



VILNIAUS UNIVERSITETAS
FIZIKOS FAKULTETAS
TEORINĖS FIZIKOS KATEDRA

Martynas Šubonis

Molekulių krūvio tankio atvaizdavimas taškiniais krūviais naudojantis optimalių
parametrų paieškos algoritmais

Pagrindinių studijų baigiamasis darbas

Fizikos studijų programa

Studentas

Martynas Šubonis

Darbo vadovas

Prof. dr. Darius Abramavičius

Konsultantas

dr. Stepas Toliautas

Recenzentas

doc. dr. Juozas Šulskus

Katedros vedėjas

Prof. habil. dr. Leonas Valkūnas

Vilnius, 2016

Turinys

Įvadas	2
1 Literatūros apžvalga	3
1.1 Molekulinės kvantinės mechanikos bazinės funkcijos	3
1.2 Sąveika tarp molekulių	6
1.2.1 Taškinių dipolių aproksimacija (TDA)	6
1.2.2 Krūvio tankio ir šuolio tankio kubų metodai	7
2 Krūvio tankio modeliavimo metodas	9
2.1 Modelinio atkaitinimo (MA) metodas	9
2.2 Optimizuojamos kaštų funkcijos įvedimas	10
2.3 Potencialo matavimo taškų $\mathbf{r}_m$ generavimas	12
3 Taikymas ir rezultatai	14
3.1 Optimizacijos algoritmo parametrai	14
3.1.1 „Temperatūra“	14
3.1.2 Krūvio pokyčio verčių Δq_i generavimas	15
3.1.3 Apribojimai	17
3.2 Vandens molekulės krūvio tankio aproksimacija taškiniais krūviais	18
3.3 Chlorofilo <i>a</i> molekulės krūvio tankio aproksimacija taškiniais krūviais	24
Išvados	29
Santrauka	30
Summary	31
Literatūra	32
Priedai	34

Išvadas

Sprendžiant kvantinės mechanikos uždavinius sudėtingoms daugiadalelinėms sistemoms, dažnai susiduriama su didelėmis skaičiavimų sąnaudomis. Priklausomai nuo nagrinėjamos molekulės ar molekulių sistemos struktūros, krūvio tankio ir jo potencialo apskaičiavimas erdvėje tampa sudėtingu uždaviniu. Todėl dažnai bandoma sukurti apytikslius metodus, kurie sumažintų skaičiavimų sąnaudas, minimaliai prarasdami tikslumą.

Vienas iš tokių metodų yra šuolio krūvio tankio apskaičiavimas įvertinant Kulono sąveiką tarp molekulių atomų taškinių krūvių verčių, kurios gaunamos elektrostatinio potencialo aproksimacijos metu [1]. Šis metodas suteikia daug informacijos apie molekulės 3-D struktūrą ir reikalauja mažų skaičiavimo sąnaudų.

Taikant tokį metodą atsiranda papildomas uždavinys – potencialo aproksimavimas. Aproksimacijos metu dirbama labai sudėtingoje kintamųjų erdvėje, kur be globalaus minimumo egzistuoja daugybė lokalių minimumų. Todėl reikalingas optimizacijos algoritmas, kuris ne tik konverguotų į globalų minimumą didelėje optimizacijos erdvėje, bet taip pat reikalautų mažų skaičiavimų sąnaudų, kas leistų jį pritaikyti didelėms molekulėms.

Šio darbo tikslas – sukurti bendrą kompiuterinį algoritmą remiantis skaitine optimizacija, kuris atvaizduotų bet kokių molekulių krūvio tankį pasirinktu taškinių krūvių rinkiniu. Algoritmo veikimo principas pagrįstas modelinio atkaitinimo metodo koncepcija [2, p. 549]. Algoritmo tikslumas ir našumas iš pradžių tikrinamas su maža molekule. Radus optimalius parametrus, algoritmo taikymo ribos praplėstos iki didelių molekulių.

Darbo uždaviniai:

1. Naudojant GAMESS [3, 4] cheminio paketo apskaičiuotus atominių orbitalių skleidinių koeficientus, apskaičiuoti vandens molekulės krūvio tankį.
2. Parašyti programą, kuri generuotų tinkamai pasiskirsčiusius matavimo taškus $\mathbf{r}_m$ aplink vandens molekulę.
3. Taikant krūvio tankio kubų (KTK) metodą ir naudojant 1 punkte apskaičiuotą krūvio tankį, apskaičiuoti „tikrąjį“ (atskaitinį) vandens molekulės potencialą matavimo taškuose $\mathbf{r}_m$.
4. Aproksimuojant atskaitinį potencialą matavimo taškuose $\mathbf{r}_m$, nustatyti atitinkamų taškinių krūvių vertes.
5. Praplėsti algoritmo taikymo ribas iki daugiadalelinių sistemų ir pakartoti 1 – 4 punktus su chlorofilo *a* molekule.

1. Literatūros apžvalga

Išspręsti stacionariąją Šriodingerio lygtį, aprašančią elektronų judėjimą molekulėje (žr. 1.1 lygtį), net ir apytiksliais metodais yra sudėtingas uždavinys [5].

$$\left[\frac{-\hbar^2}{2\mu} \nabla^2 + V(\mathbf{r}_1 \dots \mathbf{r}_N) \right] \psi(\mathbf{r}_1 \dots \mathbf{r}_N) = E\psi(\mathbf{r}_1 \dots \mathbf{r}_N) \quad (1.1)$$

Analiziškai aprašyti tokią daugiaelektroninę sistemą neįmanoma, net jeigu yra taikomas Borno ir Openheimerio artinys [6]. Dažniausiai yra naudojami įvairūs modelio suparastinimai ir aproksimacijos, kurių pagalba apskaičiuojamos banginės funkcijos ir molekulės būsenų energijos fiksuotoms atomų branduolių padėtimis. Remiantis gautomis banginėmis funkcijomis, galima įvertinti molekulės atomų krūvius, elektronų krūvių pasiskirstymą molekulėje ir molekulės dipolinį momentą.

Turint šiuos molekulių fizikinius parametrus, įvairiais metodais [1, 7] galima nustatyti molekulių tarpusavio sąveiką, kuri yra molekulinės dinamikos, molekulinį Monte Carlo simuliacijų ir kondensuotų medžiagų fizikinių savybių skaičiavimų pagrindas.

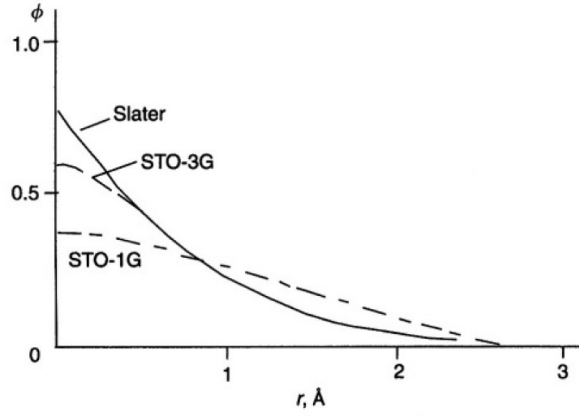
1.1 Molekulinės kvantinės mechanikos bazinės funkcijos

Bazinės atominės funkcijos yra matematinių funkcijų rinkinys, kurių tiesinė kombinacija apibūdina molekulinės orbitales [8]. Šios funkcijos yra dažniausiai (bet ne visada) centruotos ant atomų branduolių.

Elektronų pasiskirstymas aplink atomą gali būti atvaizduotas daugybe būdų. Vandenilinės funkcijos, gautos kaip Šriodingerio lygties sprendiniai vandenilio atomui, daugianariai su koreguojamais parametrais, Sleiterio funkcijos, Gauso funkcijos – visos šios funkcijos gali atvaizduoti atomines orbitales. Tačiau tik Sleiterio ir Gauso funkcijos yra „paprastos“ matematiniu požiūriu. Todėl šiuo metu jos ir yra naudojamos kaip bazinės funkcijos molekuliniuose skaičiavimuose.

Sleiterio orbitalės yra geros aproksimacijos atominėms banginėms funkcijoms ir būtų natūralus pasirinkimas bazinėms funkcijoms, jeigu nereikalautų tiek daug sąnaudų skaičiuojant daugiacentrius integralus. Daugiacentriai integralai yra lengvai analiziškai apskaičiuojami naudojant Gauso orbitales. Tačiau, norint išlaikyti tą patį tikslumą (ypač arti ir toli nuo branduolio), reikia vienos Sleiterio orbitalės vietoje naudoti nuo 3 iki keliolikos Gauso orbitalių. Abiem skaičiavimo atvejais kompiuterio resursų poreikis yra panašus.

Gauso ir Sleiterio orbitalių bendriausios formulės:



$$\begin{aligned}\phi(\text{Slater}) &= \left(\frac{\xi^3}{\pi}\right)^{1/2} e^{-\xi r} = 0.7790 e^{-1.24r} \\ \phi(\text{STO-1G}) &= \left(\frac{2\alpha}{\pi}\right)^{3/4} e^{-\alpha r^2} = 0.3696 e^{-0.4166r^2} \\ \phi(\text{STO-3G}) &= 0.4446 \left(\frac{2\alpha}{\pi}\right)^{3/4} e^{-\alpha r^2} + 0.5353 \left(\frac{2\alpha}{\pi}\right)^{3/4} e^{-\alpha r^2} + 0.1543 \left(\frac{2\alpha}{\pi}\right)^{3/4} e^{-\alpha r^2} \\ &= 0.0835 e^{-0.1689r^2} + 0.2678 e^{-0.6239r^2} + 0.2769 e^{-3.4253r^2}\end{aligned}$$

1.1 pav.: Atominės orbitalės: Sleiterio, STO-1G ir STO-3G atominių orbitalių palyginimas vandenilio atomui. Parodyta Sleiterio funkcija yra tinkamiausia iš galimų Sleiterio funkcijų aproksimacijų vandenilio atomui molekulinėje aplinkoje. Parodytos Gausinės funkcijos yra geriausios 1-G ir 3-G aproksimacijos minėtajai Sleiterio funkcijai [8].

$$\chi_{SO}(n, l, m) = R_{nl}^S(r, \xi) Y_{lm}(\vartheta, \varphi), \quad (1.2)$$

čia $\chi_{SO}(n, l, m)$ žymi Sleiterio bazinę funkciją, $R_{nl}^S(r, \xi)$ Sleiterio orbitalės radialinę dalį ir $Y_{lm}(\vartheta, \varphi)$ Sleiterio orbitalės kampinę dalį,

$$R_{nl}^S(r, \xi) = (2\xi)^{n+\frac{1}{2}} [(2n)!]^{-\frac{1}{2}} r^{n-1} e^{-\xi r}, \quad n \geq l+1; \quad (1.3)$$

$$\chi_{GO}(n, l, m) = R_{nl}^G(r, \alpha) Y_{lm}(\vartheta, \varphi), \quad (1.4)$$

čia $\chi_{GO}(n, l, m)$ žymi Gauso bazinę funkciją, $R_{nl}^G(r, \alpha)$ Gauso orbitalės radialinę dalį ir $Y_{lm}(\vartheta, \varphi)$ Gauso orbitalės kampinę dalį,

$$R_{nl}^G(r, \alpha) = (2)^{n+1} [(2n-1)!!]^{-\frac{1}{2}} (2\pi)^{-\frac{1}{4}} \alpha^{(2n+1)/2} r^{n-1} e^{-\alpha r^2}, \quad (1.5)$$

čia $(2n+1)!! = 1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1)$, n - sveikasis skaičius, kuris įgyja vertes $(l+1)$, $(l+3)$, $(l+5)$, ... Šiose išraiškose sferinės funkcijos $Y_{lm}(\vartheta, \varphi)$ yra normuotos į vienetą.

Molekulių elektroniniam skaičiavimams naudojamos Gauso orbitalės Dekarto koordinatėse:

$$\chi_{GO}(n, l, m) = N(\alpha; l, m, n) x^l y^m z^n e^{-\alpha r^2}, \quad (1.6)$$

$$N(\alpha; l, m, n) = \left(\frac{2}{\pi}\right)^{\frac{3}{4}} \left[\frac{\alpha^{l+m+n+(3/2)} 2^{2(l+n+m)}}{(2l-1)!!(2m-1)!!(2n-1)!!} \right]^{\frac{1}{2}}. \quad (1.7)$$

Kaip buvo minėta anksčiau, norint teisingai aproksimuoti elektronų pasiskirstymą arti ir toli nuo atomų branduolių, turime vienos Gauso orbitalės vietoje naudoti jų tiesinę kombinaciją. Tokiu atveju išraiška 1.6 atrodo taip:

$$\chi_{GO}(n, l, m) \approx x^l y^m z^n \sum_{i=1}^k c_i e^{-\alpha_i r^2}, \quad (1.8)$$

Stengiantis išlaikyti Gauso orbitalių tikslumą, k parenkamas ≥ 10 . Tokios tiesinės kombinacijos metu žymiai padidėja bazinių funkcijų skaičius. Tačiau nepriklausomų radialiųjų funkcijų skaičius k ir tuo pačiu bazinių funkcijų skaičius gali būti sumažintas sugrupavus artimų verčių eksponentes į bendras grupes:

$$\chi_{GO}(n, l, m) \approx x^l y^m z^n \sum_{i=1}^k c_i \left(\sum_{j=1}^{k_i} \alpha_j e^{-\alpha_j r^2} \right) \quad (1.9)$$

Naudojant vieną sugrupuotą radialiąją dalį kiekvienai užimtai atomo orbitalei, gaunamas mažiausias bazinių funkcijų rinkinys molekulių skaičiavimuose (**minimalus bazinių funkcijų rinkinys**). Bazinių funkcijų rinkinių žymėjimas – STO-3G, STO-4G, ... , STO-6G.

Platesniuose bazinių funkcijų rinkiniuose valentinių elektronų sluoksnio radialiosios dalys yra sudarytos iš dviejų grupių. Tai **valentiniai double zeta (VDZ) rinkiniai**.

