

UNIVERSITATEA BABEȘ-BOLYAI CLUJ-NAPOCA
FACULTATEA DE FIZICĂ
SPECIALIZAREA FIZICĂ TEHNOLOGICĂ

LUCRARE DE LICENȚĂ

Profesor Beu Titus-Adrian

Absolvent:

Socaciu-Lendvai Mark

2023

UNIVERSITATEA BABEȘ-BOLYAI CLUJ-NAPOCA
FACULTATEA DE FIZICĂ
SPECIALIZAREA FIZICĂ TEHNOLOGICĂ

LUCRARE DE LICENȚĂ

**Metoda Sumei-Ewald în tratarea interacțiunilor de lungă distanță în
sisteme cu condiții la limită periodice**

Profesor Beu Titus-Adrian

Absolvent:

Socaciu-Lendvai Mark

2023

Cuprins

Abstract.....	4
1. Introducere.....	5
a. Motivație.....	5
b. Structura tezei.....	6
2. Considerente teoretice	7
a. Modele moleculare	7
i. Interacțiuni moleculare	7
ii. PBC – Periodic Boundary Conditions	8
iii. Propagarea timpului	9
iv. Interacțiunea sistemului cu un termostat	9
v. Câmpuri de forțe moleculare.....	10
vi. Interacțiunile electrostatice în sisteme periodice	12
b. Modele matematice.....	13
i. Suma Ewald	13
ii. Ecuația Poisson	15
iii. Transformata Fourier și spațiul reciproc.....	17
iv. Contribuțiile spațiului reciproc și real asupra sumei Ewald	20
3. Simulări de dinamică moleculară	24
a. Generarea configurației inițiale și optimizarea metodologiei de simulare.....	24
b. Analiza datelor	27
i. Compararea evoluției MDSIM pentru mai multe configurații inițiale	27
ii. Compararea MDSIM cu NAMD	41
4. Concluzii	48
a. Dezvoltarea proiectului și implementarea funcționalităților noi.....	49
Referințe	50

Abstract

The goal of simulations and the mathematical models that underpin them is to precisely describe and recreate the natural phenomena and behaviors that have been observed. To make occurrences in our immediate environment easier to understand, the most thorough and precise description is essential. As more and more interacting aspects that can affect the final behavior under study are considered, the complexity of these models increases. As the system being investigated shrinks from macroscopic to microscopic size, the degree of complexity as well as the quantity and variety of interactions both considerably increase.

In contemporary scientific works, the field of molecular dynamics has developed into one of the most potent and trustworthy computational applications. Even though it is still relatively new, this approach has the power to forecast as well as fully explain known and observed behaviors. Despite the method's widespread use in science, it has a high level of complexity because it relies on mathematical models and simulation techniques for dynamic molecular systems. The requirement to analyze these simulation runs' output data could lead to several issues.

Implementing a straightforward program to simulate dynamic molecular systems and collecting relevant data along the way provides a good opportunity to investigate the phenomena, interactions, and evolution of system state characteristics as well as to assess the model's accuracy in comparison to other widely used models and routines. The focus is on how fundamental concepts like electrostatic interactions in periodic systems, molecular models, force fields, time propagation, system interaction with a thermostat, and numerical acceleration techniques are handled. Moreover, by creating such a model to simulate these kinds of molecular systems one can study the accuracy of the implemented algorithms compared to the already used ones and still be able to harvest valuable information from the simulation runs.

1. Introducere

a. Motivație

Scopul simulărilor și a modelelor matematice care stau la baza acestora este cel de a descrie și recrea cât mai precis fenomenele și comportamentele observate în natură. Descrierea cât mai detaliată și mai precisă este vitală pentru a facilita înțelegerea fenomenelor din lumea care ne înconjoară. Aceste modele devin tot mai complexe din momentul în care sunt considerați tot mai mulți factori de interacțiune care pot avea un efect asupra comportamentului final studiat. Atât gradul de complexitate, cât și numărul și tipul interacțiunilor crește semnificativ, cu cât sistemul studiat trece de la o mărime macroscopică la una microscopică.

Lumea Dinamicii Moleculare reprezintă un domeniu a devenit una din cele mai puternice și fiabile modalități computaționale din domeniul științific. Această metodă deși încă destul de recentă are ca abilități explicarea detaliată a comportamentelor deja cunoscute și observate, dar și anticiparea fenomenelor încă necunoscute. În ciuda faptului că această metodă este utilizată pe scară largă în lumea științifică, ea prezintă un grad de complexitate ridicat, datorită modelelor și metodologiilor matematice utilizate în simulările sistemelor moleculare dinamice. Mai multe complicații pot apărea în urma necesității de analiză a datelor de ieșire ale execuțiilor acestor simulări. [1]

Implementarea unui program simplu pentru simularea sistemelor moleculare dinamice și achiziția datelor de interes pe parcursul acestora prezintă o oportunitate bună pentru a studia atât fenomenele, interacțiunile și evoluția caracteristicilor stării sistemului, cât și pentru analizarea acurateței modelului implementat aici față de modele și rutine bine stabilite în uzul general. Tratarea aspectelor fundamentale precum modele moleculare, câmpuri de forțe, propagarea timpului, interacțiunea sistemului cu un termostat, interacțiunile electrostatice în sisteme periodice și tehnicile de accelerare numerice sunt accentuate. [2]

b. Structura tezei

Următoarea secțiune oferă o scurtă introducere a ideilor discutate în fiecare capitol.

Capitolul 2 reprezintă o introducere detaliată în conceptele teoretice din domeniul fizicii, interacțiunilor intra- și intermoleculare, cât și cele din domeniul matematicii și a metodelor de calcul numeric tratate în această lucrare. Interacțiunile considerate sunt cele din domeniile modelelor moleculare, câmpurilor de forțe, propagarea timpului, interacțiunea sistemului cu un termostat, interacțiunile electrostatice în sisteme periodice. Ele sunt menite pentru a reda cu cât mai mare acuratețe comportamentul observat al acestor sisteme moleculare și în practică. Printre modelele matematice se numără metoda Sumei Ewald pentru a trata interacțiunile electrostatice în sistemul periodic considerat în acest model și Transformata Fourier pentru a facilita calculul numeric al caracteristicilor sistemului pentru fiecare pas al simulării.

Capitolul 3 prezintă implementările conceptelor necesare simulării prezentate mai sus, atât dintr-un punct de vedere al metodelor și funcțiilor de calcul, cât și din punct de vedere al analizei datelor obținute în urma rulărilor modelului. Datele obținute din acest model sunt mai departe analizate calitativ și comparate cu date obținute din modele deja existente rulate pe aceleași situații inițiale. Prin compararea acestora se pot trage concluzii referitoare atât la implementările modelelor, cât și la acuratețea lor.

Capitolul 4 conține concluziile trase atât din realizarea modelului dinamic, cât și cele obținute în urma analizării detaliate ale rezultatelor simulării rulate. Pe baza acestora s-au formulat și obiective de viitor pentru a îmbunătăți calitatea rezultatului obținut.

2. Considerente teoretice

a. Modele moleculare

Primele metode pentru înțelegerea lumii care ne înconjoară au fost legate de observarea fenomenelor și comportamentelor diferitelor corpuri și sisteme. Cu trecerea timpului am ajuns să descriem aceste comportamente prin diferite interacțiuni dintre sistem și exterior, sau dintre particulele componente ale unui sistem. Pentru a putea trata aceste sisteme, în primul rând avem nevoie de a considera un număr mai redus de particule aparținătoare unui astfel de sistem.

i. Interacțiuni moleculare

Proprietățile și comportamentul materiei, inclusiv cel al fluidelor și gazelor, sunt determinate de interacțiunile inter- și intramoleculare. Interacțiunile intramoleculare apar în interiorul unei molecule și sunt responsabile pentru determinarea structurii și stabilității moleculei, în timp ce interacțiunile intermoleculare apar între diferite molecule și sunt responsabile pentru proprietățile macroscopice ale materiei. Un sistem considerat macroscopic este situat undeva la $N \approx N_A \approx 10^{23}$ particule. Acesta este definitiv prea numeros pentru a putea fi tratate în orice măsură, astfel se alege pentru tratare un sistem “mic” până în 10^6 particule, caz în care comportamentul este dominat de efecte de suprafață.

Acest sistem prin mărimea lui este limitat, trebuie să redea comportamentul liber, nerestrâns al unui sistem bulk macroscopic. Pentru a atinge acest efect, se alege minimizarea efectelor de suprafață care provin din mărimea sistemului considerat prin introducerea PBC – Periodic Boundary Conditions – adică condiții periodice la limită. [3]

Fluidele simple [4] sunt compuse din molecule mici, considerate sferice, care interacționează între ele prin perechi de forțe. Potențialul Lennard-Jones, unul din cele mai folosite modele moleculare, descrie forțele atractive și repulsive dintre două molecule în funcție de distanța lor de separare $V(r) \sim \frac{1}{r}$. Potențialul are doi termeni: un termen respingător cu rază scurtă care domină la separări mici și împiedică moleculele să ocupe același spațiu și un termen atractiv pe rază lungă care domină la separări mai mari și este responsabil pentru coeziunea fluidului. Potențialul Lennard-Jones este dat de ecuația:

$$V(r) = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (1)$$

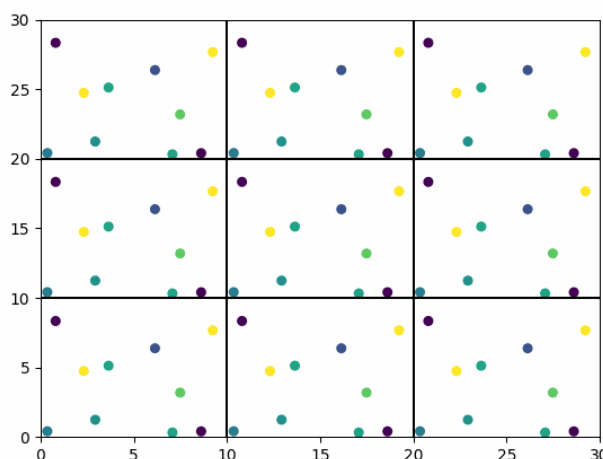
unde r este distanța de separare dintre cele două molecule, ϵ este puterea potențialului atractiv și σ este distanța la care potențialul este zero.

Potențialul Lennard-Jones este un model simplu pentru interacțiunile intermoleculare care surprinde multe dintre caracteristicile importante ale fluidelor simple, cum ar fi comportamentul lor coeziv, existența unui punct critic la care are loc o tranziție lichid-gaz și prezența fluctuațiilor și mișcarea termică în sistem [3]

ii. PBC – Periodic Boundary Conditions

Prin introducerea condițiilor periodice la limită (PBC) în tratarea sistemului, este importantă tratarea domeniului simulat. Acesta nu este nimic altceva decât o cutie cubică de laturi: $[-L, L] \times [-L, L] \times [-L, L]$, care este proiectată astfel încât să acomodeze un număr redus (10^6) de particule. Deși PBC are ca avantaj eliminarea efectelor de suprafață, prin umplerea întregului spațiu cu copii identice ale așa numitei “Cutii de Simulare”, își are și limitarea prin excluderea fluctuațiilor macroscopice $\lambda > L$.

Datorită acestei descrieri a spațiului oferite de PBC, toate particulele imagine se mișcă în sincron, care are un efect de conservare a numărului de particule N din cutia de simulări considerată, în sensul ca fiecare particulă care părăsește volumul considerat este reciclată printr-o particulă imagine, care intră în volum din latura opusă părăsirii acesteia.



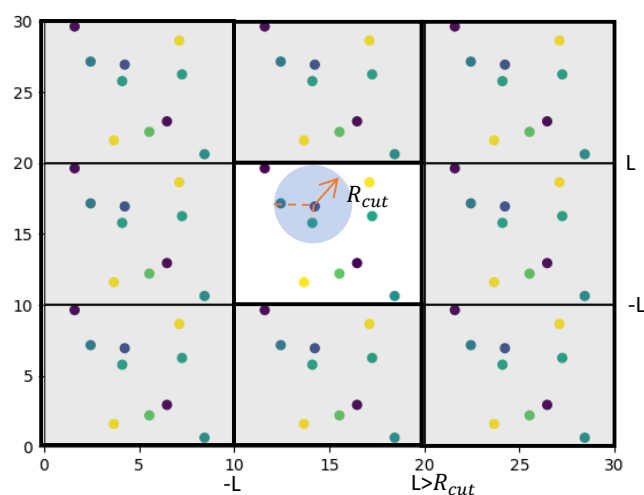
Figură 2.1 Particulele imagine considerate prin descrierea spațiului prin PBC

Modul de reciclare al acestor particule care părăsesc sistemul este realizat prin ajustarea coordonatelor în următorul fel:

$$\begin{cases} x_i < -L \rightarrow x'_i = x_i + 2L \\ x_i > L \rightarrow x'_i = x_i - 2L \end{cases} \quad (2)$$

unde conform Fig1, L reprezintă semi lărgimea uneia din laturile cubului considerat.

Un alt aspect important de considerat este reducerea numărului de interacțiuni dintre particule de la un număr teoretic infinit, la așa numita criteriul imaginii minime. Criteriul imaginii minime presupune ca pentru o particulă dată doar interacțiunea cu cea mai apropiată copie a unei alte particule trebuie considerată ca posibilă interacțiune. Această metodă constă în caracterizarea componentei de scurtă durată a potențialului cu o distanță de tăiere "cut-off distance" R_{cut} . Pentru distanțe interatomice $r_{ij} > R_{cut}$ mai mari decât această distanță de tăiere interacțiunile se neglijează. [5]



Figură 2.2 Reducerea interacțiilor de scurtă distanță

iii. Propagarea timpului

Acest aspect se referă la descrierea dependenței traiectoriei particulelor de timp în cazul simulărilor dinamice moleculare. Ecuațiile de mișcare pentru fiecare particulă din sistem pot fi exprimate folosind legile mișcării lui Newton. În simulările de dinamică moleculară, aceste ecuații sunt rezolvate numeric folosind diverse metode de integrare numerică, cum ar fi algoritmul Verlet, care este un algoritm popular utilizat în acest domeniu. Acesta calculează pozițiile și vitezele particulelor în pași de timp discreți și face acest lucru utilizând pozițiile și vitezele din pasul de timp anterior. [6]

iv. Interacțiunea sistemului cu un termostat

Din punct de vedere termodinamic un ansamblu microcanonic cu energie non-variabilă este un sistem care este caracterizat printr-un număr fix de particule, volum și energie totală. Un astfel de sistem este unul izolat din punct de vedere termodinamic, ceea ce nu corespunde cu

majoritatea circumstanțelor în care sunt experimentele efectuate. Astfel este nevoie de o tratare care consideră sistemul termodinamic în cauză un ansamblu canonic cu temperatură non-variabilă. Sistemul se află în contact constant cu un termostat (sursă de căldură), care are rolul de a menține temperatura sistemului la o valoare constantă T_0 , adecvată majorității sistemelor biomoleculare.

Prezența acestuia este importantă, termostatul fiind agentul responsabil de rescalarea vitezelor particulelor la fiecare moment. Există diferite abordări ale acestora, cum ar fi cele cu Scalarea vitezei instantanee, termostatul Berendsen sau cel Gaussian.

Scalarea vitezei instantanee: abordare deși simplă, nu este în concordanță cu realitatea

$$T_{kin} = \frac{2E_{kin}}{3Nk_B} \rightarrow f = \sqrt{\frac{T_0}{T_{kin}}} \rightarrow v'_i = f \cdot v_i \quad (3)$$

Termostatul Berendsen: evoluția vitezelor este una graduală, care depinde de o constantă de cuplare $\tau \gg \delta t$, care asigură o tendință convergentă a temperaturii sistemului.

$$\frac{dT}{dt} = \frac{1}{\tau}(T_0 - T)$$

$$f = \sqrt{1 + \frac{\delta t}{\tau} \left(\frac{T_0}{T_{kin}} - 1 \right)} \rightarrow v'_i = f \cdot v_i \quad (4)$$

v. Câmpuri de forțe moleculare

Pentru o mai bună înțelegere a interacțiunilor moleculare care se desfășoară într-un gaz sau fluid simplu, de obicei se folosește modelul câmpului de forță CHARMM, acesta fiind un câmp de forță atomic care depinde de tipurile de atomi și reziduuri moleculare care iau loc la interacțiune. Aceste interacțiuni se împart în temeni legați ce aparțin forțelor intramoleculare și în termeni nelegați, ce aparțin forțelor intermoleculare. [7]

Considerând forțele intramoleculare, acestea se pot exprima:

$$U_{bonded} = \sum_{bonds} k_B (b - b_0)^2 + \sum_{angles} k_\psi (1 + \cos(n\psi - \delta)) + \sum_{impropers} k_\psi (\psi - \psi_0)^2 + \sum_{Urey-Bradley} k_{UB} (r_{1,3} - r_{1,3;0})^2 \quad (5)$$

În această lucrare pentru a ușura partea computațională și a reduce timpii de rulare s-a ales un sistem de molecule de CO, biatomice liniare, care nu sunt supuse forțelor dihedrale, de distorsiune a unghiurilor dintre legături, doar celor de stretching.

