prostoru odpovídá úzké nábojové distribuci, což znamená, že čím větší bude  $\alpha$ , tím větší počet vektorů  $\vec{k}$  budeme k popisu distribuce potřebovat. Koeficient  $\alpha$  se tedy musí volit s ohledem na obě tyto skutečnosti.

Následující příklad je převzat z [15]. Pro kubický box se nejdříve vybere hodnota *cutoffu*  $r_c$  potenciálu v reálném prostoru, obvykle  $r_c = L/2$ . Poté se pro  $\vec{n} = (10, 10, 10)$  vypočítá průměrná Coulombova energie (přes několik kroků molekulární dynamiky), jako funkce  $\alpha$  okolo hodnoty  $\alpha = 3/r_c$ . Vhodná hodnota  $\alpha$  odpovídá hodnotě, při které dojde k vytvoření plata ve  $V^{nn}(\alpha)$ . Dále můžeme snižovat hodnoty  $\vec{n}$  a test opakovat. Bez většího testování se obecně volí  $\alpha = 6/L$  a přibližně 200–300  $\vec{k}$  vektorů.

Hodnotu parametru  $\alpha$  je také možno určit z balancování chyb reálné a reciprokné sumy, jak ukázali Kolafa a Perram v [14].

Ewaldova metoda může být rozšířena i na interakce vyšších řádů než je interakce náboj-nábojová. Jako příklad uveďme dipól-dipólovou interakci. Pro její vyjádření se v 4.7 nahradí  $q_i$  výrazem  $\vec{\mu}_i \cdot \vec{\nabla}_{\vec{r}_i}$ , kde  $\vec{\mu}_i$  je molekulární dipól a výsledný potenciál je

$$V_{vakuum}^{dd} = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \left( \sum_{|\vec{n}|=0}^{\infty} (\vec{\mu}_i \cdot \vec{\mu}_j) B(\vec{r}_{ij} + L\vec{n}) - (\vec{\mu}_i \cdot \vec{r}_{ij}) (\vec{\mu}_j \cdot \vec{r}_{ij}) C(\vec{r}_{ij} + L\vec{n}) \right) + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \left( \sum_{\vec{k} \neq 0} \frac{1}{\pi L^3} (\vec{\mu}_i \cdot \vec{k}) (\vec{\mu}_j \cdot \vec{k}) (4\pi^2/k^2) \exp(-k^2/4\alpha^2) \exp(-i\vec{k} \cdot \vec{r}_{ij}) \right) - \sum_{i=1}^N 2\alpha^3 \mu_i^2 / 3\pi^{1/2} + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (4\pi/3L^3) \vec{\mu}_i \cdot \vec{\mu}_j, \quad (4.10)$$

kde  $i$  a  $j$  probíhá přes dipóly v centrálním boxu a

$$B(r) = \operatorname{erfc}(\alpha r) / r^3 + \left(2\alpha / \pi^{1/2}\right) \exp(-\alpha^2 r^2) / r^2, \quad (4.11)$$

$$C(r) = 3 \operatorname{erfc}(\alpha r) / r^5 + \left(2\alpha / \pi^{1/2}\right) (2\alpha^2 + 3/r^2) \exp(-\alpha^2 r^2) / r^2. \quad (4.12)$$

K řešení těchto rovnic se přistupuje stejně jako v původní Ewaldově sumaci.

Z výše zmíněného popisu je zřejmé, že Ewaldovu sumaci lze použít na molekuly, které mají přiřazený větší počet parciálních nábojů (vyjadřující pólovou příslušnost molekuly) za předpokladu, že se vypořádáme s „vnitřními“ interakcemi těchto nábojů v rámci dané molekuly. Tato vlastní energie má tvar

$$V^{vlast} = \frac{1}{2} \sum_i \sum_{a=1}^{n_s} q_{ia} \left( 2\alpha q_{ia} / \pi^{1/2} + \sum_{b \neq a}^{n_s} q_{ib} \operatorname{erfc}(\alpha d_{ab}) / d_{ab} \right), \quad (4.13)$$

kde  $n_s$  je počet přiřazených parciálních nábojů na molekule  $i$  a  $d_{ab}$  je vzdálenost mezi parciálními náboji  $a$  a  $b$ . Tuto energii je nutno v případě použití parciálních nábojů odečíst.

Výraz 4.7 je odvozen pro velkou sféru boxů ve vakuu. Relativní permitivita  $\epsilon$  zkoumaného systému záleží na permitivitě prostředí  $\epsilon_s$  použitého jako okolí této sféry jako

$$\frac{1}{\epsilon - 1} = \frac{3k_B T}{4\pi\rho\mu^2 g(\epsilon_s)} - \frac{1}{2\epsilon_s + 1}, \quad (4.14)$$

kde  $\rho$  je hustota systému,  $\mu$  je velikost celkového dipólového momentu v centrálním boxu a

$$g(\epsilon_s) = \frac{\langle |\sum_{i=1}^N \vec{\mu}_i|^2 \rangle - \langle |\sum_{i=1}^N \vec{\mu}_i \rangle|^2 \rangle}{N\mu^2}. \quad (4.15)$$

Pro  $\epsilon_s = 1$  přechází 4.14 na Clausius-Mosottiho rovnost

$$\frac{\epsilon - 1}{\epsilon + 2} = \frac{4\pi\rho\mu^2}{9k_B T} g(1) \quad (4.16)$$

a pro  $\epsilon_s$  kovu

$$\epsilon = 1 + \frac{4\pi\rho\mu^2}{3k_B T} g(\epsilon_s). \quad (4.17)$$

Pro výpočet permitivity systému je doporučeno vypočítat potenciální energii 4.10 bez členu vyjadřující dipólovou korekci (obklopujícím prostředím je kov), z této energie vypočítat  $g$  a dosadit do výrazu 4.17 a to vzhledem k větším chybám, které by vznikly použitím vztahu 4.16 [15].

### 4.2.3 Particle-particle particle-mesh metoda

*Particle-particle particle-mesh* metoda (PPPM nebo P<sup>3</sup>M) představuje modifikaci Ewaldovy sumace založenou na rychlé Fourierově transformaci (FFT z angl. *fast Fourier transform*). Jako v normální Ewaldově sumaci je v prvním přiblížení potenciální energie rozdělena na dvě sumy. Výpočet krátkodosahové sumy probíhá v přímém prostoru s využitím metod jako je list sousedních atomů. Sumace zahrnující dlouhodosahové interakce probíhá v recipročním prostoru a závisí na Fourierově transformaci nábojové hustoty  $\hat{\rho}^n$

$$V_{dl}^{nm} = \frac{1}{2L^3} \sum_{\vec{k} \neq 0} \frac{4\pi}{k^2} \exp\left(\frac{-k^2}{4\alpha^2}\right) \left| \hat{\rho}^n(\vec{k}) \right|^2, \quad (4.18)$$

kde

$$\hat{\rho}^n(\vec{k}) = \int d\vec{r} \rho^n(\vec{r}) \exp(-i\vec{k} \cdot \vec{r}) = \sum_i^N q_i \exp(-i\vec{k} \cdot \vec{r}_i). \quad (4.19)$$

V PPPM metodě je nábojová hustota nejdříve interpolována na jemnou mřížku v simulačním boxu, aby mohla být použita rychlá Fourierova transformace (algoritmus, který slouží k výpočtu diskrétní Fourierovy transformace). Tato podmřížka je tvořena  $s_c^3 = (L/l)^3$  krychlemi, kde  $l$  je délka hrany těchto krychlí a  $L$  je délka hrany simulačního boxu. Hodnoty  $s_c$  obvykle nabývají hodnot 16 nebo 32 [15]. Nábojová hustota na podmřížce je definovaná v bodech  $\vec{r}_s = l(s_x, s_y, s_z)$ , kde  $s_\beta = 0, \dots, s_c - 1$  a platí pro ní

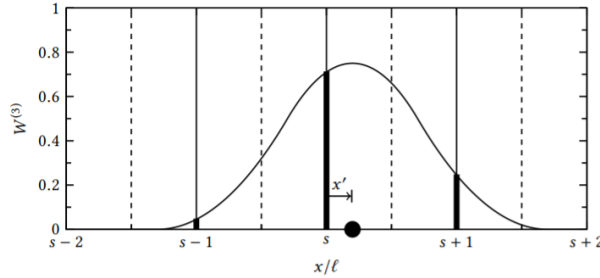
$$\rho_s^n(\vec{r}_s) = \frac{1}{l^3} \int_{V_{box}} d\vec{r} W(\vec{r}_s - \vec{r}) \rho^n(\vec{r}) = \frac{1}{l^3} \sum_{i=1}^N q_i W(\vec{r}_s - \vec{r}_i), \quad (4.20)$$

kde  $W$  je přenosová funkce, která musí zachovávat náboj, tj. součet nábojů přenesených na mřížku se musí rovnat původnímu náboji  $q_i$  a efektivně určuje počet uzlových bodů a váhu náboje, který se na podmřížku přenese. Jako příklad uvedme trojúhelníkovou funkci, která přenáší náboj na tři nejbližší uzly podmřížky v jedné dimenzi

$$W^{(3)}(x) = \begin{cases} \frac{3}{4} - (x/l)^2 & |x| \leq \frac{1}{2}l \\ \frac{1}{2} \left( \frac{3}{4} - |x|/l \right)^2 & \frac{1}{2}l \leq |x| \leq \frac{3}{2}l \\ 0 & \text{jinak,} \end{cases}$$

kde  $x$  je vzdálenost mezi nábojem a nejbližším uzlem podmřížky. Označíme-li tento bod jako  $s$  a jeho vzdálenost od náboje jako  $x' = (x - x_s)/l$ , potom platí

$$W^{(3)}(x') = \begin{cases} \frac{1}{2} \left( \frac{1}{2} + x' \right)^2 & \text{v bodě } s+1 \\ \frac{3}{4} - x'^2 & \text{v bodě } s \\ \frac{1}{2} \left( \frac{1}{2} - x' \right)^2 & \text{v bodě } s-1 \end{cases}$$



Obrázek 4.3: Přiřazení náboje na podmřížku pomocí funkce 4.2.3. Přerušovanou čarou jsou zobrazeny hranice simulačních boxů, plnou čarou uzly podmřížky uprostřed každého boxu. Zvýrazněné čáry ukazují na váhy, které jsou náboji v jednotlivých uzlech přiřazeny. Převzato z [15].

Ve třech dimenzích této váhové funkci odpovídá 27 uzlových bodů, jejichž celková váha je  $W^{(3)}(\vec{r}_s) = W^{(3)}(x') W^{(3)}(y') W^{(3)}(z')$ . Máme-li náboj převedený na podmřížku, jeho diskrétní Fourierova transformace je

$$\hat{\rho}_s^n = l^3 \sum_{\vec{r}_s} \rho_s^n(\vec{r}_s) \exp(-i\vec{k} \cdot \vec{r}_s), \quad (4.21)$$



kde  $\hat{\rho}_s^n(\vec{k})$  je definovaná pro vlnové vektory  $\vec{k} = (2\pi/L)(n_x, n_y, n_z)$ , kde  $|n_\beta| \leq (s_c - 1)/2$ . Rovnice 4.21 je ve tvaru, na který může být použit některý z algoritmů rychlé Fourierovy transformace. Vypočteme-li  $\hat{\rho}_s^n$ , pro energii dalekodosaových interakcí platí

$$V_{dl}^{nn} = \frac{1}{2V} \sum_{\vec{k} \neq 0} \hat{G}(\vec{k}) |\hat{\rho}_s^n(\vec{k})|^2, \quad (4.22)$$

kde  $\hat{G}$  je mřížková Greenova funkce a její úkol je napravit fakt, že jsme nábojovou hustotu převedli na podmřížku. Je volena tak, aby minimalizovala střední kvadratické odchytky v reciprokových silách. Elektrické pole, tj. derivaci potenciálu, můžeme získat derivací v reciprokém prostoru (vynásobením potenciálu  $i\vec{k}$ ) a následnou inverzní Fourierovou transformací

$$\vec{E}(\vec{r}_s) = -\mathfrak{F}^{-1}[i\vec{k}\hat{\rho}_s^n(\vec{k})\hat{G}(\vec{k})]. \quad (4.23)$$

Tento přístup se nazývá  $i\vec{k}$ -diferenční schéma a obsahuje tři inverzní Fourierovy transformace pro každou komponentu pole. Nakonec využijeme pole vypočítané na uzlech podmřížky k výpočtu síly působící na náboje

$$\vec{f}_i = q_i \sum_{\vec{r}_s} \vec{E}(\vec{r}_s) W(\vec{r}_i - \vec{r}_s), \quad (4.24)$$

kde  $W$ , která nyní přiřazuje síly působící na náboje, musí být stejná funkce, která byla použita k přenosu náboje na mřížku.

Nejznámější mřížkové algoritmy představují dvě metody. *Particle mesh Ewald* metoda (PME) a *smooth particle mesh Ewald* (SPME). V PME metodě se aproximuje komplexní exponenciála v 4.21 Lagrangeovými polynomy. V SPME se místo Lagrangeovy interpolace využívá aproximace kardinálními B-spline křivkami.

Normální Ewaldova sumace poskytuje nejpřesnější výsledky, je ale také výpočetně nejnáročnější  $\sim N^{3/2}$ , kde  $N$  je počet nábojů. *Mesh* metody jsou škálovány jako  $\sim N \log N$ .

V Materiálovém Studiu je možnost použití normální Ewaldovy sumace a PPPM metody. Přitom si nemůžeme volit parametry jako je šířka Gaussova rozdělení  $\alpha$  a počet reciprokových vektorů. V MS je implementován postup Karasawy a Goddarda [17], kdy

$$\alpha^2 = \frac{\frac{1}{2}k_{min}}{|\vec{R}_{L,min}|}, \quad (4.25)$$

kde  $\vec{R}_{L,min}$  je nejkratší z bazických vektorů reálné mřížky a  $k_{min}$  je velikost nejmenšího vektoru mřížky reciproké. Se specifikovaným parametrem  $\alpha$  se následně určuje počet členů, přes které se bude počítat v reálném a reciprokém prostoru pomocí specifikované tolerance  $\delta$ . Tj. zanedbávají se členy, které k sumám nepřispívají větším dílem než  $\delta$ .

Ve výše zmíněném článku se neobjevuje výraz na dipólovou korekci a prostředím obklopujícím nekonečný systém boxů je tedy kov. Pro výpočet dielektrické konstanty v MS je tedy nutné použít vztah 4.17.

### 4.3 Molekulární dynamika (MD)

V molekulární dynamice jsou konfigurace systému generovány integrací Newtonových pohybových rovnic

$$m_i \ddot{\vec{r}}_i(t) = \vec{F}_i(t), \quad i = 1 \dots N \quad (4.26)$$

kde  $N$  je počet částic v systému,  $\vec{F}_i(t)$  je síla působící na atom  $i$  v čase  $t$ . Molekulární dynamika vypočítá sílu působící na částici jako záporný gradient potenciální energie a adekvátně upraví polohy atomů podle Newtonových rovnic. Právě k této části postupu potřebujeme integrační algoritmus. Jedna skupina takových algoritmů využívá metody konečných diferencí. Základní myšlenkou je, že integraci provádíme v krocích, z nichž každý je oddělen časovým úsekem  $\Delta t$ . Celková síla působící na každou částici v čase  $t$  je vypočítána jako součet všech interakcí s ostatními částicemi (pro nás pouze párové interakce). Z této síly můžeme získat jejich zrychlení, které jsou následně použity, spolu s pozicemi a rychlostmi v čase  $t$ , k výpočtu nových souřadnic a rychlostí v čase  $t + \Delta t$ . Síla je během tohoto kroku konstantní.

Integračních algoritmů založených na metodě konečných diferencí je několik. Všechny ale vycházejí z toho, že souřadnice, rychlosti a zrychlení můžeme aproximovat Taylorovým rozvojem. Nejznámějším algoritmem v této skupině je **Verletův algoritmus**. Ten k výpočtu nových souřadnic využívá souřadnice a zrychlení v čase  $t$  a pozice z předchozího kroku, tj. v čase  $t - \Delta t$

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \vec{v}_i(t) \Delta t + \frac{\vec{F}_i(t)}{2m} \Delta t^2 + \frac{\ddot{\vec{r}}_i}{3!} \Delta t^3 + O(\Delta t^4) \quad (4.27)$$

$$\vec{r}_i(t - \Delta t) = \vec{r}_i(t) - \vec{v}_i(t) \Delta t + \frac{\vec{F}_i(t)}{2m} \Delta t^2 - \frac{\ddot{\vec{r}}_i}{3!} \Delta t^3 + O(\Delta t^4). \quad (4.28)$$

Sečtením těchto rovnic máme

$$\vec{r}_i(t + \Delta t) \approx 2\vec{r}_i - \vec{r}_i(t - \Delta t) + \frac{\vec{F}_i(t)}{m} \Delta t^2. \quad (4.29)$$

Rychlosti ve Verletově algoritmu explicitně obsaženy nejsou a mohou se dopočítat např. jako  $\vec{v}_i(t) = [\vec{r}_i(t + \Delta t) - \vec{r}_i(t - \Delta t)] / 2\Delta t$ . To znamená, že je můžeme vypočítat až když už známe souřadnice v čase  $t + \Delta t$ . Tento problém řeší **rychlostní Verletův algoritmus**, který poskytuje souřadnice, rychlosti a zrychlení ve stejném čase [13]

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \vec{v}_i(t) \Delta t + \frac{\vec{F}_i(t)}{2m} \Delta t^2 \quad (4.30)$$

$$\vec{v}_i(t + \Delta t) = \vec{v}_i(t) + \frac{\vec{F}_i(t + \Delta t) + \vec{F}_i(t)}{2m} \Delta t. \quad (4.31)$$

V prvním kroku se vypočítají pozice v čase  $(t + \Delta t)$  z 4.30. Z těchto nových pozic se vypočítají nové síly a podle 4.31 tak můžeme dopočítat i nové rychlosti.

Rychlostní Verletův algoritmus, integrační algoritmus 4. řádu, tedy umí vygenerovat pozice i rychlosti ve stejném čase. To je užitečné třeba proto, že pro daný čas můžeme získat informaci o kinetické i potenciální energii systému.

V modulu Forcite, kde byly prováděny všechny simulace dynamiky v této práci, se tento rychlostní Verletův algoritmus používá.

Volba časového kroku je důležitá v mnoha ohledech. Jedním faktorem je zachování celkové energie systému. Pokud bychom zvolili moc velký časový krok, částice by v každé iteraci mohly „skočit“ moc daleko a to by zpětně mohlo vést k tomu, že by se nacházely v příliš velké blízkosti jiné částice. To by vedlo k nefyzikálnímu navýšení potenciální, a tím i celkové energie systému.