Plečiant VDZ bazinių funkcijų rinkinys papildomas poliarizacinėmis funkcijomis. Pirmo ir antro periodo atomams tai papildoma d funkcija (**DZP bazinių funkcijų rinkinys**). Vandenilio atomui pridedama papildoma p funkcija. Tokie papildymai žymimi žvaigždutėmis (6-31G* – sunkesnių atomų poliarizacinės funkcijos).

„Triple zeta” (TZ) tai dar didesnis bazinių funkcijų išplėtimas poliarizacinėmis funkcijomis.

Išplėstųjų bazinių funkcijų sunkesnių atomų bazinių funkcijų rinkiniuose papildomai įskaitomos f ir g valentinių elektronų funkcijos, o H atomui – d funkcijos.

Išspręsti 1.1 lygtį papildomai naudojama dar viena aproksimacija – viendalelis Hartrio ir Foko artinys. Naudojant Hartrio ir Foko metodą teigiama, kad elektrono būseną molekulėje aprašo tik nuo to elektrono koordinatų priklausanti banginė funkcija. Viendalelėje kvantinėje molekulių teorijoje elektronų sistemos banginė funkcija dažniausiai pasirenkama Sleiterio determinanto pavidalo. Determinantą sudaro atskirų elektronų banginės funkcijos, priklausančios tik nuo individualaus elektrono koordinatų, bei turinčios sukinio dalį (žr. 1.10, 1.11 formules). Taip sukonstruota banginė funkcija tenkina Paulio draudimo principą: fermionų sistemos banginė funkcija turi pakeisti ženklą, kai yra sukeičiamos dviejų dalelių koordinatės.

$$\Phi = (N!)^{-1/2} |\psi_1(x_1)\bar{\psi}_1(x_1) \dots \psi_n(x_n)\bar{\psi}_n(x_n)|, \quad (1.10)$$

čia

$$\psi_i = \varphi_i \alpha \text{ ir } \bar{\psi}_i = \varphi_i \beta, \quad (1.11)$$

α ir β - sukininės orbitalės. Atskirų elektronų erdvinės banginių funkcijų dalys φ_i laikomos orto-normuotomis.

Hartrio ir Foko lygtys sprendžiamos suderintinio lauko metodu. Taikant suderintinio lauko metodą, erdvinės orbitalių dalys variaciniu principu optimizuojamos, siekiant nustatyti „geriausią“ vieno determinanto banginę funkciją. Pradinis sprendinys ($\varphi_m^{(0)}$) yra pasirenkamas laisvai (dažniausiai jis yra supaprastintos sistemos Šriodingerio lygties sprendinys). Naudojantis gautuoju sprendiniu, iš naujo įvertinamas Foko operatorius ir pakartotinai išsprendus pradinę lygtį, nustatomas naujas sprendinių rinkys ($\varphi_m^{(1)}$). Tokia procedūra tęsiama tol, kol elektronų pasiskirstymas ir sistemos energija stabilizuojasi iš anksto nurodytu tikslumu.

1.2 Sąveika tarp molekulių

1.2.1 Taškinių dipolių aproksimacija (TDA)

Klasikinis metodas įvertinti elektrinę sąveiką tarp pigmentų šuolio momentų buvo sukurtas Fiorsterio, 1948 metais [7]. Fiorsteris teigė, kad aprašant pilnutinę sąveiką tarp elektronų, joje dominuoja Kulono sąveika ir kad Kulono sąveika tarp molekulių gali būti gana tiksliai aprašyta kaip sąveika tarp dviejų šuolio dipolių:

$$V_{DA}^{\text{Pilnut.}} \approx V_{DA}^{\text{Kulono}} \approx V_{DA}^{\text{dip-dip}} = \frac{\kappa |\mu_D| |\mu_A|}{4\pi\epsilon_0 R_{DA}^3}, \quad (1.12)$$

čia μ_D ir μ_A yra šuolio donorinių ir akceptorinių dipolių amplitudės, R_{DA} yra atstumas tarp dipolių, ir κ yra orientacinis faktorius:

$$\kappa = \hat{\mu}_D \cdot \hat{\mu}_A - 3(\hat{\mu}_D \cdot \hat{R}_{DA})(\hat{\mu}_A \cdot \hat{R}_{DA}), \quad (1.13)$$

čia „ $\hat{\mu}$ “ žymi vienetinius vektorius. Pirmoji aproksimacija, kai pilnutinė elektrinė sąveika yra aprašoma tiesiog Kulonine sąveika, duoda nedideles paklaidas atliekant skaičiavimus molekulėms, kurių orbitalės stipriai tarpusavyje persikloja. Problema atsiranda atliekant antrąją, taškinių dipolių, aproksimaciją (TDA). Taškinių dipolių aproksimacijos metu vertinant pilną elektronų sąveiką mažiems atstumams - kai R_{DA} yra lygus arba mažesnis už pačių molekulių matmenis, gaunamos didelės paklaidos. Taip atsitinka todėl, kad molekulių šuolių momentai negali būti aprašyti taškiniiais krūviais. Taip pat tampa sudėtinga griežtai apibrėžti šuolių momentų centrus. Molekulės

su išplitusiais (tokios kaip karotenoidai) ar asimetriniais šuolių tankiais dar labiau padidina TDA paklaidas, todėl šis metodas yra netinkamas sąveikos tarp didelių molekulių skaičiavimams.

1.2.2 Krūvio tankio ir šuolio tankio kubų metodai

1.2.1 skyriuje buvo aptarti taškinės dipolių aproksimacijos trūkumai. Akivaizdu, kad norint įvertinti Kulono sąveiką reikia geresnių metodų negu TDA. Vienas iš tikslesnių yra šuolio tankio kubų metodas [7]. Šiame metode naudojami kvantinės mechanikos skaičiavimai nustatant banginės funkcijas pradinėms ir galutinėms molekulės būsenoms. Naudojantis banginėmis funkcijomis šuolio krūvio tankis apskaičiuojamas taip:

$$M_k(\mathbf{r}) = \int \psi_{kg}(\mathbf{r})\psi_{ke}^*(\mathbf{r})d\mathbf{s}, \quad (1.14)$$

čia g ir e žymi atitinkamai pagrindines ir sužadintas molekulės k būsenas. Sąveika tarp donorinio ir akceptorinio šuolio tankių tiksliai aprašo Kulono sąveiką (tai yra ekvivalentu visų multipolio skleidinio narių įskaitymui):

$$V_{DA}^{\text{Kulono}} = \frac{1}{4\pi\epsilon_0} \int \frac{M_D(\mathbf{r}_i)M_A(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_i d\mathbf{r}_j \quad (1.15)$$

Skaičiavimams dažniausiai naudojami kvantinės mechanikos paketai, kurie suintegruoja šuolio tankius į baigtinio dydžio tūrio elementų masyvus:

$$M_k(x, y, z) = V \int_z^{z+\delta z} \int_y^{y+\delta y} \int_x^{x+\delta x} \int_s \psi_{kg}\psi_{ke}^* ds dx dy dz, \quad (1.16)$$

čia V yra tūrio elementas. Baigtinio dydžio tūrio elementų rinkinys yra vadinamas šuolio tankio kubu (ŠTK). Analogiškus skaičiavimus galima atlikti su molekulės pasirinktos būsenos krūvio tankiu ($\int_s \psi_{ki}\psi_{ki}^* ds$). Šiuo atveju gaunamas baigtinio dydžio tūrio elementų rinkinys, kuris vadinamas krūvio tankio kubu (KTK).

Norint apskaičiuoti Kulono sąveiką tarp dviejų ŠTK (ar KTK), susumuojama Kulono sąveika tarp visų kiekvieno kubo elementų:

$$V_{DA}^{\text{ŠTK}} = \frac{1}{4\pi\epsilon_0} \sum_{i,j} \frac{M_D(i)M_A(j)}{r_{ij}}, \quad (1.17)$$

čia i ir j žymi donorinio ir akceptorinio ŠTK (x, y, z) elementus, o r_{ij} yra atstumas tarp elementų i ir j . Krūvio tankio kubo metodas leidžia dideliu tikslumu įvertinti Kulono sąveiką ir yra galimas esant bet kokiai tarpmolekulinei atskirčiai. Kadangi yra naudojamas pilnasis šuolio krūvio tankis, be jokių multipolio skleidinių, šiame metode nereikia nustatyti molekulių centrų. ŠTK metodo tikslumą riboja tik pačių banginių funkcijų tikslumas ir tūrio elementų dydis. Daugumoje taikymo sričių ŠTK metodas duoda akivaizdžiai tikslesnius rezultatus lyginant su TDA ir yra skaitmeniškai

pritaikomas didelėms, biologiškai svarbioms molekulėms, tokioms kaip chlorofilai, karotenoidai ir t.t. Reikia pastebėti, kad ŠTK metodas įvertina Kulono sąveiką, o ne pilnutinę sąveiką tarp elektronų. Todėl į orbitalių tarpusavio sanklodą turi būti atsižvelgta atskirai.

2. Krūvio tankio modeliavimo metodas

2.1 Modelinio atkaitinimo (MA) metodas

Modelinio atkaitinimo (MA) metodas susilaukė ypatingo susidomėjimo sprendžiant didelės apimties optimizacinius uždavinius, ypač tuos, kur be globalaus minimumo yra daugybė mažesnių, lokalių minimumų [9, 10]. Modelinio atkaitinimo metodu buvo efektyviai išspręsta garsioji „Keliaujančiojo pirklio problema“, surandant trumpiausią atstumą tarp N miestų, kuriuos pirklys turi aplankyti iš eilės.

Modelinio atkaitinimo metode yra elementų, kurie atspindi analogiją su termodinamika [2, p. 550]. Aukštos temperatūros skysčių molekulės laisvai juda viena kitos atžvilgiu. Jeigu skystis yra atšaldomas lėtai, šiluminis judris yra prarandamas. Lėto atšaldymo metu atomai dažniausiai sugeba suformuoti kristalines struktūras, kurios yra tvarkingos milijardus kartų didesniuose atstumuose, negu pačio atomo matmenys. Tokie kristalai yra minimalios energijos būsenoje. Įdomiausia yra tai, kad lėtai šalančios sistemos sugeba „surasti“ minimalios energijos būseną.

Dauguma optimizavimo algoritmų yra pagrįsti paprasta idėja - nuo optimizacijos pradžios judama funkcijos mažėjimo kryptimi, kol pasiekiamas minimumas. Tai dažnai veda prie lokalaus, o ne globalaus minimumo. Modelinio atkaitinimo metodas yra pagrįstas kitokia procedūra. Įvedamas Bolcmano tikimybinės pasiskirstymo daugiklis:

$$Prob(E) = e^{-\frac{E}{kT}} \quad (2.1)$$

Bolcmano faktorius išreiškia idėją, kad sistema, esanti šiluminės pusiausvyros būsenoje temperatūroje T , turi energijos tikimybinį pasiskirstymą. Net žemose temperatūrose yra tikimybė, kad sistema bus aukštos energijos būsenoje. Atsiranda atitinkama tikimybė sistemai ištrūkti iš lokalaus energijos minimo, siekiant surasti geresnį, globalesnį minimumą. Kitaip tariant, sistemos energija ne tik mažėja, bet kartais ir padidėja; bet kuo žemesnė temperatūra, tuo mažiau tikėtinas sistemos pajudėjimas energijos padidėjimo kryptimi.

1953 metais, Metropolis ir jo kolegos pirmieji pritaikė tokio pobūdžio idėjas skaičiavimuose. Tariau, kad modelinė termodinaminė sistema gali keisti savo konfigūraciją iš energijos E_1 į energiją E_2 su tikimybe $p = \exp[-(E_2/E_1)/kT]$. Jeigu $E_2 < E_1$, ši tikimybė tampa didesnė negu vienetą; tokiais atvejais pokyčio tikimybė yra tiesiog prilyginama tikimybei $p = 1$. Tai reiškia, kad kai tik sistema rasdavo žemesnės energijos būseną, ji visada pereidavo į ją.

Nors modelinio atkaitinimo metodas yra plačiau taikomas optimizuojant kombinatorikos uždavinius, jis taip pat gali būti pritaikytas uždaviniams tolydžioje, N -matėje erdvėje. Šiuo atveju yra

ieškomas minimumas (idealaus scenarijaus atveju globalus minimumas) kažkokios funkcijos $f(\mathbf{x})$ erdvėje, turinčioje lokalius minimumus. Čia $\mathbf{x}$ yra N - matis vektorius. Turime keturis modelinio atkaitinimo procedūros elementus:

- Funkcija $f(\mathbf{x})$, kurios vertė optimizuojama.
- Sistemos būseną, kuri yra taškas $\mathbf{x}$.
- Kontroluojamasis parametras T (kaip prieš tai buvo minėta, atitinka temperatūrą, kuri modelinio atkaitinimo metodo metu yra tolygiai mažinama).
- Atsitiktinių konfigūracijos pokyčių generatorius, kuris priverčia sistemą keisti būseną iš $\mathbf{x}$ į būseną $\mathbf{x} + \Delta\mathbf{x}$.

Paskutinis iš šių minėtų elementų yra sudėtingiausias. Literatūros šaltiniuose [11, 12, 13] pasiūlomi keletas metodų kaip parinkti $\Delta\mathbf{x}$, bet nei vienas iš jų nėra visiškai patikimas. Dažniausiai $\Delta\mathbf{x}$ parinkimas priklauso nuo pačio optimizacijos uždavinio ypatumų, todėl norint sukonstruoti $\Delta\mathbf{x}$ generavimo metodą, reikia atlikti papildomą uždavinio analizę.

