

Proceedings

Sixteenth ACC Cyfronet AGH HPC Users' Conference



Editors: Kazimierz Wiatr
Jacek Kitowski
Marian Bubak

Zakopane, 13-15 March 2024

Proceedings

Sixteenth
ACC Cyfronet AGH
HPC Users'
Conference

Zakopane
13-15 March 2024

Editors: Kazimierz Wiatr
Jacek Kitowski
Marian Bubak

Published in March 2024

by Academic Computer Centre Cyfronet AGH
Nawojki 11, 30-950 Kraków, P.O.Box 6, Poland

© The Authors mentioned in the Table of Contents

All rights reserved. This book or part thereof, may not be reproduced in any form or by any means electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission of the Authors and Publisher.

ISBN 978-83-61433-47-7

Cover design and book typesetting by Joanna Kasina.

Ladies and Gentlemen!

This year's High Performance Computers Users' Conference – KU KDM 2024 – is the sixteenth of our meetings bringing together scientists conducting research using Cyfronet's supercomputers.

Recent years have brought intense changes, primarily in the field of technology. Hence, Cyfronet, while continuing its mission to provide IT support for the scientific community, constantly develops its competencies. We maintain a permanent dialogue with representatives of science and with companies offering high-end solutions regarding software and IT hardware. The result is an attempt to respond to the ever-growing demand of science and innovative economy for computing power, memory resources and specialised software. The expanded and modernised IT infrastructure of Cyfronet is to serve this purpose, and that is why, at the end of 2023, we installed two high-end supercomputers at Cyfronet. They are to enable our Users to conduct research with even higher-level IT support than ever before.

The Helios supercomputer is a new system installed at ACC Cyfronet AGH, created as a result of work performed in the *National Supercomputing Infrastructure for EuroHPC - EuroHPC PL* project coordinated by Cyfronet. The supercomputer was built according to Cyfronet's design by Hewlett-Packard Enterprise based on the HPE Cray EX4000 platform. It consists of three computing partitions:

- CPU equipped with 75 264 AMD Zen4 computing cores and 200 TB of DDR5 RAM,
- GPU equipped with 440 NVIDIA Grace Hopper GH200 Superchips,
- INT for interactive work, equipped with 24 NVIDIA H100 accelerators and fast local NVMe memory.

The Helios disk subsystem consists of two types of Lustre file systems: scratch with a capacity of 1.5 PB and a speed of over 1.8 TB/s and project with a capacity of 16 PB and a speed of almost 200 GB/s. All supercomputer components are connected via the Slingshot network with a speed of 200 Gb/s. The HPE Cray EX4000 platform, based on which Helios was built, is also used in constructing the fastest supercomputers in the world (Frontier) and Europe (LUMI). Thanks to direct liquid cooling of the CPU and GPU partitions, it is possible to achieve a very low PUE (Power Usage Effectiveness) for the system, increasing its energy efficiency and reducing operating costs. Additionally, thanks to the recovery of waste heat produced by Helios, it will be possible to use it for heating. Helios reaches 35 PetaFlops of theoretical computing power, and thus, it is the fastest supercomputer in the history of Poland. Regarding computing for AI purposes, this power amounts to as much as 1.8 ExaFlops, which equals 1 800 PetaFlops.

The Faeton supercomputer is a cluster of future technologies that will be used to verify the applicability of new technologies, especially SCM (Storage Class Memory), in computing applications. Faeton consists of 64 compute servers, each equipped with two Intel Xeon Platinum 8352s processors supporting application memory encryption, 1 TB of RAM and 100 Gbps Ethernet low-latency network adapters. In addition, the four compute servers are equipped with 8 TB of Intel Optane SCM memory. Faeton also includes service servers and 12 disk servers, offering more than 1 PB of NVMe disk storage and 12 TB of SCM memory. This configuration provides an excellent environment for developing innovative software,

especially in data analytics (data science) and running applications in a high-performance, high-security cloud environment.