Considerând forțele intermoleculare, acestea se pot exprima:

$$U_{non-bonded} = U_{elec} + U_{LJ} = \sum_{atoms\ i,j} \left(\frac{q_i q_j}{\epsilon_0 r_{ij}} + \epsilon_{ij} \left(\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min,ij}}{r_{ij}} \right)^6 \right) \right) \quad (6)$$

Unde datorită regulilor de amestecare Lorentz-Berthelot avem termenii:

$$R_{min,ij} = \frac{R_{min,i} + R_{min,j}}{2},$$

$$\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j} \quad (7)$$

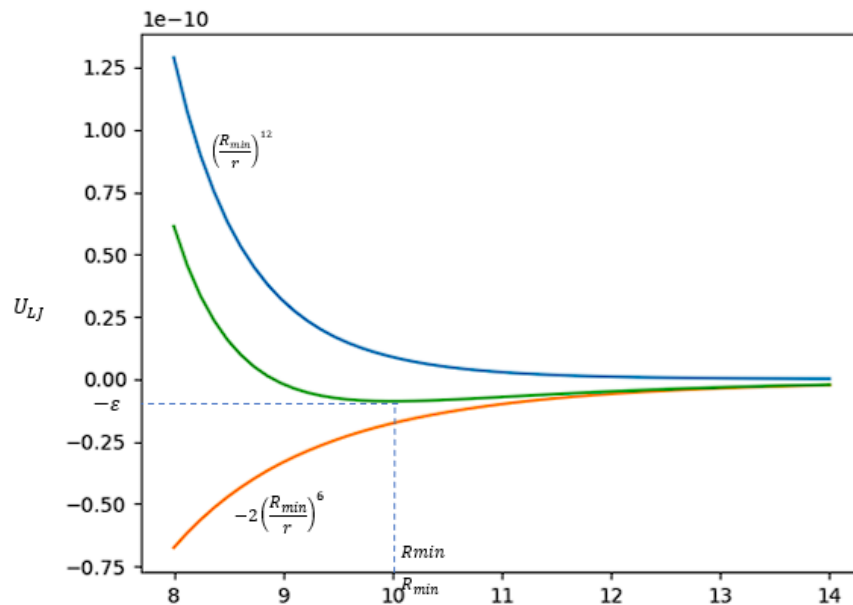
Forțele intermoleculare cele mai iminente sunt cele de tip Van der Waals, iar unul din cele mai des întâlnite abordări pentru a modela acestea este introducerea potențialului de tip Lennard-Jones 12-6.

$$U_{LJ}(r) = \epsilon \left(\left(\frac{R_{min}}{r} \right)^{12} - 2 \left(\frac{R_{min}}{r} \right)^6 \right) \quad (8)$$

$$f_{LJ} = \frac{12\epsilon}{r} \left(\frac{R_{min}}{r} \right)^6 \left(\left(\frac{R_{min}}{r} \right)^6 - 1 \right) \cdot \frac{r}{r} \quad (9)$$

Din analiza celor doi termeni din potențialul Lennard-Jones se evidențiază cel foarte simplu compoziția interacțiunii:

- $\left(\frac{R_{min}}{r} \right)^{12}$ forță de repulsie puternică pentru distanțe mici datorată suprapunerii orbitalilor electronici liberi
- $-2 \left(\frac{R_{min}}{r} \right)^6$ forță de atracție slabă pentru distanțe mari datorată dispersiei atractive a forțelor de tip Van der Waals (interacțiuni dipol-dipol).



Figură 2.3 Potențialul Lennard-Jones 12-6

vi. Interacțiunile electrostatice în sisteme periodice

Există o multitudine de sisteme fizice care prezintă interacțiuni de distanțe lungi, cum ar fi cele gravitaționale, coulombiene sau dipolare, toate având o dependență invers proporțională cu distanța. Pentru aplicația curentă se discută interacțiunea coulombiană dintr-un sistem periodic cu un număr fix de N particule, fiecare situat la pozițiile $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$ având sarcinile q_1, q_2, \dots, q_N . Cu energia totală electrostatică:

$$E_{Coul} = k_e \sum_{(i,j)}^N \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j|} = \frac{k_e}{2} \sum_{i=1}^N \sum_{j \neq i}^N \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (10)$$

Datorită tratării PBC se elimină efectele de suprafață, ceea ce are ca consecință un număr infinit de interacțiuni care contribuie la energia electrostatică a volumului considerat.

$$E_{Coul} = \frac{k_e}{2} \sum_{i=1}^N \sum_{\mathbf{n} \in \mathbb{Z}^3} \sum_{j \neq i}^N \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j + \mathbf{nL}|} \quad (11)$$

Se pot exprima :

- Potențialul electrostatic generat de toate sarcinile într-un punct al spațiului:

$$\phi(\mathbf{r}) = k_e \sum_{\mathbf{n} \in \mathbb{Z}^3} \sum_{i=1}^N \frac{q_i}{|\mathbf{r} - \mathbf{r}_i + \mathbf{nL}|} \quad (12)$$

- Potențialul electrostatic generat în punctul în care se află particula încărcată i :

$$\phi_{[i]}(\mathbf{r}_i) = k_e \sum_{\mathbf{n} \in \mathbb{Z}^3} \sum_{j \neq i}^N \frac{q_j}{|\mathbf{r}_i - \mathbf{r}_j + \mathbf{nL}|} \quad (13)$$

Astfel, din ecuațiile precedente putem rescrie ecuația (10) în următorul fel:

$$E_{Coul} = \frac{1}{2} \sum_{i=1}^N q_i \phi_{[i]}(\mathbf{r}_i) \quad (14)$$

Pentru a estima energia totală și implicit și energia electrostatică a sistemului, se dorește evaluarea componentei $\phi_{[i]}(\mathbf{r}_i)$ din ecuația (14), doar că aceasta implică sumări infinite. Această incomoditate aduce nevoia introducerii unor modele matematice care ajută atât la rezolvarea problemelor întâmpinate, cât și la realizarea simulării prin calcule numerice necesare determinării traiectoriilor.

b. Modele matematice

În multe cazuri pentru rezolvarea cât mai rapidă și mai elegantă a unei probleme originare din rezolvarea sau observarea unui sistem fizic, este necesară utilizarea unor modele matematice care pot ajuta atât la formarea și rezolvarea ecuațiilor, cât și la pașii de calcul prin introducerea unor metode numerice pentru a aproxima mai bine traiectoriile particulelor. [8]

i. Suma Ewald

Forțele cu distanță scurtă de acțiune exclud un număr foarte important de probleme care implică sarcini electrice și dipoli. În sistemele de acest tip nu poate fi neglijat rolul central al forțelor cu distanță lungă de acțiune, ele prin definiție, nu sunt anihilate de PBC și, prin urmare, se extind pe întregul sistem. Pentru a trata interacțiunile electrostatice din sistemul descris prin PBC există nevoia de a adresa dificultățile potențialului de Coulomb: [9]

- Variația rapidă pentru valori mici ale r
- Descreștere lentă pentru valori mari ale r

Prin reorganizarea sumelor de interacțiune peste imagini periodice ale sistemului, metoda Ewald este capabilă să încorporeze limite periodice într-o manieră care evită trunchierea explicită.

Adresând aceste probleme devine evident că este nevoie de o abordare asemănătoare potențialului de tip Lennard-Jones, adică despărțirea în suma a doi termeni, fiecare fiind responsabil pentru descrierea interacțiilor de distanțe lungi respectiv scurte.

$$\frac{1}{r} = \frac{f(r)}{r} + \frac{1-f(r)}{r} \quad (15)$$

În această scriere termenul $\frac{f(r)}{r}$ responsabil pentru interacțiile de distanță scurtă se neglijează pentru distanțe mai mari decât raza de tăiere $r > R_{cut}$ și are proprietăți de convergență rapidă în *Spațiul-R*.

În această scriere termenul $\frac{1-f(r)}{r}$ responsabil pentru interacțiile de distanță lungă deși prezintă o variațiune lentă și periodică, are proprietăți de convergență rapidă cu seriile Fourier, care astfel îl fac să convergă rapid prin *Spațiul k*.

Prin descompunerea distribuției de sarcină se înlocuiește distribuția de sarcină originală cu o sumă de distribuții echivalentă:

$$\rho_j(\mathbf{r}) = q_j \delta(\mathbf{r}) \quad (16)$$

$$\rho_j(\mathbf{r}) = (\rho_j(\mathbf{r}) - \rho_{j,\alpha}(\mathbf{r})) + \rho_{j,\alpha}(\mathbf{r}) \quad (17)$$

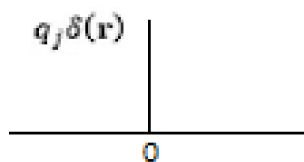
Pentru distribuția de sarcină se alege în mod convenabil distribuția Gaussiană:

$$\rho_{j,\alpha}(\mathbf{r}) = q_j \gamma_\alpha(\mathbf{r}) \quad (18)$$

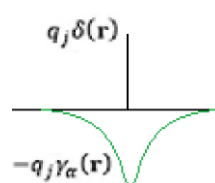
$$\gamma_\alpha(\mathbf{r}) = \frac{\alpha^3}{\pi^{3/2}} e^{-\alpha^2 r^2} \quad (19)$$

α reprezintă parametrul Ewald, care leagă distribuția Gaussiană de cea originală:

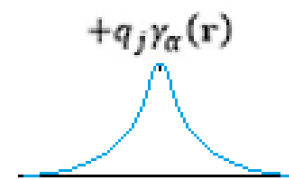
$$\gamma_\alpha(\mathbf{r}) \sim_{\alpha \rightarrow \infty} \delta(\mathbf{r}) \quad (20)$$



Figură 2.4 Distribuția originală de sarcină



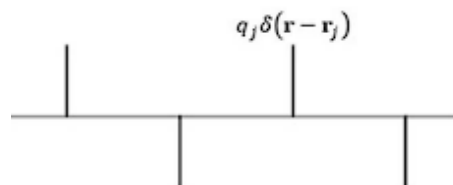
Figură 2.5 Distribuția de sarcină pentru interacțiunea de distanță scurtă



Figură 2.6 Distribuția de sarcină pentru interacțiunea de distanță lungă

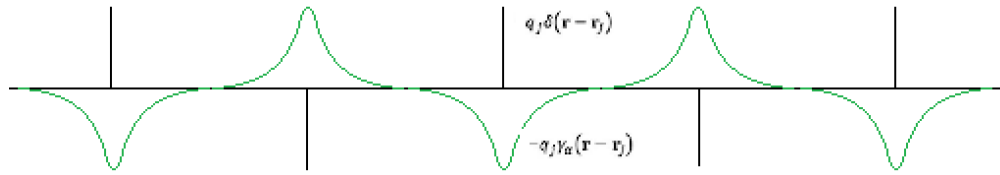
Rescriind ecuația (17) obținem:

$$\rho(\mathbf{r}) = \rho_{short-ranged}(\mathbf{r}) + \rho_{long-ranged}(\mathbf{r}) \quad (21)$$



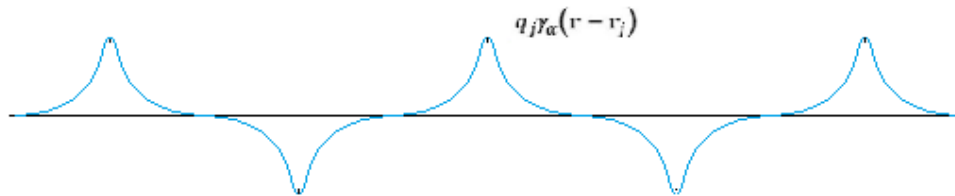
Figură 2.7 Distribuția de sarcină descompusă pentru o colecție de sarcini punctiforme

$$\rho_{short-ranged}(\mathbf{r}) = \sum_{j=1}^N q_j (\delta(\mathbf{r} - \mathbf{r}_j) - \gamma_\alpha(\mathbf{r} - \mathbf{r}_j)) \quad (22)$$



Figură 2.8 Distribuția de sarcină de distanță scurtă

$$\rho_{long-ranged}(\mathbf{r}) = \sum_{j=1}^N q_j \gamma_\alpha(\mathbf{r} - \mathbf{r}_j) \quad (23)$$



Figură 2.9 Distribuția de sarcină de distanță lungă

ii. Ecuația Poisson

Ecuația lui Poisson este o ecuație fundamentală în electrostatică, care descrie relația dintre potențialul electric și densitatea de sarcină într-o regiune a spațiului. Ecuația [10] se bazează pe principiul conservării sarcinii, care afirmă că cantitatea totală de sarcină electrică într-un sistem închis rămâne constantă. Astfel potențialul electric generat de distribuția de sarcină poate fi exprimat prin:

$$\nabla^2 \phi(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\epsilon_0} \quad (24)$$

Pentru tratarea matematică mai ușoară se va lucra cu ecuația radială a lui Poisson pentru o distribuție de sarcină simetrică sferic centrată la origine:

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\phi(r)}{dr} \right) = -\frac{\rho(r)}{\epsilon_0} \quad (25)$$

$$\frac{1}{r} \frac{d^2}{dr^2} (r\phi(r)) = -\frac{\rho(r)}{\epsilon_0} \quad (26)$$

Introducând distribuția de sarcină Gaussiană din ecuațiile (18), (19) în ecuația (26) ajungem la:

$$\frac{d^2}{dr^2} (r\phi_{j,\alpha}(r)) = -\frac{q_j \alpha^3}{\epsilon_0 \pi^{\frac{3}{2}}} r e^{-\alpha^2 r^2} \quad (27)$$

Integrând ambele părți ale ecuației anterioare de două ori cu $\int_r^\infty \dots$ respectând condiția

$\frac{d(r\phi_{j,\alpha}(r))}{dr} \rightarrow_{r \rightarrow \infty} 0$ se ajunge la:

$$\frac{d}{dr} \left(r \phi_{j,\alpha}(r) \right) = -\frac{q_j \alpha^3}{\varepsilon_0 \pi^2} \int_r^\infty r' e^{-\alpha^2 r'^2} dr' \rightarrow$$

$$\frac{d}{dr} \left(r \phi_{j,\alpha}(r) \right) = -\frac{q_j \alpha^3}{\varepsilon_0 \pi^2} \frac{1}{2\alpha^2} \int_{-\alpha^2 r^2}^\infty e^{-u} du \rightarrow$$

$$\frac{d}{dr} \left(r \phi_{j,\alpha}(r) \right) = \frac{q_j}{4\pi\varepsilon_0} \frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 r^2} \quad (28)$$

$$\rightarrow r \phi_{j,\alpha}(r) = \frac{q_j}{4\pi\varepsilon_0} \frac{2\alpha}{\sqrt{\pi}} \int_0^r e^{-\alpha^2 r'^2} dr' = \frac{q_j}{4\pi\varepsilon_0} \frac{2}{\sqrt{\pi}} \int_0^{\alpha r} e^{-u^2} du \quad (29)$$

$$\phi_{j,\alpha}(r) = \frac{k_e q_j}{r} \operatorname{erf}(\alpha r) \quad (30)$$

În ecuația (30) s-a obținut expresia potențialului generat de distribuția de sarcină Gaussiană, aceasta reprezentând un factor adițional potențialului generat de sursele punctiforme. Funcția de eroare apărută în ecuația (30) se poate exprima foarte ușor prin compararea cu ecuația (29) și se obține:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du \quad (31)$$

Din forma ecuației obținute mai sus se observă ca în cazul $\alpha \rightarrow \infty$ distribuția de sarcină se transformă în una punctiformă, care introduce potențialului clasic generat de o sarcină punctiformă situată în punctul de origine a sistemului considerat. Astfel parametrul Ewald are rolul central de a realiza trecerea de la interpretarea clasică cu distribuție discretă a sarcinii în cea Gaussiană, conform ecuației (20).