Taikant modelinio atkaitinimo metodą, dažnai susiduriama su problema nustatant temperatūros kritimo spartą. Temperatūra turi kristi pakankamai lėtai, kad sistema galėtų „rasti“ globalų minimumą, bet ne per lėtai, kad neišaugtų skaičiavimų sąnaudos. Keletas temperatūros mažinimo būdų:

- Mažinti temperatūrą T į $(1 - \epsilon)T$ naudojant m sistemos evoliucijos žingsnių, kur ϵ ir m yra empiriškai nustatomi parametrai.
- Pasirenkama konstanta K , kuri atitinka tam tikrą sistemos evoliucijos žingsnių skaičių. Tada temperatūra T yra mažinama naudojant m žingsnių iki vertės $T = T_0(1 - k/K)^\alpha$, kur k yra bendras visų įvykusių žingsnių skaičius. α yra konstanta, kurios vertės gali būti parinktos empiriškai. Optimali α vertė priklauso nuo įvairaus „gylio“ minimumų statistinio pasiskirstymo. Didesnės α vertės sąlygoja daugiau iteracijų žemesnėse temperatūrose.
- Po m sistemos evoliucijos žingsnių pakeičiame temperatūrą T į $\beta \times (f_1 - f_b)$, kur β yra empiriškai parenkamas parametras (vienetų eilės), f_1 yra žemiausia funkcijos vertė „matoma“ optimizacijos algoritmo, ir f_b yra žemiausia pasiekta funkcijos vertė visos optimizacijos metu. Temperatūros mažinimo metu nepatartina sumažinti temperatūrą T daugiau negu dalimi γ (kuri taip pat parenkama empiriškai).

2.2 Optimizuojamos kaštų funkcijos įvedimas

Visų pirma, prieš pradedant optimizaciją, reikia susikurti funkciją, kuri bus optimizuojama. Optimizacijos kontekste tokios funkcijos dažniausiai yra vadinamos kaštų funkcijomis (angliškai

„Cost functions”). Daugeliu atvejų jų forma susideda iš dviejų dalių: „tikslų” verčių, kurias mes norime pasiekti, ir „hipotezės” (verčių, kurias mes spėjame).

Paprastiausias pavidalas, kokį gali įgauti „energijos” funkcija yra:

$$E = \sum_{n=1}^N |y_n - y_{hip.}(\mathbf{x}_n)|, \quad (2.2)$$

čia y_n yra tikroji funkcija, kurią bando aproksimuoti funkcija $y_{hip.}(\mathbf{x}_n)$.

Labiausiai paplitusi kaštų funkcijos forma yra mažiausių kvadratų forma:

$$E = \sum_{n=1}^N (y_n - y_{hip.}(\mathbf{x}_n))^2, \quad (2.3)$$

kuri yra dažnai perrašoma pavidalu, atspindinčiu **vidutinį** (teigiant kad visos $\mathbf{x}_n$ vertės yra vienodai tikėtinos) hipotezės mažiausių kvadratų nuokrypį nuo tikrosios funkcijos:

$$E = \frac{1}{N} \sum_{n=1}^N (y_n - y_{hip.}(\mathbf{x}_n))^2 \quad (2.4)$$

Mažiausių kvadratų pavidalo pasirinkimas turi tvirtą teorinį pagrindą. **Maksimalaus tikėtinumo metodo** pagalba galima įrodyti, kad norint rasti labiausiai tikėtiną pasiskirstymo modelį dydžiui y_n , kurio nepriklausomos atsitiktinės matavimo paklaidos yra Gausiškai pasiskirsčiusios apie kažkokį „tikrąjį” modelį y_x , reikia minimizuoti (žr. 2.4 formulę) [2, p. 778]. Toks rezultatas gaunamas darant prielaidą, kad šių Gauso skirstinių standartiniai nuokrypiai σ yra vienodi visiems matavimo taškams.

Daugybės skaičių su mažomis atsitiktinėmis paklaidomis sumos tikimybinis pasiskirstymas, pagal **centrinę ribinę teoremą**, beveik visada konverguoja į Gauso pasiskirstymą. Todėl matavimo paklaidos yra dažnai modeliuojamos kaip Gauso triukšmas, kuris turi nulinį vidurkį ir žinomą standartinį nuokrypį [14]. Dėl šių priežasčių mažiausių kvadratų metodas yra taikomas uždaviniams, kurių paklaidos yra pasiskirsčiusios nebūtinai pagal Gauso skirstinį.

Mažiausių kvadratų metodo savybės tiesiniuose modeliuose:

- mažos skaitmeninės sąnaudos,
- tarp objektyvių tiesinių įvertinių turi mažiausią dispersiją,
- lengva įvertinti parametrų kovariaciją.

Kadangi netiesiniai modeliai dažnai gali būti aproksimuojami tiesiniais, dauguma prieš tai išvardintų savybių (tam tikru tikslumu) tinka ir netiesiniams modeliams.

Norint aproksimuoti elektrostatinį potencialą taškiniais krūviais, visų pirma reikia gauti „tikrąjį” elektrostatinį potencialą. Elektrostatinio potencialo vertė pasirinktame matavimo taške $\mathbf{r}_m$:

$$V_m = \sum_{i=1}^N \frac{kq_i}{|\mathbf{r}_m - \mathbf{r}_i|} + \sum_{l=1}^L \frac{kq_l}{|\mathbf{r}_m - \mathbf{r}_l|}, \quad (2.5)$$

čia q_i atitinka iš elektronų krūvio tankio KTK metodu gautas taškinių krūvių vertės taškuose $\mathbf{r}_i$. q_l yra branduolių, kurių koordinatės $\mathbf{r}_l$, krūvių vertės.

Taškinių krūvių, kuriais bandoma aproksimuoti „tikrąjį“ elektrostatinį potencialą, kuriamas potencialas yra:

$$V_t. = \sum_{k=1}^K \frac{kq_k}{|\mathbf{r}_m - \mathbf{r}_k|} \quad (2.6)$$

Remiantis (2.4), (2.5), (2.6) išraiškomis, galima užrašyti bet kokios molekulės potencialų kaštų funkciją:

$$E = \frac{1}{M} \sum_{m=1}^M \left(\sum_{i=1}^N \frac{kq_i}{|\mathbf{r}_m - \mathbf{r}_i|} + \sum_{l=1}^L \frac{kq_l}{|\mathbf{r}_m - \mathbf{r}_l|} - \sum_{k=1}^K \frac{kq_k}{|\mathbf{r}_m - \mathbf{r}_k|} \right)^2, \quad (2.7)$$

čia M yra pasirinktų matavimo taškų skaičius.

2.3 Potencialo matavimo taškų $\mathbf{r}_m$ generavimas

Potencialo matavimo taškų $\mathbf{r}_m$ parinkimas gali kelti sunkumų. Visų pirma, matavimo taškai turi būti sugeneruoti tinkamu atstumu nuo molekulės. Tinkamas atstumas reiškia tai, kad taškai negali būti sugeneruoti per arti ar per toli molekulės. Sunkumai atsirandantys generuojant matavimo taškus per arti molekulės:

- Taškiniais krūvais gali būti labai sudėtinga teisingai aproksimuoti molekulės kuriamą potencialą artimose zonose.
- **STO-3G** (žr. 1.1 skyrių) bazės atominės funkcijos dėl mažo Gausinių funkcijų skaičiaus gana prastai aprašo elektronų tankio pasiskirstymą arti ir toli nuo atomų branduolių.

Sugeneravus taškus pakankamai toli nuo molekulės, visa taškinių krūvių aproksimacija gali prarasti savo prasmę. Dideliuose atstumuose nuo molekulės, molekulės kuriamas potencialas gali būti aproksimuotas paprasčiausiu dipoliu. Kaip minėta 1.2.1 skyriuje, dipolinė aproksimacija prastai aprašo elektrinę sąveiką mažuose atstumuose, todėl gauti rezultatai visiškai netiks artėjant prie molekulės.

Šio darbo metu matavimo taškai $\mathbf{r}_m$ buvo generuojami ant atitinkamo spindulio R sferos paviršiaus. Iš pradžių buvo bandoma generuoti **atsitiktinius** taškus. Buvo pastebėta, kad atsitiktinis metodas ganėtinai prastai išlaiko „sferiškumą“ esant mažam matavimo $\mathbf{r}_m$ taškų skaičiui. Matavimo taškų skaičiaus parinkimas yra labai svarbus optimizuojant potencialus. Per didelis taškų

skaičius nesuteikia papildomos informacijos apie potencialo pasiskirstymą erdvėje (nepagerina algoritmo tikslumo) ir tik padidina skaičiavimų sąnaudas. Esant labai mažam matavimo taškų skaičiui, algoritmas negauna pakankamai informacijos apie potencialo pasiskirstymą ir nesukonverguoja į teisingas (realias) taškinių krūvių vertes. Vietoje atsitiktinių taškų buvo nuspręsta generuoti vienodame plote pasiskirsčiusius taškus. Šie taškai buvo generuojami metodais, pateiktais darbe [15]. Pastaba: [15] darbe yra klaida III skyriuje, skaičiuojant konstantą a . Jinai turėtų būti $a = 4\pi/N$.

Be taškų generavimo metodo aktualus yra ir taškų skaičiaus parinkimas. Darbo metu buvo stengtasi surasti optimalų matavimo taškų $\mathbf{r}_m$ skaičių, kuris pakankamai tiksliai atspindėtų molekulės potencialo pasiskirstymą erdvėje. Matavimo taškų skaičiaus parinkimas bus aptartas rezultatų skyriuje.

3. Taikymas ir rezultatai

3.1 Optimizacijos algoritmo parametrai

Šio darbo metu naudotas optimizacijos algoritmas buvo pagrįstas modelinio atkaitinimo metodo koncepcija. Kaip minėta 2.1 skyriuje, modelinio atkaitinimo metodas turi keletą parametrų, kurių parinkimas gali stipriai įtakoti optimizacijos našumą. Viena iš šio darbo užduočių buvo rasti optimalias šių parametrų vertes, kurios leistų optimizacijos algoritmą taikyti įvairioms molekulėms.

3.1.1 „Temperatūra“

Temperatūra yra vienas iš paprasčiau parenkamų simuliuoto atšaldymo metodo parametrų. Temperatūros parametrų nustatymas susideda iš dviejų etapų:

- pradinės temperatūros vertės parinkimo,
- temperatūros kritimo spartos nustatymo.

Šio darbo metu pradinė temperatūra buvo parinkta taip, kad pradinio spėjimo metu tikimybė pereiti į „energijos“ būseną, kurios „energija“ yra didesnė negu dabartinė, būtų lygi 30%. Toks rezultatas, atsižvelgiant į bet koki pradinį spėjimą, visada buvo pasiekiamas skaitmeniškai sprendžiant lygtį (3.1) temperatūros atžvilgiu:

$$e^{-\Delta E_{\text{prad.}}/T_{\text{prad.}}} = 0.3, \quad (3.1)$$

čia 30% vertė buvo parinkta empiriškai. Pradinė temperatūra turi būti pakankamai didelė, kad būtų išnaršoma visa optimizuojamos funkcijos erdvė, bet ne per didelė, kad nevyktų paprasti atsitiktiniai „vaikščiojimai“.

Temperatūra kiekvienos iteracijos metu buvo mažinama nuo T iki $(1 - \epsilon)T$. ϵ vertė priklausė nuo mūsų pasirinkto iteracijų skaičiaus, per kurį buvo manoma, sistema turėtų sukonverguoti į globalų (ar artimą globaliam) minimumą. Keletas ϵ verčių, esant skirtingiems iteracijų skaičiams, pateiktos 3.1 lentelėje.

3.1 lentelė: ϵ vertės, naudotos skirtingiems iteracijų skaičiams.

Iteracijų skaičius	ϵ
10^4	$1 \cdot 10^{-3}$
10^5	$1 \cdot 10^{-4}$
$3 \cdot 10^5$	$4.5 \cdot 10^{-5}$
$4 \cdot 10^5$	$2.5 \cdot 10^{-5}$
$5 \cdot 10^5$	$2 \cdot 10^{-5}$

Kaip ir pradinė temperatūra, ϵ vertė buvo parinkta empiriškai. Temperatūra turi būti mažinama pakankamai lėtai, kad optimizacijos algoritmas sugebėtų konverguoti į globalų minimumą. Per daug sumažinus ϵ , be reikalo padidėja skaičiavimų sąnaudos.

Siekiant pagerinti optimizacijos algoritmo gebėjimą rasti globalų minimumą, buvo įvesti „grįžimai“ į prieš tai rastus geriausius minimumus [2, p. 554]. Visos optimizacijos metu buvo išsaugomas geriausias rastas minimumas. Kai temperatūra sumažėdavo tris kartus, ir dabartinis pasiektas minimumas būdavo „prastesnis“ už geriausią rastą minimumą, optimizacijos algoritmas grįždavo į geriausio minimumo būseną. Šis metodas akivaizdžiai pagerino algoritmo rezultatus, nes taip buvo išvengta didelių stochastinių nuklydimų optimizuojamoje erdvėje.

3.1.2 Krūvio pokyčio verčių Δq_i generavimas

Kaip minėta 2.1 skyriuje, optimizuojant funkcijas su tolydžiom, N -dimensinėmis erdvėmis, iškyla didelė problema generuojant naujas „būsenas“. Šiuo atveju tos būsenos yra taškinių krūvių vertės. Kadangi simuliuoto atkaitinimo metodas yra pagrįstas stochastine erdvės paieška, metodo metu generuojamos Δq_i vertės turi būti didelės, kad algoritmas sugebėtų išnaršyti visą erdvę, kurioje gali būti globalus minimumas. Kita vertus, priartėjus prie globalaus minimumo reikalingos mažos Δq_i vertės, kad algoritmas neprašotų pasiektą minimumo ir kad galėtų konverguoti į jo „dugną“. Atsiranda problema parenkant Δq_i - per didelės vertės negali visiškai konverguoti į globalų minimumą, o per mažos vertės neišnaršo visos optimizuojamos erdvės ir taip apskritai gali nerasti globalaus minimumo.