Still attractive to Users is the Athena supercomputer, which was the fastest supercomputer in Poland until the complete installation of Helios. Athena provides the Polish scientific community and innovative economy with computing resources based on the latest processors, GPGPU accelerators, and the essential data storage sub-system based on high-speed flash memory. Athena's configuration covers servers with AMD EPYC CPUs and NVIDIA A100 GPGPU cards. The internal communication is provided by the Infiniband HDR (200 Gb/s) network. Athena's theoretical computing power is over 7.7 PetaFlops, which allowed the computer to appear in 155th place in the TOP500 list of the world's fastest supercomputers in November 2023. Athena creates extraordinary capabilities if computing for AI is considered, as it offers 240 PetaFlops for such applications. An excellent computing power-to-electricity consumption ratio also characterises the supercomputer. That was underlined in November 2023 when Athena took 22nd place in the Green500 list of the most energy-efficient supercomputers in the world (in June 2022, it was very high 9th place).

Recent Green500 lists also included another Cyfronet supercomputer – Ares, which has occupied a relatively high 114th place. Ares is built out of a set of computing servers with a computing power of over 3.5 PetaFlops for a CPU partition and 0.5 PetaFlops for the GPU one and a disc sub-system with a capacity of over 11 PB. The supercomputer has 37 824 computing cores (Intel Xeon Platinum processors) and 147.7 TB RAM. It is equipped with computationally efficient NVIDIA Tesla V100 GPGPU cards. Thanks to the innovative cooling system, the heat removed from Ares is used to heating. In the TOP500 list in November 2023, Ares occupied 404th place.

The intensity of Users' use of supercomputers is evidenced by the fact that the sum of computing tasks performed in just one year on Cyfronet's supercomputers reaches several million. However, the constantly growing demand for computing resources by Polish science drives us to develop these resources continuously. Simultaneously, we work on the facilitation of access to high-end technologies. In this light, it was essential to launch a new PLGrid User Portal earlier this year, in which, first of all, the level of security was raised, and the graphical interface was changed to a more user-friendly one.

Cyfronet is also a part of numerous essential initiatives, both on Polish and European grounds, creating new opportunities for science and society. As a coordinator of the PLGrid Consortium, it plays the leading role in two initiatives included in the Polish Roadmap for Research Infrastructures (PMIB). Within the *National Supercomputing Infrastructure for EuroHPC – EuroHPC PL*, the first stage was completed, namely the *EuroHPC PL* project, and Cyfronet functioned as a project coordinator and Consortium leader. Infrastructure and IT tools delivered thanks to EuroHPC PL allow research and development work to be conducted at a level inaccessible before. In turn, within the *National PLGrid Cloud Infrastructure for EOSC*, Cyfronet co-creates the project of the globally accessible and multidisciplinary environment, in which scientists and the innovative economy as well as the whole society will be able to conveniently publish and benefit from the data, tools and results of scientific papers for research, innovation and education.

One recent significant task has been the coordination of the Polish activity in the LUMI consortium. That effort aims to make the LUMI supercomputer, which since 2022 has been the

fastest computer in Europe and the third in the world by the TOP500 list, accessible for Polish scientists. Access for domestic researchers can be awarded via the PLGrid Portal under competitions organised by Cyfronet.

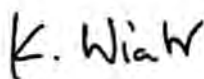
The employees of our Centre aim to search for and implement the best possible solutions to serve you – the Users of our computational and storage resources and software. That requires constant information exchange. That is why the Conference of the Users is an excellent opportunity not only to present and recognise the results of research conducted with the Cyfronet's IT support. It is also a chance for a direct debate of Users and System Administrators, and to identify new challenges of science, to face and resolve them together. The conference will focus on large-scale calculations and simulations, parallel analysis of big data, utilisation of AI methods, new algorithms in IT, and tools and techniques in HPC systems.

For over 50 years, Cyfronet has been an integral part of the ecosystem of Polish and international science. Thus, our work comes with great responsibility. With this in mind, we carefully listen to the Users' needs. We hope that KU KDM 2024 will provide us with numerous opportunities for flourishing discussions between researchers working in various scientific fields and implementing varied and interesting scientific methods. We are convinced that such an exchange of experiences will allow for a broad, interdisciplinary and multidisciplinary perspective on current research directions.

I would like to thank all the Users for their cooperation in improving our services and creating new directions for our development! I cordially invite you to participate in the Conference. I wish all the participants had fruitful discussions and new research ideas.

With best regards,

Kazimierz Wiatr

A handwritten signature in black ink that reads "K. Wiatr". The signature is written in a cursive, slightly slanted style.