Pornind de la forma potențialului generat de distribuția Gaussiană într-un punct, se poate exprima energia Coulombiană pentru o sarcină punctiformă q_i situată la r_i în prezența unui potențial creat de distribuția de sarcină $\rho_\alpha(r)$ centrată în r_j :

$$E_{ij,\alpha} = q_i \phi_{j,\alpha}(r_{ij}) = k_e \frac{q_i q_j}{r_{ij}} \operatorname{erf}(\alpha r_{ij}) \quad (32)$$

Mai departe se poate exprima și forța resimțită de o particulă i , aceasta fiind definită ca și gradientul energiei Coulombiene față de poziția considerată:

$$\begin{aligned} \mathbf{F}_{ij,\alpha} &= -\frac{\partial}{\partial \mathbf{r}_i} E_{ij,\alpha} = -q_i \frac{\partial \phi_{j,\alpha}(r_{ij})}{\partial \mathbf{r}_i} \rightarrow \\ \mathbf{F}_{ij,\alpha} &= -q_i \frac{\partial \phi_{j,\alpha}(r_{ij})}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial \mathbf{r}_i} \end{aligned} \quad (33)$$

Gradientul $\frac{\partial r_{ij}}{\partial x_i}$ introdus poate fi obținut prin diferențierea parțială pentru a obține forma simplificată a forței:

$$\begin{cases} \frac{\partial r_{ij}}{\partial x_i} = \frac{\partial}{\partial x_i} \left((x_i - x_j)^2 + \dots \right)^{1/2} = \frac{x_i - x_j}{r_{ij}} \\ -\frac{\partial \operatorname{erf}(r)}{\partial r} \frac{1}{r} = \frac{1}{r^2} \left(\operatorname{erf}(r) - \frac{2}{\sqrt{\pi}} r e^{-r^2} \right) \end{cases} \quad (34)$$

$$\mathbf{F}_{ij,\alpha} = k_e \frac{q_i q_j}{r_{ij}^2} \left(\operatorname{erf}(\alpha r_{ij}) - \frac{2}{\sqrt{\pi}} r_{ij} e^{-r_{ij}^2} \right) \cdot \frac{\mathbf{r}_{ij}}{r_{ij}} \quad (35)$$

iii. Transformata Fourier și spațiul reciproc

În metoda de însumare Ewald, distribuția periodică a sarcinii este transformată în spațiul reciproc folosind o transformată Fourier. Interacțiunile coulombiene dintre sarcini sunt apoi calculate în spațiul reciproc folosind vectorii rețelei reciproce ai cristalului. Potențialul electrostatic în spațiul reciproc rezultat este apoi transformat înapoi în spațiul real folosind o transformată Fourier inversă. Calculele în spațiul real și în spațiul reciproc sunt apoi combinate pentru a obține o soluție precisă și eficientă pentru interacțiunile coulombiene cu rază lungă. Utilizarea spațiului reciproc în acest mod permite calcule mai eficiente ale interacțiunilor electrostatice pe distanță lungă în sistemele periodice, deoarece permite utilizarea algoritmilor eficienți care sunt bine adaptați pentru a gestiona condițiile la limită periodice. [5]

Considerând sistemul bazat pe PBC stabilit în capitolele anterioare, atât distribuția de sarcini, cât și potențialul generat într-un punct al spațiului se pot rescrie în spațiul reciproc, acestea fiind reprezentate într-o manieră convenabilă prin serii Fourier:

$$\begin{cases} \phi(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{k}} \tilde{\phi}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \\ \rho(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{k}} \tilde{\rho}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \end{cases} \quad (36)$$

Unde în ecuația precedentă vectorii \mathbf{k} ai spațiului reciproc formează o bază discretă:

$$\mathbf{k} = \frac{2\pi}{L} \mathbf{n}, \quad \mathbf{n} \in \mathbb{Z}^3 \quad (37)$$

Astfel rescriind ecuația (36) în spațiul reciproc obținem:

$$\begin{cases} \tilde{\phi}(\mathbf{k}) = \int_V \phi(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} \\ \tilde{\rho}(\mathbf{k}) = \int_V \rho(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} \end{cases} \quad (38)$$

Aplicând operatorul ∇^2 asupra formei potențialului în expansiunea Fourier obținem:

$$\nabla^2 \phi(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{k}} \tilde{\phi}(\mathbf{k}) \nabla^2 e^{i\mathbf{k}\cdot\mathbf{r}} = -\frac{1}{V} \sum_{\mathbf{k}} k^2 \tilde{\phi}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \quad (39)$$

Astfel ecuația Poisson (24) se transformă în:

$$-\frac{1}{V} \sum_{\mathbf{k}} k^2 \tilde{\phi}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} = -\frac{1}{\varepsilon_0 V} \sum_{\mathbf{k}} \tilde{\rho}(\mathbf{k}) \nabla^2 e^{i\mathbf{k}\cdot\mathbf{r}} \rightarrow$$

$$\tilde{\phi}(\mathbf{k}) = \frac{1}{\varepsilon_0 k^2} \tilde{\rho}(\mathbf{k}) \quad (40)$$

Se definește funcția Green pentru spațiul complementar:

$$\tilde{\phi}(\mathbf{k}) = \tilde{g}(k) \tilde{\rho}(\mathbf{k}) \rightarrow \tilde{g}(k) = \frac{1}{\varepsilon_0 k^2} \quad (41)$$

Rezolvarea ecuației lui Poisson echivalează cu calcularea densității de sarcină transformată în spațiul Fourier, iar apoi înmulțind rezultatul obținut să se ajungă la valoarea potențialului în spațiul complementar.

Pornind de la forma densității de sarcină a unui set de sarcini periodice se poate ajunge la valoarea forma densității de sarcină pentru o distribuție Gaussiană al acesteia:

$$\rho(r) = \sum_{n \in \mathbb{Z}^3} \sum_{j=1}^N q_j \delta(r - r_j + nL) \quad (42)$$

$$\int_V \sum_{n \in \mathbb{Z}^3} f(r + nL) dr = \int_V f(r) dr \quad (43)$$

$$\tilde{\rho}(\mathbf{k}) = \int_V \rho(r) e^{-i\mathbf{k}\cdot\mathbf{r}} dr = \int_V \sum_{n \in \mathbb{Z}^3} \sum_{j=1}^N q_j \delta(r - r_j + nL) e^{-i\mathbf{k}\cdot\mathbf{r}} dr \rightarrow$$

$$\tilde{\rho}(\mathbf{k}) = \sum_{j=1}^N q_j e^{-i\mathbf{k}\cdot\mathbf{r}_j} \quad (44)$$

Se notează

$$\tilde{\rho}(-\mathbf{k}) = (\tilde{\rho}(\mathbf{k}))^* \quad (45)$$

Din (42),(43),(44) se obține:

$$\tilde{\rho}_\alpha(\mathbf{r}) = \sum_{n \in \mathbb{Z}^3} \sum_{j=1}^N q_j \gamma_\alpha(r - r_j + nL) = \frac{\alpha^3}{\pi^2} \sum_{n \in \mathbb{Z}^3} \sum_{j=1}^N e^{\left(-\alpha^2 |r - r_j + nL|^2\right)} \quad (46)$$

$$\tilde{\rho}_\alpha(\mathbf{k}) = \frac{\alpha^3}{\pi^2} \int_V \sum_{n \in \mathbb{Z}^3} \sum_{j=1}^N q_j e^{\left(-\alpha^2 |r - r_j + nL|^2\right)} e^{-i\mathbf{k}\cdot\mathbf{r}} dr \rightarrow$$

$$\tilde{\rho}_\alpha(\mathbf{k}) = \frac{\alpha^3}{\pi^2} \sum_{j=1}^N q_j \int_{\mathbb{R}^3} e^{-\alpha^2 |r - r_j|^2} e^{-i\mathbf{k}\cdot\mathbf{r}} dr \quad (47)$$

Transformata Fourier a densității de sarcină din ecuația (46) poate fi simplificată mai departe prin introducerea schimbării de variabilă $r - r_j \rightarrow r$ și scoaterea de sub integrală a factorului $e^{-i\mathbf{k}\cdot\mathbf{r}_j}$

$$\tilde{\rho}_\alpha(\mathbf{k}) = \frac{\alpha^3}{\pi^2} \sum_{j=1}^N q_j e^{-ik \cdot r_j} \int_{\mathbb{R}^3} e^{-\alpha^2 r^2 - ik \cdot r} dr \quad (48)$$

În următorul pas se alege descompunerea integralei din ecuația (47) în produsul a trei integrale 1-D în următorul fel:

$$\int_{\mathbb{R}^3} e^{-\alpha^2 r^2 - ik \cdot r} dr = \frac{\alpha}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\alpha^2 r^2 - ik_x \cdot x} dx \cdot \frac{\alpha}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\alpha^2 r^2 - ik_y \cdot y} dy \cdot \frac{\alpha}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\alpha^2 r^2 - ik_z \cdot z} dz \quad (49)$$

Se introduce schimbarea de variabilă $\xi = \alpha x + i \frac{k_x}{2\alpha}$ care permite factorizarea:

$$e^{-\alpha^2 r^2 - ik_x \cdot x} = e^{-\frac{k_x^2}{4\alpha^2}} e^{-\xi^2} \quad (50)$$

Folosind ecuația (49) integralele de 1D din ecuația (48) se pot calcula exact:

$$\begin{cases} \frac{\alpha}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\alpha^2 r^2 - ik_x \cdot x} dx = \frac{1}{\sqrt{\pi}} e^{-\frac{k_x^2}{4\alpha^2}} \int_{-\infty}^{+\infty} e^{-\xi^2} d\xi \\ \int_{-\infty}^{+\infty} e^{-\xi^2} d\xi = \sqrt{\pi} \end{cases} \quad (51)$$

$$\frac{\alpha}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\alpha^2 r^2 - ik_x \cdot x} dx = e^{-\frac{k_x^2}{4\alpha^2}} \quad (52)$$

Transformata Fourier a densității de sarcină se rescrie ca:

$$\tilde{\rho}_\alpha(\mathbf{k}) = e^{-\frac{k_x^2}{4\alpha^2}} \sum_{j=1}^N q_j e^{-ik \cdot r_j} = e^{-\frac{k_x^2}{4\alpha^2}} \tilde{\rho}(\mathbf{k}) \quad (53)$$

Potențialul scris în spațiul reciproc devine:

$$\tilde{\Phi}_\alpha(\mathbf{k}) = \tilde{g}_\alpha(k) \tilde{\rho}(\mathbf{k}) \quad (54)$$

Unde funcția Green este definită ca:

$$\tilde{g}_\alpha(k) = \frac{1}{\varepsilon_0 k^2} e^{-\frac{k_x^2}{4\alpha^2}} \quad (55)$$

Aplicând transformata Fourier inversă se obține forma potențialului în spațiul R:

$$\phi_\alpha(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{k} \neq 0} \tilde{\Phi}_\alpha(\mathbf{k}) e^{ik \cdot \mathbf{r}} \quad (56)$$

Având stabilită relația pentru potențialul generat de distribuția Gaussiană de sarcini se poate exprima și energia totală electrostatică al unui sistem alcătuit din sarcini punctiforme distribuite după o astfel de distribuție Gaussiană:

$$\begin{aligned}
E_{Coul} &= \frac{1}{2} \sum_{i=1}^N q_i \phi_{\alpha}(r_i) = \frac{1}{2V} \sum_{i=1}^N q_i \sum_{\mathbf{k} \neq 0} \tilde{\phi}_{\alpha}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{r}_i} \\
&= \frac{1}{2V} \sum_{\mathbf{k} \neq 0} \tilde{\phi}_{\alpha}(\mathbf{k}) \sum_{i=1}^N q_i e^{i\mathbf{k} \cdot \mathbf{r}_i}
\end{aligned} \tag{57}$$

$$\text{Din (45),(57)} \quad E_{Coul} = \frac{1}{2V} \sum_{\mathbf{k} \neq 0} \tilde{\phi}_{\alpha}(\mathbf{k}) (\tilde{\rho}(\mathbf{k}))^* \tag{58}$$

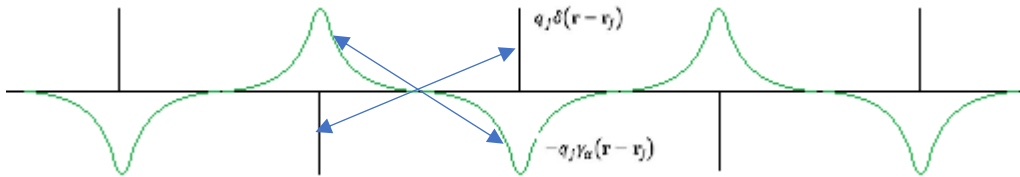
Rescriind ecuația (58) se obține contribuția spațiului reciproc la energia totală Coulombiană:

$$E_{Coul} = \frac{1}{2V} \sum_{\mathbf{k} \neq 0} \tilde{g}_{\alpha}(\mathbf{k}) |\tilde{\rho}(\mathbf{k})|^2 \tag{59}$$

iv. Contribuțiile spațiului reciproc și real asupra sumei Ewald

Contribuția *spațiului R* asupra sumei Ewald se regăsește în extinderea rezultatelor obținute pentru un set periodic de distribuții de sarcină Gaussiene într-o superpoziție de sarcini punctiforme care au un efect de distanță scurtă, dar i se supun unei distribuții de sarcină Gaussiene negative.

$$\begin{aligned}
\rho^{(r)}(r) &= \sum_{\mathbf{n} \in \mathbb{Z}^3} \sum_{j=1}^N (\rho_j(r - r_j + nL) - \rho_{j,\alpha}(r - r_j + nL)) \\
\rho^{(r)}(r) &= \sum_{\mathbf{n} \in \mathbb{Z}^3} \sum_{j=1}^N q_j (\delta(r - r_j + nL) - \gamma_{\alpha}(r - r_j + nL))
\end{aligned} \tag{60}$$



Figură 2.10 Contribuția spațiului R asupra sumei Ewald

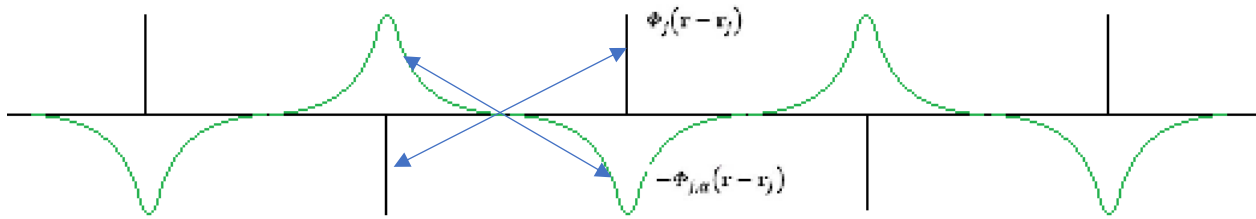
Important de reținut faptul ca pentru o valoare destul de mare aleasă pentru parametrul Ewald $\alpha \in [0,1]$ contribuțiile spațiului R referitor la toate sumele se va reduce pentru volumul inițial considerat ca și “cutie de simulare”, $\mathbf{n} = (0,0,0)$.

Acesta are ca și consecință directă considerarea sumei $\sum_{\mathbf{n} \in \mathbb{Z}^3} []$ ca fiind total redundantă pentru restul imaginilor replică a volumului inițial și posibilitatea de a folosi criteriul imaginii minime pentru interacțiile coulombiene din spațiul R.