Ši problema buvo sprendžiama įvedant dinامينius slopinimo koeficientus ir pradinę paieškos „amplitudę“. Įvedus šiuos du dydžius, Δq_i buvo generuojamas taip:

$$\Delta q_i = \alpha(k) \cdot A_0 \cdot q_i \cdot [U(0, 1) - 0.5], \quad (3.2)$$

čia $\alpha(k)$ yra nuo iteracijų skaičiaus k priklausantis slopinimo koeficientas, A_0 yra kuriamų Δq_i verčių pradinė amplitudė ir $U(0, 1)$ yra tolyginis atsitiktinių skaičių generatorius, kuris generuoja vertes nuo 0 iki 1. Išraiškos narys $[U(0, 1) - 0.5]$ leidžia keisti krūvio vertes tiek teigiamais, tiek neigiamais atsitiktiniais prieaugiais.

Išraiška (3.2) tenkina pradžioje iškeltus lūkesčius - optimizacijos pradžioje įvesta amplitudė A_0

užtikrina globalaus minimumo paiešką plačioje erdvėje. Tai leidžia algoritmui nustatyti globalaus minimo sritį optimizacijos erdvėje. Didėjant iteracijų skaičiui, $\alpha(k)$ slopinimo koeficientas pradeda vis labiau mažinti paieškos amplitudę A_0 . Galiausiai algoritmo žingsnis tampa pakankamai mažas, kad jis galėtų sukonverguoti į globalaus minimumo „dugną“.

Išraiškoje (3.2) Δq_i priklauso nuo esamos q_i vertės. Toks priklausomybės sąryšis atitinka savaiminę amplitudę, kuri papildomai moduliuoja paieškos amplitudę. Ši priklausomybė yra naudinga, kai globalus minimumas yra lokalizuotas labai mažose q_i vertėse. Jeigu optimizacijos algoritmui pasiseka ganėtinai greitai pasiekti globalaus minimumo sritį mažose q_i vertėse, nespėjus išaugti slopinimo koeficientui $\alpha(k)$, vyktų beprasmiškos šokinėjimai ir konvergavimas su mažu iteracijų skaičiumi taptų neįmanomas. Priklausomybė nuo q_i padeda to išvengti. Taip pat ji yra panaudojama keičiant pradinių spėjimų vertes. Jeigu yra manoma, kad taškinių krūvių vertės molekulėje turėtų būti didelės, q_i automatiškai padidina paieškos amplitudę, ir atvirkščiai. Kadangi išraiškoje (3.2) Δq_i priklauso nuo pačios q_i vertės ir kinta pagal ją kiekvienos iteracijos metu, ši krūvio generavimo dalis buvo pavadinta dinamine ir žymima $\Delta q_{iDynam.}$.

Dinaminė krūvių vertės generavimo dalis susiduria su vienu sunkumu. Esant labai mažoms pradinių spėjimų vertėms (taip būna dirbant su elektriškai neutraliomis molekulėmis), priklausomybė nuo q_i ryškiai sumažina paieškos amplitudę ir optimizacijos algoritmo paieškos erdvė tampa labai ribota. Siekiant to išvengti, yra įvedamas stochastinis triukšmas:

$$\Delta q_{iStoch.} = \beta(k) \cdot B_0 \cdot [U(0, 1) - 0.5], \quad (3.3)$$

čia $\beta(k)$ yra nuo iteracijų skaičiaus k priklausantis stochastinis slopinimo koeficientas ir B_0 triukšmo amplitudė.

Stochastinio triukšmo įvedimas padeda išvengti paieškos erdvės ribojimų dėl klaidingų, mažo modulio pradinių spėjimų. Triukšmo amplitudė B_0 turėtų sutampti su paieškos amplitude A_0 , bet slopinimo koeficientas $\beta(k)$ privalo gesinti stochastinį narį $\Delta q_{iStoch.}$ greičiau negu $\alpha(k)$ slopina dinaminį narį $\Delta q_{iDynam.}$. Stochastinis triukšmas yra reikalingas tik optimizacijos pradžioje, bet yra visiškai nepageidautinas vėlesniuose optimizacijos etapuose. Todėl triukšmo amplitudė yra greitai gesinama.

Išskaičius tiek dinaminę, tiek stochastinę Δq_i generavimo dalis, Δq_i išraišką galima užrašyti taip:

$$\Delta q_i = \Delta q_{iDynam.} + \Delta q_{iStoch.} = \alpha(k) \cdot A_0 \cdot q_i \cdot [U_1(0, 1) - 0.5] + \beta(k) \cdot B_0 \cdot [U_2(0, 1) - 0.5], \quad (3.4)$$

čia U_1 ir U_2 parodo, kad dinaminiam ir stochastiniam nariams naudojami atskiri tolyginiai atsitiktinių skaičių generatoriai.

Amplitudžių A_0 ir B_0 vertės parenkamos empiriškai, atsižvelgiant į molekulės sudėtingumą ir struktūrą. Taip pat jų vertės atitinkamai priklauso nuo taškinių krūvių, kuriais mes norime

aproksimuoti molekulės krūvio tankį, skaičiaus. Jeigu taškinių krūvių skaičius yra didelis, užtenka mažų amplitudžių verčių.

Šio darbo metu $\alpha(k)$ ir $\beta(k)$ pavidalas buvo pasirinktas toks:

$$\alpha(k) = \alpha^k, \quad (3.5)$$

$$\beta(k) = \beta^k, \quad (3.6)$$

čia $0.9 < \beta < \alpha < 1$. α ir β verčių parinkimas priklauso nuo iteracijų skaičiaus k . Esant mažam iteracijų skaičiui ($< 10^4$), apatinė riba yra apie 0.999. Dideliam iteracijų skaičiui ($> 10^5$) apatinė riba gali siekti 0.99999.

3.1.3 Apribojimai

Atvaizduojant molekulių krūvio tankį taškiniais krūviais, susiduriama su papildoma optimizacijos algoritmo problema. Nors optimizacijos algoritmas ir sugeba rasti optimalias taškinių krūvių vertes, dažnai gali būti netenkinama krūvio tvermė. Šio darbo metu buvo dirbama su elektriškai neutraliomis (vandens ir chlorofilo *a*) molekulėmis, todėl krūvio tvermės išlaikymas yra labai svarbus. Net maži krūvio tvermės nukrypimai gali atitikti labai stiprias molekulės perturbacijas, kurių nebuvo analizuojamos molekulės atveju.

Krūvio tvermė gali būti neišlaikoma dėl keleto priežasčių:

- sudėtingas molekulės krūvio tankis negali būti optimaliai atvaizduotas taškiniais krūviais, išlaikant krūvio tvermę,
- optimizacijos algoritmas nesugeba rasti optimalių taškinių krūvių verčių,
- per mažas matavimo taškų $\mathbf{r}_m$ skaičius nesuteikia pakankamai informacijos apie molekulę,
- matavimo taškai $\mathbf{r}_m$ yra generuojami ant geometrinės figūros paviršiaus, o ne visame tūryje.

Norint išlaikyti krūvio tvermę, reikia įvesti papildomus algoritmo apribojimus. Tai gali būti molekulės atomų sugrupavimas ir atitinkamo tarpusavio sąryšio įvedimas. Vienas iš sąryšio būdų yra priežingio krūvio ženklo priskyrimas vienam taškiniai krūviui, kitų taškinių krūvių sumos atžvilgiu:

$$q_{\mathbf{r}_k} = - \sum_{i=1}^N q_{\mathbf{r}_i} \quad (3.7)$$

Tokiu būdu ne tik išlaikoma krūvio tvermė, bet kartu ir sumažinamas optimizacijos laipsnių skaičius. Toks taškinių krūvių surišimas gerai veikia mažoms molekulėms, kur galime lengvai sugrupuoti atomus ir nuspręsti, koks sąryšio tipas geriausiai atitinka realybę.

Problema atsiranda taikant krūvių sąryšį didelėse molekulėse. Didelėse molekulėse yra gana sudėtinga nustatyti grupes ir jų dydžius. Blogai parinkus grupes, krūvio tvermė bus išlaikoma, bet dėl prastos geometrijos pasirinkimo nebus įmanoma pasiekti optimalių taškinio krūvių verčių, kurios tinkamai atvaizduotų molekulės krūvio tankį. Grupės tarpusavio sąryšiai taip pat tampa sudėtingesni, ir vieno neigiamo/teigiamo taškinio krūvio grupėje sąryšis jau nėra galimas.

Apibendrintas metodas, tinkantis bet kokio dydžio molekulei, yra papildomas narys optimizuojamoje kaštų funkcijoje. Skyriuje 2.2 išvesta kaštų funkcija:

$$E = \frac{1}{M} \sum_{m=1}^M \left(\sum_{i=1}^N \frac{kq_i}{|\mathbf{r}_m - \mathbf{r}_i|} + \sum_{l=1}^L \frac{kq_l}{|\mathbf{r}_m - \mathbf{r}_l|} - \sum_{k=1}^K \frac{kq_k}{|\mathbf{r}_m - \mathbf{r}_k|} \right)^2, \quad (3.8)$$

yra pakeičiama. Į ją yra įterpiamas naujas narys, kuris atspindi krūvio tvermę:

$$E_k = \frac{\lambda}{K} \left(\sum_{k=1}^K q_k \right)^2. \quad (3.9)$$

Išskaičius naują narį, pilna molekulės kaštų funkcija įgauna tokią formą:

$$E = \frac{1}{M} \sum_{m=1}^M \left(\sum_{i=1}^N \frac{kq_i}{|\mathbf{r}_m - \mathbf{r}_i|} + \sum_{l=1}^L \frac{kq_l}{|\mathbf{r}_m - \mathbf{r}_l|} - \sum_{k=1}^K \frac{kq_k}{|\mathbf{r}_m - \mathbf{r}_k|} \right)^2 + \frac{\lambda}{K} \left(\sum_{k=1}^K q_k \right)^2 \quad (3.10)$$

Išraiškoje (3.10) atsiranda naujas parametras λ . Šis parametras lemia, kaip stipriai optimizacijos algoritmas atsižvelgia į krūvio tvermę. Griežto λ parametro parinkimo metodo nėra, ir jis dažniausiai yra nustatomas empiriškai. Šio darbo metu buvo nuspręsta naudoti tokią λ vertę, kad pradinio spėjimo metu išraiškų (žr. 3.8 ir 3.9) vertės būtų apytiksliai vienodos. Tokiu būdu optimizacijos algoritmas vienodai atsižvelgia į potencialo aproksimavimo tikslumą ir krūvio tvermės išlaikymą.

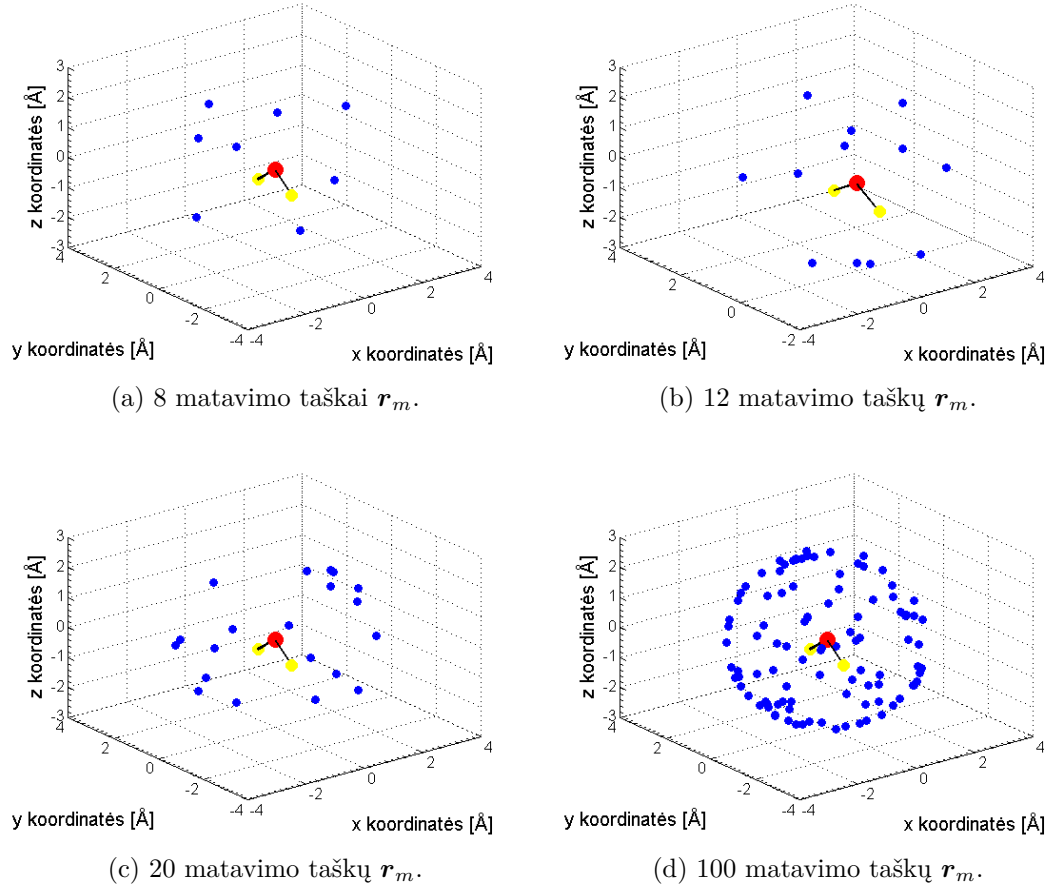
Įvedus naują narį į bendrą kaštų funkciją, gaunamas apibendrintas krūvio tvermės išlaikymas bet kokiai molekulei, dirbtinai nekonstruojant jokių tarpusavio sąryšių tarp atomų. Problema įvedant tokį narį yra ta, kad jo metu atsiranda papildomas parametras λ , kuris idealiau atveju irgi turėtų būti optimizuojamas.

3.2 Vandens molekulės krūvio tankio aproksimacija taškiniais krūviais

Vandens molekulės „tikrasis“ potencialas pasirinktuose taškuose $\mathbf{r}_m$ buvo gautas KTK metodu. Elektronų krūvio tankis erdvės taškuose $\mathbf{r}_i$ vandens molekulei buvo apskaičiuotas naudojant GAMESS [3, 4] programinį paketą ir **STO-3G** (žr. 1.1 skyrių) bazinių atominių funkcijų rinkinį.

