



IKER
GAZTE
NAZIOARTEKO
IKERKETA EUSKARAZ

I. IKERGAZTE

NAZIOARTEKO IKERKETA EUSKARAZ

2015eko maiatzaren 13, 14 eta 15
Durango, Euskal Herria

ANTOLATZAILEA:
Udako Euskal Unibertsitatea (UEU)

ZIENTZIA ZEHATZAK ETA NATUR ZIENTZIAK

**Errendimendu handiko
konputazioa fotosintesia
simulatzeko**

*J. Alberdi-Rodriguez, J. Muguerza
eta A. Arruabarrena eta A. Rubio*

321-327 or.
<https://dx.doi.org/10.26876/ikergazte.i.43>

ANTOLATZAILEA:



BABESLEAK:



LAGUNTZAILEAK:



Errendimendu handiko konputazioa fotosintesia simulatzeko

Alberdi-Rodriguez, J.^{1,2}

Muguerza, J. eta Arruabarrena, A.¹

Rubio, A.^{2,3}

¹ Konputagailuen Arkitektura eta Teknologia saila, Euskal Herriko Unibertsitatea UPV/EHU, M. Lardizabal 1, 20018 Donostia

² Nano-Bio Spectroscopy Group eta ETSF, Euskal Herriko Unibertsitatea CFM CSIC-UPV/EHU-MPC & DIPC, Joxe Mari Korta eraik., Tolosa etorb. 72, 20018 Donostia

³ Max Planck Institute for the Structure and Dynamics of Matter, Hamburg, Germany

Laburpena

Artikulu honek informatikako doktore-tesi bat laburbiltzen du, zeina informatikarien, fisikarien eta kimikarien arteko lankidetzan garatu den. DFT eta TDDFT teoriak gauzatzen dituen OCTOPUS software-paketea erabili da landareetan fotosintesia egiten duen molekularen simulazioak egiteko, Europako superkonputagailu bizkorrenetakoak erabiliz. Helburu hori lortzeko, ordea, konputazioko hainbat arazori aurre egin behar izan zaizkio, azelerazio-faktore egokiak eskuratzeko. Arazo horiek gainditzea lortu da, eta, ondorioz, inoizko TDDFT simulaziorik handienak egin ahal izan dira, milaka prozesadore eraginkorki erabiliz.

Hitz gakoak: konputagailu bidezko simulazioa, errendimendu handiko konputazioa (HPC), mekanika kuantikoa, dentsitatearen funtzionalaren teoria (DFT), denboraren menpeko dentsitatearen funtzionalaren teoria (TDDFT)

Abstract

This paper is a resume of a computer science PhD thesis, which is a collaboration between computer engineers, physicists and chemists. OCTOPUS software package, which implements DFT and TDDFT theories, is used to simulate photosynthesis or the absorption of the light in plants, using the most advanced supercomputers in Europe. However, in order to obtain high speed-ups, we have to previously solve performance problems. Once solved, we have done ever biggest TDDFT calculations, using efficiently thousand of processors.

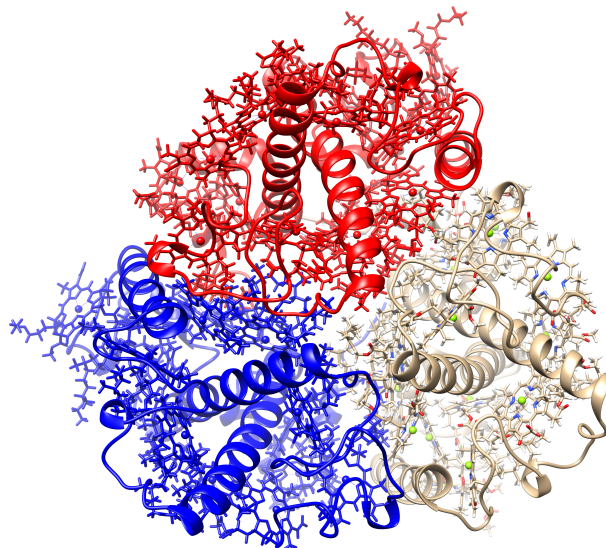
Keywords: Computer simulation, high performance computing (HPC), quantum mechanics, density functional theory (DFT), time-dependent density functional theory (TDDFT)