Pornind de la ecuația (60) se poate generaliza potențialul electric generat de o singură distribuție de sarcină Gaussiană:

$$\phi^{(r)}(r) = \sum_{\mathbf{n} \in \mathbb{Z}^3} \sum_{j=1}^N (\phi_j(r - r_j + nL) - \phi_{j,\alpha}(r - r_j + nL))$$

$$\phi^{(r)}(r) = k_e \sum_{n \in \mathbb{Z}^3} \sum_{j=1}^N \frac{q_j}{|r - r_j + nL|} (1 - \operatorname{erf}(\alpha|r - r_j + nL|)) \quad (61)$$



Figură 2.11 Potențialul generat de o singură distribuție de sarcină Gaussiană

Se introduce și funcția de eroare complementară pentru a simplifica mai departe ecuația (61)

$$\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) = \frac{1}{\sqrt{\pi}} \int_x^{\infty} e^{-u^2} du \quad (62)$$

$$\phi^{(r)}(r) = k_e \sum_{n \in \mathbb{Z}^3} \sum_{j=1}^N \frac{q_j}{|r - r_j + nL|} \operatorname{erfc}(\alpha|r - r_j + nL|) \quad (63)$$

Din exprimarea anterioară a potențialului dezvoltat de distribuția de sarcină se pot exprima și contribuțiile spațiului R la energia electrostatică a sistemului și la forța Coulombiană resimțită de o particulă i :

$$\begin{cases} E_{Coul}^{(r)} = \frac{k_e}{2} \sum_{n \in \mathbb{Z}^3} \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{|r_{ij} + nL|} \operatorname{erfc}(\alpha|r_{ij} + nL|) \\ F_i^{(r)} = k_e \sum_{n \in \mathbb{Z}^3} \sum_{j=1}^N \frac{q_i q_j}{|r_{ij} + nL|} \operatorname{erfd}(\alpha|r_{ij} + nL|) \frac{r_{ij} + nL}{|r_{ij} + nL|} \end{cases} \quad (64)$$

S-a introdus funcția auxiliară:

$$\begin{cases} \operatorname{erfd}(x) = -x^2 \frac{d \operatorname{erfc}(x)}{dx} \\ \operatorname{erfd}(x) = \operatorname{erfc}(x) + \frac{2}{\sqrt{\pi}} x e^{-x^2} \end{cases} \quad (65)$$

Contribuția spațiului k asupra sumei Ewald se regăsește în contribuțiile aduse interacțiunilor Coulombiene de lungă distanță care variază lent în fundal datorită distribuției de sarcină Gaussiană.

$$\rho^{(r)}(r) = \sum_{n \in \mathbb{Z}^3} \sum_{j=1}^N \rho_{j,\alpha}(r) = \sum_{n \in \mathbb{Z}^3} \sum_{j=1}^N q_j \gamma_{\alpha}(r - r_j + nL) \quad (66)$$

$$F_i = F_i^{(r)} + F_i^{(k)} \quad (73)$$

3. Simulări de dinamică moleculară

Sistemele moleculare dinamic reprezintă o metodă computațională practică folosită în domeniul cercetării comportamentelor și interacțiunilor inter particulare în sisteme moleculare și atomice. La bază, simulările dinamice moleculare au rezolvarea sistemelor de ecuații de mișcare ale particulelor componente, în perspectiva funcției de energie potențială (Energie Coulombiană totală). Pentru a realiza aceste simulări este necesară implementarea algoritmilor prezentați în capitolele anterioare în cod sursă rulabil pe calculatoare. [10]

a. Generarea configurației inițiale și optimizarea metodologiei de simulare

Pe lângă implementarea algoritmilor și pașilor de calcul aferenți acestora într-un limbaj de programare, pentru rularea simulării este totodată necesară specificarea configurației inițiale a sistemului studiat. În cazul lucrării de față, s-a ales pentru simplitate monoxidul de carbon (CO) ca și fluid studiat, acesta configurație liniară biatomică prezentând un factor de restrângere al tipurilor de interacțiune intermoleculară tratate, fapt care favorizează concentrarea studiului pe studierea interacțiunilor electrostatice. Pentru a transmite datele necesare configurației se alege folosirea a mai multor fișiere “standard” în cercetare. Acestea sunt “.pdb”, “.psf” și “.par”, respectiv formate de fișiere generale, dar care sunt folosite pentru a specifica atât informații suplimentare necesare rulării, cum ar fi fișierul “.dat”, sau fișiere necesare stocării măsurătorilor sau în particular a calculelor efectuate asupra stării sistemului la fiecare moment de timp considerat, cum ar fi fișierul “.out”.

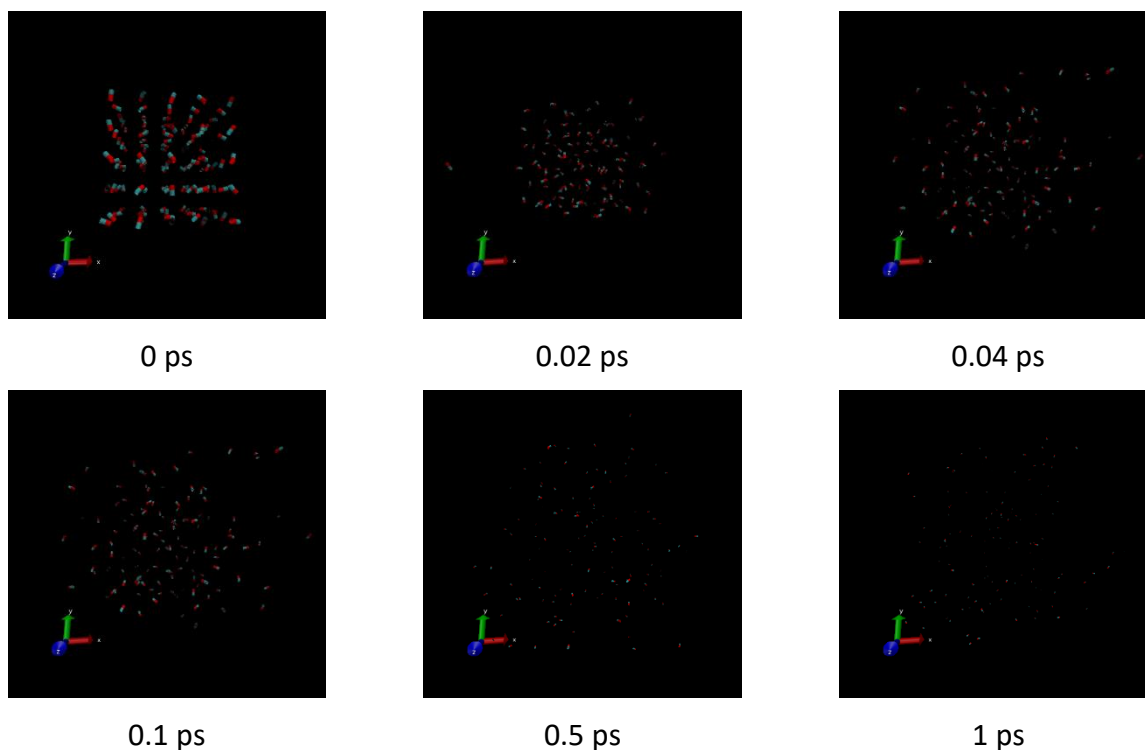
Protein Data Bank (.pdb) [11], format de fișier introdus de către organizația *Worldwide Protein Data Bank*, este fișierul folosit pentru definirea geometriei particulelor și a pozițiilor lor din cadrul simulării. Acest fișier este generat folosind un script python care are ca atribuție aranjarea particulelor pentru momentul inițial pe o rețea perpendiculară, creând un cub cu i^3 particule în el. În cazul de față s-au generat multiple fișiere pentru a analiza evoluția comportamentului sistemului pentru mai multe densități de particule. S-a considerat un volum fix în care s-a introdus un număr de $N = n^3$ particule, cu $n = \overline{5,16}$. Totodată acest fișier poate fi folosit pentru captarea detaliată a evoluției poziției particulelor din sistem, unde la rularea simulării se achiziționează pozițiile fiecărei particule din 10 în 10 frame-uri într-un fișier traiectorie “trj.pdb”.

CRYST1	17.089	17.089	17.089	90.00	90.00	90.00					
ATOM	1	C	CO	C	1	-6.764	-7.235	-6.335	0.00	0.00	C
ATOM	2	O	CO	C	1	-6.889	-6.535	-7.211	0.00	0.00	O
ATOM	3	C	CO	C	2	-6.503	-6.284	-3.416	0.00	0.00	C
ATOM	4	O	CO	C	2	-7.085	-7.250	-3.419	0.00	0.00	O
ATOM	5	C	CO	C	3	-6.806	-6.749	-0.638	0.00	0.00	C
ATOM	6	O	CO	C	3	-6.858	-6.900	0.479	0.00	0.00	O
ATOM	7	C	CO	C	4	-6.751	-7.091	2.832	0.00	0.00	C
ATOM	8	O	CO	C	4	-6.899	-6.644	3.857	0.00	0.00	O
ATOM	9	C	CO	C	5	-7.472	-6.744	6.789	0.00	0.00	C
ATOM	10	O	CO	C	5	-6.358	-6.904	6.871	0.00	0.00	O
END											

Numărul atomului Numele atomului
 Numele reziduului Numărul reziduului
 Coordonatele carteziene Factorul de ocupare Factorul de temperatură

Figură 3.1 Structura fișierului .pdb

Pentru a vizualiza traiectoriile particulelor de-a lungul simulării pe baza datelor salvate în fișierul de traiectorie, este nevoie de a importa acestea într-un program destinat vizualizării acestora, cum ar fi în cazul actual *VMD*.



Figură 3.2 Evoluția sistemului în simularea MDSIM pentru 125 particule considerate

Protein Structure File (.psf) [12] este fișierul folosit pentru definirea structurii topologice globale ale configurației inițiale, astfel el are ca rol principal definirea fiecărui atom, masă și sarcină, respectiv conexiunile atomilor între ei, astfel setând condițiile pentru descrierea câmpului de forțe prezent. Acest fișier este generat împreună cu fișierul PDB pentru a stabili legăturile și

interacțiile prezente din primul moment al simulării. Folosind acest fișier se pot realiza listele de legături și perechi de interacțiune considerate pentru rularea sistemului.

```

PSF

1 !NTITLE
REMARKS CO
      Numele      Tipul      Sarcina      Masa
250 !NATOM reziduului atomului atomului atomului
  1 CO      1 CO      C      CM      0.02100 12.01070 0
  2 CO      1 CO      O      OM     -0.02100 15.99940 0
      Numele
...      atomului

249 CO     125 CO      C      CM      0.02100 12.01070 0
250 CO     125 CO      O      OM     -0.02100 15.99940 0
      Legăturile
125 !NBOND: bonds interatomice
  1      2      3      4      5      6      7      8
  9     10     11     12     13     14     15     16
...

241      242      243      244      245      246      247      248
249      250

0 !NTHETA: angles
0 !NPHI: dihedrals
0 !NIMPHI: impropers
0 !NDON: donors
0 !NACC: acceptors
0 !NNB: nonbond exclusions

0      0      0      0      0      0      0      0
...|
0      0      0      0      0      0      0      0
0      0

```

Figură 3.3 Structura fișierului .psf

Pentru a caracteriza corect câmpul de forțe prezent este nevoie de folosirea fișierului Force Field Parameter file (.par), care conține informații referitoare la valorile constantelor de forță și echilibru necesare, culese din literatura de specialitate.

Având la dispoziție configurarea inițială a sistemului următorul pas de interes este optimizarea metodei de calcul prin alegerea directivelor de compilare ale codului C/C++ necesare. Prin analiza rapidă a modelului *MDSIM* cea mai mare parte a resurselor de calcul este risipită în efectuarea calculelor specifice metodei Ewald, prin trecerea în spațiul reciproc folosind implementarea transformatei Fourier. În momentul actual această funcționalitate include multiple bucle *for*

imbricate care asigură pași de calcul precum: generarea funcțiilor de bază Fourier într-un mod recursiv, calcularea contribuțiilor spațiului reciproc la suma Ewald, ceea ce include și trecerea valorilor calculate din spațiul reciproc în spațiul real. [9]

Fără a introduce complicații inutile implementării se pot alege menționarea anumitor directive de compilare, care au scopul de a îndruma compilatorul pentru a optimiza de sine secvența care urmează pentru diferite scenarii de aplicare. În cazul de față este benefică introducerea directivei *#pragma omp parallel for* [13], care are semnificația de a încuraja paralelizarea buclei / buclelor *for* care urmează, pentru a reduce atât timpii de execuție, cât și folosirea mai bună a resurselor mediului de rulare care sunt la dispoziție, adică rularea temporară pe mai multe fire de execuție. Prin introducerea acestei modificări minore funcționalități, se pot observa îmbunătățiri de aproximativ 8-15% în timpii de rulare obținuți pentru simulare. Această perspectivă devine importantă pentru configurații încărcate, cu număr de molecule peste 1500, unde timpii de rulare depășesc 15-20,000 secunde.

b. Analiza datelor

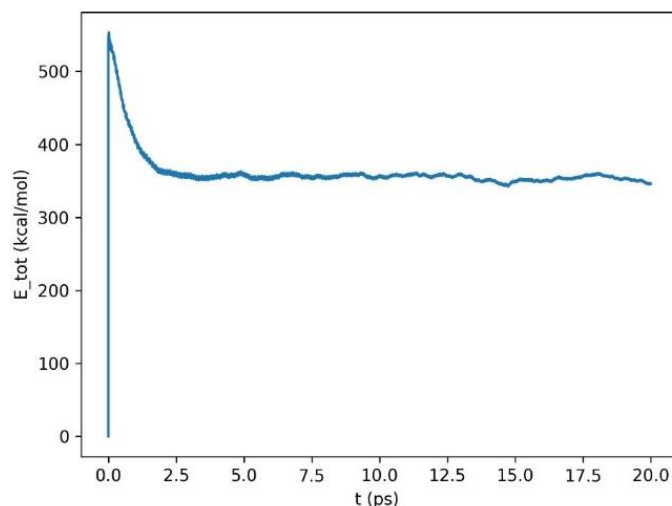
Lucrarea curentă se axează pe urmărirea comportamentului sistemului rulat în mediul de simularea MDSIM. Acesta se realizează prin urmărirea atât a traiectoriilor particulelor, cât și prin urmărirea caracteristicilor sistemului care îi descriu starea într-un mod detaliat, cum ar fi toate contribuțiile energiei Coulombiene, temperatura presiunea, dar este observată și dispersia vitezei respectiv a densității de particule. Astfel prin aceste mărimi studiate se urmărește procesul de stabilizarea a sistemului, respectiv comportamentul acestuia în procesul de oscilare în jurul punctului de echilibru, după care se compară rezultatele obținute. În primă instanță se studiază acuratețea, stabilitatea și predictibilitatea sistemului, după care se dorește analizarea valorilor de echilibru la care se stabilizează sistemul.

i. Compararea evoluției MDSIM pentru mai multe configurații inițiale

Comportamentul unui sistem de molecule introdus într-o simulare dinamică poate varia semnificativ în funcție de configurațiile inițiale considerate. Pentru a studia efectele diferitelor configurații asupra comportamentului unui sistem simulat prin modelul molecular dinamic care este în contact direct cu un termostat, se alege testarea într-un volum inițial de testare caracterizat de PBC și are un număr fix de particule.

Pentru a urmări efectul densității de particule asupra modelului studiat s-a ales rularea simulării pentru multiple valori ale volumului inițial pentru un număr relativ mic de particule. Numărul de $N=216$ de particule s-a ales convenabil, acesta prezentând timpi relativ mici de rulare și se poate obține o plajă mare de valori ale densității particulelor din volumul de simulare. Important de menționat faptul ca aceste teste s-au rulat pentru distribuții inițiale de particule uniforme pe tot volumul de simulare considerat. Adică am crescut artificial distanța dintre particule pentru a avea mai puține interacțiuni în interiorul razei de “cut-off” iar asta va influența mărimi fizice cum ar fi energia de interacțiune Lennard-Jones respectiv cea electrostatică. [14]

Pentru a putea analiza și compara stările sistemelor, acesta se va efectua când acestea deja ajung să oscileze în jurul unui punct de echilibru dat de interacțiunea cu termostatul. Având în vedere faptul că este prezentă o oscilație se vor nota valorile medii din măsurători, acestea fiind considerate ca valori de echilibru a sistemului în configurația curentă. Cum se observă din Fig,16 sistemul molecular ajunge în starea descrisă adineauri după aproximativ $t = 2.5ps$ de la pornirea simulării.