Zároveň by velký krok mohl způsobit nestabilní chování molekuly díky velkým chybám, které by při integraci nastaly [32]. Naopak s příliš malým krokem bychom mohli zbytečně ztratit čas vzhledem k velkému počtu iterací, kterými by systém musel projít.

Dalším faktorem je samozřejmě to, jak dlouho trvají děje, které chceme pozorovat. Obecně by se

Tabulka 4.1: Doporučené časové kroky podle [13].

System	Typ pohybu	Doporučený čas. krok [s]
Atomy	translace	$10^{-14}$
Rigidní molekuly	translace, rotace	$5 \cdot 10^{-15}$
Flexibilní molekuly s rigidními vazbami	translace, rotace, torze	$2 \cdot 10^{-15}$
Flexibilní molekuly s flexibilními vazbami	translace, rotace, torze, vibrace	$1 \cdot 10^{-15}$ nebo $5 \cdot 10^{-16}$

měla volit délka kroku tak, aby byla o jeden řád menší, než je délka nejkratšího pohybu v systému [13]. Tyto nejkratší pohyby jsou většinou vysokofrekvenční vibrace vazeb. Protože nás obvykle moc nezajímají, můžeme jim přiřadit jejich rovnovážnou délku a přikázat, aby ji neměnily.

V tabulce 4.1 jsou přepsány doporučené hodnoty z [13]. V simulacích provedených v této práci, byl krok volen na 1 fs a to vzhledem k informaci, že vazba C-H kmitá s periodou 10 fs [13].

Počáteční konfiguraci systému můžeme získat například z experimentálních dat. Pokud nejsou dostupná, můžeme použít konfiguraci systému, kterou jsme geometricky zoptimalizovali, tj. její potenciální energie by měla být minimální. Pokud tuto počáteční konfiguraci máme, a chceme na ní provést molekulární dynamiku, musíme nejdříve systém ‚rozhýbat‘ přiřazením počátečních rychlostí atomům systému. Ty můžeme přiřadit na základě Maxwell-Boltzmannova rozdělení

$$p(v_{ix}) = \left( \frac{m_i}{2\pi k_B T} \right)^{1/2} \exp\left( -\frac{1}{2} \frac{m_i v_{ix}^2}{k_B T} \right), \quad (4.32)$$

kde  $T$  je námi požadovaná teplota systému. Pravděpodobnost vygenerování určité hodnoty z normálního rozdělení je

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left( -\frac{(x - \langle x \rangle)^2}{2\sigma^2} \right), \quad (4.33)$$

kde  $\langle x \rangle$  značí střední hodnotu  $x$  a  $\sigma^2$  jeho varianci. Necháme-li si nyní vygenerovat dvě náhodná čísla  $a_1$  a  $a_2$  od 0 do 1, pak jim v normálním rozdělení budou náležet hodnoty [13]

$$x_1 = \sqrt{-2\ln a_1} \cos(2\pi a_2), \quad x_2 = \sqrt{-2\ln a_1} \sin(\pi a_2). \quad (4.34)$$

Počáteční rychlosti jsou přiřazeny tak, aby celková hybnost systému byla nulová (NVE soubor). Molekulární dynamika, oproti Monte Carlu, generuje v čase propojené konfigurace systému, a my tak můžeme získat časově závislé vlastnosti systému.

### 4.3.1 NVT dynamika

Newtonovy rovnice spolu s rychlostním Verletovým algoritmem zachovávají celkovou energii systému. Počet částic  $N$  a objem  $V$  boxu se taktéž zachovává. Mikrokanonický soubor tedy představuje přirozenou statistiku molekulární dynamiky. Pro popis reálného experimentu by ale mohlo být lepší vzorkovat například z kanonického (NVT) souboru. Vztahy pro transformaci termodynamických proměnných mezi jednotlivými soubory existují, všechny jsou odvozeny za předpokladu platnosti termodynamické limity (nekonečně velký systém). Zde uvedeme základní možnosti, jak přímo kontrolovat teplotu systému v průběhu simulace a tím ji přinutit, aby fluktovala okolo požadované

hodnoty. Kontrola teploty se nakonec často aplikuje i v několika krocích NVE simulací. To z toho důvodu, že může dojít k tomu, že i přesto, že numericky bude celková energie systému zachována, systém v rovnováze být nemusí.

Ekvipartiční teorém

$$\langle K \rangle_{NVT} = \frac{3}{2} N k_B T, \quad (4.35)$$

propojuje kinetickou energii  $K$  s teplotou  $T$  systému s  $N$  částicemi (pro systém bez omezení). Jednou z metod, jak udržet teplotu konstantní, je **škálování** rychlostí, tj. násobením rychlosti faktorem  $\lambda$ , potom

$$\Delta T = \sum_{i=1}^N \frac{2}{3} \frac{m_i (\lambda v_i)^2}{N k_B} - \sum_{i=1}^N \frac{2}{3} \frac{m_i v_i^2}{N k_B} = (\lambda^2 - 1) T(t) \Rightarrow \lambda = \sqrt{\frac{\Delta T + T(t)}{T(t)}}, \quad (4.36)$$

kde  $T(t)$  je teplota v čase  $t$  a čítenel v posledním výrazu představuje novou teplotu systému. Použijeme-li tuto metodu, pak v každém kroku simulace přenásobíme rychlosti faktorem  $\lambda = \sqrt{T_c/T_t}$ , kde  $T_c$  je teplota, které chceme dosáhnout a  $T_t$  je momentální teplota systému.

Jinou škálovací metodou je **Berendsenova metoda**. V ní se zavede termostat s požadovanou teplotou, který do systému buď teplotu dodává nebo odvádí. Rychlosti jsou škálovány v každém kroku tak, aby platilo [13]

$$\frac{dT(t)}{dt} = \frac{1}{\tau} (T_{ts} - T(t)), \quad (4.37)$$

kde  $T_{ts}$  je teplota termostatu a  $\tau$  je parametr popisující provázanost systému s termostatem. Škálovací faktor je [13]

$$\lambda^2 = 1 + \frac{\delta t}{\tau} \left( \frac{T_{ts}}{T(t)} - 1 \right), \quad (4.38)$$

kde  $\delta t$  značí délku kroku.

Výše zmíněné metody jsou metody přibližné. To znamená, že nedávají přesný kanonický soubor. Škálování rychlostí uměle prohlubuje rozdíly v teplotách jednotlivých komponent systému. To může vést k tomu, že celková teplota sice bude mít požadovanou hodnotu, ale jedna část systému může být chladnější než jiná.

Následující metody kanonický soubor generují. Jednou z nich je stochastická metoda srážek, **Andersenova metoda**. Ta spočívá v tom, že vybereme nějakou částici (nebo částice) a přiřadíme jí rychlost z Maxwell-Boltzmannova rozdělení při požadované teplotě. Systém necháme nějaký čas vyvíjet (v tomto čase se systém bude vyvíjet mikrokanoonicky) a přiřazení rychlostí provedeme znovu. Provádíme tedy několik mikrokanoonických simulací, pokaždé s mírně odlišnou energií. Frekvence, se kterou bychom měli měnit rychlosti je [13]

$$\nu = \frac{2a\kappa}{3k_B M^{1/3} N^{2/3}}, \quad (4.39)$$

kde  $a$  je konstanta,  $\kappa$  je tepelná vodivost a  $M$  je hustota počtu částic.

Další metodou je metoda **Nosé-Hooverova termostatu**. V této metodě se přímo do pohybových rovnic zavádí nový stupeň volnosti  $s$ , který představuje tepelnou lázeň. Tomuto stupni volnosti je přiřazena hmotnost  $Q$  a potenciál  $\phi$ . Nosé ukázal, že vhodnou volbou tohoto potenciálu, vygenerujeme řešením rozšířených pohybových rovnic kanonický soubor.

Tato potenciální energie je  $(f+1)k_B T \ln s$ , kde  $f$  je počet stupňů volnosti systému a  $T$  kýžená teplota [13], kinetická energie je  $(Q/2)(ds/dt)^2$ , kde  $Q$  je ‚hmotnost‘ našeho nového stupně volnosti a jeho hodnota udává míru propojenosti lázně se systémem. Je-li  $Q$  velké, pak dochází k pomalé výměně

energie (pro  $Q \rightarrow \infty$  dostáváme NVE dynamiku), je-li příliš malé, pak bude docházet k rychlým oscilacím v energii a museli bychom volit zbytečně malý integrační krok. Nosé navrhl  $Q \sim f k_B T$ , kde  $f$  je počet stupňů volnosti systému a  $T$  je požadovaná teplota [13]. V MS se zadává hodnota  $Q_{pomr}$ , definovaná jako  $Q_{pomr} = Q/(\tau^2 f k_B T)$ , kde  $\tau = 1$  ps.

Nosé-Hooverův termostat potřebuje k ustavení teploty relativně hodně času. Na druhou stranu se ale jedná o ‚důvěryhodný‘ termostat, a proto se většinou aplikuje některá z předešlých metod ze začátku simulace, kdy chceme systém rychle uvést do rovnováhy a poté se přistupuje k Nosé-Hooverově metodě.

Překonání pomalé konvergence Nosé-Hooverova termostatu je připisováno **Nosé-Hoover-Langevinově (NHL) metodě**, která do pohybových rovnic přidává Langevinův třecí člen a parametr zahrnující náhodné fluktutace v systému. Velká hodnota tohoto parametru vede k rychle fluktuujícím  $s$ , což vede k útlumu vlastních oscilací této proměnné. Pokud je tento koeficient nulový, získáme Nosé-Hooverův termostat. Právě tento člen poskytuje možnost rychle konvergujícího kvalitního termostatu [7].

V kanonickém souboru jsou fluktuační teploty [7]

$$\sigma(T_t) = T \sqrt{\frac{2}{f}}, \quad (4.40)$$

kde  $\sigma$  je standardní odchylka,  $T_t$  je okamžitá teplota systému,  $T$  je požadovaná teplota a  $f$  je počet stupňů volnosti. Parametry termostatů by se měly volit tak, aby výsledné fluktuační teploty odpovídaly těm z rovnice 4.40.

### 4.3.2 Potenciál středního pole a stochastická dynamika

Potenciál středního pole vyjadřuje, jak se mění volná energie s některou ze svých proměnných (např. vzdálenost dvou atomů, torzní úhel). Myšlenka je taková, že v sobě zahrnuje vliv prostředí, ve kterém se molekula nachází a my bychom tak nemuseli simulovat velké molekulové systémy. Tyto potenciály mohou být odvozeny z molekulární dynamiky i Monte Carla například pomocí *umbrella* vzorkování.

- Metoda **Umbrella vzorkování** narušuje potenciální funkci [13]

$$\mathcal{V}' = \mathcal{V}(\vec{r}^N) + W(\vec{r}^N), \quad (4.41)$$

kde  $W(\vec{r}^N) = k_W (\vec{r}^N - \vec{r}_0^N)^2$  je váhová funkce. Pro konfigurace, které jsou daleko od rovnovážného  $\vec{r}_0^N$  nabývá váhová funkce velkých hodnot a modifikovaná  $\mathcal{V}'(\vec{r}^N)$  bude usměrňována směrem od  $\vec{r}_0^N$ . Výsledná distribuce není Boltzmannovská, ale výsledné střední hodnoty mohou být získány jako [13]

$$\langle A \rangle = \frac{\langle A(\vec{r}^N) \exp(+W(\vec{r}^N)/k_B T) \rangle_W}{\langle \exp(+W(\vec{r}^N)/k_B T) \rangle_W}, \quad (4.42)$$

kde index  $W$  znamená, že se jedná o střední hodnoty vzhledem k pravděpodobnosti  $P_W(\vec{r}^N)$ , která je odvozena z funkce  $\mathcal{V}'(\vec{r}^N)$ .

Pokud bychom chtěli například získat potenciál středního pole přes radiální distribuční funkci  $g(r)$ , kdy volná energie  $A(r)$  je  $A(r) = -k_B T \ln g(r) + konst$ , vypočítali bychom nejdříve  $g(r)$  s  $\mathcal{V}'(\vec{r}^N)$  a následně ji upravili tak, aby dávala  $g(r)$  při  $\mathcal{V}(\vec{r}^N)$ . Z ní bychom potom vypočítali volnou energii jako funkci vzdálenosti.

Tato metoda se hodí k překonávání větších potenciálových bariér.

Prostředí ovlivňuje dynamické chování molekuly přes náhodné srážky a třecí členy.

Ve stochastické dynamice je startovacím bodem Langevinova pohybová rovnice. Síly působící na částici zde mají tři zdroje. Prvním jsou interakce mezi částicemi. Tato složka síly,  $\vec{F}_i$  je modelována potenciálem středního pole.

Druhým zdrojem je pohyb částic prostředím. Jedná se o třecí sílu  $\vec{F}_{tec} = -\xi\vec{v}$ , kde  $\xi$  je třecí koeficient a  $\vec{v}$  je rychlost. Třecí koeficient je vztažen ke srážkové frekvenci  $\gamma$ , platí  $\gamma = \xi/m$ , kde  $m$  je hmotnost částice a  $\gamma^{-1}$  je rychlostní relaxační čas. Pro sférickou částici je třecí koeficient vztažen s difúzní konstantou  $D$ ,  $\xi = k_B T/D$ . Pro větší molekuly, kdy požadujeme třecí koeficient pro každý atom, se hodnoty přiřazují většinou na základě toho, jak moc jsou v kontaktu s okolním prostředím.

Třetím zdrojem jsou náhodné fluktuace způsobené interakcí s okolím. Tento příspěvek k síle označíme  $\vec{R}(t)$ . Langevinovu pohybovou rovnici pro částici  $i$  potom můžeme zapsat jako [13]

$$m_i \frac{d^2 x_i(t)}{dt^2} = \vec{F}_i(x_i(t)) - \gamma_i \frac{dx_i(t)}{dt} m_i + \vec{R}_i(t). \quad (4.43)$$

Předpoklady, které jsou kladeny na třecí koeficienty a náhodné síly jsou různé. Například, že srážková frekvence není časově ani pozičně závislá. Náhodné síly obvykle nejsou vůbec korelovány s rychlostmi, souřadnicemi ani silami na částicích a volí se tak, aby se dosáhlo Gaussovy distribuce s nulovou střední hodnotou [13].

#### 4.4 Monte Carlo simulace

Monte Carlo (MC) simulace generuje konfigurace systému náhodnými změnami v polohách částic. V molekulárních simulacích se využívá tzv. *importance sampling*, které umožňuje generaci nízkoenergetických konfigurací systému. Z těchto souřadnic můžeme vypočítat potenciální energii konfigurace a z ní jiné vlastnosti. Monte Carlo vzorkuje pouze  $3N$  prostor souřadnic. To nám stačí díky tomu, že funkční předpis potenciální energie v sobě obvykle neobsahuje závislost na rychlostech, což vede k možnosti separace integrálů v kanonické partiční funkci

$$\begin{aligned} Q_{NVT} &= \frac{1}{N!} \frac{1}{h^{3N}} \int \int d\vec{p}^N d\vec{r}^N \exp\left(-\frac{\mathcal{H}(\vec{p}^N, \vec{r}^N)}{k_B T}\right) = \\ &= \frac{1}{N!} \frac{1}{h^{3N}} \int d\vec{p}^N \exp\left(-\frac{|\vec{p}_i|^2}{2mk_B T}\right) \int d\vec{r}^N \exp\left(-\frac{\mathcal{V}(\vec{r}^N)}{k_B T}\right) = \\ &= \frac{1}{N!} \left(\frac{2\pi mk_B T}{h^2}\right)^{3N/2} \int d\vec{r}^N \exp\left(-\frac{\mathcal{V}(\vec{r}^N)}{k_B T}\right), \end{aligned}$$

kde poslední integrál přes souřadnice se někdy nazývá konfiguračním integrálem.

V ideálním plynu mezi sebou částice neinteragují,  $\exp(-\mathcal{V}(\vec{r}^N)/k_B T) = 1$  a konfigurační integrál odpovídá objemu  $V^N$  a pro ideální plyn tak dostáváme

$$Q_{NVT}^{id} = \frac{V^N}{N!} \left(\frac{2\pi k_B T m}{h^2}\right)^{3N/2}. \quad (4.44)$$

Dospěli jsme k tomu, že partiční funkci reálného systému můžeme zapsat jako součin příspěvků od ideálního plynu a ostatních interakcí mezi částicemi [13]

$$Q_{NVT} = Q_{NVT}^{id} Q_{NVT}^{ost}, \quad (4.45)$$

kde

$$Q_{NVT}^{ost} = \frac{1}{V^N} \int d\vec{r}^N \exp\left(-\frac{\mathcal{V}(\vec{r}^N)}{k_B T}\right). \quad (4.46)$$

Důsledkem je, že termodynamické veličiny můžeme zapsat pomocí těchto dvou partičních sum a nepotřebujeme k tomu informaci o hybnostním podprostoru.

Na „simulačním“ poli se nejčastěji využívá Metropolisova varianta Monte Carla. Tato metoda generuje tzv. Markovův řetězec, to znamená, že [13]

- Každý z pokusů závisí pouze na předešlém pokusu.
- Každý z pokusů náleží konečné množině možných výstupů.

Pravděpodobnost přechodu ze stavu  $m$  do stavu  $n$  označme  $\pi_{mn}$ , kde  $\pi_{mn}$  je obecně matice  $N \times N$  a nazývá se matice přechodu. Pravděpodobnost, že se systém nachází v nějakém stavu je dána vektorem pravděpodobností  $\vec{\rho} = (\rho_1, \dots, \rho_m, \rho_n, \dots, \rho_N)$ . Pokud  $\vec{\rho}(1)$  reprezentuje počáteční, náhodně vybranou konfiguraci, pak pravděpodobnost druhého stavu je  $\vec{\rho}(2) = \vec{\rho}(1) \pi$ , pravděpodobnost třetího je  $\vec{\rho}(3) = \vec{\rho}(2) \pi = \vec{\rho}(1) \pi \pi$  atd.

Rovnovážné rozdělení systému dostaneme aplikací nekonečného počtu matice přechodu  $\vec{\rho}_{im} = \lim_{N \rightarrow \infty} \vec{\rho}(1) \pi^N$  [13]. Limitní distribuce je nezávislá na náhodně zvolené počáteční konfiguraci a pravděpodobnost každého stavu je dána Boltzmannovým rozdělením. Pokud máme limitní distribuci, pak aplikací další matice přechodu získáme opět stejnou limitní distribuci.

Dalším pojmem je stochastická matice nebo *underlying matrix* Markovova řetězce. Vyjadřuje pravděpodobnost, se kterou se přejde do nového stavu. Pokud je pravděpodobnost přijetí pokusu o přestup z  $m$  do  $n$   $p_{mn}$ , potom je pravděpodobnost tohoto přechodu dána jako součin pravděpodobnosti, že vybereme stavy  $m$  a  $n$  ( $\alpha_{mn}$ , stochastická matice) a pravděpodobnosti přijetí pokusu ( $p_{mn}$ ) [13]

$$\pi_{mn} = \alpha_{mn} p_{mn}. \quad (4.47)$$

Stochastická matice je obvykle symetrická, tj. pravděpodobnost přechodu z  $m$  do  $n$  a naopak je stejná.