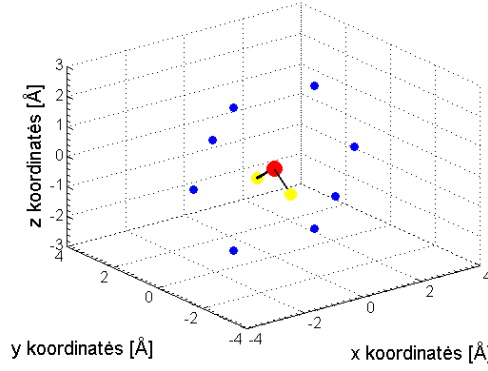
Tūris aplink vandens molekulę buvo padalintas mažais, $\sigma^3 = (0.053 \text{ Å})^3$ tūrio kubeliais. Atomų branduoliai buvo laikomi taškiniais krūviais su teigiamomis vertėmis, atitinkančiomis jų atominį skaičių (vandens molekulės geometrija pateikta priede, 2 lentelėje).

Matavimo taškai $\mathbf{r}_m$ buvo sugeneruoti ant sferos paviršiaus, kurios spindulys $R = 2.75 \text{ Å}$. Vertė 2.75 Å atitinka vandens molekulės diametrą. Toks atstumas nepakiūna į labai artimą zoną, bet nėra per toli nuo pačios molekulės. Pati sfera buvo centruota aplink deguonies atomą. Iš pradžių buvo bandoma generuoti **atsitiktinius** taškus (3.1 pav.).

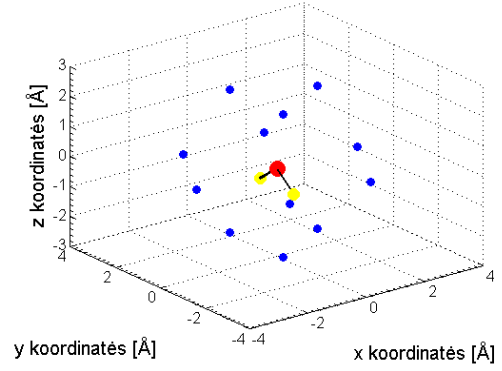


3.1 pav.: Ant sferos paviršiaus **atsitiktinai** sugeneruoti matavimo taškai $\mathbf{r}_m$. Sferos spindulys $R = 2.75 \text{ Å}$. Sfera centruota aplink deguonies atomą.

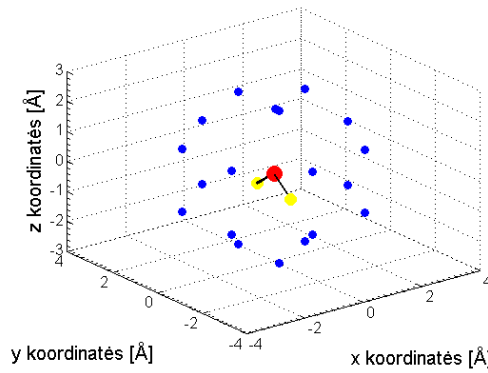
Norint išlaikyti „sferiškumą“ esant mažam matavimo taškų $\mathbf{r}_m$ skaičiui, buvo atsisakyta atsitiktinio metodo. Taškai $\mathbf{r}_m$ buvo sugeneruoti remiantis [15] metodu (3.2 pav.).



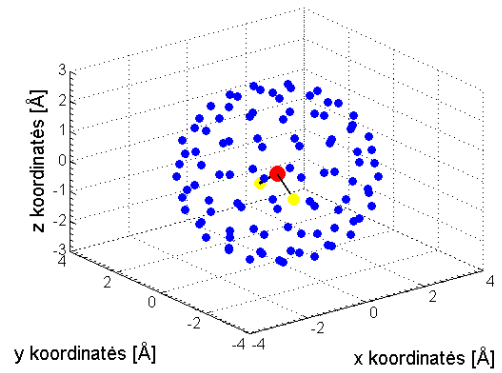
(a) 8 matavimo taškai $\mathbf{r}_m$.



(b) 12 matavimo taškų $\mathbf{r}_m$.



(c) 20 matavimo taškų $\mathbf{r}_m$.



(d) 100 matavimo taškų $\mathbf{r}_m$.

3.2 pav.: Ant sferos paviršiaus sugeneruoti vienodai pasiskirstę matavimo taškai $\mathbf{r}_m$. Sferos spindulys $R = 2.75 \text{ Å}$. Sfera centruota aplink deguonies atomą.

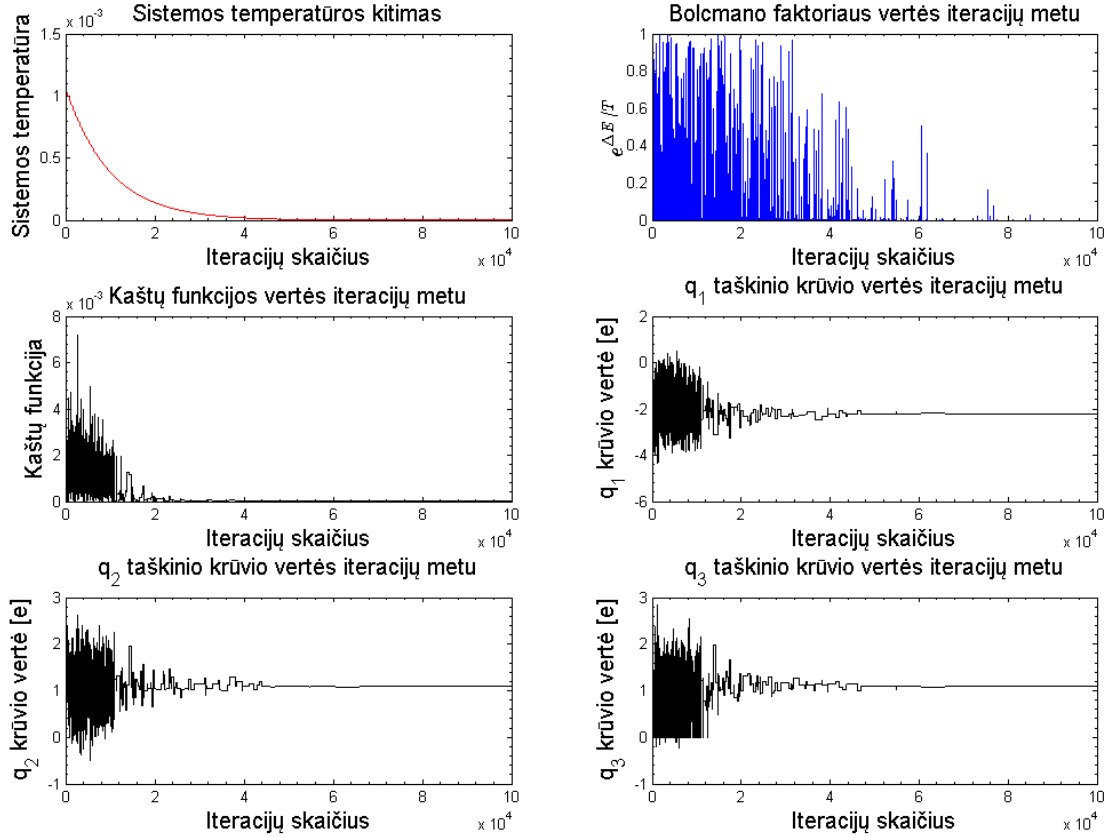
Šio darbo metu aproksimuoti vandens molekulei buvo pasirinkti trys q_k taškiniai krūviai, kurių koordinatės sutapo su vandenilių ir deguonies branduolių koordinatėmis $\mathbf{r}_l$. Vandens molekulės kaštų funkcija įgijo tokį pavidalą:

$$E = \frac{1}{M} \sum_{m=1}^M \left(\sum_{i=1}^N \frac{kq_i}{|\mathbf{r}_m - \mathbf{r}_i|} + \sum_{k=1}^3 \frac{kq_l}{|\mathbf{r}_m - \mathbf{r}_l|} - \sum_{k=1}^3 \frac{kq_k}{|\mathbf{r}_m - \mathbf{r}_{l=k}|} \right)^2, \quad (3.11)$$

Optimalios taškinių krūvių vertės (3.2 lentelė, 3.3 pav.) buvo gautos naudojant krūvių sąryšio apribojimą:

$$q_{r_O} = - \sum_{i=2}^3 q_{r_{H_i}}. \quad (3.12)$$

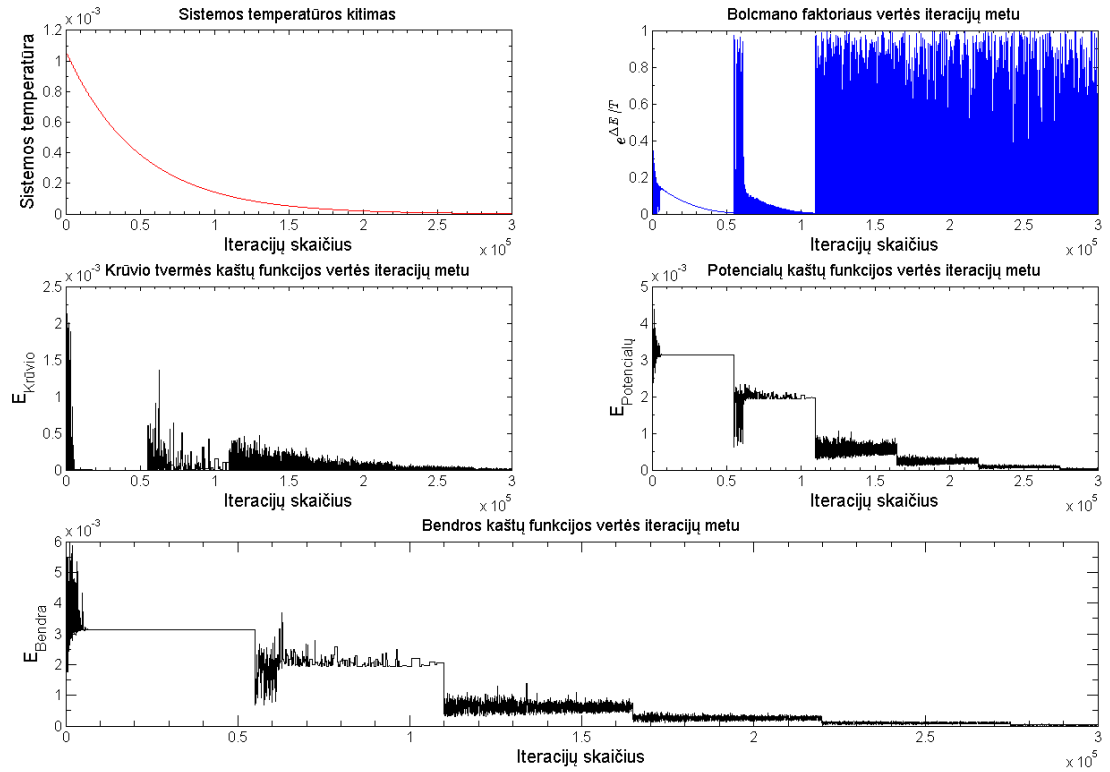
Taškinio krūvio, centruoto ant O atomo, vertė buvo pasirinkta kaip neigiama suma dviejų taškinių krūvių, centruotų ant H_1 ir H_2 atomų. Optimizacijos metu buvo optimizuojamos tik $q_{r_{H_i}}$ krūvių vertės, o q_{r_O} buvo keičiamas kiekvienos iteracijos metu pagal sąryšį (3.12).



3.3 pav.: Algoritmo parametrų pokyčiai iteracijų metu. $\epsilon = 3 \cdot 10^{-4}$, naudota 100 matavimo taškų $\mathbf{r}_m$. Taškinis krūvis q_1 yra lokalizuotas ant deguonies atomo, o taškiniai krūviai q_2 ir q_3 - ant vandenilių atomų. A_0 ir $B_0 = 10$, $\alpha = 0.99999$, $\beta = 0.99987$. Pradinio spėjimo vertės: -1, 0.5 ir 0.5. Bolcmano faktoriaus vertės gali būti lygios 0 esant labai mažoms temperatūroms **arba radus geresnę sistemos būseną**.

Esant 100 matavimo taškų $\mathbf{r}_m$, optimizacijos algoritmas su 3.3 pav. pateiktais parametrais leido pasiekti 94.56% potencialų aproksimavimo tikslumą.

Apart sąryšio apribojimo, optimizacijos algoritmas taip pat buvo patikrintas su λ parametru koreguojama kaštų funkcija (3.10). Šiuo atveju algoritmui nepavyko pasiekti tokių tikslų verčių, net esant didesniai iteracijų skaičiui (3.4 pav.). Pasiektas potencialų aproksimavimo tikslumas buvo 92.63%.



3.4 pav.: Algoritmo parametrų pokyčiai iteracijų metu. $\lambda = 11.6$ ir $\epsilon = 2 \cdot 10^{-5}$, naudota 100 matavimo taškų $\mathbf{r}_m$. Taškinis krūvis q_1 yra lokalizuotas ant deguonies atomo, o taškiniai krūviai q_2 ir q_3 - ant vandenilių atomų. A_0 ir $B_0 = 2$, $\alpha = 0.99999$, $\beta = 0.9987$. Pradinio spėjimo vertės: -1.6, 0.8 ir 0.8.