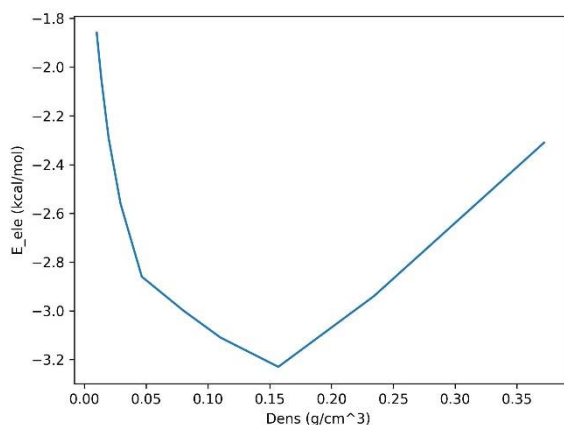
Figură 3.4 Evoluția în timp a Energiei de totale Coulombiene a sistemului studiat pentru latura de 25Å a volumului de simulare

Tabel 3.1 Evoluția sistemului pentru diferite volume de simulare considerate

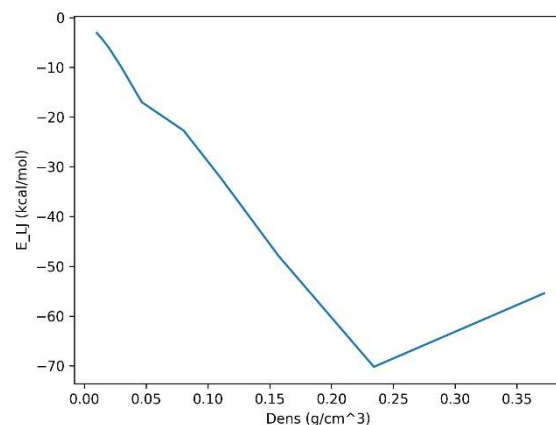
L_{Box} [Å]	E_{tot} [$\frac{kcal}{mol}$]	E_{bnd} [$\frac{kcal}{mol}$]	E_{LJ} [$\frac{kcal}{mol}$]	E_{ele} [$\frac{kcal}{mol}$]	E_{kin} [$\frac{kcal}{mol}$]	Press [bar]	ρ [$\frac{g}{cm^3}$]
10	$1.70 \cdot 10^{19}$	$8.68 \cdot 10^{18}$	0.01	932.54	$8.36 \cdot 10^{18}$	$3.87 \cdot 10^{20}$	10.04
15	$9.93 \cdot 10^{25}$	$5.06 \cdot 10^{25}$	560206.02	673.19	$4.87 \cdot 10^{26}$	$6.68 \cdot 10^{26}$	2.97
20	$2.00 \cdot 10^{42}$	$9.97 \cdot 10^{42}$	0.00	929.36	$1.01 \cdot 10^{42}$	$4.68 \cdot 10^{43}$	1.25
25	354.98	25.77	-55.44	-2.31	386.97	4577.23	0.64
30	342.10	28.89	-70.23	-2.94	386.39	1099.91	0.37
35	366.11	30.86	-47.86	-3.23	386.35	523.33	0.23
40	379.99	28.88	-32.19	-3.11	386.42	318.56	0.15
45	393.85	33.24	-22.72	-3.00	386.32	211.59	0.11
50	394.50	27.96	-17.00	-2.86	386.41	150.12	0.08
60	409.41	35.27	-9.68	-2.56	386.38	84.91	0.04
70	409.56	31.43	-5.96	-2.29	386.39	52.96	0.02
80	414.03	33.83	-4.12	-2.05	386.38	35.29	0.019
90	412.51	31.08	-3.08	-1.86	386.38	24.78	0.013
100	414.01	31.55	-2.21	-1.69	386.37	17.99	0.010

Analizând datele obținute în tabelul de mai sus se observă ca sistemul construit din 216 molecule de CO prezintă o stabilitate sporită pentru latura volumului de simulare mai mare de $L_{Box} = 20\text{Å}$ ceea ce se traduce într-o densitate generală a particulelor de $\rho < 1.25 \text{ g/cm}^3$, unde densitatea în condiții normale ale CO este de $\rho_0 = 1.165 \text{ g/cm}^3$. Acest fenomen de instabilitate se poate explica cu faptul ca distanța intermoleculară e apropiată, comparativă cu lungimea legăturii CO, $R_{CO} = 1.128 \text{ Å}$ iar distanțele intermoleculare în aceste cazuri de densitate moleculară ridicată se situează în intervalul [1.66-4.16] Å

Mărimile fizice precum energia Coulombiană totală, energia de legătură, energia cinetică totală a particulelor sunt constante în pentru toate valorile ale densității, ceea ce sugerează că sistemul este unul stabil și previzibil în ceea ce privește simularea interacțiunilor moleculare care depind în primul rând de interacțiunea sistemului cu un termostat.



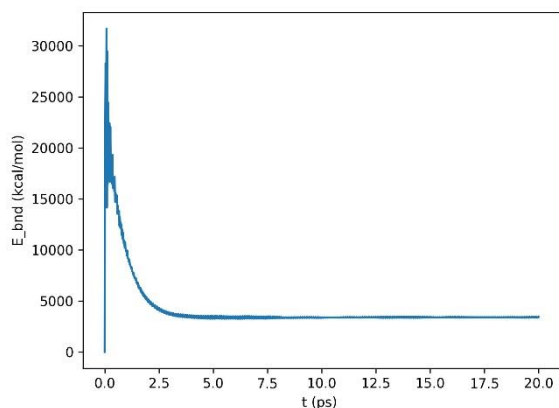
Figură 3.5 Evoluția Energiei Electrostatice în funcție de densitatea particulelor din configurația inițială



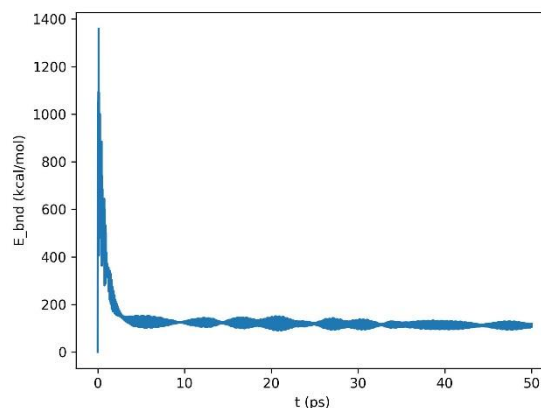
Figură 3.6 Evoluția Energiei de interacțiune Lennard-Jones în funcție de densitatea particulelor din configurația inițială

Mai departe se observă o abatere de la cea așteptată a evoluției energiei de interacțiune Lennard-Jones respectiv a celei electrostatice cu descreșterea densității particulelor din volumul considerat. Acest comportament se explică prin analiza efectului densității particulelor asupra parametrului Ewald α . În modelul curent s-a folosit pentru calculul interacțiunilor $\alpha = 0.238$, ceea ce influențează în mod negativ acuratețea interacțiunilor de scrută durată. Acest coeficient oferă precizie mai mare pentru valori mai mari, iar în cazul modelului el trebuie să fie calibrat pentru fiecare valoare a densității particulelor (distanța dintre particule în configurația originală). Conform teoriei: $\alpha \sim \frac{1}{d}$, unde d este distanța intermoleculară din configurație, iar acesta sugerează ca $\alpha \sim \rho$, deci pentru a menține acuratețea modelării, este nevoie de a ajusta valoarea lui $\alpha(\rho)$ pentru a evita defazajele. [2]

În urma acestor constatări urmează studiul sistemului pentru un volum de simulare constant cu latura de $L_{Box} = 100\text{\AA}$, pentru multiple configurații inițiale ale sistemului molecular aranjate în așa fel ca ele să creeze o rețea de molecule, în jurul centrului volumului de simulare, cu distanța dintre particule menținută constantă pentru fiecare configurație. Generând configurația inițială în așa fel ca distanța intermoleculară să fie constantă în fiecare caz, iar acesta să derive din densitatea CO la condiții normale. Astfel deși s-a dorit realizarea experimentelor pentru o durată de $t = 50ps$, măsurătorile obținute urmăresc starea sistemului doar pentru durata de timp amintită mai sus. Configurațiile testate au variat în privință de număr de molecule considerate de la $N_{min} = 125$ particule până la $N_{max} = 4096$ particule. Și în acest caz s-a studiat porțiunea în care sistemul deja s-a stabilizat și oscilează în jurul punctului de echilibru dat de interacțiunea cu termostatul cu care este în contact.



Figură 3.7 Evoluția în timp a Energiei de legătură din sistemul studiat pentru 4096 particule



Figură 3.8 Evoluția în timp a Energiei de legătură din sistemul studiat pentru 125 particule

Din dependențele prezentate în graficele de mai sus, se observă ca sistemul ajunge într-o stare stabilă după aproximativ $t = 2.5ps$ de la pornirea simulării. Ca și în cazul precedent s-au calculat valorile medii ale energiilor și presiunii înregistrate în această zonă și s-au introdus în tabelul de mai jos:

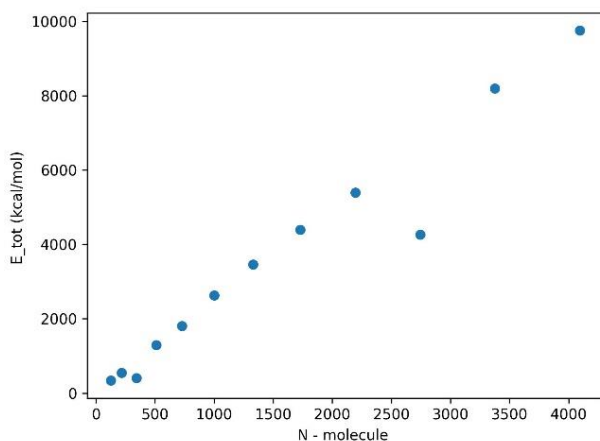
Tabel 3.2 Evoluția sistemului pentru diferite grade de umplere a volumului de simulare considerat

N	$E_{tot} \left[\frac{kcal}{mol} \right]$	$E_{bnd} \left[\frac{kcal}{mol} \right]$	$E_{LJ} \left[\frac{kcal}{mol} \right]$	$E_{ele} \left[\frac{kcal}{mol} \right]$	$E_{kin} \left[\frac{kcal}{mol} \right]$	Press [bar]
125	339.77	118.28	-1.37	-0.98	223.85	10.35
216	550.67	169.69	-3.98	-1.68	386.64	17.89
343	405.08	24.17	-3.60	-1.69	386.21	17.97
512	1293.58	397.38	-15.83	-4.01	916.05	42.47
729	1804.04	535.47	-29.94	-5.72	1304.22	60.64
1000	2633.86	910.28	-57.85	-7.82	1789.26	82.93
1331	3459.00	1187.29	-99.18	-10.41	2381.30	110.66
1728	4396.58	1483.95	-165.46	-13.53	3091.63	144.50
2197	5396.74	1745.37	-261.81	-17.20	3930.38	185.84
2744	4268.31	592.70	-235.20	-17.22	3928.03	193.57
3375	8198.69	2809.86	-623.43	-26.37	6038.64	293.18
4096	9753.63	3395.05	-944.69	-31.76	7335.02	369.62

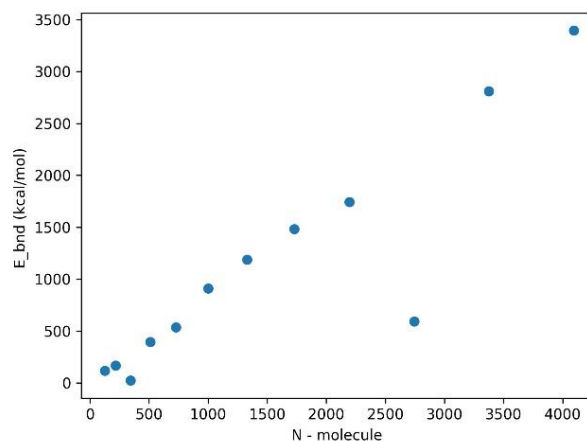
Analizând rezultatele obținute din tabelul anterior se poate observa o evoluție aproape liniară a sistemului în funcție de gradul de umplere al acestuia cu particule uniform distribuite în jurul centrului volumului de simulare considerat. Singura anomalie de la această regulă o prezintă

măsurătoarea obținută pentru $N = 2744$ particule. Cel mai probabil în acest caz au apărut situații consecutive în care particulele au reușit să interacționeze în așa fel încât să orienteze evoluția sistemului către o stare mult mai echilibrat dispersată decât în restul cazurilor.

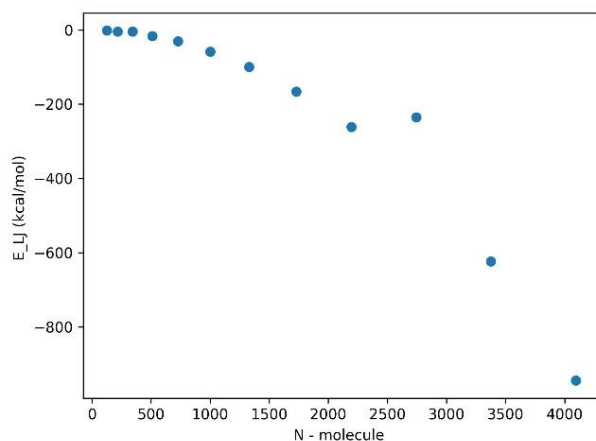
Această dependență liniară se poate observa foarte clar în graficele atașate mai jos , cu excepția energiei de interacțiune Lennard-Jones, care după definiție prezintă o dependență invers proporțională cu distanța intermoleculară, care scade cu umplerea volumului cu particule, ceea ce rezultă creșterea în modul a potențialului de interacțiune.



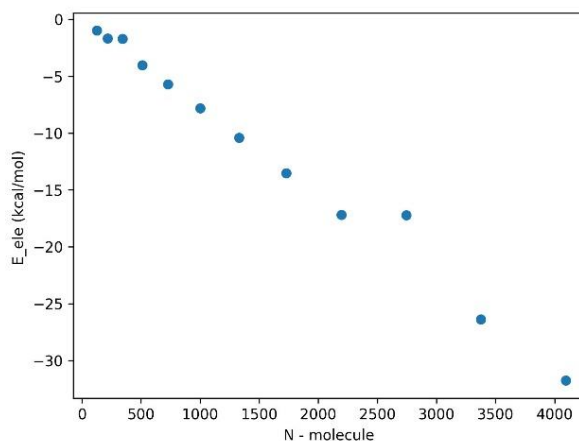
Figură 3.9 Energia totală coulombiană a sistemului în funcție de numărul de particule



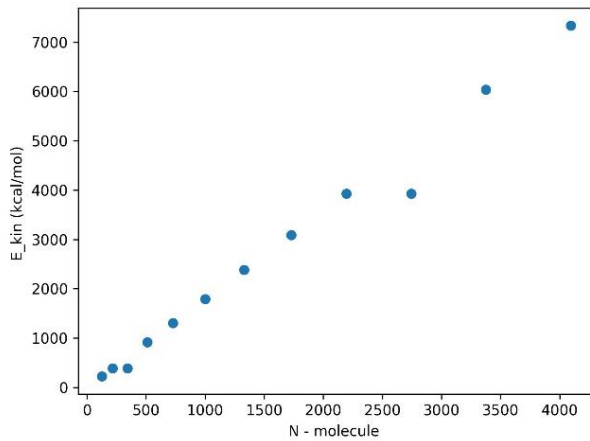
Figură 3.10 Energia de legătură a sistemului în funcție de numărul de particule



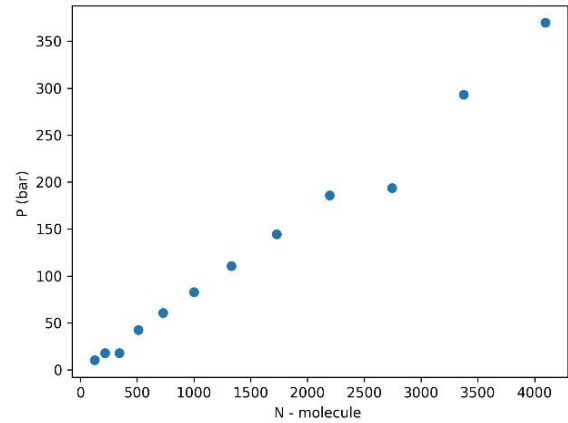
Figură 3.11 Energia de interacțiune Lennard-Jones a sistemului în funcție de numărul de particule



Figură 3.12 Energia electrostatică de interacțiune a dinamic în funcție de numărul de particule



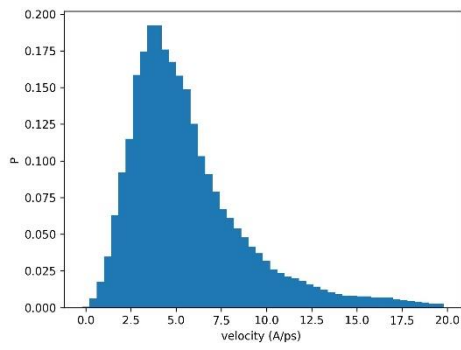
Figură 3.13 Energia cinetică totală a sistemului în funcție de numărul de particule



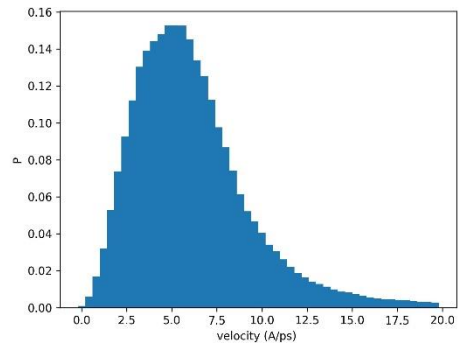
Figură 3.14 Presiunea sistemului în funcție de numărul de particule

Pentru a studia mai departe atât predictibilitatea, cât și coerența sistemului s-au făcut analize care vizează mai mult evoluția traiectoriei și pozițiilor particulelor. Astfel s-au făcut analize pe distribuțiile de viteze, distribuțiile radiale și coeficientul de difuzie.