Pokud je stav  $n$  více pravděpodobný než  $m$  (v limitní distribuci, tj. Boltzmannův faktor pro  $n$  má vyšší hodnotu), potom [13]

$$\pi_{mn} = \alpha_{mn} \quad (\rho_n \geq \rho_m), \quad (4.48)$$

pro opačnou situaci

$$\pi_{mn} = \alpha_{mn} (\rho_m / \rho_n) \quad (\rho_n < \rho_m). \quad (4.49)$$

Shrňme tedy, že pokud má nový stav  $n$  vyšší energii než původní stav  $m$ , porovnáme Boltzmannovy faktory  $\exp(-\Delta \mathcal{V}(\vec{r}^N) / k_B T)$ , kde  $\Delta \mathcal{V} = \mathcal{V}(\vec{r}^N)_n - \mathcal{V}(\vec{r}^N)_m$ , s náhodně vybraným číslem od 0 do 1. Pokud je Boltzmannův faktor větší než toto náhodné číslo, stav přijmeme. V opačném případě stav odmítneme.

Pokud má nový stav  $n$  energii nižší než stav  $m$ , potom stav přijímáme.

Během každé iterace se vygeneruje nová konformace a to běžně tak, že se náhodně vybere jedna částice a změní se jí souřadnice podle [13]

$$x_{nová} = x_{stará} + (2\xi - 1) \delta r_{max} \quad (4.50)$$

$$y_{nová} = y_{stará} + (2\xi - 1) \delta r_{max} \quad (4.51)$$

$$z_{nová} = z_{stará} + (2\xi - 1) \delta r_{max}, \quad (4.52)$$

kde  $\xi$  je náhodně vygenerované číslo od 0 do 1 a  $\delta r_{max}$  maximální dovolený posun ve všech směrech. Přepočítá se potenciální energie. K tomu nemusíme počítat energii celého systému, ale k předchozí přičíst pouze část, která se týká námi změněné částice. Následuje výše zmíněný proces porovnávání Boltzmannových faktorů.

Volba  $\delta r_{max}$  se většinou provádí tak, aby bylo  $\sim 50\%$  pokusů přijatých [13]. Zvolíme-li velké  $\delta r_{max}$ , pak bude většina pokusů odmítnuta, protože tento přístup může snadno vést k překryvům částic. Zvolíme-li ho příliš malý, pak bude sice většina pokusů přijata, ale pokryv fázového podprostoru by byl příliš časově náročný.  $\delta r_{max}$  se tedy obvykle upravuje automaticky v průběhu simulace tím, že si zaznamenáváme, kolik pokusů jsme přijali, kolik odmítli a měníme ho tak, abychom se pohybovali okolo výše zmíněných 50%.

#### 4.4.1 'Biased' MC

V některých simulacích pro nás může mít určitá část systému větší význam než jiná. Máme-li například jednu molekulu rozpuštěnou v rozpouštědle, které je simulováno větším počtem molekul, bude nás nejvíce zajímat oblast, kde dochází k nejnápadnějším interakcím rozpouštědla s molekulou. Jednou z metod, jak v tomto případě prozkoumat fázový prostor, může být tzv. preferenční vzorkování. V této metodě zavedeme určitý *cutoff* okolo rozpuštěné molekuly s tím, že molekulám uvnitř tohoto objemu častěji nabízíme změnu. Pokud v nějakém kroku vybereme pro změnu molekulu v *cutoff* objemu, potom jí změníme souřadnice. Pokud v tomto objemu není, pak se vygeneruje náhodné číslo a porovná se s pravděpodobnostním parametrem  $p$ . Pokud je  $p$  větší než vygenerované číslo, zkusíme s molekulou pohnout, v opačném případě zamítneme. Čím menší  $p$  tedy zvolíme, tím 'rigidnější' oblast za *cutoffem* bude. Alternativně můžeme pravděpodobnost volby molekuly realizovat např. jako  $p \sim r^{-n}$  [13].

*Force-bias* vzorkování upřednostňuje pohyb částice ve směru sil, které na ní působí. Vybereme částici, vypočítáme působící sílu a změníme jí souřadnice ve směru, který vybereme z normální distribuce mající pík ve směru této síly. Jinou metodou využívající znalost působících sil je *smart MC*, ve které má polohový vektor změny souřadnic dvě odlišné složky. Jedna je přímo ve směru působící síly, druhá je volena náhodně z normálního rozdělení s nulovou střední hodnotou a variancí  $A$ , kde  $A$  je volitelný parametr.

#### 4.4.2 *J-walking* a Multikanonické MC

Z výše uvedeného postupu můžeme dospět k otázce, jak si MC poradí se systémy, které mají vysokoenergetické potenciálové bariéry. V tomto případě by mohlo dojít k tomu, že se zmapuje pouze malá část energeticky výhodného fázového prostoru, a my nemáme šanci dosáhnout na předpoklad ergodicity. Proto se simulace, ve kterých narazíme na tento problém nazývají kvazi-ergodické.

Zde uvedeme dvě metody, které se snaží tento problém zjemnit. První je tzv. *J-walking* (*Jump-walking*). Jak název napovídá, metoda umožňuje systému občas si vyskočit. To zařizuje změnou teploty. Nejdříve spustíme simulaci s vyšší teplotou (*J-walker*) s tím, že si zapamatujeme buď všechny, nebo reprezentativní vzorek konfigurací systému. Poté spustíme simulaci s nižší teplotou a v náhodných krocích se snažíme zabudovat některou z vysokoenergetických konfigurací. Výběrová pravidla jsou přitom stejná jako v předchozím případě [13]. Abychom měli šanci, že se nějaký ze stavů *J-walkeru* přijme, musíme tedy zajistit, aby se distribuce energií pro jednotlivé teploty překrývaly. Další možností je využití **Multikanonického MC**. V kanonickém souboru je pravděpodobnost vygenerování dané konfigurace [13]

$$P_{kan}(T, E) \propto n(E) w_B(E, T), \quad (4.53)$$



kde  $n(E)$  je hustota stavů a  $w_B = \exp(-E/k_B T)$  je Boltzmannův faktor. V tzv. multikanonickém souboru (jedná se o uměle vytvořený soubor) je tato pravděpodobnost, v určitých energetických intervalech, na energii nezávislá [13]

$$P_{multi}(E) \propto n(E) w_{multi}(E) = konst. \Rightarrow w_{multi}(E) \propto n(E)^{-1}, \quad (4.54)$$

kde  $w_{multi}(E)$  je multikanonický váhový faktor. Největší problém je s tím, jak právě tento váhový faktor volit. Většinou je odvozen iterativně. Uvedme příklad z [13].

Nejdříve je provedena simulace při vysoké teplotě  $T_0$  tak, abychom měli jistotu, že se při ní mohou překonat všechny potenciálové bariéry. Vytvoříme celou  $S(E)$ , kde každý prvek odpovídá energii v rozmezí  $E + \delta E$ , kde  $\delta E$  má hodnotu např. 1 kcal/mol, a všem elementům přiřadíme nulové hodnoty. Z původní simulace vytvoříme histogram tak, že přiřazujeme počet vygenerování konfigurace s danou energií v  $E + \delta E$  do nové cely  $H(E)$ , tedy  $H(E) \propto n(E) \exp(-E/k_B T)$ . Celou  $S(E)$  je aktualizována na

$$S(E) = S(E) + \ln H(E) \quad (4.55)$$

pro všechny energie v rozpětí  $E_{min}$  a  $E_{max}$ , což jsou minimální a maximální energie vygenerované během první simulace a hodnota  $H(E)$  je větší než zadaná tolerance (stav musel být vygenerován s určitou četností). Následně se vypočítají dva parametry

$$\beta(E) = \begin{cases} 1/k_B T_0 & E \geq E_{max} \\ 1/k_B T_0 + \frac{S(E') - S(E)}{E' - E} & E_{min} \leq E \leq E_{max} \\ \beta(E_{min}) & E < E_{min} \end{cases}$$

$$\alpha(E) = \begin{cases} 0 & E \geq E_{max} \\ \alpha(E') + (\beta(E') - \beta(E)) E' & E < E_{max} \end{cases},$$

kde  $E'$  je energie následujícího prvku v celé  $S(E)$ , tj.  $E$  a  $E'$  jsou sousedi. Multikanonický váhový faktor se potom vypočítá

$$w_{multi}(E) = \exp(-\beta(E) E - \alpha(E)). \quad (4.56)$$

Následně jsou spuštěny nové simulace, které toto rozdělení iterativně vylepšují až do bodu, kdy je  $H(E)$  dostatečně plochá v rozsahu energií, které nás zajímají. Když získáme toto finální rozdělení, je pro simulaci jednodušší překonávat vysoko-energetické bariéry. Nakonec můžeme z jedné takové simulace získat kanonickou distribuci při libovolné teplotě

$$P_{kan}(T, E) \propto P_{multi}(E) w_{multi}^{-1} \exp\left(\frac{-E}{k_B T}\right) \quad (4.57)$$

a tu použít k výpočtu středních hodnot chtěných veličin.

### 4.4.3 Modul Amorphous Cell

Modul Amorphous Cell (AC) dostupný v MS slouží, mimo jiné, ke konstrukci struktur s periodickými okrajovými podmínkami. Molekuly v něm ‚rostou‘ fragment za fragmentem. Fragmentem myslíme určitou část molekuly, obvykle definovanou torzním úhlem. To, na kolik fragmentů naši molekulu rozdělíme, si můžeme zvolit. Ostatní stupně volnosti, jako jsou např. délky vazeb, přitom držíme konstantní, což je ospravedlněno tím, že při pokojové teplotě má jejich rozdělení jeden ostrý pík. Pokud molekula neobsahuje žádný torzní úhel, pak se k ní AC chová jako k rigidní molekule. Jedná se většinou o molekuly rozpouštědla, v této práci o molekuly vody. V tomto případě se berou v úvahu

pouze pozice a orientace jednotlivých molekul. AC tyto ‚malé‘ molekuly buduje jako první. Vytvoří tedy krychli (nebo jiný geometrický objekt) na základě zadaných informací o hustotě výsledné struktury a množství molekul, které chceme vybudovat, a následně zabuduje všechny molekuly vody. Až poté se snaží ‚naskládat‘ do vymezeného prostoru větší molekulu.

V AC mohou torzní úhly nabývat různých hodnot spojitě, což je dáno potenciálem torzních úhlů v předpisu silového pole. Předtím, než dojde k umístění nového fragmentu, si AC navrhne určitý počet kandidátů. Každý z nich potom projde kontrolou na blízké kontakty, tj. nepřekrývají-li se atomy v rámci zadaného van der Waalova poloměru. Takových podmínek může být více, například na překryv prstenců cyklických molekul atd. Pro všechny kandidáty, které prošli tímto výběrem, se vypočítá energie  $E_i$ , kterou by systém měl, kdybychom ho do něj vložili. Tato energie se počítá jako příspěvek všech nevazebných interakcí mezi atomy, které chceme umístit a těmi již umístěnými. Na základě těchto energií se vypočítají Boltzmannovy faktory  $w_i = \exp\left(-\frac{E_i}{k_B T}\right)$ . Z nich vypočteme pravděpodobnosti kandidáta  $P_i = \frac{w_i}{\sum_{j=1}^M w_j}$ , kde  $M$  je počet kandidátů. Nakonec vypočítáme kumulativní pravděpodobnost  $Q_i = \sum_{j=1}^i P_j$ ,  $Q_0 = 0$  a  $Q_r = 1$ . Vygenerujeme náhodné číslo od 0 do 1 a pokud toto číslo náleží do intervalu mezi hodnotami  $Q_{k-1}$  a  $Q_k$ , vybíráme kandidáta  $k$ .

Pokud se nepodaří umístit ani jednoho z kandidátů díky překryvu atomů, pak je celá molekula z boxu odstraněna a pokusíme se o vybudování znovu. Počet, kolikrát se může AC pokusit o vybudování je dán počtem *loading* kroků, který zadává uživatel.

AC tedy buduje molekulu fragment za fragmentem. Nastaven je tak, že se dívá jeden krok dopředu, což odpovídá výše popsanému postupu. Počet kroků, na které se dívá, se ale dá navýšit. To je užitečné, pokud budujeme větší struktury, protože navýšením kroků, na které se díváme, zvyšujeme pravděpodobnost, že struktura bude vybudována úspěšně. Tj., může se stát, že AC nezvolí zrovna to nejméně energetické umístění jednoho segmentu v zájmu toho, že bude moci umístit další. Pořád se bude budovat jeden fragment za druhým, ale s tím, že si AC navrhne i možnosti umístění dalšího segmentu a má je na paměti.

Je pravděpodobné, že se po vygenerování výsledné struktury budou v systému stále nacházet nefyzikální překryvy jednotlivých složek. To lze ale většinou vyřešit geometrickou optimalizací.

#### 4.4.4 Modul Conformers

Modul Conformers poskytuje možnost prozkoumat konfigurační prostor molekuly pomocí tří metod, které spadají do dvou skupin. *Systematic grid scan search* patří do systematických metod. Tato metoda systematicky mění torzní úhly. Počet úhlů a periodu, se kterou bude ke změnám docházet si můžeme volit. Využití této metody je vhodné pouze pokud se zajímáme o malý počet torzních úhlů. V jiném případě vede k velmi vysokému počtu možností (kombinatorická exploze), které metoda musí projít. Další dvě metody spadají do stochastické kategorie. Jsou to metody náhodného vzorkování a Boltzmannova skoku.

Metoda **náhodného vzorkování** narušuje počáteční konformaci tím, že náhodně mění hodnoty všech torzních úhlů v molekule. Například torznímu úhlu  $A$  je přiřazena hodnota v rozmezí od  $(A - 0, 5w)$  do  $(A + 0, 5w)$ , kde  $w$  je námi specifikovaná limitní hodnota torzního úhlu [7]. Každá z takto vygenerovaných konformací je pak uložena jako kandidát. Proces opakujeme, dokud nedostaneme požadovaný počet kandidátů. To, jestli se narušení struktury bude provádět vždy na původní konformaci nebo na předchozí konformaci, si můžeme zvolit. Můžeme tedy buď prozkoumat okolí konfiguračního prostoru původní molekuly nebo naopak větší část fázového prostoru.

Druhou metodou je metoda **Boltzmannova skoku**. Tak jako v předešlé metodě narušuje torzní úhly v zadané toleranci, rozdíl je v tom, že tentokrát vygenerované konformace procházejí Metro-polisovými výběrovými pravidly.

Nevýhodou je, že Conformers neumožňují generaci konfigurací molekuly v rozpouštědle.

## 4.5 Radiální distribuční funkce a tepelná kapacita

Radiální distribuční funkce  $g(r)$  udává pravděpodobnost nalezení částice ve vzdálenosti  $r$  od jiné částice v porovnání s distribucí ideálního plynu. Pro krátké vzdálenosti je vzhledem k odpudivým silám nulová. Pro kapalinu dochází ve větších vzdálenostech ke konvergenci  $g(r)$  k jedničce, tedy ke stejné hodnotě jakou má ideální plyn, což svědčí o její neuspořádanosti na větších vzdálenostech. Pro výpočet radiální distribuční funkce ze simulace rozdělíme sousedy každé částice do binů. To znamená, že počítáme sousedy například ve vzdálenostech 2,5 Å až do 2,75 Å atd. Počet těchto sousedů je potom na konci simulace v každém binu zprůměrován.

Pro výpočet izochorické tepelné kapacity

$$C_V = \left( \frac{\partial U}{\partial T} \right)_V, \quad (4.58)$$

můžeme buď provést sérii výpočtů při různých teplotách a pak derivovat vnitřní energii buď numericky nebo získáním analytické funkce fitem a její následnou derivací. Tepelnou kapacitu ale můžeme vypočítat i z jedné simulace, platí [13]

$$C_V = \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2} = \frac{\langle (E - \langle E \rangle)^2 \rangle}{k_B T} \quad (4.59)$$

a můžeme ji tedy vypočítat ze zaznamenaných hodnot energie na konci simulace.

## Kapitola 5

### Nastavení počítačové simulace

Molekuly navržené pro procesy přepracování jaderného paliva by měly operovat v kyselém prostředí radioaktivních izotopů. Zkoumá se proto, mimo jiné, jejich chování v prostředí s kyselinou a odolnost vůči produktům gama záření.

Některé ze systémů, jejichž simulace je provedena v této práci proto byly ve skutečném experimentu ozařovány  $\gamma$ -fotony. Zahrnout do simulace přímo produkty radiolýzy prostředí, ve kterém se molekuly nacházejí, není v rozsahu této práce možné vzhledem k jejich malé koncentraci [35]. Jedinou možností, jak tento fakt do simulace molekulární dynamiky zahrnout tak zůstává prostřednictvím zvýšené teploty.

Pro vodu, která absorbovala dávku 150 kGy se teplota celého systému změnila o  $\sim 36$  K nebo  $\sim 0,004$  eV, a tato informace v žádné ze simulací zahrnuta není.

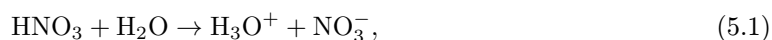
V této práci jsou simulovány systémy, které reprezentují experimenty [35] a [36].

V prvním případě se jedná o simulace molekul TMDGA, TEDGA, Me-TEDGA a Me<sub>2</sub>-TEDGA, které jsou zobrazeny na obrázku 5.1. Jejich obecné vlastnosti jsou zapsány v tabulce 5.1. Tyto molekuly byly simulovány ve vodném prostředí. V případě molekuly TEDGA byly provedeny simulace odrážející vliv koncentrace kyseliny a teploty v rozpětí, ve kterém proběhlo experimentální měření. Pro všechny tyto molekuly byly navíc provedeny simulace s koncentracemi 1, 2, 3 a 4 mol·L<sup>-1</sup> HNO<sub>3</sub> za pokojové teploty, které nelze přímo porovnat s experimentem a byly spuštěny za účelem získu nejpravděpodobnější konformace ligandů.

Voda byla simulována se zanedbáním její autoprotolýzy, kdy dochází ke vzniku hydroxoniového kationtu H<sub>3</sub>O<sup>+</sup> a hydroxidového aniontu OH<sup>-</sup>. Koncentrace těchto iontů jsou velmi nízké, na  $\sim 6 \cdot 10^8$  molekul vody připadá jeden pár iontů [18] a vzhledem k typickému počtu  $\sim 1110$  molekul vody simulovaných v této práci, se jedná o informaci, kterou je nutné zanedbat.