1 Sarrera eta motibazioa

Informatikaren aurrerapenak medio, zientzialariak eta ingeniariak gauza dira prozesu fisiko oso konplexuen portaera simulazio-tekniken bidez analizatzeko. Izan ere, konputagailu bidezko simulazioak errealtatea aurreikusteko tresna ezinbestekoak dira gaur egungo industria eta ikerketa-laborategietan, problema berrientzat soluzio berritzaileak sortzeko. Eta horretarako prest ditugu superkonputagailuen konputazio-ahalmen itzela, zeinetan ohikoa baita 100.000 prozesadore baino gehiago erabiltzea aplikazio oso konplexuak ebazteko. Zehazki, artikulu honetan fisika kuantikoko simulazioei buruz arituko gara. Teoria jakin batzuetan oinarrituta (aurrerago azalduko ditugun DFT eta TDDFT), fotosintesiaren prozesua aztertu dugu; izan ere, oraindik ez da erabat ulertzen nola xurgatzen duten argia landareek. Hala, lan honen helburua jakintza horretan pauso bat aurrera egitea da, eta, horretarako, OCTOPUS (Andrade *et al.*, 2012, 2015) software-paketea erabili dugu.

Landareen fotosintesiaren arduraduna *Light Harvesting Complex II* (LHC-II) molekula da (ikus 1 irudia). Konplexu erraldoi horrek hiru aldiz errepikatzen den egitura simetrikoa du eta, orotara, 17.000 atomo baino gehiagoz osatua da. Hiru aldeko simetria horrengatik, trimero deitzen zaio eta bere ataletako bakoitzari, monomero. Monomero bakoitzaren zatirik garrantzitsuenak proteinak, karotenoideak (xantofilak eta klorofilak) eta ur molekulak dira.

1 irudia: LHC-II molekula. Landareetan fotosintesia egiten duen molekula. Monomero bakoitza kolore batez adierazita dago, eta hiru monomeroek trimero bat osatzen dute. Lan honen helburua molekula erraldoi hau simulatzea izan da, mekanika kuantikoa erabiliz.



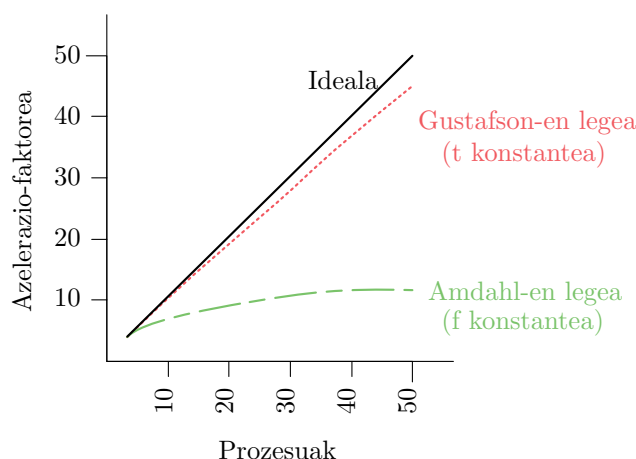
LHC-II bezalako molekula handiak simulatu ahal izateko, bai konputagailu konplexuak eta programa aurreratutak behar dira, eta hori dena biltzen duen terminoa *High Performance Computing* (HPC) da. HPC arloan, azelerazio-faktorea da kontuan izan beharreko neurri garrantzitsuenetariko bat. Oro har, konputazio-baliabideak ugaltu ahala, programen exekuzio-denborak murrizten dira; idealki, prozesadore kopurua bikoiztuta exekuzio-denbora erdira jaitsi beharko litzateke, hau da, programak bi aldiz azkarrago exekutatzeko lortu beharko genuke. Programa bat zenbat aldiz azkarrago exekutatzen den zehazten duen neurri horri azelerazio-faktore edo *speed-up* deritza, eta azelerazio-faktore ideala 2 irudiko marra beltz jarraituan ikus daiteke.

Programa baten exekuzioa hainbat prozesutan banatu daiteke, zeinek programa osoaren zati bana exekutatzen duten. Prozesu horiek aldi berean exekutatzen badira, programa paraleloa dela esaten da. Baina, zoritxarrez, programa guztietan badaude paraleloan exekutatu ezin daitezkeen zatiak, nahitaez seriean (prozesadore bakarrean) exekutatu behar direnak. Serieko zatiek errendimendua mugatzen dute, Amdahl-en legeak adierazten duen moduan (Amdahl, 1967). 2 irudiko beheko marra berdeak adierazten du programa batek lortuko lukeen errendimendua %90 paraleloan eta %10 seriean exekutatu balitz, benetan urria. Baina, zorionez, Gustafson-en legeak muga hori gaindigarria dela iragartzen du (Gustafson, 1988). Zientzialariak beti daude sistema geroz eta handiagoak simulatzeko gogoz, eta, problemen tamaina hazten dugunean, serieko atala bere horretan mantendu ohi da, paraleloan egin daitezkeen lan-karga askoz ere handiagoa izanik. Era horretan, baliabide berriei etekin hobea atera dakieke, 2 irudiko tarteko marra gorriak adierazten duen moduan. Problema erreala bi hurbilpen horien artean egon ohi dira.