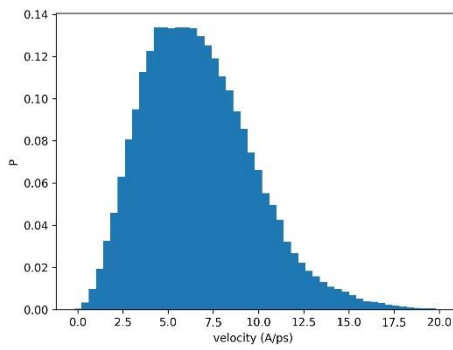
În privința distribuției particulelor pe viteze s-au obținut următoarele histograme:



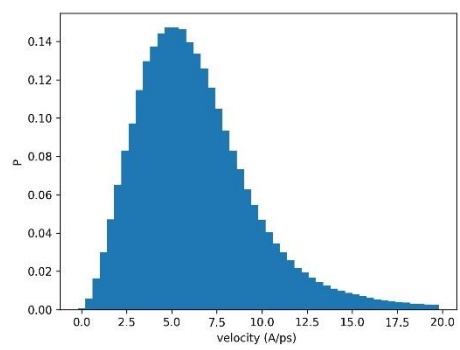
Figură 3.15 Distribuția fracționară a vitezei particulelor pentru 125 molecule



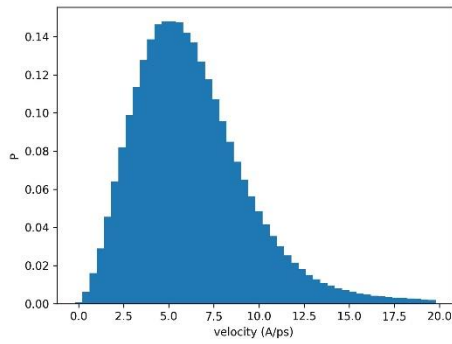
Figură 3.16 Distribuția fracționară a vitezei particulelor pentru 216 molecule



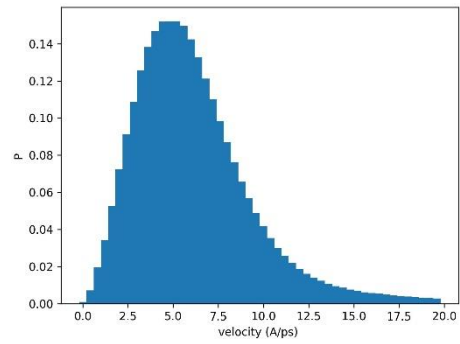
Figură 3.17 Distribuția fracționară a vitezei particulelor pentru 343 molecule



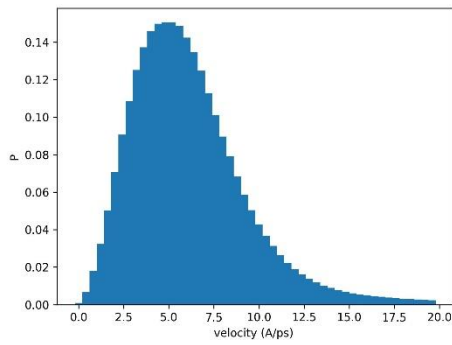
Figură 3.18 Distribuția fracționară a vitezei particulelor pentru 512 molecule



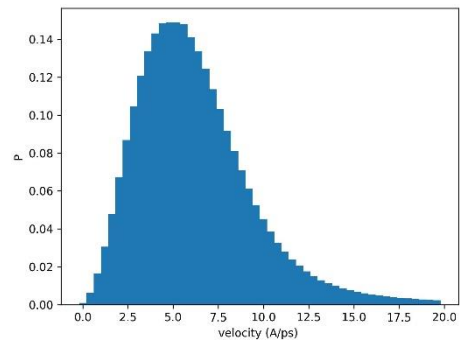
Figură 3.19 Distribuția fracțională a vitezei particulelor pentru 729 molecule



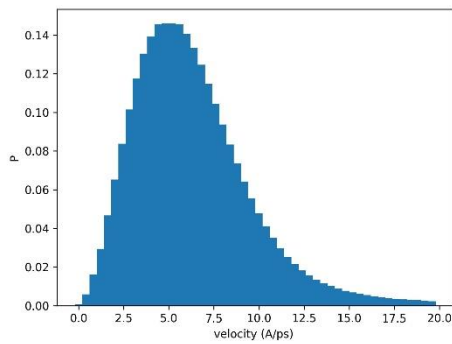
Figură 3.20 Distribuția fracționară a vitezei particulelor pentru 1000 molecule



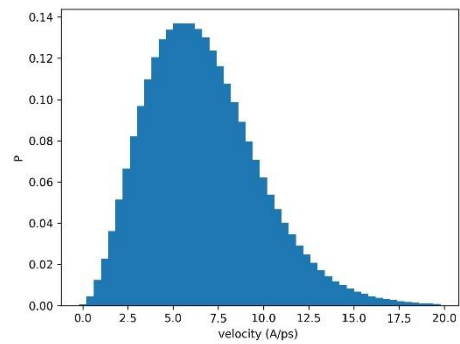
Figură 3.21 Distribuția fracționară a vitezei particulelor pentru 1331 molecule



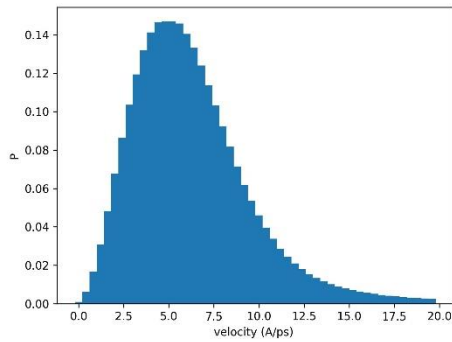
Figură 3.22 Distribuția fracționară a vitezei particulelor pentru 1728 molecule



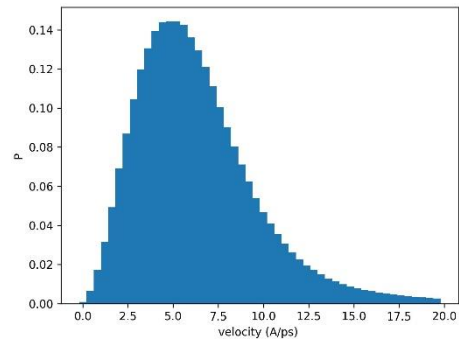
Figură 3.23 Distribuția fracționară a vitezei particulelor pentru 2197 molecule



Figură 3.24 Distribuția fracționară a vitezei particulelor pentru 2744 molecule



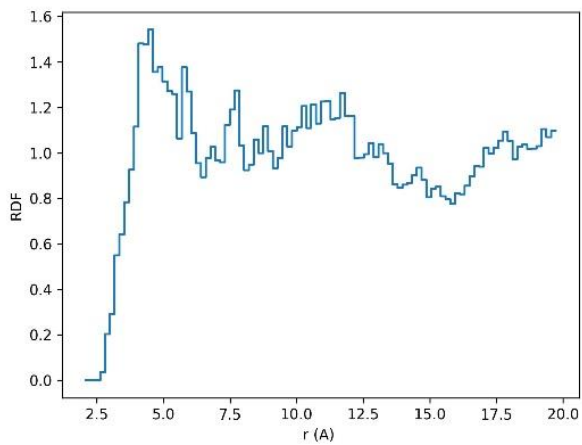
Figură 3.25 Distribuția fracționară a vitezei particulelor pentru 3375 molecule



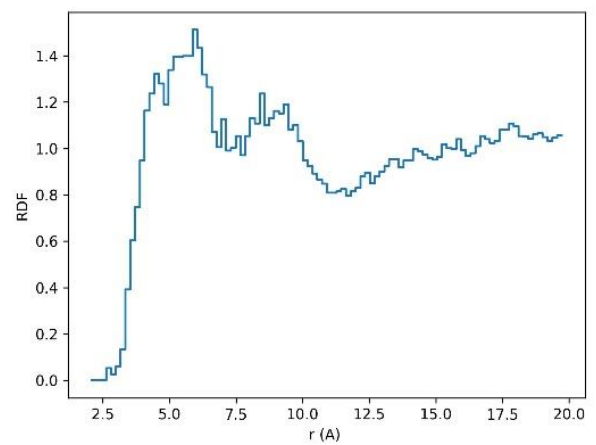
Figură 3.26 Distribuția fracționară a vitezei particulelor pentru 4096 molecule

Din histogramele de mai sus se observă ca în fiecare caz distribuția de viteze urmează o distribuție Poisson, care își atinge maximul la $v \approx 5 \text{ A/ps}$. Însă important de menționat faptul ca acest maxim se shiftează spre valori mai mari ($3.4 \rightarrow 5.8 \text{ A/ps}$) cu cât avem un număr de particule mai mare, iar semilărgimea curbei de distribuție prezintă și aceasta o creștere semnificativă cu numărul de molecule.

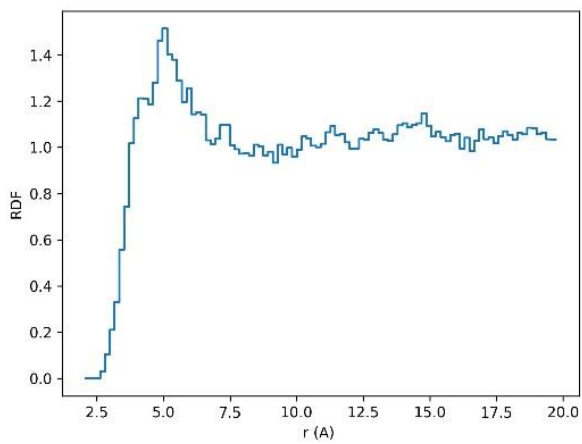
După analiza distribuției de viteze este importantă analiza distribuției radiale ale particulelor pe baza datelor obținute din fișierele de traiectorie generate în urma rulării simulărilor.



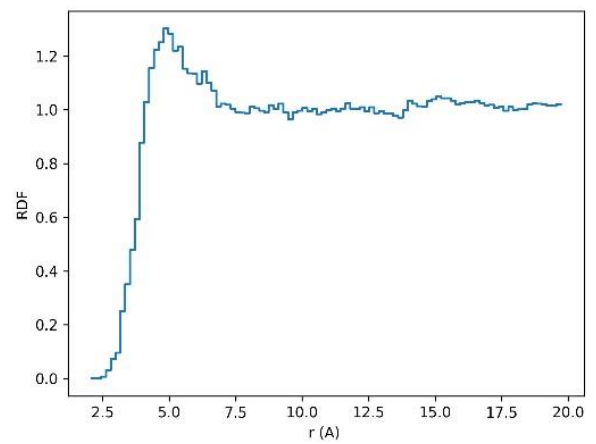
Figură 3.27 Factorul de distribuție radială în funcție de distanță pentru 125 molecule



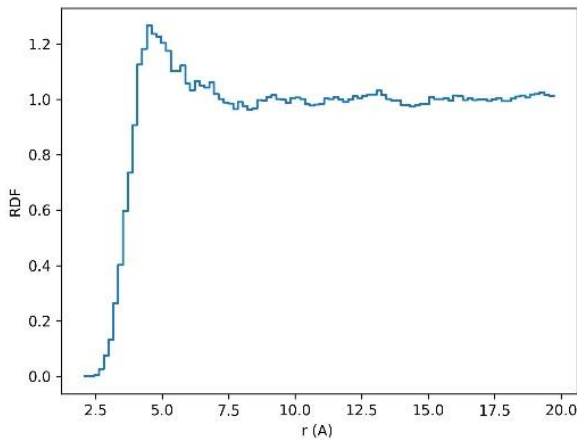
Figură 3.28 Factorul de distribuție radială în funcție de distanță pentru 216 molecule



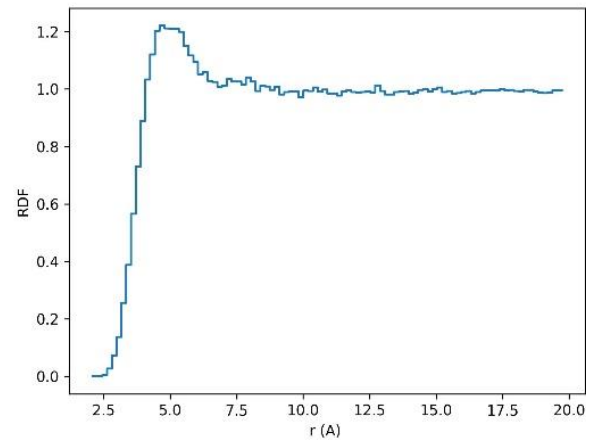
Figură 3.29 Factorul de distribuție radială în funcție de distanță pentru 343 molecule



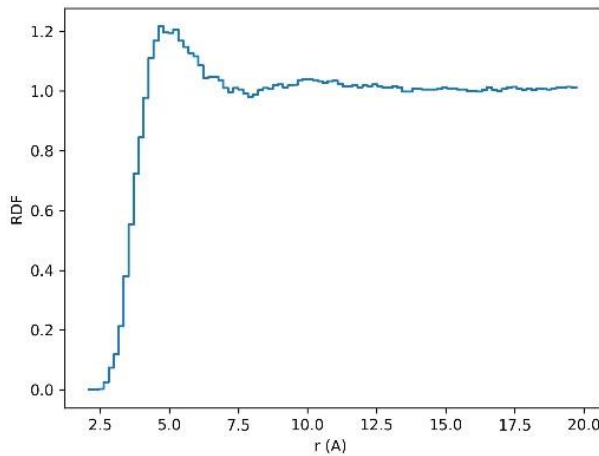
Figură 3.30 Factorul de distribuție radială în funcție de distanță pentru 512 molecule



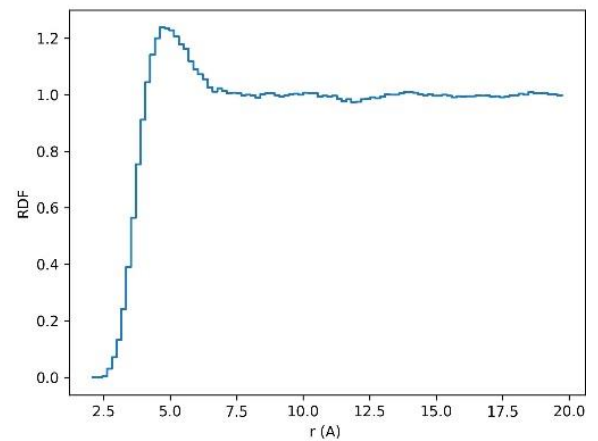
Figură 3.31 Factorul de distribuție radială în funcție de distanță pentru 729 molecule



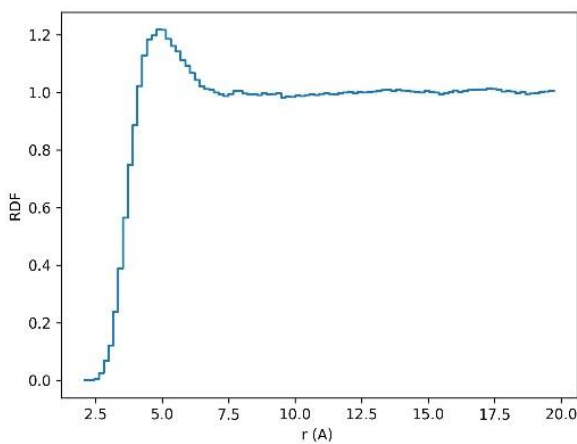
Figură 3.32 Factorul de distribuție radială în funcție de distanță pentru 1000 molecule



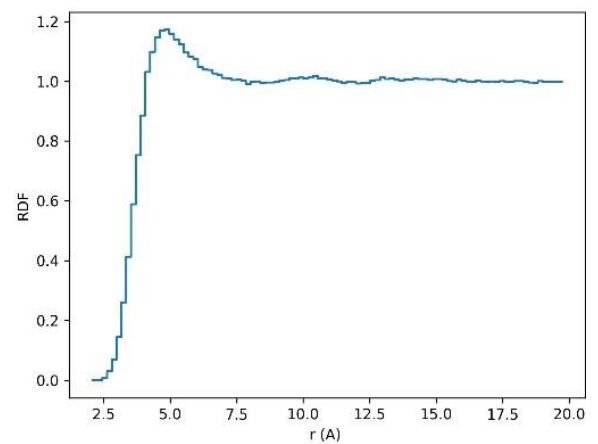
Figură 3.33 Factorul de distribuție radială în funcție de distanță pentru 1331 molecule