Dále jsou zanedbány možné izotopické změny ve složení vody. V případě vodíku se v přírodě vyskytují dva stabilní izotopy, protium <sup>1</sup>H a deuterium <sup>2</sup>H, přičemž protium má v přírodě zastoupení  $\sim 99,9$  %. V případě kyslíku se přírodně vyskytují tři stabilní izotopy, <sup>16</sup>O, <sup>17</sup>O a <sup>18</sup>O. <sup>16</sup>O má zastoupení  $\sim 99,7$  % všech přírodně se vyskytujících kyslíků a voda tedy byla simulována jako molekuly <sup>1</sup>H<sub>2</sub> <sup>16</sup>O.

Molekula TEDGA byla při experimentech rozpuštěna v kyselině dusičné a chlorovodíkové. HNO<sub>3</sub> i HCl jsou silné kyseliny a byla předpokládána jejich úplná disociace



Speciální nastavení atomových typů bylo provedeno pouze pro chlorový aniont, kdy byl nastaven typ cl-, vybraný z listu parametrů silového pole COMPASS. Kromě vody byly ponechány obecné

atomové typy přiřazené COMPASSu.

V druhé experimentální práci [36] bylo zkoumáno radiolytické chování molekuly CMPO v čistém dodekanu, v dodekanu, který byl v kontaktu s roztokem kyseliny dusičné a v dodekanu, který byl s  $\text{HNO}_3$  přímo smíchán a následně byly obě fáze odděleny odstředivkou.

Dodekan ( $\text{C}_{12}\text{H}_{26}$ ) má 355 strukturálních izomerů. V této práci byl simulován pouze v lineární variantě. Nebylo simulováno rozhraní dodekan/kyselé prostředí. Kyselina byla do souboru vložena přímo v molekulární formě, což odpovídá případu přímého smíchání kyseliny s dodekanem.

Dodekan je hydrofobní molekula a autorka v literatuře nenašla experimentální práci, která by se věnovala extrakci  $\text{HNO}_3$  do čistého dodekanu. Z experimentů ale plyne, že vliv kyseliny na chování molekuly CMPO je výrazný i v případě smíchání a následném oddělení obou fází a určitá koncentrace kyseliny tedy musí v dodekanu být přítomná. Vliv kyselého prostředí se projevuje u různých extraktantů odlišnými způsoby, se zvyšováním koncentrace  $\text{HNO}_3$  ale dochází k zesílení jejího efektu na danou molekulu, dokud nedojde k určité saturaci.

Bylo proto čerpáno z článku [37], ve kterém je, mimo jiné, referována extrakce  $\text{HNO}_3$  do roztoku s 5% TBP (tributyl fosfát ( $\text{CH}_3\text{CH}_2\text{CH}_2\text{CH}_2\text{O}$ ) $\text{PO}$ ) v dodekanu. TBP je polární molekula a na extrakci kyseliny bude mít zásadní vliv, jedná se tak o horní odhad koncentrace  $\text{HNO}_3$  v dodekanu. V této práci je molekula CMPO simulována v čistém dodekanu a v dodekanu po kontaktu s  $5 \text{ mol}\cdot\text{L}^{-1}$ , což dle [37] bylo aproximováno  $\sim 0,16 \text{ mol}\cdot\text{L}^{-1}$   $\text{HNO}_3$  v dodekanu. Molekula CMPO je zobrazena na obrázku 7.8.

Velikosti simulovaných systémů a délka jednotlivých simulací je zaznamenána v tabulce 5.2. Systémy označené v této tabulce jako 1, 3, 4, 5, 6 byly simulovány pro všechny diglykolamidy, přičemž se v souboru vyskytovala vždy pouze jedna z těchto molekul. Hustota v posledním sloupci tabulky odpovídá po řadě hustotám systémů s molekulami TMDGA, TEDGA, Me-TEDGA a  $\text{Me}_2$ -TEDGA. Jejich koncentrace v experimentu byla  $0,05 \cdot \text{L}^{-1}$ . Teplota byla nastavena na  $25^\circ\text{C}$ .

Systémy označené číslem 2 a 6 byly simulovány s molekulou TEDGA v teplotním rozsahu  $25^\circ\text{C}$ ,  $45^\circ\text{C}$  a  $65^\circ\text{C}$ .

Systémy 9 a 10 byly simulovány s jednou molekulou CMPO při teplotě  $25^\circ\text{C}$  a systém odpovídá koncentraci  $0,1 \text{ mol}\cdot\text{L}^{-1}$  CMPO v roztocích.

Jako produkční běh je označen časový interval, ze kterého byla získána data pro následné zpracování, tj. v tabulce 6 ns/ 5ns znamená, že se čerpalo z posledních 5 ns 6-ti ns simulace. Všechny simulace byly provedeny v kanonickém souboru kontrolovaném NHL termostatem s Q-poměrem 0,01 a rozpadovou konstantou 1 ps. Integrační krok byl volen na 1 fs.

Tabulka 5.1: Obecné informace o molekulách TMDGA, TEDGA, Me-TEDGA,  $\text{Me}_2$ -TEDGA [10] a CMPO [36].

Molekula	vzorec	Molární hmotnost	Název
TMDGA	$\text{C}_8\text{H}_{16}\text{N}_2\text{O}_3$	188, 23	<i>N, N, N', N'</i> -tetrametyl-diglykolamid
TEDGA	$\text{C}_{12}\text{H}_{24}\text{N}_2\text{O}_3$	244, 34	<i>N, N, N', N'</i> -tetraetyl-diglykolamid
Me-TEDGA	$\text{C}_{13}\text{H}_{26}\text{N}_2\text{O}_3$	258, 36	2-(2-(diethylamino)-2-oxoetoxy)- <i>N, N</i> -dietylpropanamid
$\text{Me}_2$ -TEDGA	$\text{C}_{14}\text{H}_{28}\text{N}_2\text{O}_3$	272, 39	2,2'-oxybis( <i>N, N</i> -dietylpropanamid)
CMPO	$\text{C}_{24}\text{H}_{42}\text{NO}_2\text{P}$	407, 58	octylphenyl- <i>N, N</i> -diisobutylcarbamoylmethyl phosphine oxide

## 5.1 Volba silového pole

Silové pole bylo vybíráno nejdříve pro molekuly ve vodném prostředí, v tomto případě je tedy nutné zvolit takové, které co nejlépe popisuje vodu.

Tabulka 5.2: Údaje o velikostech a časovém rozpětí simulovaných systémů. Křížkem je značen počet.

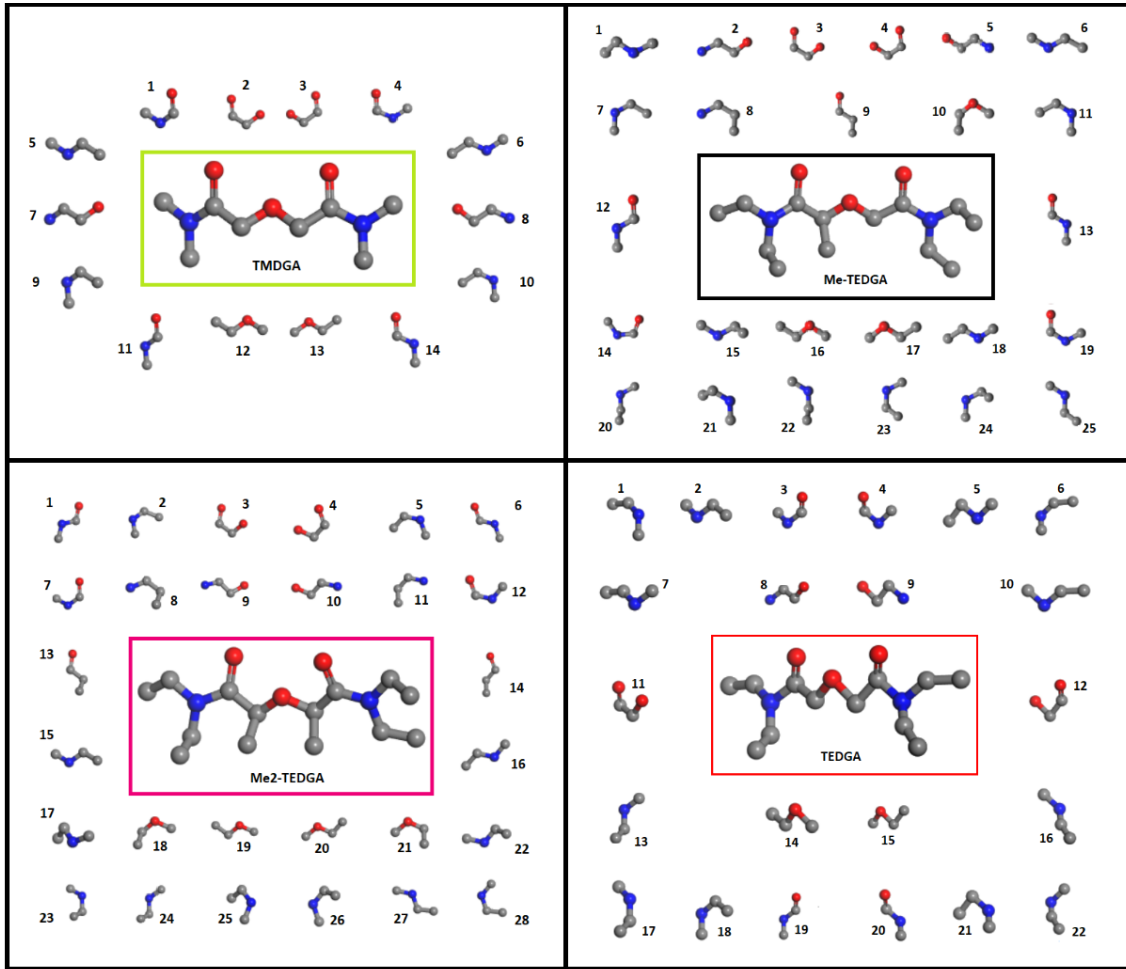
	prostředí	# voda	# $\text{H}_3\text{O}^+$	# $\text{NO}_3^-$	celkový čas/ produkční běh	hustota systému [ $\text{g}\cdot\text{cm}^{-3}$ ]
1	voda	1110			6 ns/ 5 ns	1,009; 1,012; 1,013; 1,014
2	0,5 M $\text{HNO}_3$ /voda	1100	10	10	6 ns/ 5 ns	1,044
3	1 M $\text{HNO}_3$ /voda	1090	20	20	6 ns/ 5 ns	1,072; 1,075; 1,076; 1,077
4	2 M $\text{HNO}_3$ /voda	1070	40	40	6 ns/ 5 ns	1,135; 1,138; 1,139; 1,140
5	3 M $\text{HNO}_3$ /voda	1050	60	60	6 ns/ 5ns	1,198; 1,201; 1,202; 1,203
6	4 M $\text{HNO}_3$ /voda	1030	80	80	6 ns/ 5 ns	1,261; 1,264; 1,265; 1,266
				# $\text{Cl}^-$		
7	5 M $\text{HCl}$ /voda	1009	101	101	6 ns/ 5 ns	1,195
8	8 M $\text{HCl}$ /voda	951	159	159	6 ns/ 5 ns	1,304
		# dodekan	# $\text{HNO}_3$			
9	dodekan	59			10 ns/ 6 ns	1,041
10	dodekan/0,16 M $\text{HNO}_3$	59	2		10 ns/ 6 ns	1,051

V rámci MS je člen s přímou referencí na výpočet vodíkových vazeb obsažen pouze ve předpisu silového pole Dreiding prostřednictvím Lennard-Jonesova potenciálu 12-10. Lennard-Jonesův potenciál 12-6 je zde popsán jako člen zahrnující van der Waalsovou interakci.

Byla porovnána celkem tři silová pole, Dreiding, COMPASS a pccf, prostřednictvím radiálních distribučních funkcí.

Pro každý ze tří předpisů SP bylo vybudováno 30 mikrostavů s 1110 molekulami vody v modulu Amorphous Cell, které byly geometricky zoptimalizovány. Následně byl vybrán ten s minimální energií a na něm byla spuštěna NVT dynamika při 298 K, která byla kontrolována NHL termostatem. Elektrostatické síly byly počítány Ewaldovou sumací, van der Waalsovy *atom-based* sčítací metodou (párové interakce). Délka simulací byla 500 ps s krokem 1 fs. V žádném ze tří případů nebyl volen speciální typ parametrů silového pole, náboje byly přiřazeny automaticky. Následně byly vypočítány radiální distribuční funkce  $g_{\text{OO}}(r)$ ,  $g_{\text{OH}}(r)$  a  $g_{\text{HH}}(r)$ . Výsledek je zobrazen na obrázku 5.2 a popsán v tabulce 5.3.

Prvním krokem k ověření výsledků byl výpočet koordinačního čísla vody, tj. plochy pod prvním



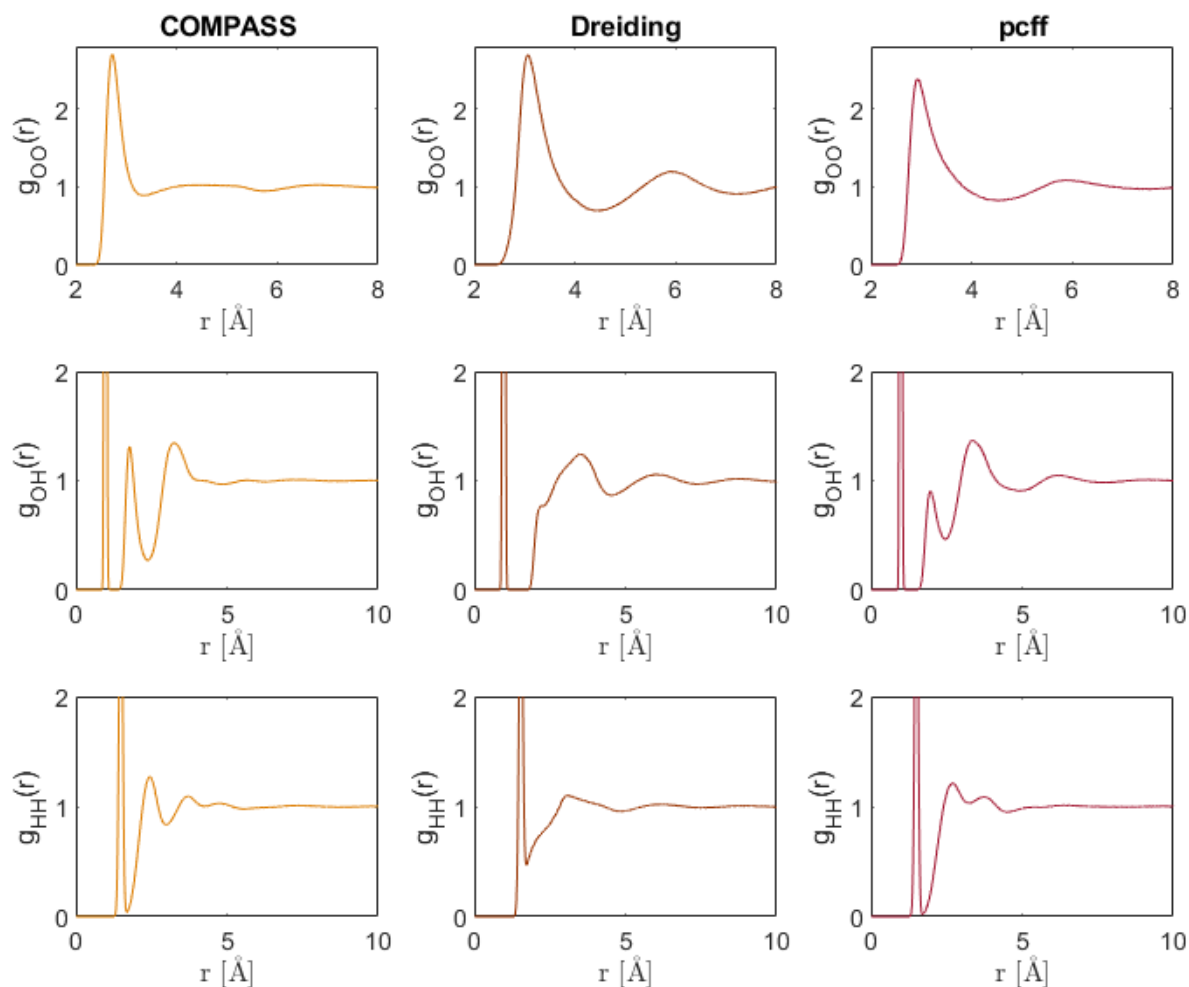
Obrázek 5.1: Molekuly TMDGA, Tedga, Me-Tedga a Me<sub>2</sub>-Tedga zobrazené spolu se zadanými torzními úhly. Šedou barvou jsou vyznačeny uhlíky, červenou kyslíky a modrou dusíky. Molekuly jsou zobrazeny bez příslušných vodíků.

píkem v  $g_{OO}$  [18]

$$N_c = 4\pi\rho \int_0^{r_1} r^2 g_{OO}(r) dr, \quad (5.3)$$

kde  $r_1$  je poloha prvního minima  $g_{OO}$  a  $\rho \sim 33,37 \cdot 10^{27} \text{ m}^{-3}$  je objemová hustota vody. Bylo získáno  $N_c(\text{COMPASS}) \sim 4,5$ ,  $N_c(\text{Dreiding}) \sim 11,5$  a  $N_c(\text{pcff}) \sim 12,5$ . Voda by měla nabývat hodnoty  $N_c$  v rozpětí 4–5 [20], kdy  $N_c = 4$  odpovídá tetrahedrálnímu uspořádání ledu. Rozhodně není přípustná hodnota  $N_c(\text{pcff})$  ani  $N_c(\text{Dreiding})$ . Tyto výsledky jsou zajímavé tím, že pcff a COMPASS mají stejný funkční předpis, liší se ale svou parametrizací.

Pro porovnání radiální distribuční funkce  $g_{OO}$  byl vybrán experiment využívající rentgenovou difrakci. Byl proveden v roce 2000 na synchrotronu ALS (Advanced Light Source) v LBNL. Bližší informace jsou v [22]. Porovnání těchto výsledků s výsledky COMPASSu je na obrázku 5.3. Pro experimentální data  $g_{OH}$  a  $g_{HH}$  byl využit experiment neutronové difrakce [23] také z roku 2000.



Obrázek 5.2: Radiální distribuční funkce  $g_{OO}$ ,  $g_{OH}$  a  $g_{HH}$  pro vodu. V prvním sloupci jsou výsledky z COMPASSu, ve druhém z Dreidingu, ve třetím z pcff.

Data byla získána z [24].

Po analýze dat  $g_{OO}$  docházíme k závěru, že polohy minim a maxim na ose  $x$  relativně dobře souhlasí s COMPASSem. Výraznější je rozdíl hodnot  $g_{OO}$ . COMPASS nevytváří vodu tak strukturovanou, jak by měl.

