

Posudek práce

předložené na Matematicko-fyzikální fakultě
Univerzity Karlovy

- | | |
|--|---|
| <input type="checkbox"/> posudek vedoucího | <input type="checkbox"/> posudek oponenta |
| <input type="checkbox"/> bakalářské práce | <input type="checkbox"/> diplomové práce |

Autor/ka: Martin Crhán

Název práce: *Full Centroid Molecular Dynamics through Machine Learning*

Studijní program a obor: Teoretická fyzika [FTFP]

Rok odevzdání: 2023

Jméno a tituly vedoucího/opponenta: Prof. RNDr. Petr Slaviček, Ph.D.

Pracoviště: Ústav fyzikální chemie, VŠCHT Praha

Kontaktní e-mail: petr.slavicek@vscht.cz

Odborná úroveň práce:

- vynikající velmi dobrá průměrná podprůměrná nevyhovující

Věcné chyby:

- téměř žádné vzhledem k rozsahu přiměřený počet méně podstatné četné závažné

Výsledky:

- originální původní i převzaté netriviální kompilace citované z literatury opsané

Rozsah práce:

- veliký standardní dostatečný nedostatečný

Grafická, jazyková a formální úroveň:

- vynikající velmi dobrá průměrná podprůměrná nevyhovující

Tiskové chyby:

- téměř žádné vzhledem k rozsahu a tématu přiměřený počet četné

Celková úroveň práce:

- vynikající velmi dobrá průměrná podprůměrná nevyhovující

Slovní vyjádření, komentáře a připomínky vedoucího/oponenta:

Diplomová práce Bc. Martina Crhána nadepsaná *Full Centroid Molecular Dynamics through Machine Learning* byla vypracována na Fyzikálním ústavu Univerzity Karlovy pod vedením dr. Ondřeje Maršálka. Práce je zaměřena na tzv. *centroid molecular dynamics*, což je přístup usilující o získání dynamických informací z molekulárně-dynamických simulací s dráhovými integrály. Ústřední myšlenkou stávající práce je nahradit kvantové výpočty s dráhovými integrály simulacemi klasickými na před-vypočítaném povrchu volné energie. K získání kvalifikovaného odhadu toho povrchu (resp. jeho derivací) jsou přitom využity techniky strojového učení.

Práce vyžadovala porozumění dosti rozsáhlých oblastí, autor musel nastudovat a zpracovat partie z klasické a kvantové mechaniky, molekulární teorie, statistické mechaniky a molekulárních simulací, jakož i rychle se rozvíjející oblast strojového učení. To se podařilo a autor tyto oblasti zpracoval ve srozumitelně sepsaném úvodním textu, ke kterému nemám podstatnějších námitek. Důraz je v literárním úvodu kladen na fyzikální partie (kvantová mechanika s dráhovými integrály, teorie lineární odezvy, metody molekulových simulací), oblast strojového učení je naopak probrána velmi stručně.

Jako v každém větším textu, i v předložené práci může čtenář nalézt drobnější detaily k vylepšení. Např. v posledním členu rovnice 2.4 nejspíš chybí nábojové číslo atomového jádra, chybička bude nejspíš také v posledním členu rovnice 2.7, občas není úplně jednotné značení různých matematických objektů, jako jsou vektory nebo operátory (srov. např. 2.19 a 2.20 či 2.113), ne vždy jsou všechny symboly představeny, jakkoliv jejich význam plyne z kontextu (např. Q v 2.88). Ze stylistického pohledu bych ocenil, kdybych se již na začátku práce dozvěděl, co je konečným cílem celé studie. To je jasně vysvětleno až v oddíle 3.4. Práci by v úvodní části prospěl výraznější grafický doprovod, který by vizualizoval představované koncepty (při diskuzi technik založených na dráhových integrálech se řada schémat přímo nabízí).

Ve výsledkové části diskutuje autor jednak modelové případy v jedné a ve dvou dimenzích se známým numericky přesným či analytickým řešením. Autor ukazuje, že navržené schéma v těchto případech většinou dobře souhlasí s adiabatickou implementací CMD, má ovšem určité interpretační výhody. Větší rozdíly jsou patrné pro realistické molekulární simulace kapalně vody a těžké vody.

Práce je sepsaná anglicky, srozumitelně a jasně. Text umožňuje práci reprodukovat, metodika je popsána dostatečně detailně. Převzaté zdroje jsou správně citovány, jen jako drobnost bych zmínil jistou nejednotnost ve formátu citací, např. ve jménech autorů. V úvodní části by možná mohly být citace někdy i častější – autor se odkazuje na absolvované přednášky, které ale pro čtenáře nejsou relevantním zdrojem. Práci by slušel seznam zkratk a použitých symbolů, jde ale opět pouze o drobnost. Autor jasně odděluje svůj přínos od prací dřív publikovaných.

Práce se mi celkově líbila. Konstatuji, že jde o kvalitní a původní práci. Doporučuji práci bez váhání k obhajobě a navrhuji hodnotit stupněm výborně.

Případné otázky při obhajobě a náměty do diskuze:

V rozpravě by se autor mohl zaměřit na následující otázky:

1. Bylo by možné navrženou implementaci CMD použít pro modelování chemických reakcí? Měla by to nějakou výhodu oproti běžně užívaným technikám? Bylo by možné využít stávající typ neuronové sítě nebo by bylo třeba sáhnout k výraznějším modifikacím?
2. Mezi A-CMD a interpolovanou CMD byly někdy vidět určité rozdíly. V případě simulace vody autor naznačuje, že výsledky s navrženou modifikací jsou fyzikálně smysluplnější. Poprosil bych autora o hlubší diskuzi, byť třeba spekulativní, o původu těchto změn. Je možné, že jde o důsledek nedostatečné konvergence, ať už z pohledu počtu „chodců“ v simulacích s dráhovými integrály nebo třeba v kontextu samotných neuronových sítí? Nemůže hrát roli jiné termostatování v obou typech simulací?
3. V návaznosti na předchozí otázku: Mohl by autor navrhnout nízko-dimenzionální model vykazující podobně silný rozdíl mezi oběma přístupy?
4. Na str. 45 autor mluví o možnosti, že k systému může být připojen slabě vázaný termostat, který nenarušuje dynamiku systému, ale garantuje asymptoticky kanonické rozdělení. Mohl by autor specifikovat, jak správně termostat nastavit?

5. Zkoušel autor kvantifikovat podobnost výsledků získaných pomocí různých přístupů?
6. Vnímá autor navržený přístup jako aproximaci k adiabatické CMD, zefektivnění této metody nebo její vylepšení?

Práci

doporučuji

nedoporučuji

uznat jako diplomovou/bakalářskou.

Navrhuji hodnocení stupněm:

výborně **velmi dobře** **dobře** **neprospěl/a**

Místo, datum a podpis vedoucího/oponenta: Praha, 28. 8. 2023, Petr Slaviček