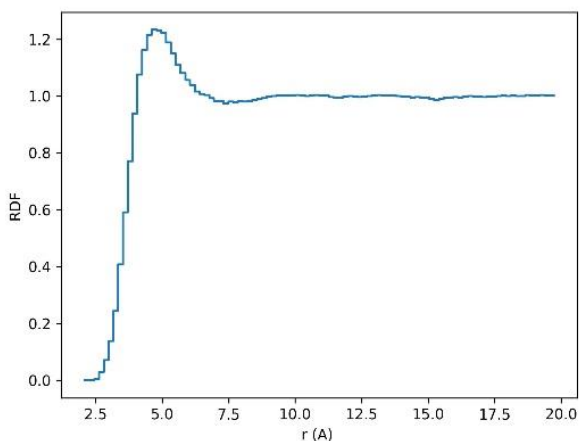
Figură 3.34 Factorul de distribuție radială în funcție de distanță pentru 1728 molecule



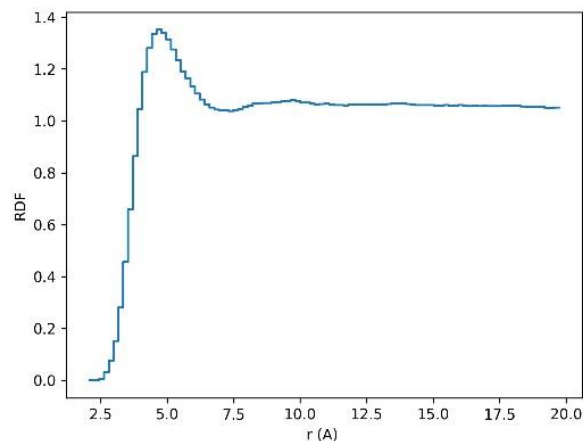
Figură 3.35 Factorul de distribuție radială în funcție de distanță pentru 2197 molecule



Figură 3.36 Factorul de distribuție radială în funcție de distanță pentru 2744 molecule

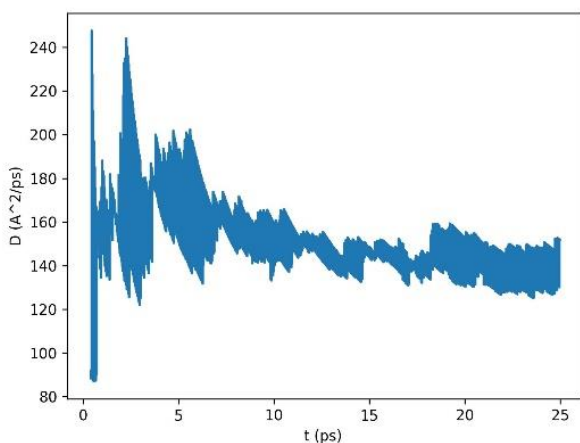


Figură 3.37 Factorul de distribuție radială în funcție de distanță pentru 3375 molecule

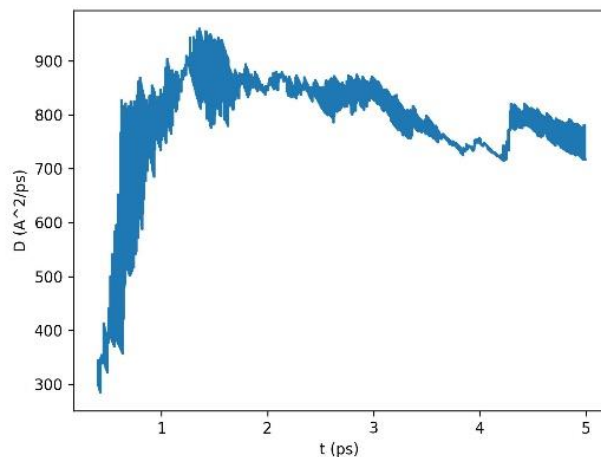


Figură 3.38 Factorul de distribuție radială în funcție de distanță pentru 4096 molecule

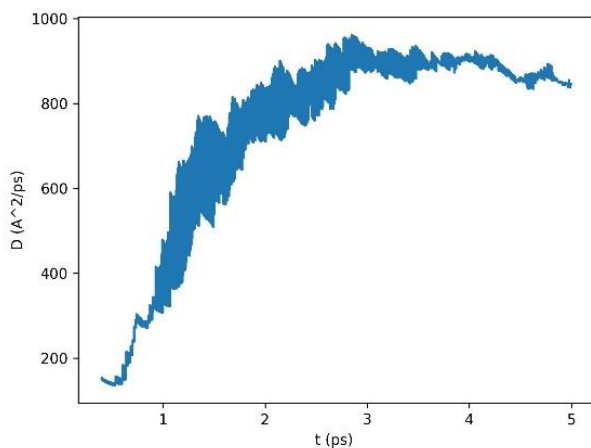
Gazele în sine sunt tratate ca fluide simple, astfel și ele sunt caracterizate de absența unei structuri permanente, dar se evidențiază anumite corelații structurale bine definite care reflectă o organizare prezentă la nivel molecular. Astfel se analizează comportamentul factorului de distribuție radială al unui sistem isotopic, care este definit ca o mediere asupra mai multor pași din simulare. Evoluția acestui factor de distribuție se resimte din factorul de structură al sistemului studiat. Se observă un peak major al factorului foarte aproape de punctul de separare atomic după care se observă oscilații cu scădere exponențială, comportament specific fluidelor. Se studiază și evoluția coeficientului de difuzie al gazului / fluidului studiat pornind de la legea lui Flick și folosind formula lui Einstein pentru a o aproxima pe aceasta.



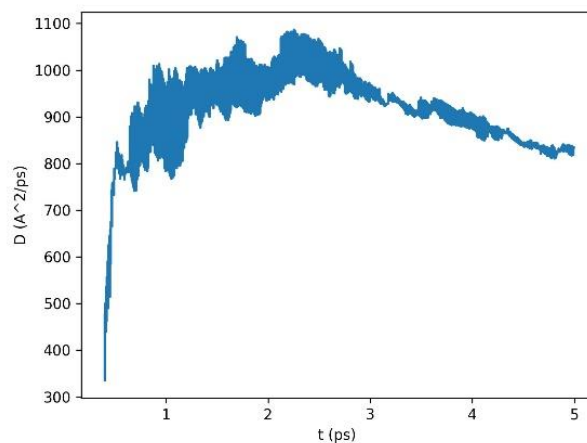
Figură 3.39 Evoluția coeficientului de difuzie pentru 125 molecule



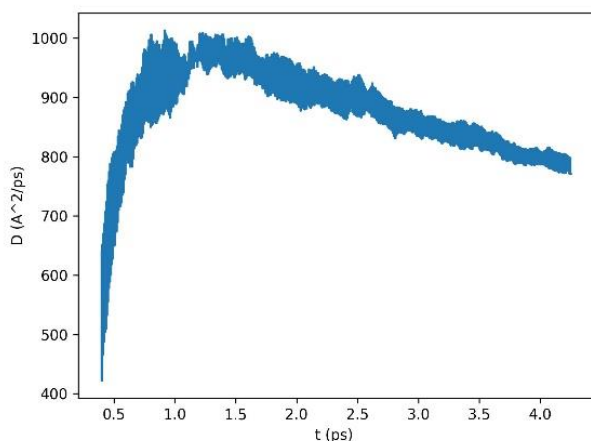
Figură 3.40 Evoluția coeficientului de difuzie pentru 216 molecule



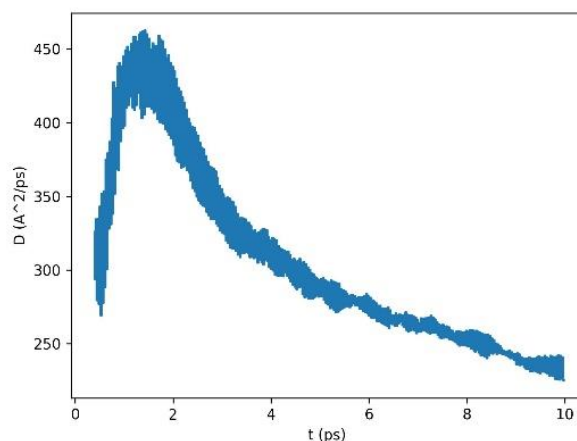
Figură 3.41 Evoluția coeficientului de difuzie pentru 343 molecule



Figură 3.42 Evoluția coeficientului de difuzie pentru 512 molecule



Figură 3.43 Evoluția coeficientului de difuzie pentru 3375 molecule



Figură 3.44 Evoluția coeficientului de difuzie pentru 4096 molecule

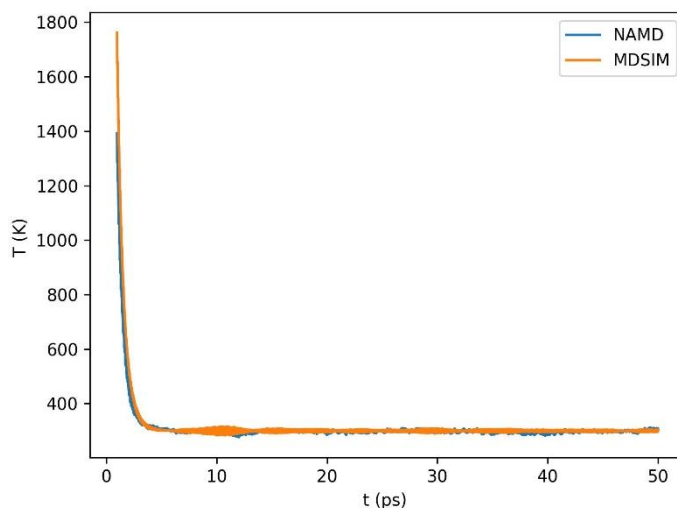
Din cauza modului în care PBC este tratat, pentru a observa difuzia cu adevărat, trebuie desfăcută împachetarea, și sistemul finit tratat pentru un timp suficient de scurt $t < 5ps$. Astfel comportamentul acestui coeficient va fi o creștere până la o plafonare, după care urmează o scădere introdusă de decorelarea particulelor. Acest comportament teoretic este susținut și de graficele obținute după măsurătorile simulărilor. Această decorelare este insesizabilă pentru un număr mic de particule tratate, în cazul actual $N < 500$ molecule, iar ea devine mult mai vizibilă pentru un număr mai mare de particule, cum se observă și în cazul sistemelor caracterizate de $N > 3000$ molecule, unde apare o decorelare mai accentuată și pentru timpi mai mici decât cele $5ps$ considerate inițial.

ii. Compararea MDSIM cu NAMD

NAMD este un software dezvoltat pentru aplicații în simularea de înaltă performanță a sistemelor moleculare dinamice, iar avantajele i se remarcă prin algoritmi implementați. Conceput pentru studiul sistemelor biomoleculare mari, NAMD folosește algoritmi avansați, cum ar fi tehnici de paralelizare scalabile, calcule eficiente de forță și rafinament adaptiv al rețelei. Aceste caracteristici permit NAMD să simuleze cu acuratețe interacțiuni moleculare complexe la detalii atomice, oferind perspective neprețuite asupra proceselor biologice. Cu accent pe noile modele matematice și algoritmi implementați, respectiv eficiența computațională rezultată, NAMD a devenit un instrument de bază pentru cercetătorii din domeniul biologiei computaționale.

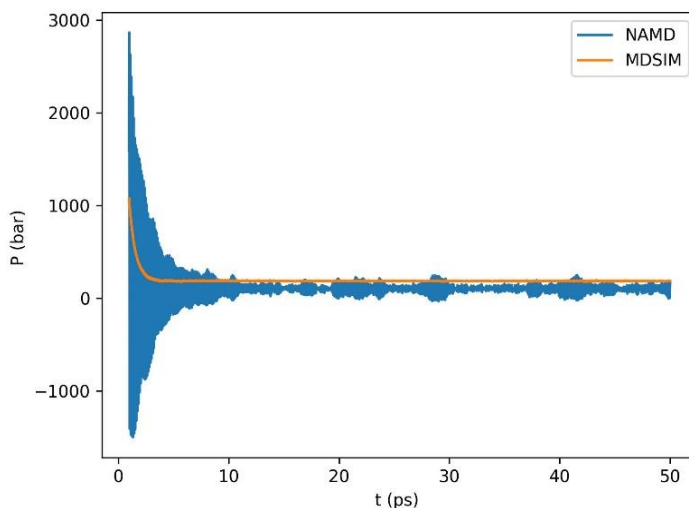
Prin urmare NAMD reprezintă etalon de comparație a MDSIM, pentru a verifica nu numai coerența și predictibilitatea acestuia, dar și pentru a-i compara acuratețea față de un program utilizat pe scară largă în acest domeniu. Pentru a realiza această comparație, s-au efectuat rulări ale simulării cu NAMD pentru studiul sistemului prin modelul molecular dinamic pentru un volum de simulare constant cu latura de $L_{Box} = 100\text{\AA}$ pentru multiple numere de molecule aranjate în mod egal în jurul centrului volumului de simulare în așa fel, încât densitatea de particule să fie menținută aceeași pentru fiecare caz. Generând configurația inițială în așa fel ca distanța intermoleculară să fie constantă în fiecare caz, iar acesta să derive din densitatea CO la condiții normale. În aceste condiții sistemul a fost lăsat să ruleze timp de $t = 50ps$, iar măsurătorile obținute urmăresc starea sistemului. Configurațiile testate au variat în privință de număr de molecule considerate de la $N_{min} = 125$ particule până la $N_{max} = 4096$ particule. Și în acest caz s-a studiat porțiunea în care sistemul deja s-a stabilizat și oscilează în jurul punctului de echilibru dat de interacțiunea cu termostatul cu care este în contact.

Ca în cazurile anterioare s-a urmărit evoluția sistemului prin măsurarea mărimilor fizice precum energia Coulombiană totală, energia de legătură, energia cinetică totală, energia electrostatică, energia de interacțiune Lennard-Jones respectiv temperatura și presiunea și s-au realizat comparații între evoluțiile configurației sistemului atât pentru rularea cu NAMD cât și cea cu MDSIM.



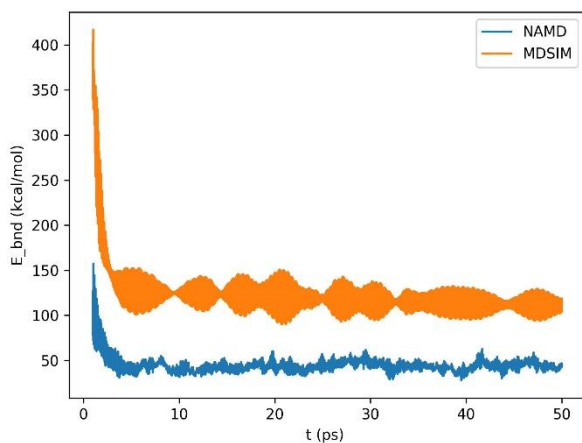
Figură 3.45 Evoluția temperaturii sistemului de 1000 molecule

Având în vedere faptul ca modelul molecular este menținut în contact cu un termostat, o primă comparație a celor două rulări o reprezintă evoluția temperaturii cinetice a moleculelor din clusterul studiat. O primă diferență semnificativă ce se poate evidenția este porțiunea din primele $t = 2 - 2.5ps$ din simulare, unde pentru MDSIM se observă un peak al temperaturii mai pronunțat de $T_{MDSIM} \approx 1800K$, respectiv pentru NAMD se observă un peak doar de $T_{NAMD} \approx 1400K$, în porțiunea în care sistemul ajunge să oscileze în jurul stării de echilibru dictate de termostat. Totodată se observă ca această temperatură de echilibru este cu aproximativ $\Delta t \approx 1ps$ mai lent atinsă în cazul MDSIM decât NAMD. Totuși greu de observat dar modelul MDSIM oferă o stabilitate mult mai mare a valorilor, amplitudinea oscilațiilor este mai mică decât în cazul NAMD pentru perioade de timp mai îndelungate.

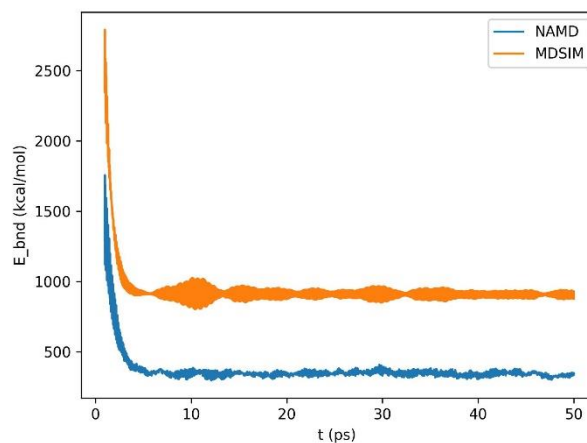


Figură 3.46 Evoluția presiunii sistemului de 2197 molecule

Presiunea exercitată de moleculele unui gaz menținut la o temperatură constantă este dată de ciocnirile moleculelor cu pereții, în cazul curent imaginari, ai volumului considerat. În contrast cu comportamentul observat în cazul evoluției temperaturii în cele două modele, în cazul presiunii MDSIM este cel care dă o evoluție mult mai coerentă care se stabilizează mult mai repede comparativ cu NAMD. MDSIM prezintă un peak de la început de aproximativ jumătate din înălțimea celui întâmpinat la NAMD, iar starea de echilibru se atinge considerabil mai repede $\Delta t \approx 4ps$. Totodată este de menționat ca NAMD prezintă oscilații serioase în primele 10ps din rulare și din calcule chiar redă valori incompatibile cu realitatea, ajungând la presiuni negative pentru câteva picosecunde. Acest comportament observat în ambele cazuri este datorat procesului de stabilizare a sistemului.