Oro har, beraz, prozesadore asko dituzten konputagailuak problema konplexuak eraginkorki ebazteko erabil daitezke; esaterako, 64.000 prozesadoreko makina bat problema jakin bat 64.000 aldiz azkarrago ebazteko (edo gauza bera dena, prozesadore batek zazpi urtetan egingo litzatekeen kalkuluak ordu bakar batean egitea). Lan honetan munduko superkonputagailu azkarrenetako batzuk erabili ditugu; besteak beste: Alemaniako Juqueen konputagailua (458.752 prozesatze-nukleo edo *core*), Italiako Fermi konputagailua (163.840 nukleo), Alemaniako Hydra (65.320 nukleo), Frantziako Curie konputagailua (11.520 nukleo), eta, Espainian, Bartzelonako MareNostrum konputagailua (48.896 nukleo). Juqueen eta Fermi IBMren Blue Gene/Q arkitekturako superkonputagailuak dira, hau da, propio diseinatutako makinak

konputazio-abiadura handienak lortzeko. Hydra, Curie eta MareNostrum, aldiz, cluster motakoak dira, hau da, konputagailu “arrunt” ugari sare bizkor batekin lotutako makinak. Puntako-puntako superkonputagailuei buruzko informazio gehigarria E. Strohmaier, J. Dongarra, H. Simon eta M. Meuer-ek urtean bitan argitaratzen duten TOP500¹ zerrendan aurki daiteke, non, LINPACK (Dongarra eta Luszczek, 2011) proba-bankua exekutatzen munduko superkonputagailu azkarrenak ageri baitira.

2 irudia: Programa paralelo baten aurreikusitako errendimendu teorikoa: Ideala, Amdahl-en legea eta Gustafson-en legea. Zehazki, Amdahl eta Gustafson-en legeen arteko diferentzia ageri da, $f = 0.9$ kasurako (paraleloan exekutatzen den programaren zatia). Argi dago Gustafson-en legearen arabera azelerazio-faktore handiak lor daitezkeela, Amdahl-en aurreikuspen ezkorrekikiko.



2 Arloko egoera eta ikerketaren helburuak

OCTOPUS software-paketea simulazioak *ab-initio* egiteko erabili da, hots, beste inolako hurbilpenik erabili gabe, teoria fisikoetan bakarrik oinarrituta. Funtsean, OCTOPUS kodean *Density Functional Theory* (DFT) (Hohenberg eta Kohn, 1964; Kohn eta Sham, 1965) eta *Time-Dependent Density Functional Theory* (TDDFT) (Runge eta Gross, 1984; Fiolhais *et al.*, 2003; Marques *et al.*, 2012) teoriak erabiltzen dira atomo eta molekulen simulazioak egiteko. Teoria horiek mekanika kuantikoaren birformulazio bat dira, eta bai denborarekiko independente (DFT) nahiz menpeko (TDDFT) izan, dentsitate elektronikoan oinarritzen dira, frogatu ahal izan baita sistema atomiko baten propietate guztiak erator daitezkeela bere dentsitate elektronikoa ebatzita.

Kodean, orokorrean, elektroioak mekanika kuantikoari jarraiki simulatzen dira, eta nukleoa klasikoki tratatzen da. Teoria horiek mekanika kuantikoaren konputazioa ahalbideratu dute, bestela zuzenean ebatzi ezin diren ekuazio konplexuegiak sortzen baitira, baita gaur egungo superkonputagailu bizkorrenarentzat ere. DFT eta TDDFT metodoak iteratiboak dira, hau da, sistema iterazioz iterazio ebazten da, konbergentzia lortu arte. Iterazio bakoitzean funtzio berak exekutatzen dira, eta bukaera erabakitzen da simulazioko beste iterazio bati ekitea ala ez. Exekutatu behar den funtzio horietako bat Poisson ebazlea da, zeinak potentzial elektrostatikoa kalkulatzeko duen dentsitate elektronikotik abiatuta.