Na obrázku 5.4 jsou porovnány experimentální výsledky (barevně) pro  $g_{OH}$  a  $g_{HH}$  s výsledky z COMPASSu (přerušované). Nejdříve si všimneme prvního píku, který popisuje intramolekulární vzdálenosti.  $g_{OH}$  má první maximum v  $\sim 0,97$  a oproti experimentu je střední hodnota délky vazby OH mírně kratší. Zásadnější rozdíl je v šířce píku. COMPASS vazbám O-H nedává volnost, jakou by měl. Podobně jsou na tom intramolekulární vzdálenosti H-H. Docházíme tedy k závěru, že COMPASS má, v případě vody, příliš vysoké hodnoty silových konstant u *bond-stretch* a *angle-bending* členů nebo přiřazuje nadhodnocený náboj kyslíku a podhodnocený vodíku.

Maxima i minimum v následujícím průběhu  $g_{OH}$  COMPASS posunul směrem k menšímu  $r$  a to



Tabulka 5.3: Polohy minim a maxim jednotlivých radiálních distribučních funkcí spolu s jejich hodnotami v těchto bodech.

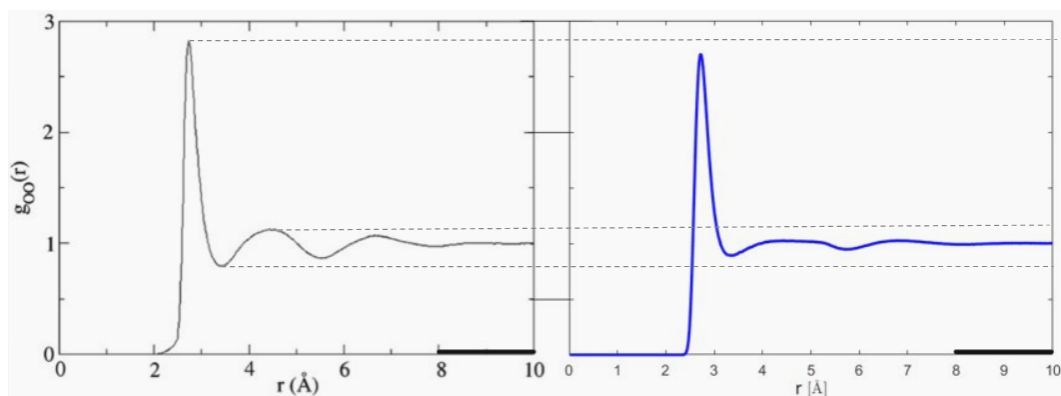
O-O						
	COMPASS		Dreiding		pcff	
	r [Å]	g(r)	r [Å]	g(r)	r [Å]	g(r)
<b>1.min</b>	3,35	0,89	4,39	0,69	4,55	0,83
<b>2.min</b>	5,73	0,95	7,13	0,91	7,55	0,97
<b>3.min</b>	8,29	0,99	9,99	0,97	10,17	0,99
<b>4.min</b>	-	-	-	-	-	-
<b>1.max</b>	2,73	2,70	3,07	2,69	2,91	2,38
<b>2.max</b>	4,43	1,03	5,89	1,19	5,89	1,09
<b>3.max</b>	6,83	1,03	8,65	1,05	8,81	1,02
<b>4.max</b>	-	-	-	-	-	-

O-H						
	COMPASS		Dreiding		pcff	
	r [Å]	g(r)	r [Å]	g(r)	r [Å]	g(r)
<b>1.min</b>	1,17	0,00	1,17	0,00	1,15	0,00
<b>2.min</b>	2,39	0,27	4,55	0,87	2,45	0,46
<b>3.min</b>	4,83	0,97	7,47	0,97	4,85	0,91
<b>4.min</b>	6,31	0,99	-	-	7,65	0,98
<b>1.max</b>	0,97	37,72	0,97	32,45	0,99	36,14
<b>2.max</b>	1,77	1,31	3,47	1,24	1,95	0,90
<b>3.max</b>	3,25	1,35	5,97	1,06	3,37	1,37
<b>4.max</b>	5,76	1,00	8,71	1,02	6,29	1,05

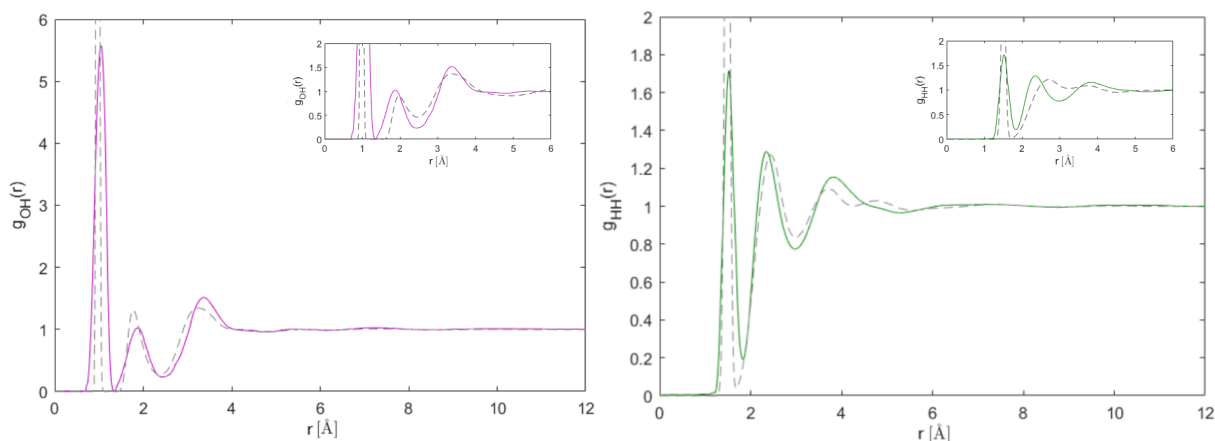
H-H						
	COMPASS		Dreiding		pcff	
	r [Å]	g(r)	r [Å]	g(r)	r [Å]	g(r)
<b>1.min</b>	1,69	0,04	1,75	0,48	1,71	0,02
<b>2.min</b>	3,01	0,83	4,81	0,95	3,25	1,03
<b>3.min</b>	4,23	1,00	7,63	0,99	4,51	0,9494
<b>4.min</b>	5,55	0,98	9,97	0,99	-	-
<b>1.max</b>	1,49	3,87	1,55	3,65	1,51	3,87
<b>2.max</b>	2,43	1,27	3,07	1,1	2,69	1,22
<b>3.max</b>	3,71	1,09	6,19	1,02	3,71	1,09
<b>4.max</b>	4,75	1,03	9,05	1,01	-	-

přibližně o hodnotu  $\sim 0,1$  Å, což se vzhledem k právě diskutovanému prvnímu píku nejeví špatně a nasvědčuje to tomu, že problémem jsou spíše výše uvedené silové konstanty než přiřazené náboje. V případě druhého píku  $g_{OH}$  je nadhodnocen i počet molekul ve druhé koordinační sféře. Naopak ve třetí sféře se jeví jako slabě podhodnocen. V  $g_{HH}$  se pak vyskytuje zvláštní pik v  $r \sim 4,75$  Å, který lze nejspíš opět přisoudit malé volnosti, kterou COMPASS přiřazuje vazbě O-H.

V pravých horních rozích grafů na obrázku 5.4 jsou zaneseny výsledky experimentu spolu s výsledky pcff. Jeho volbu pro simulaci vody zamítáme vzhledem k průběhu  $g_{OO}$ .



Obrázek 5.3: Na levé straně je zobrazena experimentálně získaná RDF  $g_{OO}$  [22], na pravé straně  $g_{OO}$  z COMPASSu. Černá úsečka na ose x od 8-10 odpovídá stejné vzdálenosti  $r$  na obou grafech.



Obrázek 5.4: Na levé straně je zobrazena  $g_{OH}$ , na pravé  $g_{HH}$ . Přerušovanou čarou jsou označeny výsledky COMPASSu, barevně jsou vyneseny experimentální hodnoty [24]. V pravých horních rozích jsou porovnány výsledky z pcff (přerušovaně) a experimentu (barevně).

Nicméně z podobnosti funkčních závislostí  $g_{OH}$  a  $g_{HH}$  pcff a experimentu, se můžeme domnívat, že popis ne vazebných sil v tomto typu předpisu silového pole relativně dobře odpovídá vodnému prostředí a dává lepší výsledky než Dreiding, který v sobě člen, snažící se přímo o popis vodíkových vazeb, zahrnuje.

Dále byly zdefinovány atomové typy silového pole. V COMPASSu byl pro kyslík zvolen typ  $o2^*$ ,  $sp^3$  hybridizovaný kyslík ve vodě. Pro vodík typ  $h1o$ , vodík vázaný ke kyslíku nebo fluoru. Na stejném stavu jako v předešlé simulaci byla spuštěna dynamika se stejnými, výše uvedenými parametry. Nebyla pozorována žádná změna v radiálních distribučních funkcích.

Výpočet proveden s SP COMPASS a jeho obecnými parametry dává pro tepelnou kapacitu  $C_V$  hodnotu  $\sim 107,8 \text{ Jmol}^{-1}\text{K}^{-1}$ . Pro simulaci s definovanými atomovými typy vyjde  $C_V \sim 81,2 \text{ Jmol}^{-1}\text{K}^{-1}$ . Správná hodnota by měla odpovídat  $\sim 74,9 \text{ Jmol}^{-1}\text{K}^{-1}$  [25]. Můžeme tedy říci, že výběr atomového typu, jiné sady parametrů, vedl ke změně v energii, která je poznat na hodnotě

$C_V$ .

Výběr atomových typů byl proveden i pro Dreiding. Kyslík byl zdefinován jako typ O...J,  $sp^3$  hybridizovaný kyslík a vodík jako typ H...A, vodík ve vodíkových vazbách. Opět bez zřetelného zlepšení RDF.

Diskutovaná simulace nicméně byla pouze 500-ti ps a byla proto spuštěna 2 ns simulace vody v COMPASSu se zvolenými atomovými typy.

Nedošlo k žádným změnám v RDF,  $C_V$  mírně klesla na  $\sim 79 \text{ Jmol}^{-1}\text{K}^{-1}$ .

Výsledkem je, že silové pole COMPASS je v rámci MS nejlepší volbou co se týče popisu vody a bylo použito ve všech následujících výpočtech spolu se zmíněnými definovanými atomovými typy.

SP COMPASS byl použit i ve výpočtech, kde voda obsažena není, protože se jedná o nejlepší možnou volbu v rámci MS. Nebude-li uvedeno jinak, pak v těchto simulacích nebyly nastaveny žádné speciální parametry silového pole.

## Kapitola 6

### Konformační analýza

Pro konformační analýzu byl vybrán prostor torzních úhlů, které byly průběžně měřeny z kartézských souřadnic atomů v simulaci. Volbou tohoto měřítka byly zanedbány délky vazeb i úhly mezi třemi atomy.

K vybudování simulovaných systémů a získání počátečních konformací daných molekul byl použit modul Amorphous Cell. Nebude-li uvedeno jinak, pak počty molekul a celková hustota systému byla počítána z údajů o koncentracích uvedených v popisech jednotlivých experimentů. Všechny vybudované buňky byly kubické a u všech byla aplikována podmínka na blízké kontakty v podobě 40 % van der Waalsových poloměrů jednotlivých atomů. Elektrostatické interakce byly počítány Ewaldovou sumací, van der Waalsovy sčítáním přes páry atomů s *cutoffem* 16 Å (obvyklá délka hrany simulačního boxu  $\sim 32$  Å). Pro všechny systémy bylo použito silové pole COMPASS. Výsledné struktury byly geometricky optimalizovány algoritmem SMART (automatická posloupnost kroků metod minimalizace nejstrmějšího sestupu, ABNR a kvazi-Newtonovy metody) s omezením maximálního počtu iterací na 7500 bez možnosti optimalizace cely.

Pro každou z molekul TMDGA, TEDGA, Me-TEDGA a Me<sub>2</sub>-TEDGA bylo vybudováno 300 mikrostavů s cílem uspokojivého pokrytí fázového souřadnicového podprostoru. Z nich byly získány informace o hodnotách torzních úhlů daných molekul a napsán krátký kód v Matlabu, který

- Hledá stejné konformace v rámci zadané odchylky 20° pro každý z úhlů.
- Hledá zrcadlové obrazy molekul v rámci zadané odchylky 20°.
- Hledá molekuly, které jsou ze symetrického hlediska identické v rámci zadané odchylky. Například pro molekulu TMDGA to znamená, že dojde k vzájemnému přecházení úhlů 1 a 4, 2 a 3 atd. a k porovnání s hodnotami původními. Dále následuje první krok. Tento druh porovnání lze použít pouze pro molekuly, které by v určitém případě mohly mít dvojčetnou osu symetrie.

Mikrostavy molekul, které v jednotlivých krocích odhalí, doporučí k vymazání.

Ve druhém kroku program hledá enantiomery. Ty můžeme z našich úvah vyloučit pouze pokud je prostředí, ve kterém se nacházejí, achirální. Tříatomová molekula vody z definice vlastnost chiralitity mít nemůže. Mimo definice existuje názor, že voda chirální je [26], nicméně v této práci je považována za prostředí achirální.

To ale mimo jiné znamená, že by AC měl enantiomerní páry generovat se stejnou pravděpodobností, k čemuž nedošlo. To lze ale přiřadit příliš malému počtu vygenerovaných konformací, což vyplývá i z tabulky 6.1. Z ní vyplývá, že AC nemá problém s generováním odlišných nízkoenergetických struktur a není důvod se domnívat, že by se v tomto vzorku museli nacházet enantiomery.

V tabulce 6.1 jsou pro každou z molekul uvedeny výsledné počty mikrostavů, které můžeme z

dalších výpočtů vynechat.

Tabulka 6.1: Počet stavů, které můžeme z jednotlivých kroků programu vymazat, při tolerančním úhlu 20°. Jako „fin“ je označen zbylý počet mikrostavů, se kterými musíme počítat.

	1.krok	2.krok	3.krok	4.krok	fin
<b>TMDGA</b>	86	3	6	6	<b>199</b>
<b>TEDGA</b>	174	0	0	0	<b>60</b>
<b>Me-TEDGA</b>	2	0	0	0	<b>298</b>
<b>Me2-TEDGA</b>	4	0	0	0	<b>296</b>

V dalším kroku byly na 60 strukturách TEDGy ve vodě spouštěny 1 ns simulace s cílem zjistit, zda po tomto časovém intervalu nebudou nabývat stejných konformací. Pro tyto simulace byl vzhledem ke krátkému časovému rozmezí simulace volen Berendsenův termostat. K určení středních hodnot, kterých může každý zvolený torzní úhel nabývat, byl napsán kód, jehož úkolem je především identifikace intervalů, ve kterém se střední hodnoty mají počítat. Problém je ilustrován na obrázku 6.1. Kód realizuje následující kroky

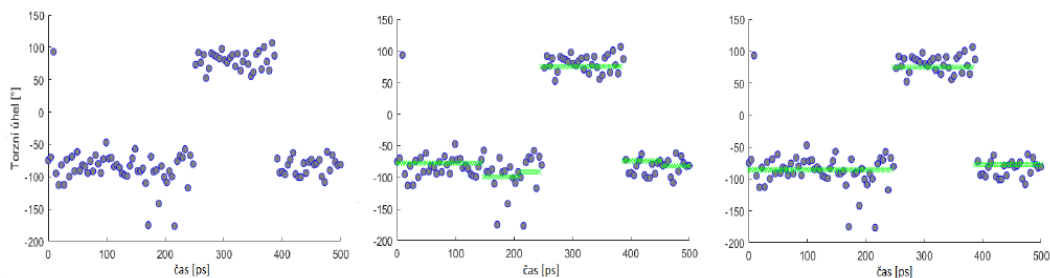
- Hodnoty torzního úhlu z prvních 20 snímků proloží lineární funkcí a zapamatuje si absolutní hodnotu jejího sklonu. Pokud tato hodnota není větší než zadaný parametr, pak k původním 20-ti snímkům připojí dalších dvacet a operace se opakuje. Hodnota sklonu  $a$  byla empiricky volena na  $a=80$ .
- Pokud se v následujícím kroku stane, že je hodnota sklonu větší než  $a$ , zafixuje se snímek s nejvyšším pořadovým číslem (pro nás nyní snímek 60), zleva se začne zkracovat interval a hledá se maximální hodnota sklonu lineárního fitu. Po nalezení mezního snímku se tento zafixuje a interval se začne prodlužovat doprava se stejným cílem.
- V určených intervalech se vypočítá střední hodnota a střední kvadratická odchylka. Tímto postupem se získá graf analogický příkladu uvedenému na obrázku 6.1 uprostřed.
- Nalezne intervaly, ve kterých se střední hodnoty v rámci svých odchylek rovnají, přepočítá střední hodnotu a odchylku pro soubor všech těchto dat a daným intervalům tuto hodnotu přiřadí. Výsledkem je třetí graf na obrázku 6.1.
- Získané střední hodnoty porovná v rámci střední kvadratické odchylky pro všechny simulace.

Ze 60-ti simulací pro TEDGu nabývalo stejných hodnot všech torzních úhlů 58 souborů (bez uvážení závislosti na příslušných intervalech). Zbylé dva soubory se lišily v jednom torzním úhlu, přičemž mu přiřazovaly pouze jednu hodnotu. V ostatních simulacích tento torzní úhel nabýval hodnot dvou, přičemž vždy obsahovaly hodnotu obsaženou v daných dvou souborech.

Vzhledem k tomu, že u některých úhlů byly pozorovány periodické změny např. mezi dvěma možnými hodnotami jednoho torzního úhlu, byl učiněn předpoklad, že molekuly mají nízké potenciálové bariéry úměrné tepelnému pohybu a soubory, u kterých nedošlo k úplné shodě ve všech úhlech jsou pouze projevem nedostatečně dlouhého časového vývoje systému.

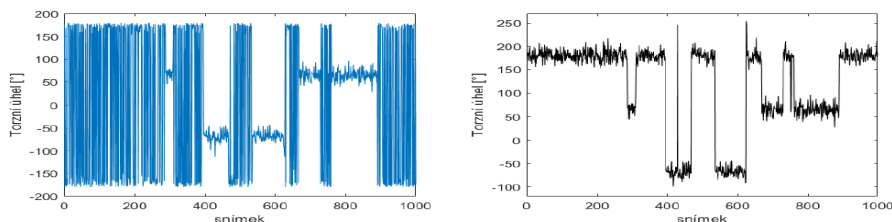
Z výše uvedené analýzy vyplývá, že již pro krátké simulace (a malé molekuly) je vliv počátečního konformačního stavu zanedbatelný a nejpravděpodobnější konformace, kterých molekula nabývá, je možné hledat v jedné delší simulaci.

