



Oponent:

Ing. et Ing. Štěpán Sršeň, Ph.D.

Vysoká škola chemicko-technologická v Praze

Fakulta chemicko-inženýrská

Ústav fyzikální chemie

Technická 5, 166 28 Praha

Student: Bc. Jakub Martinka

Témá práce: Neadiabatická molekulová dynamika fotochemických procesů akcelerovaná pomocí strojového učení

### **Hodnocení oponenta k diplomové práci Bc. Jakuba Martinky**

Práce se zabývá neadiabatickou molekulovou dynamikou na případu molekuly vinyl bromidu a zkoumá její možné urychlení pomocí metod strojového učení. Kandidát se nejprve věnuje popisu teoretických metod, který následují kapitoly s výsledky neadiabatické dynamiky a strojového učení. Oceňuji, že se jedná se o aktuální a ambiciózní téma vyžadující znalosti ze dvou odlišných oborů. Zatímco dynamika v základním elektronovém stavu typicky představuje pro strojové učení jednoduchý problém, u neadiabatické dynamiky je tomu právě naopak.

Formálně je práce na vysoké úrovni pouze s malým množstvím překlepů a gramatických chyb, kterým se v textu o tomto rozsahu v podstatě nelze vyvarovat. Práce se čte dobře, má dostatečný rozsah a je logicky členěna. Množství citací je dostatečné, nicméně lehkou výtku si zaslouží jejich občasné vzdálené umístění od citovaného obsahu, někdy i dalece pod ním. Jakkoliv je odborné názvosloví obtížně přeložitelné do češtiny, je potřeba dát si pozor na polovičaté a krkolomné překlady: zde bych uvedl jako příklad v práci často používaný výraz „jádrová *ridge* regrese“. Překlad pro *ridge regression* zní „hřebenová regrese“, nicméně i plně anglický výraz by byl v pořádku.

Teoretická sekce je psána erudovaně a v dostatečném rozsahu. Pouze občas student popisuje zbytečně složité koncepty, které bez dalšího kontextu nevedou k lepšímu porozumění, ale spíše naopak k zhoršení plynulosti textu a drobným chybám v rovnicích. Větší prostor by mohl být věnován multireferenčním metodám, které jsou vhodnější pro neadiabatickou dynamiku, klidně na úkor rozsáhlého popisu metod jednoreferenčních. Z textu mi také není plně zřejmý způsob zahrnutí gradientů do minimalizované kritériální funkce použité regresní metody.

Ve výsledkové sekci autor studuje disociaci vinyl bromidu po fotoexcitaci. Oceňuji, že se kandidát podrobně zabývá možnými zdroji chyb. Zároveň oceňuji, že student prezentuje absorpční spektra na absolutní škále, což není v komunitě úplná samozřejmost. Co se týká aplikace strojového učení na neadiabatickou dynamiku, tak student zkoumá možnosti nasazení v rámci omezeného



konfiguračního prostoru, přičemž výsledky nejsou prozatím úplně povzbudivé. Nicméně jedná se složitý problém, na kterém pracují i přední výzkumné skupiny v oboru.

Otázky k obhajobě:

1. Student používá bázi modifikovanou zahrnutím difuzních funkcí na atomech uhlíku. Standardně bývají pro vyvážený popis difuzní funkce zahrnuty pro všechny těžké atomy. Dojde ke změně výsledků, pokud je difúzní funkce přidána i atomu bromu?
2. Kandidát uvádí jako referenci metodu spřažených klastrů EOM-CCSD(T) s poruchovými triple excitacemi. Existence takové metody mi není známa a manuálu uvedeného výpočetního programu také ne. Zároveň neiterativní poruchové metody standardně neposkytují oscilátorové síly, které v práci uvedené jsou. Nejedná se o vývojářskou verzi programu? Prosím o vyjádření k této metodě.
3. Student sám správně píše, že metoda ADC(2) není úplně vhodná pro popis konických intersekcí a disociace, přesto je tato metoda použita právě pro popis těchto fenoménů v rámci neadiabatické molekulové dynamiky. Zároveň obě použité metody, ADC(2) a CASSCF, poskytují dle výsledků odlišné pořadí stavů než referenční metoda spřažených klastrů. Co to znamená pro důvěryhodnost výsledků dynamických simulací? Existuje v literatuře experimentální či teoretická doba disociace pro srovnání se získanou hodnotou 20-25 fs? Nebylo by pro tuto molekulu možné použít, alespoň pro menší množství trajektorií, multireferenční metodu s větší mírou zahrnutí dynamické korelace jako například CASPT2?
4. Chyby predikce jsou celkově vysoké vzhledem k použitému množství trénovacích dat. Zároveň autor uvádí zhoršující se chybu predikce s počtem trénovacích dat. Takový výsledek je lehce znepokojivý. Je stejný trend pozorovatelný i pro základní stav? Zkoušel kandidát trénovat model s ještě menším množstvím dat? Tato data by mohla by pomoci objasnit příčinu například v zahrnutí geometrií z oblastí konických intersekcí, kde nejsou fitované hyperplochy hladké.

Na závěr bych rád konstatoval, že délka tohoto posudku odráží zájem oponenta, a nikoliv nekvalitu práce.

Diplomovou práci Bc. Jakuba Martinky doporučuji k obhajobě s hodnocením A „výborně“.

V Praze dne 6. září 2022

Ing. et Ing. Štěpán Sršeň, Ph.D.