Milaka atomoko sistemekin maila kuantikoan simulazioak egiteko gaitasuna izan ezker, fisikako, kimikako eta biologiako fenomeno asko ulertzea posible izango litzateke. Nahiz eta azken urteetan HPC arloan egin diren hobekuntzak handiak izan, TDDFTrekin egin daitezkeen simulazioen tamainak mugatua izaten jarraitzen du, eta erronka handia dira oraindik. Tamalez, hainbat arazo aurkitu dira milaka atomo simulatu ahal izateko, eta arazo horiei nola aurre egin diegun azalduko dugu hurrengo atalean.

¹<http://www.top500.org>

3 Lan-ildoak eta emaitzak

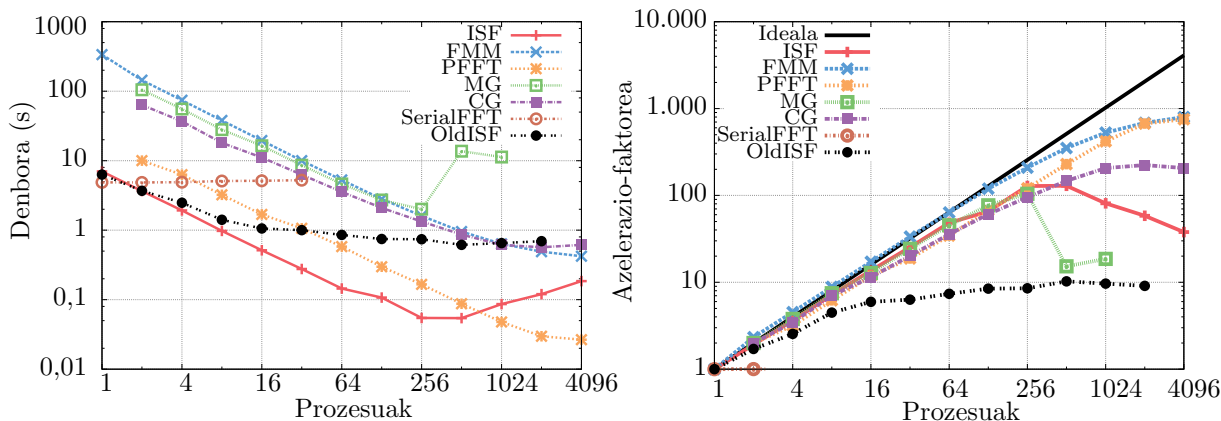
OCTOPUS kodearen errendimendua hobetze aldera, hainbat azterketa egin dira. Hasierako azterketa horien ondorioz, exekuzio paraleloan hainbat hutsune aurkitu ziren simulazioan; aipagarriena, Poisson ebazlearen exekuzio ez optimoa izan zen (Alberdi-Rodriguez, 2010). Funtzio hori ia seriean exekutatzen zen eta, ondorioz, exekuzio-denborak ez ziren murrizten prozesu kopurua handitu ahala. 1 taulan ikus daiteke Poisson funtzioa exekutatze behar zen denborak gero eta pisu handiagoa hartzen zuela prozesu kopurua handitzen zenean (3 irudiak portaera bera berresten du; bai ezkerrean zein eskuinean “OldISF” marra horizontal bihurtzen da). Horregatik, Poisson ebazleak galarazten zuen errendimendu optimoa lortzea prozesu asko erabiltzen zirenean.

1 taula: Poisson funtzioaren exekuzio-denbora, iterazio-denboraren ehuneko gisa, 1365 atomoko sistema baterako, Blue Gene/P makinan exekutatuta.

Prozesuak	512	1024	2048	4096	8192
Iterazioaren exekuzio-denbora (s)	44,89	34,03	19,61	11,74	8,81
Poisson funtzioaren exekuzio-denbora (s)	1,14	1,27	2,72	3,21	4,25
Poisson funtzioaren ehunekoa (%)	3	4	14	27	48

Hori horrela, hainbat aukera aztertu dira Poisson funtzio optimizatuagoa lortzeko, eta, azkenean, Poisson ebazle azkarragoa eta eskalagarriagoa (prozesadore kopuru handietan ongi moldatuko dena) izatea lortu da. Orotara, zazpi metodo aztertu dira, 3 irudian ikus daitekeen bezala. Lau metodo aurrez inplementatuta zeuden: Interpolating Scaling Function zaharra (OldISF), Conjugate Gradients (CG), Multigrid (MG) eta serieko Fast Fourier Transform (serialFFT); eta hiru metodo berri gauzatu dira Poisson ebazle gisa: aurrez lehenetsia zen ISF metodoaren bertsio berria (Genovese *et al.*, 2006), Parallel Fast Fourier Transform (PFFT) (Pippig, 2013) eta Fast Multipole Method (FMM) (Kabadshow eta Dachsel, 2010). Lortutako errendimenduari dagokionez, FMM metodoak, paralelizazio maila ona erakutsi duen arren, exekuzio-denbora altuak behar ditu. PFFT metodoa, aldiz, magnitude-ordena bat azkarragoa da aurrekoa baino, pareko paralelismo maila lortuz. Hots, bi metodo horiek azelerazio-faktore parekoa dute. Bestalde, ISF metodoa aurreko bi metodoak baino azkarragoa da, baina, prozesu asko erabili behar direnean, errendimendua jaitsi egiten da.