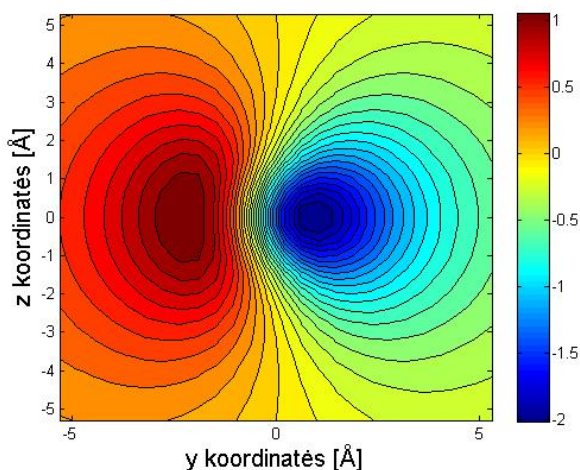
Rezultatai gauti optimizacijos algoritmu su 3.3 pav. pateiktais parametrais buvo naudojami kaip etalonas, norint surasti optimalų matavimo taškų skaičių vandens molekulei. Tikslumas buvo vertinamas lyginant taškinių krūvių vertes.

3.2 lentelė: Taškinių krūvių q_i vertės gautos naudojant skirtingą matavimo taškų skaičių.

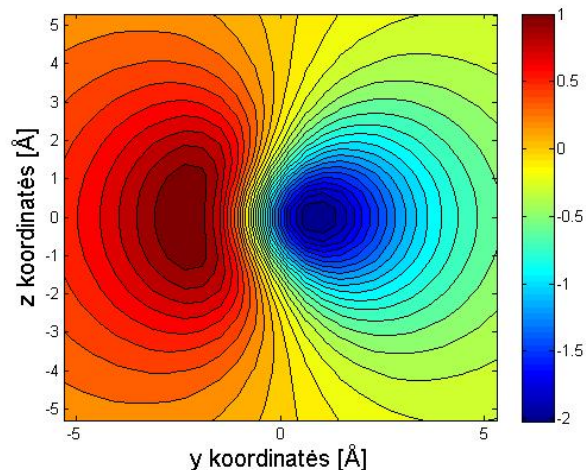
Matavimo taškų skaičius	$q_1[e]$	$q_2[e]$	$q_3[e]$	Paklaida
100	-2.2146	1.1073	1.1073	
20	-2.2203	1.1101	1.1102	0.256%
12	-2.2376	1.1186	1.1190	1.03%
8	-2.2723	1.1360	1.1362	2.6%

Iš 3.2 lentelės matyti, kad norint gauti pakankamai informacijos apie molekulės potencialo pasiskirstymą erdvėje nereikia labai didelio matavimo taškų skaičiaus. Skirtumas tarp tikslumų naudojant 100 ir 20 matavimo taškų yra vos 0.256%. Sumažinus matavimo taškų skaičių iki 8 taškų, paklaida padidėjo 2.6%.

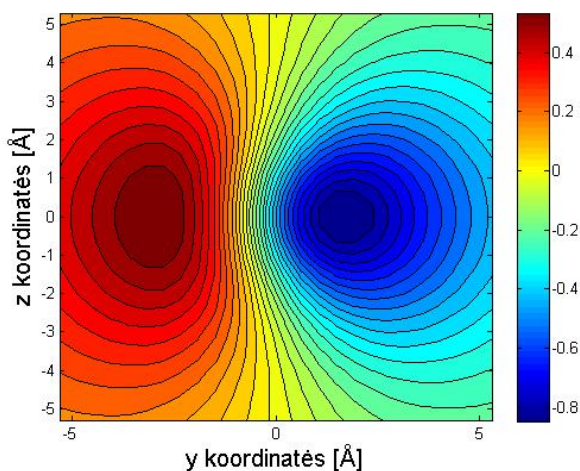
Akivaizdesnis vandens molekulės krūvio tankio aproksimacijos taškinais krūviais rezultatas yra pateiktas 3.5 pav. Čia naudotos taškinių krūvių vertės, gautos optimizacijos algoritmo su 3.3 pav. nurodytomis parametru vertėmis.



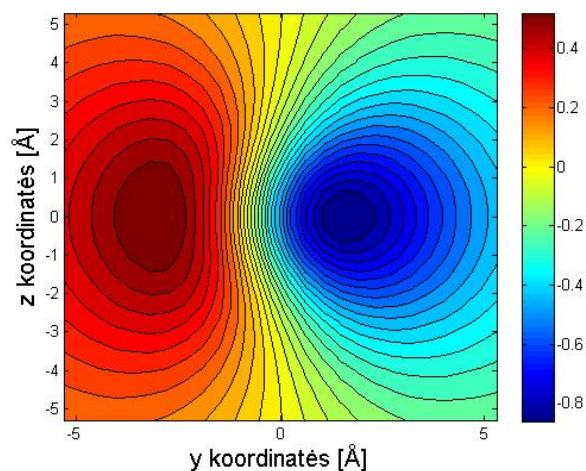
(a) „Tikrojo“ potencialo vertės, $x = 2.11$ Å.



(b) Taškinių krūvių potencialo vertės, $x = 2.11$ Å.



(c) „Tikrojo“ potencialo vertės, $x = -3.17$ Å



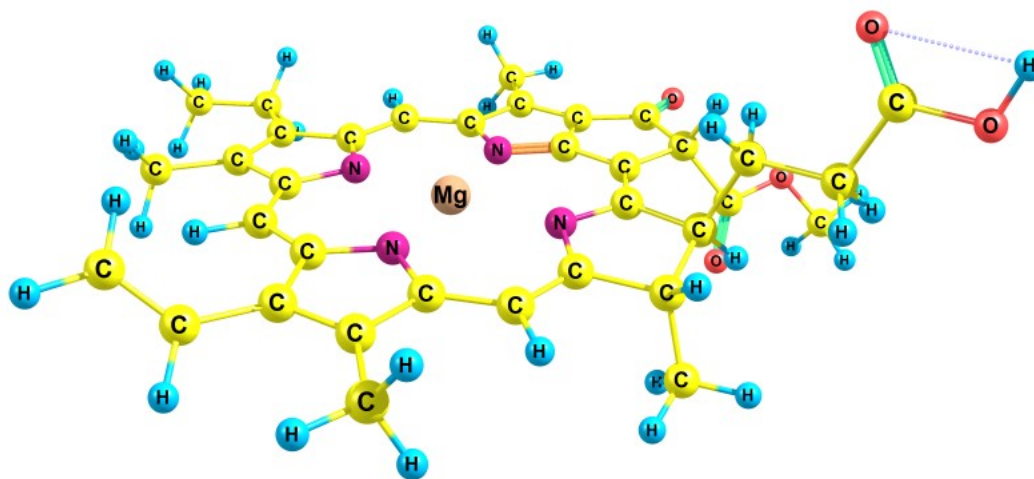
(d) Taškinių krūvių potencialo vertės, $x = -3.17$ Å.

3.5 pav.: KTK metodu suskaičiuoto atskaitinio potencialo ir taškinių krūvių kuriamo potencialo vertės yz plokštumoje. Spalva žymi potencialo dydį [V] vienetais. Kontūrai buvo braižomi yz plokštumoje, nes x ašis buvo statmena molekulės plokštumai.

Iš 3.5 pav. matyti, kad atvaizduoti vandens molekulės krūvio tankį taškiniais krūviais pavyko gana dideliu tikslumu. Pastebimi labai maži kontūrų neatitikimai yra tik artimose zonose, kur dėl kvantinės mechanikos reiškinių toks atvaizdavimo metodas iš principo nėra tikslus.

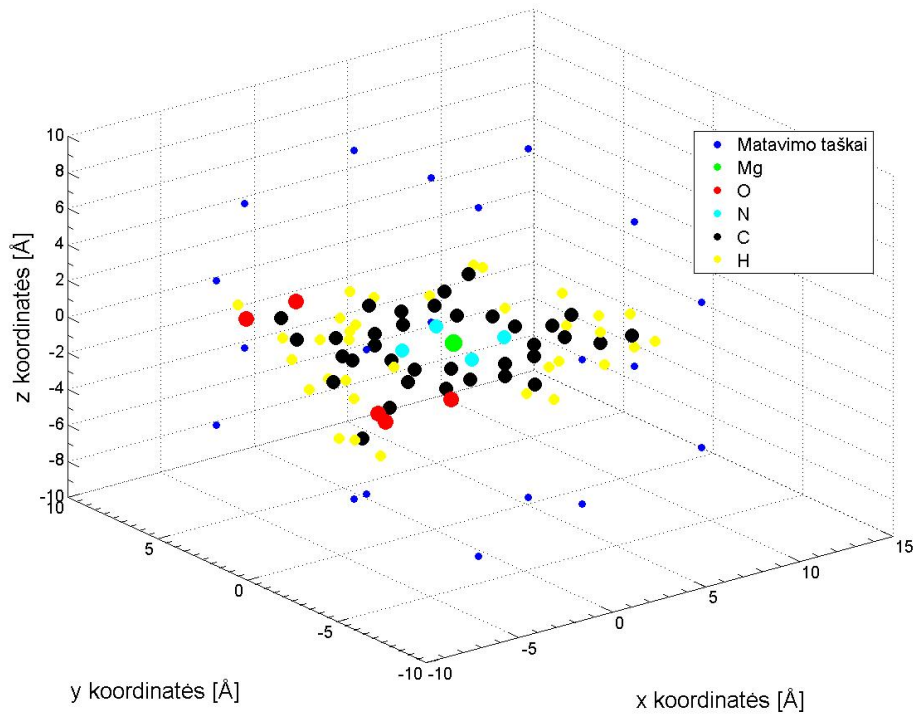
3.3 Chlorilo *a* molekulės krūvio tankio aproksimacija taškiniiais krūviais

Kaip ir vandens molekulei, chlorilo *a* molekulės atskaitinis potencialas pasirinktuose taškuose $\mathbf{r}_m$ buvo gautas naudojantis KTK metodu. Elektronų krūvio tankis erdvės taškuose $\mathbf{r}_i$ buvo apskaičiuotas naudojantis GAMESS programiniu paketu ir **STO-3G** bazinių atominių funkcijų rinkiniu. Tūris aplink molekulę buvo padalintas $\sigma^3 = (0.211 \text{ Å})^3$ tūrio kubeliais. Atomų branduoliai buvo laikomi taškiniiais krūviais su teigiamomis vertėmis, atitinkančiomis jų atominį skaičių. Chlorilo *a* molekulės geometrija pateikta priede, 1 lentelėje.



3.6 pav.: Chlorilo *a* struktūra.

Chlorilo *a* molekulei matavimo taškai $\mathbf{r}_m$ buvo generuojami tuo pačiu metodu kaip ir vandens molekulei. Tik šios molekulės atveju sferos centras buvo ant magnio atomo, o sferos spindulys - 10.46 Å. Chlorilo *a* atveju pasirinktas matavimo taškų $\mathbf{r}_m$ skaičius 20 (3.7 pav.).



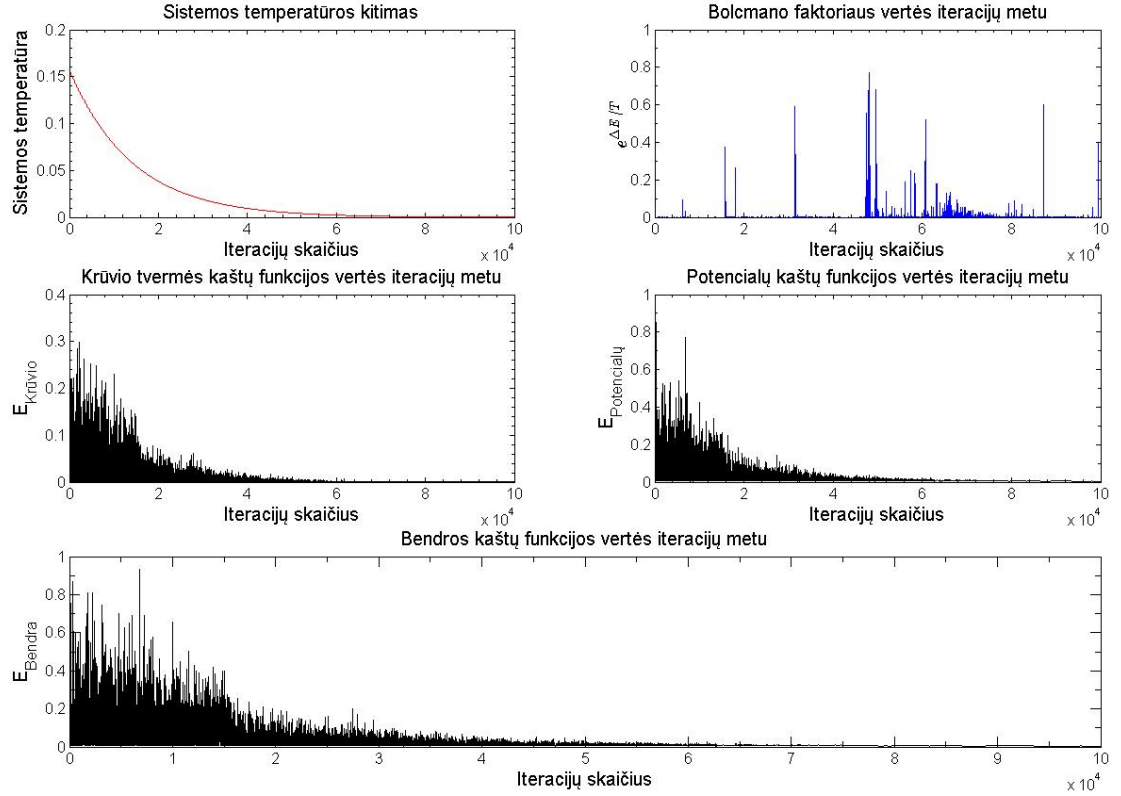
(a) 20 matavimo taškų $\mathbf{r}_m$.

3.7 pav.: Ant sferos paviršiaus sugeneruoti vienodai pasiskirstę matavimo taškai $\mathbf{r}_m$. Sferos spindulys 10.46 Å. Sfera centruota aplink magnio atomą.