Pro samotnou konformační analýzu není ale předchozí program vhodný především vzhledem ke dvěma skutečnostem



Obrázek 6.1: Na prvním grafu je ilustrován průběh dat. Na druhém a třetím grafu jsou zeleně naznačeny vypočtené střední hodnoty v různých fázích zpracování.

- Úhly jsou měřeny ve stupních, v rozmezí  $\pm 180^\circ$ . Pokud se jedná o úhel, jehož střední hodnota se pohybuje okolo těchto krajních hodnot, pak je pro správné určení střední hodnoty nutný přesun dat. V této verzi kódu je přesun dat korigován zavedením podmínek na velikosti úhlů, čímž se do dat mohou zavést hrubé chyby. Příklad je znázorněn na obrázku 6.2.



Obrázek 6.2: Zavedení hrubé chyby nesprávným převodem dat. Nalevo jsou původní data, napravo jsou zobrazena po příslušném přesunu.

- Kód ze své podstaty nemůže odhalit dvě různé střední hodnoty, mezi kterými může úhel s vysokou frekvencí překmitávat.

Byl proto napsán nový kód, jehož kroky budou popsány v následující podkapitole. Všechny použité funkce, které budou v textu přímo pojmenovány, jsou funkce vestavěné v Matlabu. Pokud bude některá z nich popsána, pak je čerpáno výhradně z [33].

## 6.1 Program pro konformační analýzu

V první řadě je učiněn předpoklad o normálním rozdělení dat. Je třeba poznamenat, že žádná z dat testovaných v této práci úspěšně neprojdou standardním testem normality jako je Kolmogorův-Smirnovův test, který porovnává distribuční funkce experimentálních dat a předpokládaného rozdělení. Jeho statistikou je absolutní hodnota rozdílu těchto distribucí.

Za obecnou charakteristiku většiny testovaných dat lze považovat vysokou hodnotu koeficientu špičatosti. Oproti normálnímu rozdělení se tak většina dat vyskytuje v blízkosti střední hodnoty.

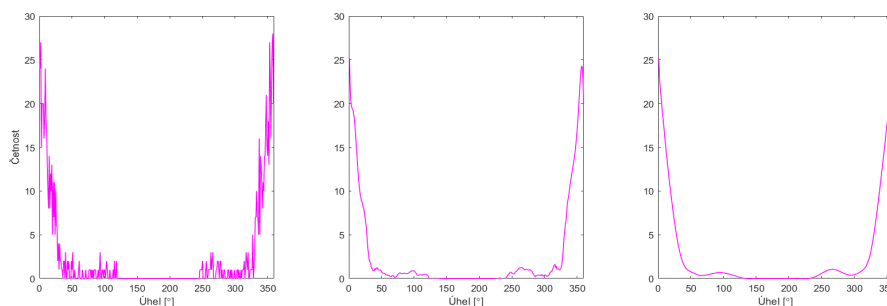
Není ale důvod se domnívat, že by hodnoty torzních úhlů měly nabývat jiného rozdělení než normálního.

Vstupní informací programu je textový soubor s uloženými daty o časovém průběhu torzních úhlů

zadejnovaných pro danou molekulu, přičemž každý sloupec odpovídá jednomu torznímu úhlu. Kód pracuje v následujících krocích:

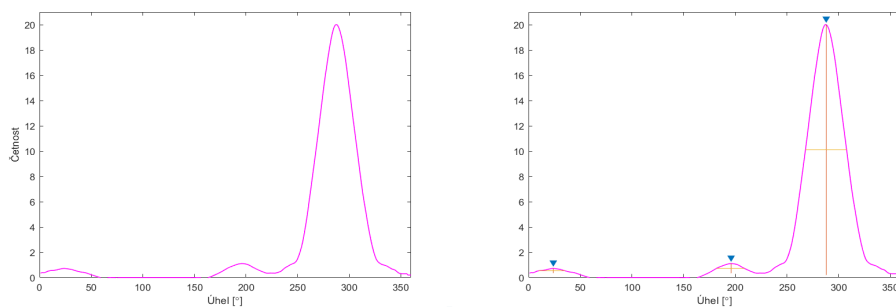
- Pro jednotlivé úhly vytvoří histogramy s šířkou binu  $1^\circ$ .
- Získané histogramy vyhladí ve dvou krocích
  1. Savitzky-Golay filtrem s krokem 41 dat a polynomem sedmého stupně. Jeho princip je následující: prvních 41 hodnot se proloží polynomem sedmého stupně metodou nejmenších čtverců. Z výsledného předpisu funkce se přiřadí funkční hodnota prostřednímu bodu, nyní tedy bodu dvacátému. Takto algoritmus přiřadí hodnoty všem hodnotám kromě prvních a posledních dvaceti. V Matlabu se k jejich určení používá metoda přesunu dat, kdy se před první naměřená data přidá, v našem případě 20, konečných hodnot v převráceném pořadí (konec  $\rightarrow$  konec-20). Stejně tak se ke konci připojí hodnoty počáteční. Tím je umožněno dopočítat funkční hodnoty všech požadovaných dat.
  2. Na takto získaných datech proběhne další vyhlazení metodou klouzavého průměru s oknem 17 dat, tj. po intervalech 17-ti hodnot se počítá jejich průměr, který se opět přiřadí vždy prostřední hodnotě v okně. Krajní body jsou počítány jako průměr z nižšího počtu dat, okno je automaticky zmenšováno. První hodnota bude v našem případě vypočítána jako průměr prvních dvou, druhá jako průměr prvních tří atd.

V obou případech byly metody implementovány prostřednictvím funkce *smooth* vestavěné v Matlabu. Výsledek takového vyhlazení je zobrazen na obrázku 6.3.



Obrázek 6.3: Na obrázku je ilustrován výsledek uvedeného vyhlazení. Na prvním z nich jsou data z histogramu (spojení funkčních hodnot odpovídajících pravým hranám jednotlivých binů). Na druhém jsou zobrazena po použití Savitzky-Golay filtru, na posledním po další úpravě metodou klouzavého průměru.

- Ve vyhlazených datech nalezneme první minimum a vzhledem k jeho x-ové souřadnici přesune původní hodnoty z histogramu. Tato data jsou opět vyhlazena postupem popsaným v předešlém kroku. Výsledek je zobrazen v první části obrázku 6.4. Následně použije funkci *findpeaks* se zavedenými podmínkami na prominenci a šířku píku (šířka je měřena v polovině prominencí) pro nalezení maxim průběhu četnosti, jak je ilustrováno na druhém obrázku v obrázku 6.4.
- Následující krok vychází z pozorování zaznamenaného na obrázku 6.5. Jedná se o četnosti jednoho z torzních úhlů molekuly TMDGA, přičemž první ze zaznamenaných průběhů odpovídá 5 ns simulaci, druhý průběh posledních pěti nanosekund z celkového simulačního času 40 ns.



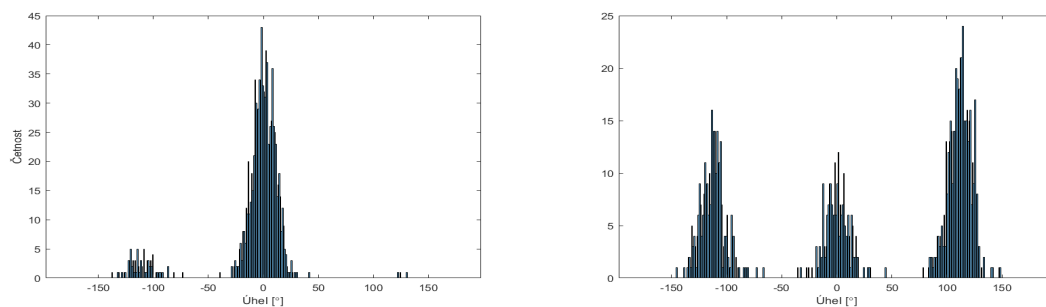
Obrázek 6.4: Vlevo jsou data z první části obrázku 6.3 po přesunu vzhledem k nalezenému prvnímu minimu a následném vyhlazení. Vpravo je zobrazen výstup funkce *findpeaks*. Oranžovou barvou je zakreslena prominence píku, žlutou jeho šířka.

Výsledek odpovídá růstu entropie systému a kdybychom měli k dispozici data z takto dlouhých simulací, neměli bychom problém detekovat všechny píky, které se v datech skutečně vyskytují. To ale z časových důvodů není možné.

Problém je v tom, že by nám v důsledku procesu vyhlazení dat mohl „uniknout“ malý pík v levé části obrázku 6.5 s hodnotou  $\sim 125^\circ$ , čímž bychom ovšem přišli o důležitou informaci a je nutné ho zachytit jiným způsobem. Kód tyto hodnoty přidává k sadě původně nalezených maxim uměle.

Označíme-li hodnotu maxima píku v prvním z histogramů na obrázku 6.5 pohybujících se okolo  $-115^\circ$  jako  $m_1$ , pak pokud se v intervalu  $(-m_1 - 15, -m_1 + 15)$  vyskytuje alespoň jedna nenulová hodnota histogramu a tato hodnota neleží příliš blízko u dříve nalezených maxim, je přidání nové hodnoty  $m_2 = -m_1$  povoleno. Vzhledem k udanému intervalu může dojít k dodatečné detekci pouze pro torzní úhly projevující symetrii vzhledem k počátku.

Tento krok je ospravedlněn tím, že pracujeme s daty získanými ze simulace, která je energeticky v rovnováze a není důvod se domnívat, že by nerovnovážné kmity torzního úhlu vedly ke vzniku píku (maxima) v jeho úhlové distribuci.



Obrázek 6.5: Vlevo je zaznamenána četnost torzního úhlu z 5 ns simulace. Napravo je četnost získaná z posledních 5 ns z celkové délky 40 ns simulace.

- Počet výše nalezených maxim a případných dodatečných hodnot nalezených v předchozím kroku slouží jako informace o počtu Gaussových křivek, jejichž součtem bude histogram proložen. x-ové souřadnice těchto hodnot slouží jako počáteční odhad jejich středních hodnot.

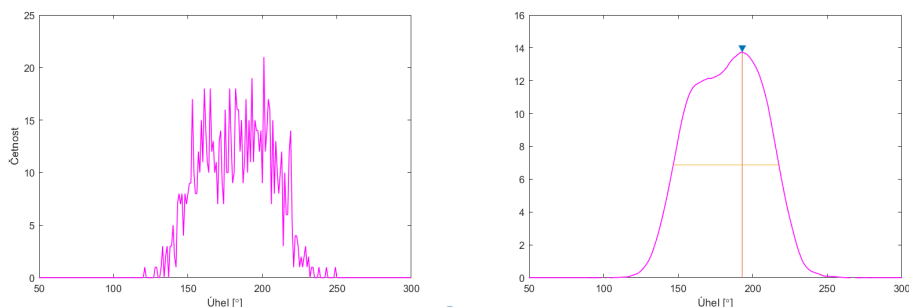


Pro samotné fitování je použita funkce *fit* s Gaussovským modelem (*Gaussian Models*), který je v Matlabu implementován v následujícím tvaru

$$y = \sum_{i=1}^n a_i \exp \left[ - \left( \frac{x - b_i}{c_i} \right)^2 \right], \quad (6.1)$$

kde  $a_i$ ,  $b_i$  a  $c_i$  jsou parametry, které se fitují. Vzhledem k lehce odlišnému tvaru této funkce a předpisu normálního rozdělení, které je jako funkce dále v kódu zdefinováno, je nutné přepočítat rozptyl jako  $\sigma^2 = c_i^2/2$ .

- Vzhledem k cyklické povaze dat je do matice obsahující nafitované parametry přidán stejný set  $a_i$ ,  $b_i$ ,  $c_i^2/2$  odpovídající největší střední hodnotě  $b_i$ , přičemž hodnota  $b_i$  je změněna na  $b_i = b_i - 360$ .
- Vyhodnotí průsečíky vždy dvou po sobě následujících křivek (seřazeny dle velikosti  $b_i$ ) s omezením na jejich x-ovou souřadnici. Ta musí ležet v intervalu omezeném danými hodnotami  $b_i$ . Pokud se nepodaří v některém z těchto intervalů průsečík nalézt, jedná se o případ, kdy byla chybně přidána hodnota na základě předpokládané symetrie torzního úhlu. V takové situaci se tato hodnota vymaže a proběhne opětovný fit a určení průsečíků.
- Průsečíky v tomto momentě slouží pro vymezení jednotlivých píků, které budou dále jednotlivě fitovány dalším součtem Gaussových křivek. Tento krok je nezbytný ze dvou důvodů. Prvním z nich je opět vyhlazení dat, při kterém mohlo dojít k potlačení některého z píků, což je zobrazeno na obrázku 6.6. K tomu dochází v případech, kdy se dvě různá maxima nacházejí vzájemně relativně blízko a je důsledkem použití metody klouzavého průměru. Ve druhé řadě se jedná o píky reprezentující více hodnot úhlů, u nichž v průběhu simulace došlo k takovému rozložení, při kterém nedojde k zvýraznění této skutečnosti v podobě většího počtu maxim.



Obrázek 6.6: Ztráta informace o jednom z maxim, která je důsledkem vyhlazení dat. Nalevo jsou data z histogramu, napravo jsou tato data po vyhlazení. Modrou šipkou je zvýrazněno nalezené maximum.

Dalším krokem je postupný výběr dat, které náležejí jednotlivým píkům. K určení počtu křivek, kterými budou data proložena, se v prvním kroku použije funkce *findchangepts*. Ta pracuje v následujících krocích

1. Vybere bod a rozdělí data na dvě části (funguje pro libovolný počet bodů menší než počet dat) a vypočítá v nich zadanou statistiku, například průměr.

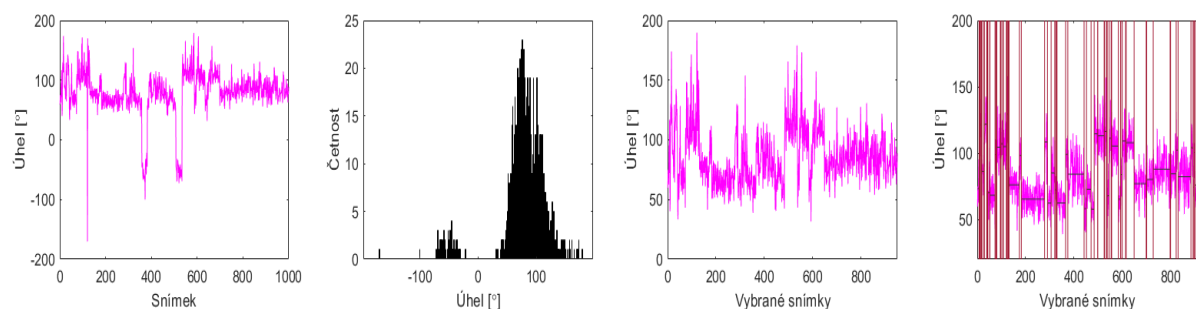
2. V každém bodě změří odchylky od těchto referenčních hodnot, které postupně sčítá.
3. Posouvá bod  $k$  zvolený v prvním kroku tak, aby variance dat (přenasobená normalizačním faktorem)

$$J = \sum_{i=1}^{k-1} \left( x_i - \frac{1}{k-1} \sum_{r=1}^{k-1} x_r \right)^2 + \sum_{i=k}^N \left( x_i - \frac{1}{N-k+1} \sum_{r=k}^N x_r \right)^2 \quad (6.2)$$

byla minimální.  $x_i$  jsou jednotlivé body a  $N$  je jejich celkový počet.

Za její statistiku je v tomto kroku volen průměr a je použita ve formě upravující minimální možnou odchylku dat na hodnotu  $J = 1000$ .

Tento krok je zobrazen na obrázku 6.7.



Obrázek 6.7: V prvním grafu je zaznamenán celý časový průběh torzního úhlu, ve druhém je příslušná četnost. Na třetím obrázku jsou zobrazena data „přiblížená“ na základě hodnot průsečíků a odpovídají většímu píku na druhém obrázku. V poslední části je zobrazen výstup funkce *findchangepts*.

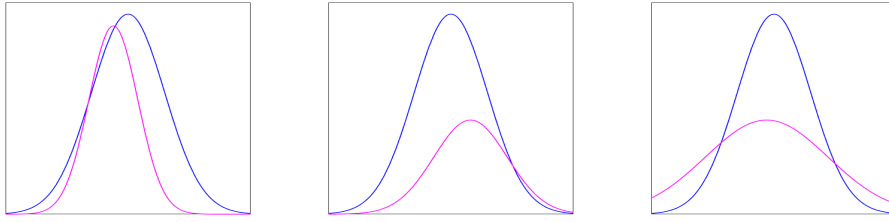
Jako výstup z této funkce jsou získány indexy bodů na koncích nalezených intervalů. V nich se vypočítá průměrná hodnota a směrodatná odchylka  $s$ .

Pokud se průměry v jednotlivých intervalech liší méně než o  $0,2s$ , potom je jedna z hodnot vymazána. Počet hodnot  $n$ , které projdou touto podmínkou, je brán jako horní odhad pro počet komponent následujícího fitu.

- Každý z píků je proložen součtem  $n$  Gaussových křivek. Z jejich funkčních předpisů jsou vypočítány průsečíky, které musí splňovat podmínku, která je znázorněna na obrázku 6.8. Oba dva průsečíky dané komponenty musejí ležet vždy na jedné straně od střední hodnoty funkce druhé. Pokud některá z funkcí tuto podmínku nesplní, pak je sníženo  $n = n - 1$  a proces se opakuje dokud podmínku nesplní všechny komponenty. Pokud  $n > 1$ , pak se dojde k nahrazení původní hodnoty nově nalezenými.

Omezení původních  $n$  hodnot v předchozím kroku (hodnotou  $0,2s$ ) a právě zmíněnou podmínkou je jediným ošetřením, které brání přefitování dat. Tato kombinace funguje velice dobře pro většinu dat. V některých případech se situace ale vyhodnotí příliš přesně a výsledky po kontrole neodpovídají úhlům, které bychom osobně přiřadili, což je důležitý faktor. Tento problém v kódu prozatím vyřešen nebyl a v takových případech je nutné hodnoty v průběhu kódu vymazat ručně.

Pro data zpracovávána v této práci nebyl pozorován žádný pík, který by bylo třeba fitovat více než dvěma křivkami. Zpracování dat proto probíhalo přímo s nastavenou horní hodnotou  $n = 2$ , což je nejrychlejší přístup k problému.