3 irudia: Poisson ebazleen exekuzio-denborak eta azelerazio-faktoreak prozesu kopuruaren arabera; 4.019.679 sareta-puntuoko sistema bat Blue Gene/P superkonputagailuan. Prozesu gutxi erabiltzen direnean ISF metodoa da azkarrena. Aldiz, eskalagarritasun hobea dela eta, PFFT azkarrena da prozesu askotan.



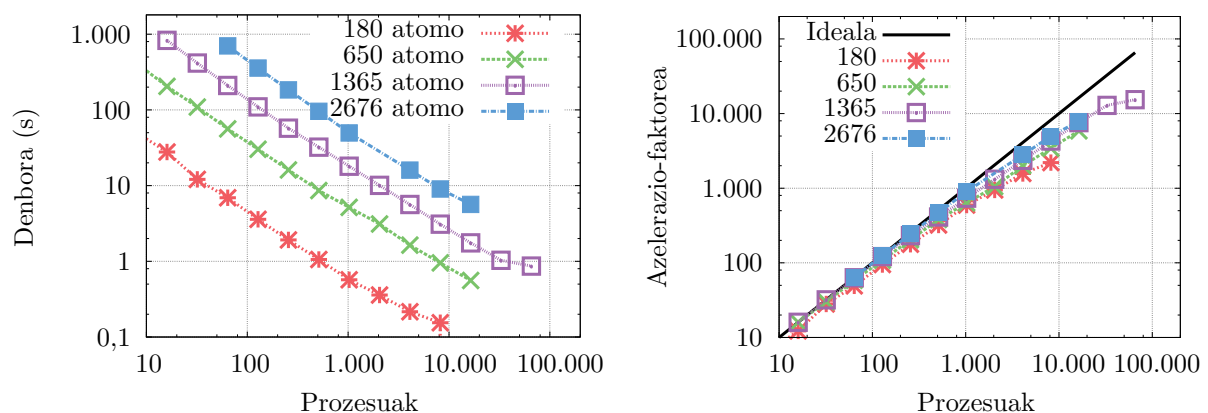
Esfortzu berezia egin da FMM metodoa gure aplikaziora egokitzen. Modu diskretuan lan egiteko pentsatuta dago, eta nolabait jarraitua den dentsitatera egokitu behar izan dugu. Hau da, FMM metodoak sareta bateko puntuetan balio zehatzak jorratzen ditu; aldiz, dentsitate elektronikoa banatuago dago, eta hortik moldaketaren beharra. Alde teknikoari dagokionez, Scafacos (Bolten *et al.*; Arnold *et al.*, 2013) liburutegiari deia egiten zaio eta lortutako emaitzari geuk garatutako zuzenketa aplikatu. Bestalde, ISF eta PFFT metodoak gauzatzerakoan, datu-transferentzia eraginkorrak lortzeko ahalegin berezia egin dugu, sarrerako dentsitate elektronikoa eta irteerako potentzial elektrostatikoa modu desberdinean irudikatatuak baitaude liburutegi horietan eta jatorrizko kodean. Bi sareta desberdin horien artean (eta gainera,

puntuak modu desberdinean banatuak prozesuen artean) datu-mugimendua era azkar eta eskalagarrian gauzatu dugu. ISF metodoa, berez, Poisson ebazlea da eta liburutegiari deia egitea nahikoa da emaitza lortzeko (kontuan izanda aipatu berri dugun datu-transferentzia). PFFT aldiz, FFT liburutegi eraginkor bat besterik ez da, eta Poisson ebazle bezala funtziona dezan bi dei egin behar dira eta, tartean, datuak eskalatu behar dira (aurrez kalkulatuak balioekin biderketa bat aplikatuz).