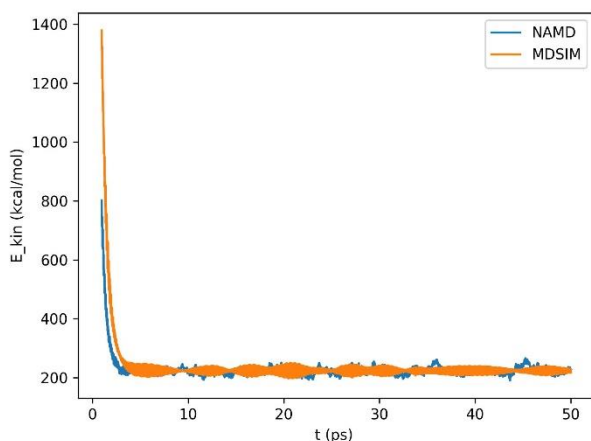


Figură 3.47 Evoluția energiei de legătură pentru sistemul de 125 molecule

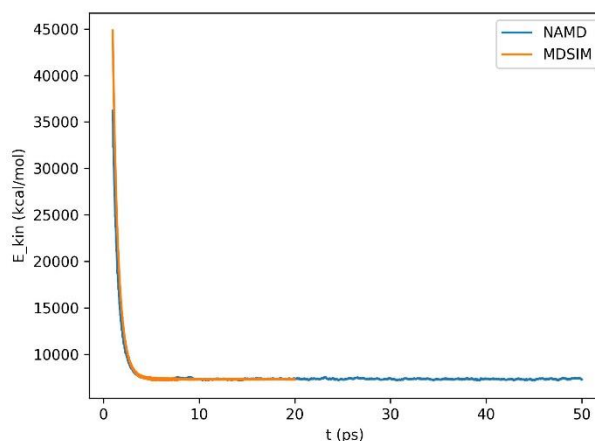


Figură 3.48 Evoluția energiei de legătură pentru sistemul de 1000 molecule

Analizând evoluția energiei de legătură din sistemele moleculare se observă o diferență care persistă și este și într-o creștere proporțională. Observând în detaliu cele două dependențe, NAMD prezintă valori mai mici decât MDSIM, mai ales în cazul stării de echilibru. Deși pentru sistemul de $N = 125$ de particule cele două valori sunt destul de compatibile, dar în cazul $N = 1000$ de particule, diferența dintre media de valori din oscilările din jurul stării de echilibru. Analizând mai departe comportamentele din starea de echilibru, modelul MDSIM pentru puține molecule prezintă oscilații cu amplitudini considerabile, dar care scad cu creșterea numărului de particule.



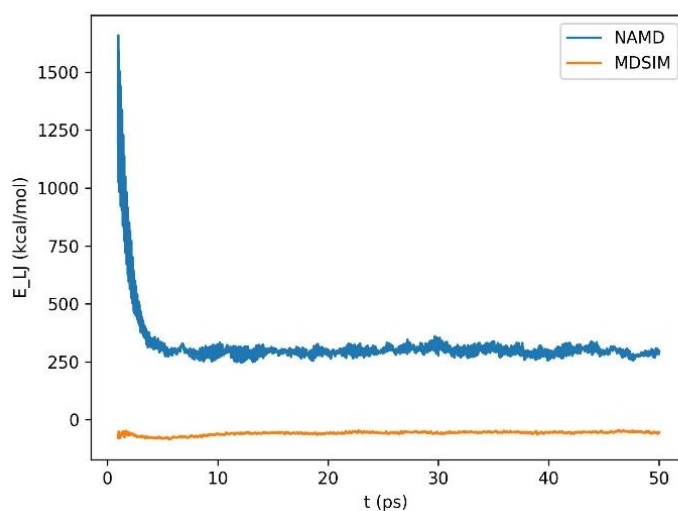
Figură 3.49 Evoluția energiei cinetice a sistemului pentru 125 molecule



Figură 3.50 Evoluția energiei cinetice a sistemului pentru 4096 molecule

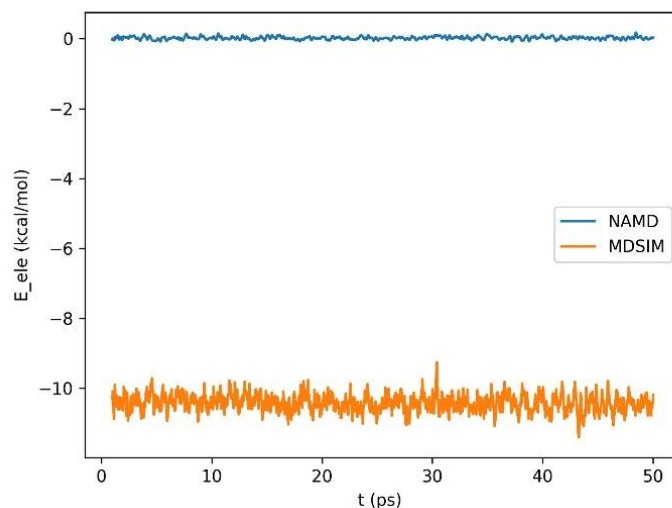
În cazul energiei cinetice a sistemului molecular se observă o suprapunere aproape perfectă între valorile măsurate pentru rulările efectuate cu MDSIM și NAMD pentru un număr de $N = 125$ molecule, cu precizia ca modelul MDSIM ca în cazul presiunii și temperaturii prezintă o evoluție

mult mai coerentă decât NAMD, care prezintă oscilații mai accentuate în jurul poziției de echilibru. În contrast, pentru sistemul cu $N = 4096$ particule, suprapunerea dispare iar în locul acesteia apare o shiftare spre valori mai mari cu ordinul zecilor de unități. Acest trend este în concordanță cu ce s-a observat în cazul energiei de legătură pentru sisteme mai mari rulate în modelul MDSIM. Ca în cazurile anterioare, și acum se observă un peak mult mai pronunțat pentru ambele configurații pentru modelul MDSIM. Se menționează ca pentru sistemul cu $N = 4096$ molecule de CO simularea a fost rulată doar pentru $t = 20ps$ pentru modelul MDSIM din cauze practice care implică timpi de simulare îndelungați.



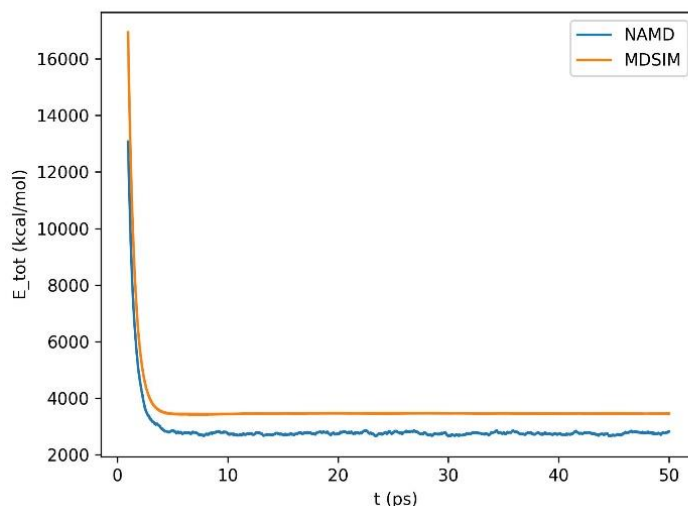
Figură 3.51 Evoluția energiei de interacțiune Lennard-Jones pentru sistemul cu 1000 molecule

Tendențele de evoluție privind valorile energiei de interacțiune Lennard-Jones sunt destul de asemănătoare pentru toate configurațiile testate, prezentând o suprapunere destul de apropiată între tendințele de evoluție obținute pentru măsurători. Modelul NAMD prezintă cel mai accentuat peak în comparație cu modelul MDSIM (multiple ordine de mărime mai mare), dar atinge oscilația în jurul stării de echilibru mai repede ca și MDSIM. Tendința de stabilitate și concordanță mare între cele două modele este datorată interacțiilor de scurtă durată care au un efect mic asupra valorii acestui potențial de interacțiune.



Figură 3.52 Evoluția energiei de interacțiune electrostatică pentru sistemul de 1331 molecule

Din primul moment se observă diferențele semnificative care vizează valorile și formele celor două dependente date de cele două modele de simulare. Această diferență semnificativă își are originile în modul de calcul diferit al celor două, NAMD prezentând o oscilare centrată în jurul energiei de 0 (probabil datorată unui proces de rescalare constantă), iar MDSIM prezintă oscilații cu amplitudini mai mari ca în cazul celuilalt model și o deplasare graduală a punctului de echilibru care este în legătură directă cu numărul de particule considerate .



Figură 3.53 Evoluția energiei totale Coulombiene a sistemului de 1331 molecule

Însumând toate mărimile energetice prezentate mai sus se ajunge la un rezultat destul de interesant pentru a compara în totalitate modelele MDSIM și NAMD. Comparând evoluția nivelului total de energie a celor două sisteme se observă o suprapunere și concordanță semnificativă între ele în ciuda diferitelor instabilități și diferențe punctate deja. Această observație sugerează ca deși descompus pe componente

cele două sisteme prezintă diferențe pe alocuri, în totalitate aceste diferențe se anulează între ele pentru a obține două sisteme care prezintă o evoluție aproape identică în timp care se menține coerent pentru multiple configurații inițiale.

4. Concluzii

În urma analizei prezentate în capitolul anterior, programul MDSIM prezintă atât calități cât și lipsuri în aplicații din viața reală. O primă direcție de analiză a modelului molecular o reprezintă compararea comportamentului sistemului față de comportamentul așteptat pe baza descrierii teoretice a sistemului, precum acesta este prezentat în primul capitol. Multiplele rulări ale simulării moleculare pentru diferite configurații moleculare prezintă o evoluție coerentă a comportamentului modelului pentru majoritatea configurațiilor testate, în regim de densități comparative ale unui gaz. Pe lângă această evoluție coerentă, comportamentele observate sunt în concordanță cu așteptările pe baza ecuațiilor teoretice care descriu comportamentul sistemului în prim plan. Având în vedere faptul că acest model molecular este destinat pentru a simula comportamentul unui fluid simplu, el nu este deloc destinat pentru aplicarea unei simulări care prezintă o configurație inițială în care distanțele intermoleculare sunt considerabil inferioare celor caracteristice gazelor.

Acuratețea și aplicabilitatea modelului se poate determina și punându-l pe acesta într-o comparație directă cu modelul NAMD, un algoritm deja utilizat la scară largă în domeniul simulării sistemelor moleculare dinamice. Acest pas oferă o posibilitate de a analiza și compara atât acuratețea modelului, cât și stabilitatea și tendința de evoluție al acestuia în timp, având ca și termen de comparație direct NAMD. În urma rulărilor a multiple simulări pentru diferite configurații cele două modele au fost comparate din toate punctele de vedere. În mai multe cazuri s-a observat o stabilitate mult mai pronunțată a evoluției în timp a modelului MDSIM comparativ cu NAMD, dar și o stabilizare în jurul poziției de echilibru mai rapidă. Dar s-au mai observat și câteva diferențe în valorile acestor poziții de echilibru în care ajunge sistemul în cazul unor componente energetice, dar însumând toate componentele, nivelul energetic total al aceleiași sistem molecular supus unei simulări prin ambele modele este aproape identic.

În momentul de față datorită limitelor computaționale, modelul MDSIM devine unul foarte solicitant și impractic de folosit pentru aplicații reale de interes în sisteme care sunt compuse din zeci de mii de molecule sau care tind să imite un sistem macroscopic. Aceste limitări provin în primul rând din implementarea de execuție secvențială la momentul actual, ceea ce duce la timpi de execuție deloc practici, în cazul actual pentru configurația de $N = 4096$ molecule timpul de execuție a simulării a depășit 19 ore.

Într-un final, modelul MDSIM reprezintă un compromis favorabil pentru a studia configurații cu un număr redus de particule, unde accentul principal cade pe studiul evoluției caracteristicilor energiilor de interacțiune dintre particule și stabilitatea și predictibilitatea modelului sunt importante.

a. Dezvoltarea proiectului și implementarea funcționalităților noi

Analizând greutățile întâmpinate în timpul rulării modelelor moleculare se evidențiază foarte proeminent necesitatea introducerii unei metode mai rapide și mai eficiente în folosirea tuturor resurselor de calcul care stau la dispoziție.

Pentru a adresa aceste remarci se poate opta pentru a implementa dedicat paralelizarea calculelor din buclele imbricate pentru a fi calculate în interiorul diferitelor fire de execuție (*thread*). Se dorește descompunerea manuală a pașilor de calcul pentru a fi executate separat în diferite fire de execuție, ceea ce necesită identificarea elementelor de buclă care execută calcule independente (caracteristicile moleculelor recalculat nu influențează restul moleculelor tratate într-un alt fir de execuție).

Există posibilitatea implementării unui algoritm orientat spre metode de calcul numeric, Fast-Fourier Transformat *FFT*. Acest algoritm își are ca avantaj principal reducerea semnificativă (ordine de mărime) a timpului de execuție prin excluderea buclelor imbricate care sunt necesare transformării sistemului în spațiul reciproc. În schimb, trecerea se face prin proiectarea fiecărei particule cu sarcină pe o rețea dreptunghiulară uniformă cu un număr de noduri ales convenabil pentru a ușura calculele, care în spațiul reciproc se reduc la manipulări algebrice. Pe lângă eficiența sporită, FFT este un algoritm stabil din punct de vedere numeric, care sporește acuratețea rezultatului obținut. Important de menționat ca în cazul FFT implementarea PBC în spațiul reciproc este mai complex, iar datorită structurii, acest algoritm necesită resurse de proces considerabil mai mari decât algoritmul tradițional.

Referințe

- [1] M. P. T. D. J. Allen, *Computer Simulation of Liquid*, Oxford: Clarendon Press, 1991.
- [2] D. Rapaport, *The Art of Molecular Dynamics Simulation*, Cambridge: Cambridge University Press, 2004.
- [3] A. R. Leach, *Molecular Modelling. Principles and applications*, Prentice Hall, 2001.
- [4] D. S. B. Frenkel, *Understanding Molecular Simulation. From Algorithms to Applications*, Academic Press, 2002.
- [5] R. Sadus, *Molecular Simulation of Fluids. Applications to Physical Systems*, Elsevier, 1999.
- [6] E. J. Hockney R.W., *Computer Simulation Using Particles*, Bristol: IOP, 1989.
- [7] M. D. a. C. Holm, *J. Chem. Phys*, p. 109, 1998.
- [8] „PDB File Format,” [Interactiv]. Available: <http://www.wwpdb.org/documentation/file-format-content/format33/v3.3.html>.
- [9] „RCSB Protein Data Bank (archive),” [Interactiv]. Available: <https://www.rcsb.org/>.
- [10] M. K. S. Z. G. Griebel, *Numerical Simulation in Molecular Dynamics. Numerics, Algorithms, Parallelization, Applications*, Springer, 2007.
- [11] T. Beu, *Introduction to Numerical Programming: A Practical Guide for Scientists and Engineers Using Python and C/C++*, CRC Press / Taylor % Francis, 2014.
- [12] T. Beu, „MDSquad - Molecular Dynamics made accessible,” 2021. [Interactiv]. Available: <https://phys.ubbcluj.ro/~titus.beu/MDSquad/index.html>.
- [13] T. Beu, „Introduction to Numerical Programming - A Practical Guide for Scientists and Engineers Using Python and C/C++,” 2021. [Interactiv]. Available: <https://phys.ubbcluj.ro/~titus.beu/INP/index.html>.
- [14] S. S. B. Stenberg, „An Exact Ewald Summation Method in Theory and Practice,” *The Journal of Physical Chemistry A*, vol. 124, nr. 19, pp. 3943-3946, 2020.
- [15] J. L. J. Z. S. R. X. C. C.-L. J. D. Wanga, „Ewald summation for ferroelectric perovskites with charges and dipoles,” *Computational Materials Science*, nr. 162, pp. 314-321, 2019.

[16] A. D. M. J. K. Vanommeslaeghe, *J. Comput. Chem.*, nr. 371, p. 31, 2010.