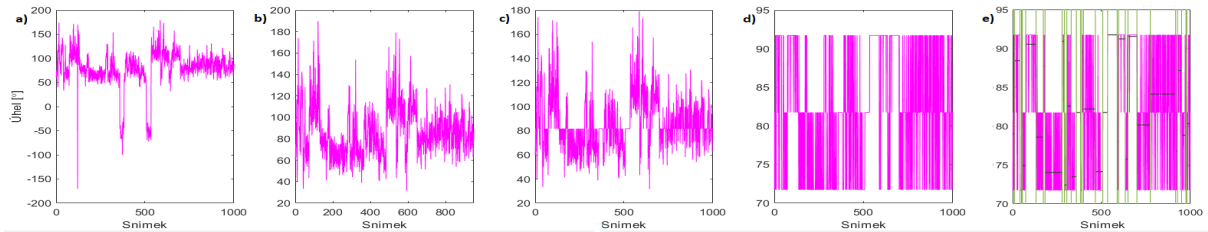
Obrázek 6.8: První dvě situace jsou z hlediska podmínky v pořádku a v tomto případě tedy  $n = 2$ . Ve zbylých dvou případech podmínka splněna není a v této ukázce  $n = 1$ .

- Posledním krokem je správné přiřazení nalezených středních hodnot původním datům. To v první řadě probíhá opět pomocí průsečíků. Tentokrát se vypočtou průsečíky mezi všemi nalezenými funkcemi a hodnoty se přiřadí na základě toho, která z funkcí má v daných intervalech nejvyšší funkční hodnoty. Tento přístup ale obecně nemůže být aplikován na nově nalezené funkce původně náležející jednomu píku. Obecně je následující způsob aplikován pro všechny případy, kdy je hodnota průsečíku daných dvou křivek větší než nastavený parametr. Ten je závislý na konkrétní sadě dat, většinou byla nastavena na 4. Postup je následující

1. Pro dvě po sobě následující střední hodnoty (dle velikosti) spočítá průměr  $u$  a průměr rozdílu těchto hodnot  $v$ .

Datům, která nenáležejí původnímu píku nalezenému v prvním kroku je přiřazena hodnota  $u$ . Ta je dále považována za počátek, kolem kterého se vytvoří „vyprázdněná“ oblast, tj. úhlům, které jsou v intervalu  $(u - v, u + v)$  se přiřadí opět hodnoty rovny  $u$ . Pokud jsou zbývající data větší než  $u$ , nabudou hodnoty  $u + 10$ , v opačném případě  $u - 10$ .

Na takto upravená data je opět aplikována funkce *findchangepts*, tentokrát se zadaným maximálním počtem bodů, ve kterých může ke změnám dojít. Hodnota je nastavena na 30. Postup je zobrazen na obrázku 6.9.



Obrázek 6.9: a) Původní data, b) vybraná data, c) vybraná data s „vyprázdněnou“ oblastí, d) data z bodu c) po přiřazení zmíněných hodnot, e) aplikace funkce *findchangepts*. Šedou barvou jsou zobrazeny průměrné hodnoty v příslušných intervalech.

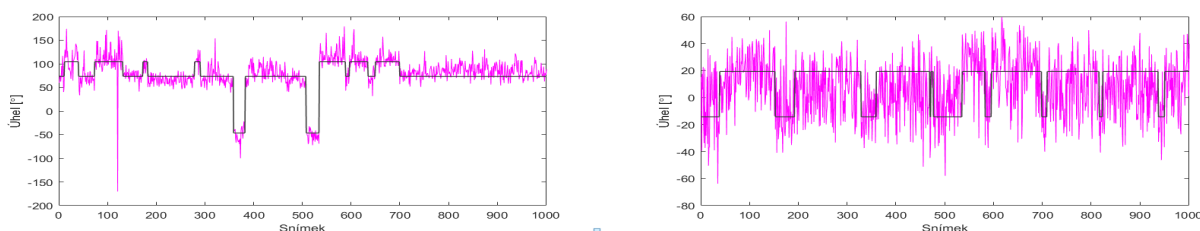
2. V intervalech z funkce *findchangepts* vypočítá průměrné hodnoty (obrázek 6.9 e)), které příslušným datům přiřadí.
3. Na takto upravený vektor aplikuje shlukovací algoritmus *k-means*, jehož úkolem je rozdělit data do  $k$  skupin. Počet shluků musí být zadán a pro nás představuje počet funkcí, který byl danému píku přiřazen. V našem příkladě je tedy nyní  $k = 2$ . Počáteční odhady pozice těžišť jednotlivých shluků zadány být nemusejí. My ale tuto informaci máme a funkce si jako vstup bere vektor středních hodnot daných křivek.

Pro všechny body vypočítá kvadrát Euleidovské vzdálenosti k jednotlivým těžištím centroidů a přiřadí je k těm, pro které nabývá nejmenší hodnoty. Následně přepočítá souřadnice těžišť a opakuje předchozí krok. Algoritmus končí, pokud se nezmění počet bodů přiřazených jednotlivým shlukům a těžiště tak zůstanou konstantní.

Tento krok je finální pro přiřazení středních hodnot původním datům.

- S informací o celkovém počtu Gaussových křivek a počátečními odhady středních hodnot nakonec přefituje celý soubor dat. Tyto finální hodnoty se přepíše do vektoru obsahující informace o jejich přiřazení.

Příklad konečného nařizování dat z obrázku 6.9 je v levé části obrázku 6.10.



Obrázek 6.10: Příklady nařizovaných dat.

- Chyby měření jsou počítány ze směrodatné odchylky

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N |x_i - \mu|}, \quad (6.3)$$

kde  $N$  je počet naměřených dat  $x_i$  a  $\mu$  je jim přiřazená střední hodnota.

- Nyní máme k dispozici matici, jejíž sloupce představují zadané torzní úhly a řádky jednotlivé záznamy snímků simulace. Každý řádek tak představuje informaci o celkové konformaci molekuly v podprostoru torzních úhlů v daném čase.
- Nalezne počet změn konformace v celém průběhu simulace.
- Nalezne různé konformace a vypočítá četnost, se kterou se v simulaci vyskytly.
- Pokud by molekula mohla nabývat dvojčetné symetrie, pak kód symetrické konformace nalezne (v rámci směrodatné odchylky) a dále s nimi nepracuje. Jejich četnosti přiřazené v předchozím kroku se přičítají k četnostem symetrických obrazů.  
K provedení tohoto kroku je nezbytné zadat vektor symetrických úhlů.
- Pokud je zadáno, nepočítá se s enantiomery (v rámci směrodatné odchylky). U takových nalezených molekul se potom četnost opět přičítá konformaci, která ve výsledných datech zůstane.
- Vypočítá počet překmitů mezi různými hodnotami v rámci jednoho torzního úhlu.
- Určí, kolikrát se přímo za sebou daná konformace v simulaci nacházela spolu s informací, kolikrát tento případ nastal.

## Kapitola 7

### Výsledky a jejich diskuze

Původním cílem provedené konformační analýzy bylo určení nejpravděpodobnějších konformací různých molekul v závislosti na jejich rozpouštědle. Ty by sloužily jako vstup do kvantové mechanických výpočtů chemické stability a alespoň částečně by se v nich tak odrážel vliv prostředí a teploty.

Z 5 ns simulace systému vody s molekulou TMDGA, která je nejmenší molekulou zkoumanou v této práci, bylo získáno 151 různých konformací, které mezi sebou vzájemně překmitly 785-krát. Nejvyšší četnost jedné konformace je 64, což svědčí o extrémní flexibilitě této molekuly. Pro stejný systém byla spuštěna i 40 ns simulace. Při analýze posledních 5 ns této simulace se došlo k závěru, že nedochází k upřednostnění některé z možných hodnot jednotlivých úhlů a naopak se projevuje jejich entropie. Z tohoto vzorku vyjde 157 různých konformací, s počtem překmitnutí 766. Nejvyšší četnost je 59.

U obou souborů se vyskytuje  $\sim 50$  konformací, které se v simulaci vyskytly pouze jednou, což ovšem může mít výpovědní hodnotu pouze ve druhém případě delší simulace. Seřadíme-li konformace dle jejich četnosti, pak ve dvou souborech konformací není pozorována žádná korelace.

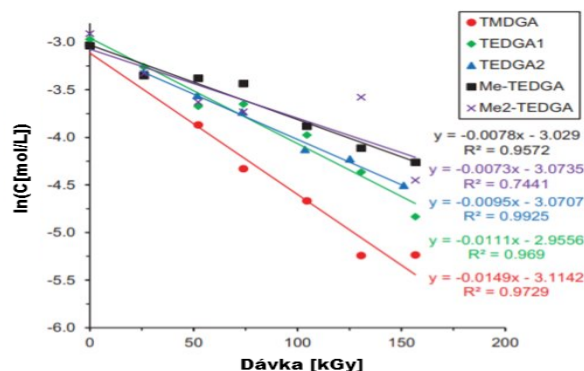
Obecně tedy četnost není dostatečným indikátorem k určení reprezentativní konformace a pro usnadnění jejího výběru je u každé z nich vypsán počet snímků, ve kterých se daná konformace vyskytla beze změny, spolu s informací, kolikrát se tak stalo, což ukazuje na energeticky výhodnější stav systému.

V dalším textu bude zkoumána možnost korelace přímo mezi stabilitou molekul a výsledky konformační analýzy. Analýza je založena na následující úvaze. Síla vazby je dána hodnotou elektronové hustoty. Čím větší je elektronová hustota, tím je silnější vazba a tím „stabilnější“ (tj. méně fluktuující) by měla být v molekulární dynamice. To samé platí i o celkovém počtu změn konformace molekuly v simulaci, který bude značen jako  $L$ . Počet možných konformací nemá z tohoto hlediska výpovědní hodnotu, nicméně pro úplnost bude uváděn, značen bude jako  $K$ . Tučným písmem budou uvedeny výsledky experimentů, pod nimi uvedeny výsledky konformační analýzy. Časové průběhy nejsou uváděny v sekundách, ale ve snímcích. Nebude-li uvedeno jinak, pak časový interval mezi dvěma snímky je 5 ps.

- **Experimentálně bylo zjištěno, že stabilita molekul TMDGA, TEDGA, Me-TEDGA a Me<sub>2</sub>TEDGA ozářených ve vodném prostředí roste s rostoucí hmotností. Dle stability Me<sub>2</sub>-TEDGA  $\geq$  Me-TEDGA  $>$  TEDGA  $>$  TMDGA [35].** V tomto případě máme jedinou možnost jak se pokusit zachytit tento fakt daty z konformační analýzy, konkrétně počtem změn konformací během simulace  $L$ . Tyto hodnoty jsou zapsány v tabulce 7.1. Na obrázku 7.1 je zobrazen pokles koncentrace ligandů v závislosti na dávce.

Tabulka 7.1: Hodnoty  $L$  a  $K$  pro molekuly TMDGA, TEDGA, Me-TEDGA a Me<sub>2</sub>-TEDGA ve vodném prostředí.

	TMDGA	TEDGA	Me-TEDGA	Me <sub>2</sub> -TEDGA
$L$	766	370	310	315
$K$	157	178	131	203



Obrázek 7.1: Přírodní logaritmus koncentrace molekul  $C$  rozpuštěných ve vodě v koncentraci  $0,05 \text{ molL}^{-1}$  v závislosti na absorbované dávce [35]. Rozpadové konstanty odpovídají hodnotám sklonu zobrazených lineárních fitů ( $1/\text{kGy}$ ).

Hodnota  $L$  dobře zachytila výrazně nižší stabilitu TMDGy oproti ostatním třem molekulám. Hodnoty 310 a 315 pro Me-TEDGA a Me<sub>2</sub>-TEDGA ukazují, v souladu s experimentálními daty [35], na podobnou míru stability těchto dvou molekul. Hodnota 370 pro TEDGA naznačuje lehce větší stabilitu oproti jejímu methyl- a dimethyl- derivátu. Za předpokladu exponenciálního průběhu rozpadu nejsou zachovány poměry mezi experimentálními daty a hodnotami  $L$ , nicméně byl získán kvalitativní odhad.

- **Experimentálně byly proměřeny rozpadové konstanty molekuly TEDGA v závislosti na prostředí.**

Data pocházejí z článku [35] a jsou zapsána v tabulce 7.2 spolu s hodnotami  $L$  a  $K$ .

S rostoucí teplotou a koncentrací kyseliny dusičné klesá stabilita molekuly TEDGA. Zajímavé

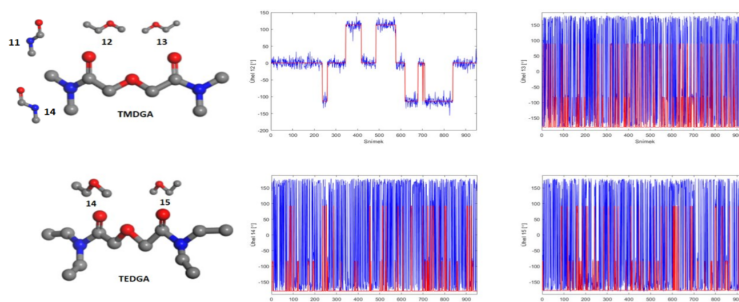
Tabulka 7.2: Naměřené rozpadové konstanty [ $\text{h}^{-1}$ ] pro původně  $0,05 \text{ mol} \cdot \text{L}^{-1}$  roztoky molekuly TEDGA s různými koncentracemi  $\text{HNO}_3$  a  $\text{HCl}$  při daných teplotách  $T$  [35]. Za experimentálními hodnotami jsou zapsány hodnoty  $L$  a  $K$  ze simulace.

$T$ [ $^{\circ}\text{C}$ ]	0,5 M $\text{HNO}_3$	$L$	$K$	4 M $\text{HNO}_3$	$L$	$K$	5 M $\text{HCl}$	$L$	$K$
25	$(1,1 \pm 0,1) \cdot 10^{-3}$	459	245	$(1,2 \pm 0,2) \cdot 10^{-3}$	515	283	$(7,0 \pm 0,4) \cdot 10^{-2}$	756	499
45	$(8,5 \pm 0,1) \cdot 10^{-3}$	640	342	$(8,8 \pm 0,5) \cdot 10^{-3}$	659	317	<b>8 M HCl</b>		
65	$(5,0 \pm 0,6) \cdot 10^{-2}$	737	478	$(7,8 \pm 0,6) \cdot 10^{-2}$	809	463	$(5,2 \pm 0,2) \cdot 10^{-2}$	707	432

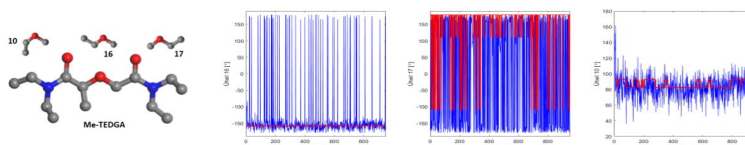
chování je pozorováno pro prostředí s kyselinou chlorovodíkovou, kdy při koncentraci  $8 \text{ mol}\cdot\text{L}^{-1}$  dochází, oproti případu s  $5 \text{ mol}\cdot\text{L}^{-1}$ , ke stabilizaci molekuly. I tento jev je správně replikován vývojem hodnoty parametru  $L$ . Opět ale nedochází k zachování poměrů vzhledem k rozpadovým konstantám.

- **Produkty rozpadu vazby  $\text{O}_{\text{ether}}\text{-C}_{\text{ether}}$  byly po ozáření ve vodě identifikovány u všech molekul TEDGA, Me-TEDGA,  $\text{Me}_2\text{-TEDGA}$ . Nebyly identifikovány pro molekulu TMDGA [35].**

Torzní úhly zdefinované pro jednotlivé molekuly přes vazbu  $\text{O}_{\text{ether}}\text{-C}_{\text{ether}}$  spolu s jejich časovými průběhy ze simulace jsou zobrazeny na obrázcích 7.2, 7.3 a 7.4.



Obrázek 7.2: Molekula TMDGA a TEDGA a časové průběhy torzních úhlů, zdefinovaných přes vazbu  $\text{O}_{\text{ether}}\text{-C}_{\text{ether}}$ . Pro molekulu TMDGA se jedná o úhly 12 a 13.

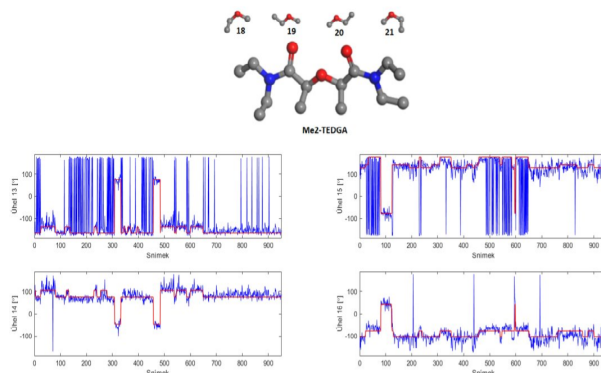


Obrázek 7.3: Molekula Me-TEDGA a časové průběhy torzních úhlů, zdefinovaných přes vazbu  $\text{O}_{\text{ether}}\text{-C}_{\text{ether}}$ .

Pro molekulu TEDGA se jedná o úhly, které v simulaci mají nejvyšší frekvenci změn ze všech zdefinovaných úhlů, totéž platí o úhlu 16 molekuly Me-TEDGA. O molekule  $\text{Me}_2\text{-TEDGA}$  nic takového tvrdit nelze, podobné průběhy nabývají i jiné úhly. Ani u jedné z těchto molekul se ale nevyskytuje úhel s výrazně vyšší frekvencí změn. To neplatí pro molekulu TMDGA, u které jako u jediné dochází ke zdatelnému rozštěpení úhlů 11 a 14 z hodnoty  $\sim 180^\circ$  na hodnoty  $\sim \pm 165^\circ$ . Je tedy možné se domnívat, že rozpad vazby  $\text{O}_{\text{ether}}\text{-C}_{\text{ether}}$  bude potlačen ve prospěch rozpadu vazby  $\text{C}_{\text{amid}}\text{-N}$ . Produkty rozpadu této vazby byly pozorovány pro všechny studované molekuly [35].

Pro všechny hydrofilní diglykolamidy byly provedeny simulace ve vodném prostředí a v kyselém prostředí s rozsahem koncentrací  $\text{HNO}_3$  1, 2, 3 a  $4 \text{ mol}\cdot\text{L}^{-1}$ . Simulace byly provedeny kvůli původní myšlence o nalezení nejpravděpodobnější konformace v těchto roztocích a neexistují k nim experimentální výsledky.

Vše z nich ale záznam o zajímavém chování molekuly TMDGA, který zde pro zajímavost uvedeme.

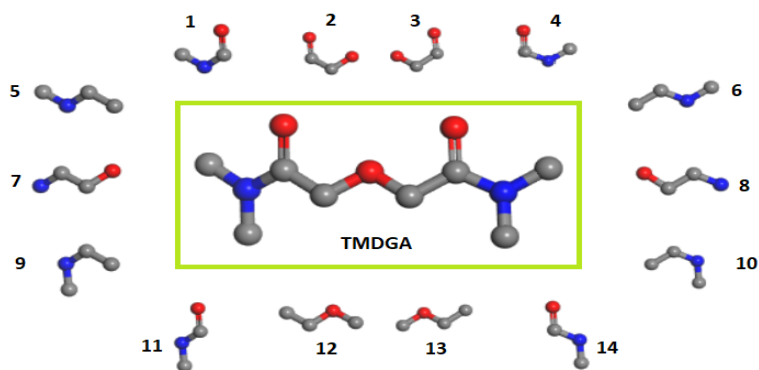


Obrázek 7.4: Molekula Me<sub>2</sub>-TEDGA a časové průběhy torzních úhlů, zdefinovaných přes vazbu O<sub>ether</sub>-C<sub>ether</sub>.