Dirbant su chlorofilo *a* molekule buvo išbandyti abu algoritmo apribojimai. λ parametru koreguojamos kaštų funkcijos atveju, buvo naudojamas 35 taškinių krūvių, lokalizuotų ant chlorofilo *a* molekulės anglies atomų, rinkinys. Kaštų funkcijos pavidalas buvo:

$$E = \frac{1}{20} \sum_{m=1}^{20} \left(\sum_{i=1}^N \frac{kq_i}{|\mathbf{r}_m - \mathbf{r}_i|} + \sum_{l=1}^L \frac{kq_l}{|\mathbf{r}_m - \mathbf{r}_l|} - \sum_{k=1}^{35} \frac{kq_k}{|\mathbf{r}_m - \mathbf{r}_{C_k}|} \right)^2 + \frac{\lambda}{35} \left(\sum_{k=1}^{35} q_k \right)^2 \quad (3.13)$$

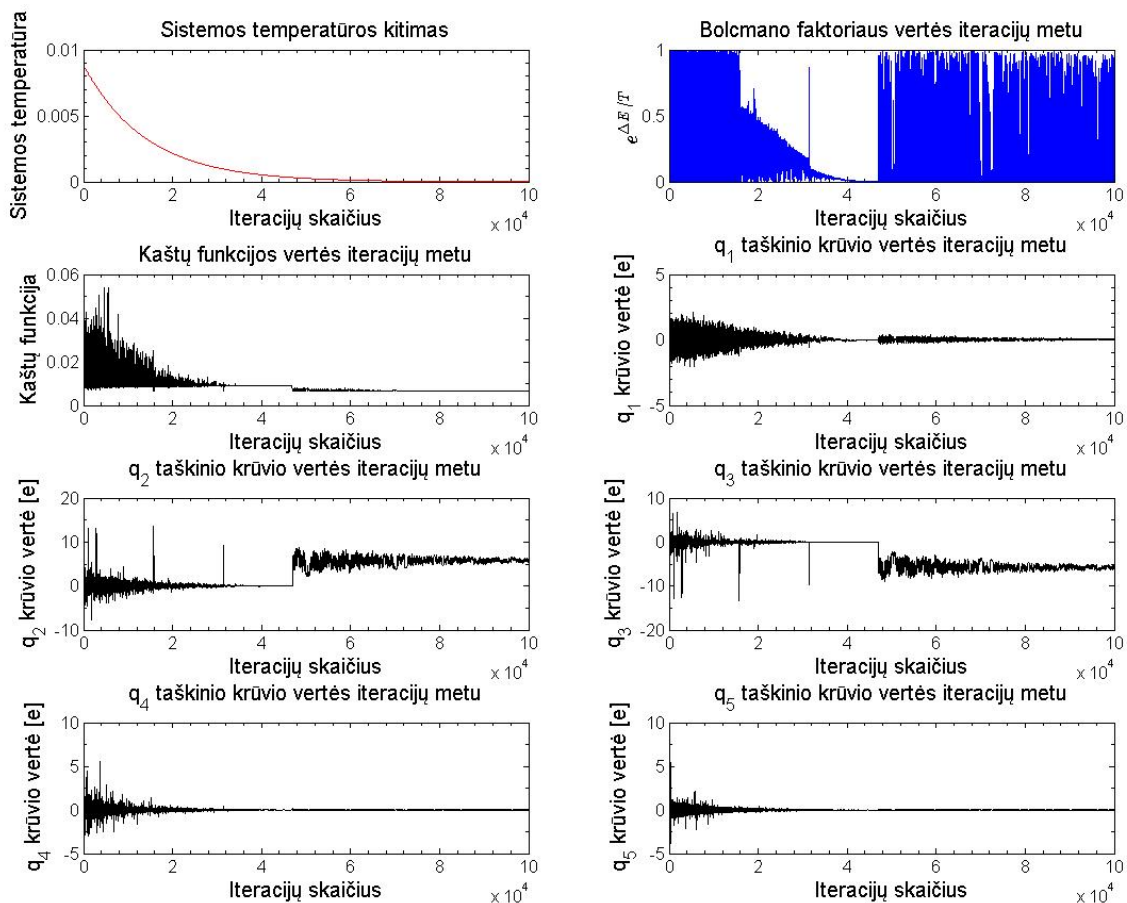
Šios optimizacijos metu (3.8 pav.) potencialų aproksimavimo paklaidos siekė 165.11% (optimizacijos metu gautos taškinių krūvių vertės pateiktos priede, 3 lentelėje). Algoritmas nesugebėjo rasti tokių taškinių krūvių verčių, kurios potencialų vertes aproksimuotų taip tiksliai kaip vandens molekulės atveju.



3.8 pav.: Algoritmo parametrų pokyčiai iteracijų metu. $\lambda = 0.03$ ir $\epsilon = 7 \cdot 10^{-5}$, naudota 20 matavimo taškų $\mathbf{r}_m$. A_0 ir $B_0 = 10$, $\alpha = 0.99999$, $\beta = 0.9987$.

Taškinių krūvių sąryšio metu, buvo naudojamas ant magnio ir azoto atomų lokalizuotas penkių taškinių krūvių rinkinys. Sąryšio tipas buvo:

$$q_{\mathbf{r}_{Mg}} = - \sum_{i=2}^5 q_{\mathbf{r}_{N_i}}. \quad (3.14)$$



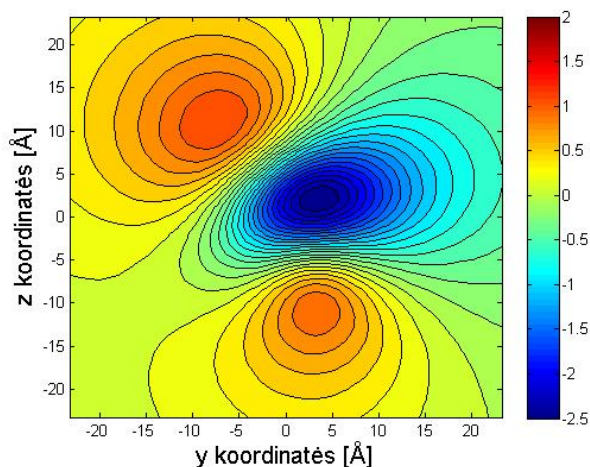
3.9 pav.: Algoritmo parametrų pokyčiai iteracijų metu. $\epsilon = 7 \cdot 10^{-5}$, naudota 20 matavimo taškų r_m . A_0 ir $B_0 = 2$, $\alpha = 0.99997$, $\beta = 0.9999$.

Optimizuojant potencialus su įvestu sąryšiu (3.9 pav.) aproksimavimo paklaidos siekė 198.68%. Šios optimizacijos metu susidūrėme su 3.1.3 skyriuje minėtomis problemomis. Gauti rezultatai atspindi per mažo taškinio krūvių rinkinio trūkumus atvaizduojant didelių molekulių krūvio tankius. Penki taškiniai krūviai yra per daug geometriškai riboti, kad galėtų sėkmingai aproksimuoti molekulės kuriamą potencialą. Norint gauti tikslesnius rezultatus reikia naudoti didesnius taškinio krūvių rinkinius, ko eigoje sąryšio apribojimas tampa sunkiai realizuojamas.

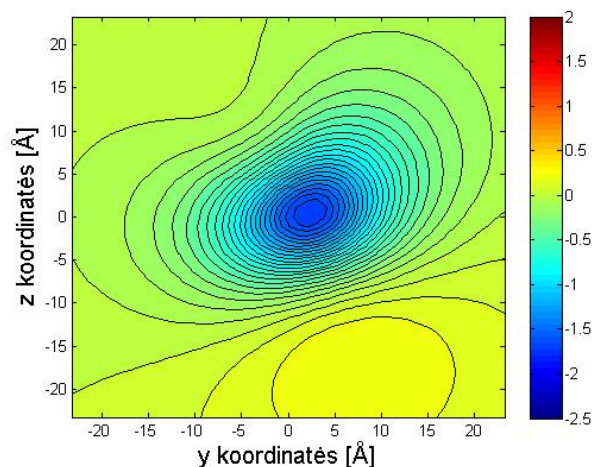
Akivaizdesni chlorofilo *a* molekulės krūvio tankio aproksimacijos taškiniais krūviais rezultatai pateikti 3.10 pav. Čia naudotos taškinio krūvių vertės, gautos optimizacijos algoritmo su 3.8 pav. nurodytais parametrais.

Atvaizduoti chlorofilo *a* krūvio tankį taškiniais krūviais buvo daug sudėtingesnis uždavinys negu vandens molekulės atveju. Didelėse molekulėse kvantinės mechanikos reiškiniai gali stipriai įtakoti optimizacijos procesą ne tik artimose zonose. Norint gauti tikslesnius krūvio tankio atvaizdavi-

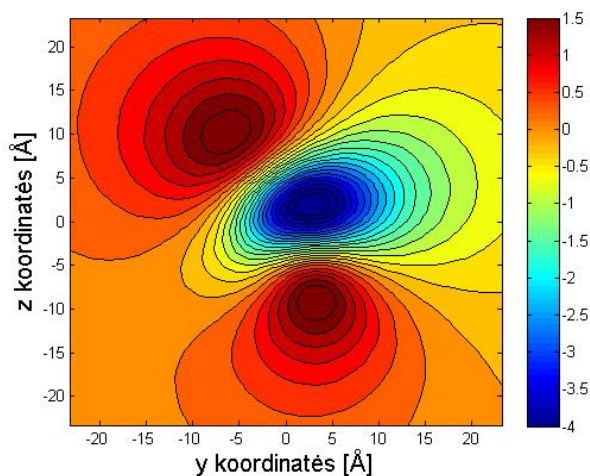
mus, reikia naudoti didesnius taškinių krūvių rinkinius ir išvalgiai parinkti jų išsidėstymą erdvėje. Išankstinės žinios apie molekulę (pradinių spėjimų įvertinimas, tikslesnių sąryšių įvedimas ir t.t) tokiose situacijose gali pagerinti optimizacijos rezultatus.



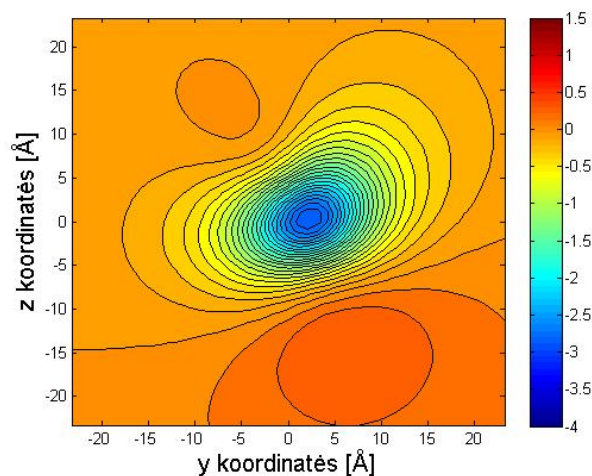
(a) „Tikrojo“ potencialo vertės, $z = 10.58 \text{ Å}$.



(b) Taškinių krūvių potencialo vertės, $z = 10.58 \text{ Å}$.



(c) „Tikrojo“ potencialo vertės, $z = 8.46 \text{ Å}$



(d) Taškinių krūvių potencialo vertės, $z = 8.46 \text{ Å}$.

3.10 pav.: KTK metodu suskaičiuoto atskaitinio potencialo ir taškinių krūvių kuriamo potencialo vertės xy plokštumoje. Spalva žymi potencialo dydį [V] vienetais. Kontūrai buvo daromi xy plokštumoje, nes z ašies atžvilgiu chlorofilo a molekulė buvo beveik plokščia.

Išvados

Molekulių krūvio tankio atvaizdavimas taškiniais krūviais yra sudėtingas uždavinys, kurio tikslumas stipriai priklauso nuo pasirinkto taškinių krūvių skaičius ir jų išdėstymo erdvėje. Šio darbo metu buvo parodyta, kad:

- Mažų molekulių krūvio tankis gali būti atvaizduotas taškiniais krūviais gana dideliu tikslumu, o labai maži neatitikimai atsiranda tik artimose molekulės zonose.
- Bandant atvaizduoti didelių molekulių krūvio tankį, reikia daug didesnio taškinių krūvių rinkinio.

Aproksimuojant didelių molekulių krūvio tankį taškiniais krūviais iškyla tiek fundamentalios, tiek optimizavimo problemos. Tokių molekulių krūvio tankis yra labai sudėtingos formos, ir mažuose atstumuose (palyginamuose su pačios molekulės matmenimis) jį teisingai įvertinti gali tik kvantinės mechanikos skaičiavimai. Tolstant nuo molekulės krūvio tankio aproksimacija taškiniais krūviais tampa įmanoma, bet norint gauti šių taškinių krūvių vertes reikalingas efektyvus optimizacijos algoritmas. Šiame darbe buvo parodyta, kad:

- Optimizacijos algoritmas, pagrįstas modelinio atkaitinimo metodo koncepcija, gali sėkmingai atlikti mažų molekulių krūvio tankio aproksimaciją taškiniais krūviais.

Norint taikyti tokio pobūdžio optimizacijos algoritmą didelių molekulių krūvio tankio atvaizdavimui taškiniais krūviais, reikia atlikti papildomų tyrimų. Pirmiausiai reiktų patikrinti, kaip algoritmas sugeba aproksimuoti molekulės kuriamą potencialą esant dideliame taškinių krūvių rinkiniui (kai taškinių krūvių skaičius yra lygus arba didesnis už molekulės atomų skaičių). Priklausomai nuo gautų rezultatų tikslumo, būtų galima daryti išvadas apie pačio algoritmo efektyvumą. Turint tokią informaciją, kiti žingsniai būtų optimizacijos algoritmo tobulinimas arba atvaizdavimo tikslumo vertinimas didesniais atstumais.

Molekulių krūvio tankio atvaizdavimas taškiniais krūviais naudojantis optimalių parametrų paieškos algoritmais

Santrauka

Šiame darbe tiriamas molekulių krūvio tankio atvaizdavimas taškiniais krūviais naudojantis optimalių parametrų paieškos algoritmais. Apžvelgiamas modelinio atkaitinimo optimizacijos metodas bei jo galimos pritaikymo sritys.