Cyfronet Director

Organization

KU KDM 2024 was organized by the Academic Computer Centre Cyfronet AGH, Nawojki 11, 30-950 Kraków, Poland.

Organizing Committee

Kazimierz Wiatr, Joanna Kasina, Kamil Mucha, Robert Pająk

KU KDM'24 Program Committee

Kazimierz Wiatr (chairman)	Academic Computer Centre Cyfronet AGH
Krzysztof Boryczko	Institute of Computer Science AGH
Marian Bubak	Institute of Computer Science AGH / Academic Computer Centre Cyfronet AGH
Joanna Dulińska	Cracow University of Technology
Łukasz Dutka	Academic Computer Centre Cyfronet AGH
Armen Edigarian	Jagiellonian University
Andrzej Eilmes	Jagiellonian University
Marek Gorgoń	Department of Automatics and Robotics AGH
Ernest Jamro	Institute of Electronics AGH / Academic Computer Centre Cyfronet AGH
Zbigniew Kąkol	Department of Solid State Physics AGH
Marek Kisiel-Dorohinicki	Institute of Computer Science AGH
Jacek Kitowski	Institute of Computer Science AGH / Academic Computer Centre Cyfronet AGH
Antoni Ligęza	Department of Applied Computer Science AGH
Grzegorz Mazur	Jagiellonian University
Janusz Mrozek	Jagiellonian University
Janusz Orkisz	Cracow University of Technology
Łukasz Rauch	Department of Applied Computer Science and Modelling AGH
Irena Roterman-Konieczna	Jagiellonian University Medical College
Paweł Russek	Institute of Electronics AGH / Academic Computer Centre Cyfronet AGH
Marek Skomorowski	Jagiellonian University
Renata Słota	Institute of Computer Science AGH
Marek Stanuszek	Cracow University of Technology
Mariusz Sterzel	Academic Computer Centre Cyfronet AGH
Tomasz Szmuc	Department of Applied Computer Science AGH
Piotr Tworzewski	Jagiellonian University

Sponsoring Institutions

Hewlett Packard Enterprise
Advanced Micro Devices
Intel

DataDirect Networks
Lenovo Global Technology Poland

Table of Contents

Invited Talks

Stepping along the Stabilization Pathways in Hybrid Perovskites: Insights from Neutron Spectroscopy & Materials Modeling 11
Kacper Druzbicki

Stationary States in Fixed Potentials 12
Bartłomiej Dybiec

Contributed Papers

Theoretical Studies of the Propylene Oxide Opening Reaction Mechanism in the Copolymerization of CO₂ with Epoxides Process Catalyzed by a Cobalt(III) Complex with Salen Ligand: Unraveling the Computational Complexity 13
O. Żurowska, A. Michalak

A Study of the Properties of Unsymmetrically Substituted 4,5-dialkoxy-2-nitroanilines by Spectroscopic and Theoretical Methods 15
K. Synowiec, M. Boczar, J. Grolik

Investigating the Mechanical Characteristics of Al₂O₃ through Density Functional Theory and Molecular Dynamics 17
M. Fathalian, E. Postek, M. Tahani, T. Sadowski

Multi-scale Modeling of Texture Evolution in Deformed Metals 19
B. Sulkowski

Studies on the Interaction of Nanoplastics with Lipid Systems 21
F. Włodek, W. Kulig, J. Korchowicz, A. Stachowicz-Kuśnierz

Theoretical Study of Copper Complexes in Artificial Photosynthesis 23
A. Miklas, M. Z. Brela

Hydrolysis as an Alternative Path of Nylon 6 and PHB Degradation: a Computational Study 25
Y. Didovets, M. Z. Brela

A Detail That Matters? Rotational Isomerism in Some Ortho-substituted Benzophenones 27
A. J. Kalka, M. Z. Brela, A. M. Turek