Jako jediná z testovaných molekul TMDGA zcela změnila své konformační chování již při přechodu z vodného prostředí do prostředí s 1 M HNO<sub>3</sub>. Chování ostatních molekul lze popsat zvýšenou frekvencí změn mezi jednotlivými úhly, nedošlo u nich ale ke změně jejich hodnot a to ani při nejvyšších koncentracích kyseliny.

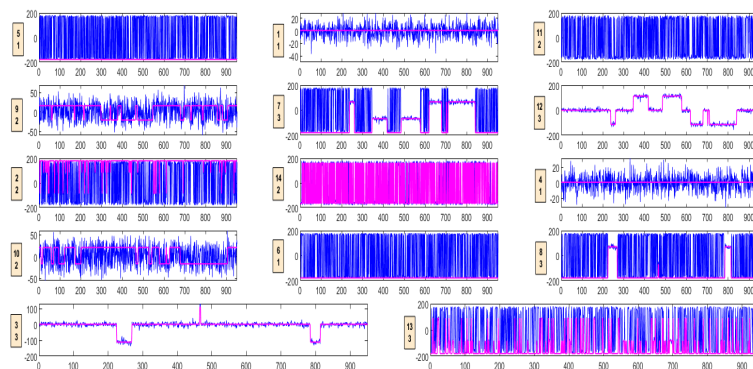
Naopak, z obrázků 7.6 a 7.7 vyplývá, že pro TMDGA nedochází ke změně pouze v případě úhlu 13.

Budeme-li předpokládat, že lze z reakce dané molekuly na kyselé prostředí soudit na sílu jejích vazeb ve vodném prostředí, což se vzhledem k výše zmíněnému chování molekuly TEDGA nejvíce jako chybný předpoklad, můžeme učinit závěr, že nejstabilnější jsou vazby C<sub>amid</sub>-C<sub>ether</sub> odpovídající úhlům 3, 7 a 8. To by podporovalo fakt, že se v ozářených vodných roztocích neobjevují produkty rozpadu této vazby. Nejcitlivější by měly být opět vazby C<sub>amid</sub>-N.

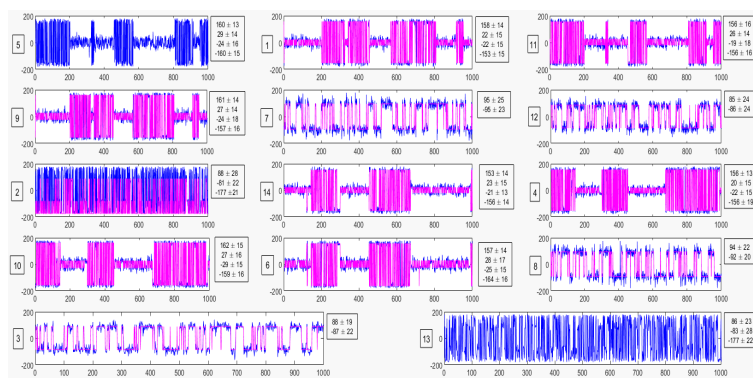


Obrázek 7.5: Molekula TMDGA se zdefinovanými torzními úhly. Šedou barvou je zobrazen uhlík, modrou dusík a červenou kyslík. Vodíky nejsou pro přehlednost zobrazeny.





Obrázek 7.6: Časové průběhy torzních úhlů molekuly TMDGA ve vodě. Růžově je zvýrazněn fit. Pro úhel 11 není fit zobrazen kvůli přehlednosti a odpovídá fitu úhlu 14. Vlevo od každého grafu je číslo úhlu odpovídající obrázku 7.5, pod ním je počet nalezených úhlů.



Obrázek 7.7: Časové průběhy torzních úhlů molekuly TMDGA v roztoku 1 M  $\text{HNO}_3$ . Růžově je zvýrazněn fit. Pro úhel 5 není fit zobrazen kvůli přehlednosti a vizuálně odpovídá fitu úhlu 1. Totéž platí pro úhel 13 a 2. Vlevo od každého grafu je číslo úhlu odpovídající obrázku 7.5. Napravo jsou střední hodnoty těchto úhlů se směrodatnou odchylkou.

- Molekula CMPO rozpuštěná v dodekanu je méně stabilní, pokud tento roztok není v kontaktu s  $\text{HNO}_3$ . Za přítomnosti kyseliny se mění rozpadové produkty. [36].

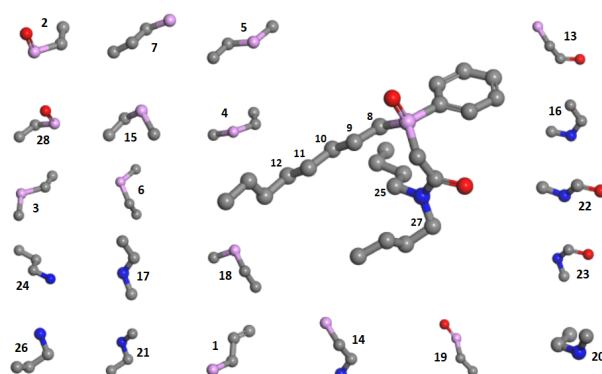
Hodnoty  $L$  potvrzují zvýšení stability CMPO v dodekanu a v kontaktu s  $\text{HNO}_3$ , kdy  $L=532$ , zatímco bez kyseliny je  $L=711$ . To lze připsat celkem šesti z definovaných úhlů (z 28), u kterých došlo ke snížení degenerace jejich hodnot. Jedná se o úhly 7, 8, 9, 10, 12 a 27 (obrázek 7.8) uhlovodíkových řetězců a jako příklad je na obrázku 7.9 zachycen úhel 10 molekuly CMPO v prostředí bez vlivu a s vlivem  $\text{HNO}_3$ .

V simulaci prostředí s kyselinou ale dochází kromě útlumu výše zmíněných úhlů také ke zvýšení degenerace úhlů 16, 20, 21 a 22. Některé z těchto úhlů jsou bez kyseliny a s kyselinou zaznamenány na obrázku 7.9.

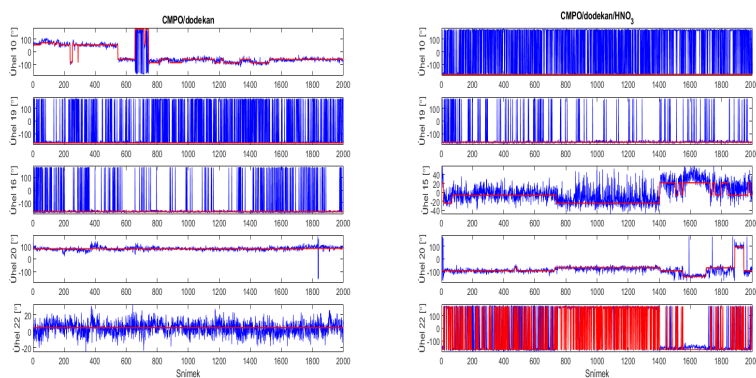
Tyto výsledky vedou k hypotéze, že v kyselém prostředí dojde k oslabení vazby N- $\text{C}_{\text{butyl}}$  a vazby definované prostředními atomy torzního úhlu 26 (obrázek 7.8).

Experimentálně se zjistilo, že za přítomnosti kyseliny dusičné začne docházet ke zvyšování produkce monoisobutylu CMPO (obrázek 7.10 b)) [36], které odpovídají rozpadu vazby N-C<sub>butyl</sub>. Zároveň klesá koncentrace molekuly OPPA (obrázek 7.10 a)), což může naznačovat snížení variance úhlu 19 v kyselém prostředí (obrázek 7.9), a to na úkor zvýšené produkce OPPAA (obrázek 7.10 d)), která odpovídá přetržení vazby N-C<sub>karbonyl</sub>. Tento produkt není pomocí torzních úhlů přímo zachycen a pokud se nejedná o případ, kdy k jeho produkci dochází ve více rozpadových krocích, pak je nutné konstatovat, že konformační analýza může pouze poukázat na nestabilní oblasti molekuly.

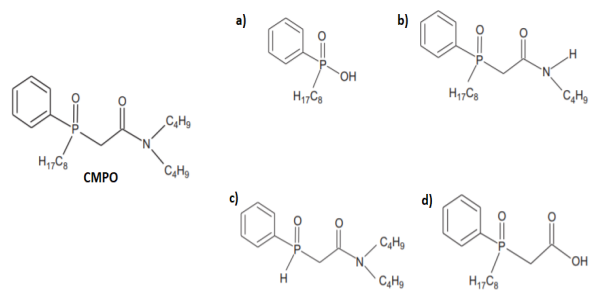
U ostatních torzních úhlů není pozorována žádná změna v degeneraci ani rozptylech.



Obrázek 7.8: Molekula CMPO a definované torzní úhly. Vodíky zobrazeny nejsou. Šedou barvou je značen uhlík, červenou kyslík, modrou dusík a růžovou fosfor.



Obrázek 7.9: Časové průběhy torzních úhlů 10, 16, 19, 20, 22 v prostředí bez kyseliny a v prostředí s kyselinou. Modře jsou označena data získaná ze simulací, červeně jejich fit.



Obrázek 7.10: Molekula CMPO (vlevo) a některé její hlavní radiolytické produkty. a) octylphenylphosphinová kyselina (OPPA); b) octylphenyl-*N*-isobutylcarbamoylmethyl phosphine oxid (monoisobutyl CMPO); c) phenyl-*N,N*-diisobutylcarbamoylmethyl phosphine oxide; d) octylphenyl-*N,N*-diisobutylcarbamoylmethyl phosphine oxide (OPPAA). Převzato z [36].

## Kapitola 8

### Závěr

V této práci byl nastíněn princip kapalina-kapalinové extrakce. Dále byly shrnuty základní myšlenky počítačové simulace fyzikálního systému. Nejdříve byl přiblížen pojem silového pole, které je základním stavebním kamenem každé simulace. Následně byly ve čtvrté kapitole uvedeny principy simulace Newtonovské molekulární dynamiky a stochastické metody Monte Carlo.

V páté kapitole bylo shrnuto nastavení simulací provedených v této práci. V šesté byl popsán proces vlastní konformační analýzy, která je hlavním bodem práce. Napsaný kód je třeba v budoucnosti výrazně vylepšit, především je nutné nalézt adekvátní podmínky pro vyřešení problému přefitování dat.

I přes nedostatky této analýzy se podařilo dojít k závěrům o vztahu zadaného parametru  $L$ , počtu změn v konformaci molekuly v čase a její celkové stability. Čím větší hodnota  $L$ , tím nestabilnější je daná molekula. To se podařilo ověřit na všech provedených simulacích v této práci, jak je zdokumentováno v sedmé kapitole.

Naopak se nepodařilo prokázat přímý vztah mezi torzními úhly, u kterých dojde se změnou prostředí ke zvýšení degenerace jejich hodnot a nestabilitou příslušných vazeb. To může být způsobeno i důsledkem kaskády chemických reakcí. Závěrem zatím zůstává, že nestabilní chování torzních úhlů pouze orientačně ukazuje na místa molekuly, která jsou nejnáchylnější k rozpadu.

Pro ověření vztahu parametru  $L$  k celkové stabilitě molekuly, a pro ověření korelace chování jednotlivých torzních úhlů a stability vazeb budou provedeny další simulace.

## Kapitola 9

### Přehled použité literatury

## Literatura

- [1] Spadaro J. V., Langlois L., Hamilton B. (2000). Greenhouse gas emission of electricity generation chains: Assessing the Difference. IAEA, Vídeň, Rakousko.
- [2] Bateman, H. (1910). The solution of a system of differential equations occurring in the theory of radioactive transformations. *Proc. Cambridge Philos. Soc.*, 15, pp. 423-427.
- [3] Radioactivity.eu.com. (2019). Radioactivity : Spent fuel composition. [online] [http://www.radioactivity.eu.com/site/pages/Spent\\_Fuel\\_Composition.htm](http://www.radioactivity.eu.com/site/pages/Spent_Fuel_Composition.htm). Převzato 5.4.2019.
- [4] Uh.edu. (2018). Index of /~lcr3600/simulation. [online] <http://www.uh.edu/~lcr3600/simulation/>. Převzato 4.8.2018.
- [5] Kopecky, J. (1997). Atlas of neutron capture cross section. *International Nuclear Data Committee*.
- [6] Gérard Mourou – Nobel Lecture. (2019) NobelPrize.org. [online] Dostupné na <https://www.nobelprize.org/prizes/physics/2018/mourou/lecture/>
- [7] Dassault Systèmes BIOVIA, Materials Studio 8.0, San Diego: Dassault Systèmes.
- [8] The Hyperion Cluster - Info For Users. [online] <http://hyperion.fjfi.cvut.cz/info/>. Převzato 4.8.2018.
- [9] Choppin, G., Liljenzin, J.-O., Rydberg, J. (2013) Radiochemistry and Nuclear Chemistry, 4th edition. Academic Press.
- [10] Modolo G. et al. GENIORS; 1St Half Yearly Partner Activity Report.
- [11] Paiva, A. P., Malik, P. (2004) Recent advances on the chemistry of solvent extraction applied to the reprocessing of spent nuclear fuels and radioactive wastes. *J. Radioanal. Nucl. Chem.*, 261(2), pp.485-496.
- [12] Sun, H. (1998). COMPASS: An ab Initio Force-Field Optimized for Condensed-Phase Applications Overview with Details on Alkane and Benzene Compounds. *J. Phys. Chem. B*, 102(38), pp.7338-7364.
- [13] Leach, A. (2001). Molecular modelling. Harlow: Pearson Prentice Hall.
- [14] Kolafa, J., Perram, J. W. (1992). Cutoff Errors in the Ewald Summation Formulae for Point Charge Systems. *Mol. Simul.*, 9(5), pp.351-368.
- [15] Allen, M. P., Tildesley, D. J. (2017). Computer Simulation of Liquids. Oxford University Press.

- [16] Leeuw, S. W., Perram, J. W., Smith, E. R. (1980). Simulation of Electrostatic Systems in Periodic Boundary Conditions. I. Lattice Sums and Dielectric Constants. *Proc. R. Soc. Lond. A*, 373, pp.27-56.
- [17] Karasawa, N. and Goddard, W. A. (1989). Acceleration of Convergence for Lattice Sums. *J. Phys. Chem.*, 93, pp.7320-7327.
- [18] Franks, F. (2000). *Water*. New York: Springer.
- [19] Chaplin, M. (2018). Water models. [online] [http://www1.lsbu.ac.uk/water/water\\_models.html](http://www1.lsbu.ac.uk/water/water_models.html). Převezato 16.7.2018.
- [20] Head-Gordon, T., Hura, G. (2002). Water Structure from Scattering Experiments and Simulation. *Chem. Rev.*, 102(8), pp.2651-2670.
- [21] Head-Gordon, T. and Johnson, M. (2006). Tetrahedral structure or chains for liquid water. *Proc. Natl. Acad. Sci. U.S.A.*, 103(21), pp.7973-7977.
- [22] Hura, G., Sorenson, J., Glaeser, R. and Head-Gordon, T. (2000). A high-quality x-ray scattering experiment on liquid water at ambient conditions. *J. Chem. Phys.*, 113(20), pp.9140-9148.
- [23] Soper, A. (2000). The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 MPa. *J. Chem. Phys.*, 258(2-3), pp.121-137.
- [24] Isis.stfc.ac.uk. (2018). ISIS External databases. [online] <https://www.isis.stfc.ac.uk/Pages/External-databases.aspx>. Převezato 17.6.2018.
- [25] Engineeringtoolbox.com. (2018). Water - Heat Capacity (Specific Heat). [online] [https://www.engineeringtoolbox.com/specific-heat-capacity-water-d\\_660.html](https://www.engineeringtoolbox.com/specific-heat-capacity-water-d_660.html). Převezato 18.7.2018.
- [26] Scolnik, Y., Portnaya, I., Cogan, U., Tal, S., Haimovitz, R., Fridkin, M., Elitzur, A., Deamer, D. and Shinitzky, M. (2006). Subtle differences in structural transitions between poly-l- and poly-d-amino acids of equal length in water. *Phys. Chem. Chem. Phys.*, 8(3), pp.333-339.
- [27] Gasteiger, J. and Marsili, M. (1978). A new model for calculating atomic charges in molecules. *Tetrahedron Lett.*, 19(34), pp.3181-3184.
- [28] Rappe, A. and Goddard, W. (1991). Charge equilibration for molecular dynamics simulations. *J. Phys. Chem. A*, 95(8), pp.3358-3363.
- [29] Is.muni.cz. (2018). [online] [https://is.muni.cz/el/1431/jaro2015/F8510/um/F8510\\_prednaska1.pdf](https://is.muni.cz/el/1431/jaro2015/F8510/um/F8510_prednaska1.pdf). Převezato 25.7.2018.
- [30] 'ABNR', (2018). [online] <http://klingon.uab.es/prat/Thesis/node40.html>. Převezato 26.7.2018
- [31] Frenkel D., Smit B. (1996). *Understanding Molecular Simulation*. Academic Press.
- [32] Kim S. (2014). Issues On The Choice Of A Proper Time Step In Molecular Dynamics. *Phys. Procedia*, 53, pp.60-62.
- [33] The MathWorks Documentation. [online] <https://www.mathworks.com/help/documentation-center.html>
- [34] MATLAB R2018a, The MathWorks, Natick, MA, 2018

- [35] Wilden, M., Mincher, B. J., Mezyk, S. P., Twight, L., Rosciolo-Johnson, K. M., Zarzana, C. A., Case, M. E., Hupert, M., Stärk, A., Modolo, G. (2018). Radiolytic and hydrolytic degradation of the hydrophilic diglycolamides. *SOLVENT EXTR ION EXC*, 36(4), pp.347-359.
- [36] Mincher, B. J., Mezyk, S. P., Elias, G., Groenewold, G. S., Riddle, C. L., Olson, L. G. (2013). The Radiation Chemistry of CMPO: Part 1. Gamma Radiolysis. *SOLVENT EXTR ION EXC*, 31(7), pp.715-730.
- [37] Ochkin, A. V., Afonina, M. A., Merkushev, A. O., Nekhaevskii, S. Y. (2010). Extraction of nitric acid by tributyl phosphate solution in n-dodecane. *Russ. J. Phys. Chem. A*, 84(9), pp. 1526-1531.