Azterketa horren ondorioa garbia da, ISF metodoa aukeratu behar da prozesu gutxi erabili behar direnean, eta PFFT metodoa, ordea, prozesu askoren kasuan. Gainera, metodo berriek guztiek aurreko errore maila berdindu edo hobetzen dute: hortaz, errendimenduarekin batera, zehaztasuna mantendu edo irabazi da. Beraz, emaitza garrantzitsua lortu dugu, Poisson ebazleak azelerazio-faktore handiak lortzeko oztopoa ez izatea. García-Risueño *et al.* (2014) egileen lanean sakonago azaltzen dira xehetasun guztiak.

Prozesadoreen erabilera eraginkorraz gain, programek erabiltzen duten memoria kopurua ere oso aintzat hartzeko arloa da. Izan ere, gaur egungo superkonputagailuek milaka edota milioika prozesadore dituzten arren, prozesadore bakoitzak eskura duen memoria kopurua urria izan ohi da (1-8 GB inguru besterik ez). Programak behar duen memoria kopurua simulatu behar den sistema atomikoaren arabera da eta independentea behar luke erabilitako prozesu kopuruarekin. Errealitatean, ordea, posible da horrela ez izatea, eta memoria-beharrak gero eta handiagoak izatea prozesadore kopurua handitu ahala. OCTOPUS kodea prozesadore kopuru handietan exekutatu ahal izateko, hainbat memoria- eta errendimendu-arazo konpondu behar izan dira. Hobekuntza horietako adibide gisa, geuk garatutako sareta-puntuen banaketa berria aipatuko dugu. OCTOPUS-ek espazio errealean egiten ditu simulazioak eta, horretarako, espazioa hainbat puntutan banatzen du, hiru dimentsiotako sareta osatuz. Paraleloan exekutatu nahi denean, sareta-puntuak prozesadoreen artean banatu behar dira. Gure lanaren aurretik, banaketa hori seriean egiten zen eta memoria asko eskatzen zuen. Orain, berriz, banaketa paraleloan egiten da, ParMETIS liburutegia erabiliz (Karypis eta Kumar, 1996). Exekuzioa azkarragoa izateaz gain, memoria gutxiago behar da. Hobekuntza horren eta beste hainbaten informazio zehatza Alberdi-Rodríguez *et al.* (2014) lanean aurkitu daiteke.

4 irudia: 180, 650, 1365 eta 2676 atomoko sistemak Blue Gene/Q superkonputagailuan exekutatuta (prozesu bat PUZ bakoitzean exekutatuta). OCTOPUS kodeak eskalagarritasun handia erakusten du.



Optimizazio guztien ondoren, frogatu ahal izan dugu errendimendu handia lor dezakegula prozesadore asko erabilita, 4 irudian ikus daitekeen moduan. Hainbat proba egin dira LHC-II molekularen zati desberdinekin, hain zuzen ere, 180, 650, 1365 eta 2676 atomoko molekulekin (452.878, 1.282.324, 2.371.809 eta 4.106.680 sareta-puntu hurrenez hurren). Sistema txikiena kodearen errendimendu handiaren erakusle da; 4 prozesutan banatuta exekutatu daiteke eta 8192 prozesura arte exekutatu daiteke eraginkor. Hau da, azelerazio-faktore edo *speed-up* handia lortzen du prozesu askorekin (2205eko azelerazio-faktorea du 8192 prozesurekin). Era berean, sistema handienak portaera bera erakusten du, nahiz eta, jakina, exekuzio-denbora handiagoak behar dituen, simulatu beharreko espazioa (sareta-puntu gehiago ditu) eta kontuan izan beharreko elektroi kopurua handiagoak baitira. Aipagarria da, halaber, 1365 atomoko sistema 65.536 prozesutan exekutatu ahal izan dugula, errendimendu handia mantenduz. Tarteko 650 atomoko sistemak ere aurrekoen portaera bera du. Erakusten diren emaitzak Blue Gene/Q superkonputagailuan lortu dira; bestelako arkitekturetan ere pareko errendimenduak lortu ditugu.

OCTOPUS kodea optimizatu eta errendimendu handia frogatu ondoren, proiektuaren amaieran fisikaren ikuspuntutik interesgarriak diren simulazioak egin ahal izan ditugu, lehen aldiz. Nahiz eta LHC-II molekula osorik exekutatzeko erronka bat izaten jarraitzen duen, monomero osatua, eta, dimero eta trimero sinplifikatuak (proteinarik gabekoak) simulatzeko lortu dugu. Guk dakigula, sistema horiek dituzten 5759, 4050 eta 6075 atomo kopururekin sekula egin diren simulaziorik handienak dira. Simulazioetan frogatu ahal izan dugu teoria bat datorrela errealitatearekin; hots, simulazioan eta errealitatean molekula hauen xurgapen-espektroak baliokideak direla.