The System for Automatic Skin Cancer Detection – Achievements and Challenges	29
<i>R. Frączek, M. Karwatowski, J. Grzeszczyk, J. Caputa, P. Pindel, D. Łukasik, M. Wielgosz, P. Russek, A. Dąbrowska-Boruch, E. Jamro, M. Pietroń, S. Koryciak, K. Wiatr</i>	
Transcriptomics Atlas Pipeline: HTC with Salmon	31
<i>P. Kica, S. Licholai, M. Malawski</i>	
Simulation of Neurotransmitter Flow in a Presynaptic Bouton of a Neuron with Cellular Automaton	33
<i>A. Bielecki, M. Gierdziewicz</i>	
Magnetic Resonance T_1 - and T_2 -weighted Data Synthesis from Heterogeneous Sources: a Federated Learning Approach	35
<i>J. Fiszer, D. Ciupek, M. Malawski, T. Pięciak</i>	
Deep Learning Classification of Cytology Images with Uncertain Training Data	37
<i>J. Krupiński, S. Mazurek, K. Strzałka, M. Wielgosz, J. Caputa, R. Frączek, M. Karwatowski, J. Grzeszczyk, D. Łukasik, P. Russek, A. Dąbrowska-Boruch, E. Jamro, M. Pietroń, S. Koryciak, K. Wiatr</i>	
Comparison of One and Two-Step Deep Learning Models for Cytology Image Classification	39
<i>J. Krupiński, S. Mazurek, K. Strzałka, M. Wielgosz, J. Caputa, R. Frączek, M. Karwatowski, J. Grzeszczyk, P. Russek, A. Dąbrowska-Boruch, E. Jamro, M. Pietroń, S. Koryciak, K. Wiatr</i>	
Integrating Data Repositories with HPC Resources for Execution of Simulation Models	41
<i>T. Zhyhulin, K. Zajac, P. Nowakowski, M. Malawski, M. Kasztelnik, P. Poleć, M. Bubak</i>	
Research towards Virtual Human Twins	43
<i>M. Bubak, M. Kasztelnik, M. Malawski, J. Meizner, P. Nowakowski, P. Poleć</i>	
Numerical Simulations of Short and Long Gamma-Ray Bursts on the Ares Supercomputer	45
<i>G. Urrutia, A. Janiuk</i>	
Viscous Torque in Turbulent Magnetized AGN Accretion Disks and Its Effects on EMRI's Gravitational Waves	47
<i>A. Janiuk, F. H. Nouri</i>	
Analysis of Accelerated Particles in PIC Simulations by Using Neural Networks	49
<i>G. Torralba Paz, A. Bohdan, J. Niemiec</i>	

Particle-In-Cell Simulations of Merger Shocks in Galaxy Clusters	51
<i>O. Kobzar, J. Niemiec, T. Amano, M. Hoshino, S. Matsukiyo, Y. Matsumoto, M. Pohl</i>	
Performance of LLAMA Large Language Model in Argument Detection	53
<i>M. Karwatowski, M. Pietroń, K. Faber, D. Żurek</i>	
Performance of Benchmarking Continual Curriculum Learning	55
<i>D. Żurek, M. Pietroń, K. Faber, M. Karwatowski</i>	
Comparison of GPU and CPU Implementations of New Variants of SDLS Algorithms for LABS Problem	57
<i>D. Żurek, M. Pietroń, K. Faber, M. Karwatowski, K. Piętak</i>	
Ab Initio Molecular Dynamics Simulations of Aqueous LiTFSI Solutions – Structure, Hydrogen Bonding, and IR Spectra	59
<i>A. Eilmes, P. Wróbel, P. Kubisiak</i>	
Interaction Spheres for Molecular Interactions – Application in Drug Design	61
<i>R. Kurczab</i>	
Dynamic Disulfide Bonds in Molecular Dynamics Simulation	63
<i>P. Smardz, P. Krupa</i>	
Application of Fragment Molecular Orbital Method to Investigate G Protein Coupled Receptors	65
<i>P. Śliwa</i>	
Theoretical Description of Peptide Bonds in Selected Systems Based on the ETS-NOCV Method	67
<i>M. Rózga, A. Michalak</i>	
Analytical Inversion Regional Modeling System for Estimation of National-Scale CO ₂ Fluxes – Results of the CoCO ₂ Project	69
<i>M. Galkowski, P. Sekula, M. Zimnoch</i>	
Big Data in Proton Beam Monitoring: LGAD Detectors Study	71
<i>L. Grzanka, N. Minafra, R. McNulty, T. Nowak, J. Swakoń, P. Rzeźnik</i>	
Pedestrian Crossing Intention Detection within Challenging Road Conditions	73
<i>S. Mazurek, M. Sakhai, M. Wielgosz</i>	
Author Index	75