Šio darbo tikslas – sukurti bendrą kompiuterinį algoritmą remiantis skaitine optimizacija, kuris atvaizduotų bet kokių molekulių krūvio tankį pasirinktu taškinių krūvių rinkiniu. Optimizacijos algoritmas pagrįstas modelinio atkaitinimo metodo koncepcija. Darbe susitelkta ties optimizacijos algoritmo rašymu ir jo parametrų analize. Optimizacijos objektas – „tikrasis“ molekulės potencialas pasirinktuose matavimo taškuose, kuris buvo aproksimuojamas taškinių krūvių rinkinio kuriamu potencialu. Matavimo taškams kurti buvo parašyta programa, kuri generavo vienodame plote pasiskirsčiusius taškus ant pasirinkto spindulio sferos paviršiaus. Pati sfera centruota ant nagrinėjamos molekulės. Optimizuojamos kaštų funkcijos pavidalas pasirinktas remiantis mažiausių kvadratų metodu.

Buvo pateikta kaip nuo papildomų algoritmo parametrų ir apribojimų priklauso optimizacijos rezultatai mažoms bei didelėms molekulėms. Darbo metu sėkmingai aproksimuotas vandens molekulės krūvio tankis trimis taškiniais krūviais. Pasiektas 94.56% tikslumas, esant $R = 2.75$ Å atstumui nuo deguonies atomo. Aproksimuojant chlorofilo *a* molekulės krūvio tankį taškiniais krūviais $R = 10.46$ Å atstumu nuo magnio atomo, susidurta su aproksimacijos tikslumo problemomis, kurios parodė mažų taškinių krūvių rinkinių optimizacinį ribotumą.

Šiame darbe parodyta, kad mažų molekulių krūvio tankis gali būti atvaizduotas taškiniais krūviais gana dideliu tikslumu, o labai maži neatitikimai atsiranda tik artimose molekulei srityse. Taip pat parodėme, kad modelinio atkaitinimo metodu pagrįstas optimizacijos algoritmas gali dideliu tikslumu atvaizduoti mažų molekulių krūvio tankius taškiniais krūviais. Norint taikyti tokio pobūdžio optimizacijos algoritmą didelių molekulių krūvio tankio atvaizdavimui taškiniais krūviais, reikia atlikti papildomų tyrimų.

Representation of molecular charge density by point charges by using numerical optimization algorithms

Summary

This thesis is about molecular charge density representation by point charges using numerical optimization algorithms. The work covers simulated annealing optimization and its applications.

The aim of this study was to create a generic computer algorithm based on numerical optimisation, which would display molecular charge of any molecule type using selected set of point charges. The optimisation algorithm is based on the concept of model simulated annealing method. The study is concentrated on writing the optimisation algorithm and the analysis of its parameters. The object of optimisation – “real” potential of molecule in chosen measurement points, which was approximated using the potential created by the set of point charges. The program was created to create measuring points, which generated equidistributed points on a surface of a sphere of chosen radius. The sphere itself was centred on the molecule being examined. The form of optimised cost function was chosen using least squares method.

It was shown how optimisation results depend on additional parameters and restrictions of the algorithm for small and large molecules. During the work molecular charge density of water molecule was approximated using three point charges. At $R = 2.75$ Å distance from oxygen atom 94.56% accuracy was achieved. Approximation accuracy problems were encountered while approximating molecular charge density of chlorophyll *a* molecule using point charges at $R = 10.46$ Å distance from a magnesium atom, which showed optimisational limitations of sets of small point charges.

In this study it was shown, that charge density of small molecules can be represented by point charges with high accuracy, and very small inconsistencies appear only in close areas of the molecule. Also, it was shown that optimisation algorithm based on simulated annealing method can represent density of small molecules’ charge using point charges with high accuracy. In order to apply this optimisation algorithm for representation of molecular charge density of large molecules by point charges, further research is necessary.

Literatūra

- [1] M. E. Madjet, A. Abdurahman, and T. Renger. Intermolecular Coulomb Couplings from Ab Initio Electrostatic Potentials: Application to Optical Transitions of Strongly Coupled Pigments in Photosynthetic Antennae and Reaction Centers. *The Journal of Physical Chemistry B*, 110(34):17268–17281, 2006.
- [2] William H. Press, Saul A. Teukolsky, William T. Vetterling and Brian P. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge university press, 2007.
- [3] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis and J. A. Montgomery J. General Atomic and Molecular Electronic Structure System. *Journal of Computational Chemistry*, 14(11):1347–1363, 1993.
- [4] M. S. Gordon, M. W. Schmidt. Advances in electronic structure theory: Gamess a decade later. In C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria, editors, *Theory and Applications of Computational Chemistry: the first forty years*, chapter 41, pages 1167–1189. Elsevier, Amsterdam, 2011.
- [5] Juozas Šulskus. *Kompiuterinės Chemijos Įvadas*. Vilniaus Universitetas, 2011.
- [6] Ramamurti Shankar. *Principles of Quantum Mechanics*. Springer Science & Business Media, 2012.
- [7] The Krueger Laboratory, Hope College. Description of transition density cube method. http://www.chem.hope.edu/~krieg/TDC/background/tdc_background.htm. Accessed: 2016-04-19.
- [8] Errol G. Lewars. *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics*. Springer Science & Business Media, 2010.
- [9] S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi. Optimization by Simmulated Annealing. *Science*, 220(4598):671–680, 1983.
- [10] S. Kirkpatrick. Optimization by Simulated Annealing: Quantitative Studies. *Journal of Statistical Physics*, 34(5-6):975–986, 1984.
- [11] David Vanderbilt and Steven G. Louie. A Monte Carlo Simulated Annealing Approach to Optimization over Continuous Variables. *Journal of Computational Physics*, 56(2):259–271, 1984.

- [12] Ihor O. Bohachevsky, Mark E. Johnson and Myron L. Stein. Generalized Simulated Annealing for Function Optimization. *Technometrics*, 28(3):209–217, 1986.
- [13] A. Corana, M. Marchesi, C. Martini, and S. Ridella. Minimizing Multimodal Functions of Continuous Variables with the “Simulated Annealing” Algorithm. *ACM Transactions on Mathematical Software (TOMS)*, 13(3):262–280, 1987.
- [14] Lennart Lindegren, Lund Observatory, Lund University. Least squares estimation. <http://www.astro.lu.se/Education/utb/ASTM21/ASTM21-statistics-Ch6.pdf>. Accessed: 2016-04-23.
- [15] Markus Deserno, Max Planck Institute for Polymer Research. How to generate equidistributed points on the surface of a sphere. https://www.cmu.edu/biolphys/deserno/pdf/sphere_equi.pdf, 2004. Accessed: 2016-04-24.

Priedai

1 lentelė: Chlorofilo *a* molekulės atomų koordinatės.

	x [Å]	y [Å]	z [Å]
<i>N</i>	−0.75879	0.95895	−0.29585
<i>C</i>	−1.9622	0.29009	−0.22764
<i>C</i>	−3.1332	1.256	−0.33658
<i>C</i>	−2.4466	2.5453	−0.8558
<i>C</i>	−0.97178	2.272	−0.57583
<i>H</i>	−2.7753	3.4322	−0.28952
<i>C</i>	−2.6967	2.7951	−2.353
<i>H</i>	−2.3983	1.9181	−2.95
<i>H</i>	−3.7663	2.9839	−2.541
<i>H</i>	−2.1266	3.6651	−2.7146
<i>C</i>	0.001204	3.2653	−0.65224
<i>H</i>	−0.36661	4.2686	−0.87548
<i>N</i>	2.0563	1.9739	−0.24053
<i>C</i>	1.3981	3.1452	−0.48173
<i>C</i>	2.3444	4.2441	−0.54212
<i>C</i>	3.5971	3.705	−0.30471
<i>C</i>	3.4051	2.262	−0.14574
<i>C</i>	1.9803	5.6786	−0.77478
<i>H</i>	1.4132	5.8052	−1.7124
<i>H</i>	1.3457	6.0716	0.03834
<i>H</i>	2.8698	6.3212	−0.83186
<i>C</i>	4.409	1.3041	−0.007545
<i>H</i>	5.4298	1.683	−0.020193
<i>N</i>	3.0754	−0.75568	0.13674
<i>C</i>	4.2694	−0.096964	0.11821
<i>C</i>	5.3755	−1.0375	0.24961
<i>C</i>	4.8084	−2.2949	0.3506
<i>C</i>	3.3633	−2.0967	0.276
<i>C</i>	6.8096	−0.58603	0.25922
<i>H</i>	7.0663	−0.060827	−0.67658
<i>H</i>	7.5155	−1.4127	0.37573
<i>H</i>	6.9955	0.12554	1.0819
<i>C</i>	2.397	−3.1161	0.33366
<i>H</i>	2.7631	−4.1372	0.45139
<i>N</i>	0.41216	−1.7061	0.097968

<i>C</i>	1.0065	−2.9588	0.25514
<i>C</i>	−0.011655	−3.9926	0.31511
<i>C</i>	−1.2197	−3.3081	0.18924
<i>C</i>	−0.9074	−1.9267	0.060674
<i>C</i>	0.20009	−5.4655	0.47559
<i>H</i>	−0.76493	−5.9902	0.4839
<i>H</i>	0.72735	−5.6995	1.4161
<i>H</i>	0.80915	−5.879	−0.34616
<i>C</i>	−2.6703	−3.4727	0.14533
<i>C</i>	−3.2621	−1.9907	0.005594
<i>C</i>	−2.0359	−1.0823	−0.067327
<i>O</i>	−3.351	−4.4755	0.20367
<i>Mg</i>	1.2139	0.13435	−0.073298
<i>C</i>	4.8495	4.4589	−0.23441
<i>C</i>	5.9042	4.2147	0.56178
<i>H</i>	4.9017	5.3375	−0.88832
<i>H</i>	6.7873	4.8567	0.52692
<i>H</i>	5.9126	3.3966	1.285
<i>C</i>	5.3965	−3.6796	0.50601
<i>H</i>	4.9713	−4.1275	1.4232
<i>H</i>	5.0094	−4.3055	−0.31886
<i>C</i>	6.9172	−3.8456	0.5552
<i>H</i>	7.1681	−4.9131	0.65865
<i>H</i>	7.364	−3.3193	1.4129
<i>H</i>	7.403	−3.4811	−0.36335
<i>H</i>	−3.875	−1.8321	0.90854
<i>C</i>	−4.1947	−1.9186	−1.1909
<i>O</i>	−3.9336	−1.4262	−2.2678
<i>O</i>	−5.3752	−2.5002	−0.90024
<i>C</i>	−6.3025	−2.5979	−1.9923
<i>H</i>	−7.1887	−3.0972	−1.5826
<i>H</i>	−5.8721	−3.1905	−2.8132
<i>H</i>	−6.562	−1.5995	−2.3762
<i>H</i>	−3.8581	0.87697	−1.0737
<i>C</i>	−3.8428	1.4399	1.0215
<i>H</i>	−4.0959	0.45851	1.4508
<i>H</i>	−3.1514	1.9095	1.7413
<i>C</i>	−5.1223	2.2699	0.91903
<i>H</i>	−5.8408	1.8085	0.21808
<i>H</i>	−4.9349	3.2804	0.51751
<i>C</i>	−5.8269	2.4313	2.2479
<i>O</i>	−5.4883	1.9335	3.2971
<i>O</i>	−6.9239	3.2202	3.2202
<i>H</i>	−7.3159	3.2617	3.0258

2 lentelė: Vandens molekulės atomų koordinatės.

	x [Å]	y [Å]	z [Å]
O	0	−0.000003	−0.0673252
H	0	0.7616662	−0.6663385
H	0	0.7616662	−0.6663363

3 lentelė: Taškinių krūvių vertės, gautos chlorofilo a molekulės potencialo aproksimacijos metu.

q [e]	x [Å]	y [Å]	z [Å]
0.0773	−1.9622	0.29009	−0.22764
−0.345	−3.1332	1.256	−0.33658
18.62	−2.4466	2.5453	−0.8558
−26.609	−0.97178	2.272	−0.57583
2.3805	−2.6967	2.7951	−2.353
1.63	0.001204	3.2653	−0.65224
−0.00013	1.3981	3.1452	−0.48173
−0.0051	2.3444	4.2441	−0.54212
−0.0396	3.5971	3.705	−0.30471
−3.3263	3.4051	2.262	−0.14574
2.1828	1.9803	5.6786	−0.77478
−0.3519	4.409	1.3041	−0.007545
5.947	4.2694	−0.096964	0.11821
$4.805e - 05$	5.3755	−1.0375	0.24961
$−3.416e - 05$	4.8084	−2.2949	0.3506
$−1.057e - 05$	3.3633	−2.0967	0.276
$−1.8e - 05$	6.8096	−0.58603	0.25922
$3.41e - 05$	2.397	−3.1161	0.33366
$−6.567e - 05$	1.0065	−2.9588	0.25514
$−6.611e - 05$	−0.011655	−3.9926	0.31511
−0.00012	−1.2197	−3.3081	0.18924
$−1.474e - 05$	−0.9074	−1.9267	0.060674
$−3.052e - 05$	0.20009	−5.4655	0.47559
$−5.153e - 05$	−2.6703	−3.4727	0.14533
0.00012	−3.2621	−1.9907	0.005594
$−4.545e - 05$	−2.0359	−1.0823	−0.067327
$−3.018e - 05$	4.8495	4.4589	−0.23441
$5.472e - 05$	5.9042	4.2147	0.56178
$−4.589e - 05$	5.3965	−3.6796	0.50601
−0.00015	6.9172	−3.8456	0.5552
0.000541	−4.1947	−1.9186	−1.1909
−0.00093	−6.3025	−2.5979	−1.9923
0.00036	−3.8428	1.4399	1.0215
−0.00082	−5.1223	2.2699	0.91903
0.00014	−5.8269	2.4313	2.2479