4 Ondorioak

Proiektu honetan kode zientifiko baten optimizazioan lan egin dugu. Nagusiki alderdi informatiko batetik landu den arren, proiektua jakintza-arlo desberdinen arteko lankidetzaren emaitza da: informatikarien, fisikarien eta kimikarien artekoa hain zuzen ere. OCTOPUS software-paketea sistema atomikoen portaera simulatzeko tresna aurreratua da eta mekanika kuantikoko teoriak (DFT eta TDDFT) erabiltzen ditu. Zoritxarrez, tamaina errealeko sistemak simulatzeko, exekuzio-denbora oso handiak behar dira. Beraz, irtenbide bakarra superkonputagailuak erabiltzea da. Makina horiek dituzten ehunka milako prozesadoreetan errendimendu handia lortzea prozesu konplexua da, eta hori izan da proiektu honen helburua. Hori horrela, hainbat arlotan jorratu da optimizazioa. Batetik, Poisson ebazlearen errendimendu mugatu gainditu dugu, aukera berri eta eraginkorrak gauzatu. Bestetik, memoria erabilera, datu-banaketa eta beste hainbat atal ere optimizatu eta birmoldatu ditugu.

Hobekuntza horien guztien ondorioz, inoiz egin diren DFT eta TDDFT simulazio handienak egitea lortu dugu, tamaina errealeko simulazioak ahalbidetatuz. Simulazio horiek aukera emango digute, lehen aldiz, fotosintesian gertatzen diren lehenengo femtosegundoetako erreakzioak ulertzeko.

5 Etorkizunerako planteatzen den norabidea

LHC-II molekularen simulazioa egiteko gai garestia, simulazioen emaitzak, fisikako/kimikako alderditik aztertu behar dira orain. Oinarritzko egoera bakar batetik abiatuta, kitzikapen desberdinak aplikatuko dizkiogu sistemari, eta bere portaera ulertzen saiatuko gara. Lan honek guztiak denbora beharko du simulatu eta aztertu ahal izateko. Izan ere, simulazio horiek aurrera atera ahal izateko milaka prozesadore erabili behar dira, hainbat egunez gainera. Gerora, fisika berri bat aztertzen ari garela kontuan izanda, lan nekeza da zer aurkituko dugun aurreikusten eta aukera berrietara irekita egon beharko dugu.

Era berean, aplikazioan egin diren hobekuntzek atea ireki dute tamaina horretako beste hainbat sistema simulatu ahal izateko. Software librea izaki, fisikarien eskura dago kodea, eta ziur gaude beste hainbat fenomeno ulertzeko baliagarria izango dela.

Erreferentziak

Alberdi-Rodriguez, Joseba. 2010. *Analysis of performance and scaling of the scientific code Octopus*. LAP LAMBERT Academic Publishing.

Alberdi-Rodriguez, Joseba, Micael J.T. Oliveira, PABLO García-Risueño, Fernando Nogueira, Javier Nogueira, Javier Muguerza, Agustin Arruabarrena, eta Angel Rubio. 2014. Recent memory and performance improvements in Octopus code. In *Computational Science and Its Applications ICC- SA 2014*, ed. by Beniamino Murgante, Sanjay Misra, Ana Maria A.C. Rocha, Carmelo Torre, Jorge Gustavo Rocha, Maria Irene Falco, David Taniar, Bernady O. Apduhan, eta Osvaldo Gervasi, volume 8582 of *Lecture Notes in Computer Science*, 607–622. Springer International Publishing.

Amdahl, Gene M. 1967. Validity of the single processor approach to achieving large scale computing capabilities. In *AFIPS spring joint computer conference*, volume 30, 483–485. AFIPS Press.

Andrade, X., J. Alberdi-Rodriguez, D.A. Strubbe, M.J.T. Oliveira, F. Nogueira, A. Castro, J. Muguerza, A. Arruabarrena, S.G. Louie, A. Aspuru-Guzik, eta others. 2012. Time-dependent density-functional theory in massively parallel computer architectures: the octopus project. *Journal of Physics: Condensed Matter* 24.233202.

Andrade, Xavier, David Strubbe, Umberto De Giovannini, Aak Hjorth Larsen, Micael J. T. Oliveira, Joseba Alberdi-Rodriguez, Alejandro Varas, Iris Theophilou, Nicole Helbig, Matthie J. Vestraete, Lorenzo Stella,

- Fernando Nogueira, Alan Aaspuru-Guzik, Alberto Castro, Miguel A. L. Marques, eta Angel Rubio. 2015. Real-space grids and the octopus code as tools for the development of new simulation approaches for electronic systems. *Phys. Chem. Chem. Phys.* –.
- Arnold, Axel, Florian Fahrenberger, Chistian Holm, Olaf Lenz, Matthias Bolten, Holger Dachsel, Rene Halver, Ivo Kabadshow, Franz Gähler, Frederik Heber, Julian Iseringhausen, Michael Hofmann, Michael Pippig, Daniel Potts, eta Godehard Sutmann. 2013. Comparison of scalable fast methods for longrange interactions. *Phys. Rev. E* 88.063308.
- Bolten, M., F. Fahrenberger, R. Halver, F. Heber, M. Hofmann, I. Kabadshow, O. Lenz, M. Pippig, eta G. Sutmann. ScaFaCoS, C subroutine library. <http://scafacos.github.com/>.
- Dongarra, Jack, eta Piotr Luszek. 2011. Linpack benchmark. In *Encyclopedia of Parallel Computing*, ed. by David Padua, 1033–1036. Springer US.
- Fiolhais, C., F. Nogueira, eta M. A. L. Marques (eds.) 2003. *A Primer in Density Functional Theory*. Springer, Series: Lecture Notes in Physics, Vol. 620, 1st edition.
- García-Risueno, P., J. Alberdi-Rodriguez, M. J. T. Oliveira, X. Andrade, M. Pippig, J. Muguerza, A. ARRUBARRENA, eta A. RUBIO. 2014. A survey of the performance of classical potential solvers for charge distributions. *Journal of Computational Chemistry* 35.427–444.
- Genovese, L., T. Deutsch, A. Neelov, S. Goedecker, eta G. Beylkin. 2006. Efficient solution of Poisson's equation with free boundary conditions. *J. Chem. Phys.* 125.074105. Genovese, L. and Deutsch, T. and Neelov, A. and Goedecker, S. and Beylkin, G.
- Gustafson, John L. 1988. Reevaluating Amdahl's law. *Communications of the ACM* 31.532–533.
- Hohenberg, P., eta W. Kohn. 1964. Inhomogeneous Electron Gas. *Phys. Rev.* 136.B864–B871.
- Kabadshow, I., eta H. Dachsel. 2010. The Error-Controlled Fast Multipole Method for Open and Periodic Boundary Conditions. In *Fast Methods for Long-Range Interactions in Complex Systems*, IAS Series, Volume 6, Forschungszentrum Jülich, Germany. CECAM.
- Karypis, George, eta Vipin Kumar. 1996. Parallel multilevel k-way partitioning scheme for irregular graphs. In *Proceedings of the 1996 ACM/IEEE conference on Supercomputing (CDROM)*, Supercomputing '96, Washington, DC, USA. IEEE Computer Society.
- Kohn, W., eta L. J. Sham. 1965. Self-Consistent Equations Including Exchange and Correlation effects. *Phys. Rev.* 140.A1133–A1138.
- Marques, Miguel A.L., Neepa T. Maitra, eta Fernando M.S. Nogueira. 2012. *Fundamentals of Time-Dependent Density Functional Theory*, volume 837. Springer Berlin Heidelberg.
- Pippig, Michael. 2013. PFFT - An extension of FFTW to massively parallel architectures. *SIAM J. Sci. Comput.* 35.C213 – C236.
- Runge, Erich, eta E. K. U. Gross. 1984. Density-functional theory for time-dependent systems. *Phys. Rev. Lett.* 52.997–1000.

6 Eskerrak eta oharrak

Proiektu hau gauzatzeko nahitaezkoa izan da M.J.T. Oliveira, J. Jornet-Somoza, B. Milne, P. García-Risueño, F. Nogueira eta X. Andrade-ren ekarpena eta laguntza. Lan hau Eusko Jaurlaritzako Hezkuntza, Unibertsitate eta Ikerkuntza sailak (IT-395-10 eta IT578-13 proiektuak), European Research Council Advanced Grant DYName (ERC-2010-AdG-267374), European Community FP7 project CRONOS (grant number 280879-2), COST Actions CM1204 (XLIC), MP1306 (EUSpec) eta Spanish Grant (FIS2013-46159-C3-1-P) proiektuek finantzatu dute. JAR-ek Euskal Herriko Unibertsitatearen beka jaso du bere doktore-tesia egin ahal izateko. Simulazioak aurrera atera ahal izateko, Red Española de Supercomputación (RES), Rechenzentrum Garching (RZG) eta Partnership for Advanced Computing in Europe (PRACE) erakundeek konputazio-orduak erabili ditugu.