Stepping along the Stabilization Pathways in Hybrid Perovskites: Insights from Neutron Spectroscopy & Materials Modeling

Kacper Druzbicki^{1,2}

¹ Centre of Molecular and Macromolecular Studies, Polish Academy of Sciences

² Materials Physics Center, CSIC-UPV/EHU

Hybrid Organic-Inorganic Perovskites (HOIPs) have emerged as a promising and top-rated class of materials for efficient solar energy conversion and a growing range of optoelectronic applications, including light-emitting diodes and photodetectors [1]. To date, many studies aimed at determining the structural properties of HOIPs using diffraction and optical spectroscopy continue to reveal evident inconsistencies and challenges [2-3]. To a significant extent, these difficulties can be traced back to the extraordinary softness of the metal-halide framework [4], which results in pronounced deformations of the inorganic scaffold and the emergence of nano-size domains of different origins [5-6]. As a result, a consistent understanding of these materials at the atomic scale has yet to be reached.

This situation is particularly severe in the case of the quintessential HOIP methylammonium lead iodide (MAPbI₃), where the ferroelastic relations between thermodynamically stable phases result in intrinsic crystal twinning, severely impeding a thorough understanding of its structure using diffraction techniques. In this contribution, we present our recent efforts to explore the local structure around the organic cations, exploiting the exceptional sensitivity of Inelastic Neutron Scattering (INS) to hydrogen motions [7].

We also provide an overview of our methodology, combining broadband high-resolution INS experiments with state-of-the-art ab initio molecular dynamics beyond the harmonic picture to interrogate local structure in the archetypal case of MAPbI₃ across its pressure-temperature phase diagram and under chemical pressure-induced by cation engineering [8-12].

References

1. S. Brittman, *MRS Commun.*, 2015, 5, 7.
2. K. Druzbicki et al., *Polymers*, 2021, 13, 1440.
3. W. Kong et al., *Phys. Chem. Chem. Phys.*, 2015, 17, 16405.
4. A. C. Ferreira et al., *Phys. Rev. Lett.*, 2018, 121, 085502.
5. J. Breternitz, *Z. Kristallogr. – Cryst. Mater.*, 2022, 237, 135.
6. F. Ambrosio et al., *Z. Kristallogr. – Cryst. Mater.*, 2022, 237, 135.
7. E. M. Mozur et al., *Annu. Rev. Mater. Res.*, 2021, 51, 269.
8. K. Druzbicki et al., *J. Phys. Chem. Lett.*, 2016, 7, 4701.
9. K. Druzbicki et al., *J. Phys. Chem. Lett.*, 2021, 12, 3503.
10. P. Marín-Villa et al., *J. Phys. Chem. Lett.*, 2022, 13, 8422.
11. K. Druzbicki et al., *Cryst. Growth Des.*, 2023, 24, 391.
12. K. Druzbicki et al., *J. Phys. Chem. Lett.*, 2024, in preparation.

Stationary States in Fixed Potentials

Bartłomiej Dybiec

Institute of Theoretical Physics, Jagiellonian University

Many situations in natural sciences can be successfully investigated adopting the stochastic level of description, considering the system of interest as a dynamical system responding to external perturbations represented by a noise. In the simplest situations this noise is assumed to be white and Gaussian. The white type of the noise is the consequence of the large number of independent interactions bounded in time. Its Gaussian character arises due to the assumption that the interactions are bounded in their strength. In many far from equilibrium cases the second assumption fails; the interactions still can be of the white type (i.e. independent) but described by heavy-tailed distributions frequently of the alpha-stable Lévy type. Such heavy-tailed fluctuations, among others, are abundant in heartbeat dynamics, neural networks, search on a folding polymers, animal movement, financial time series, and even in spreading of diseases and dispersal of banknotes.

Under the action of the Gaussian white noise, the form of a stationary state reflects the shape of the potential, because stationary states are of the Boltzmann–Gibbs type. Consequently, the number of modes in the stationary state is the same as the number of minima of the potential. The situation is very different in the non-equilibrium regime, e.g., under action of the Lévy noise. Non-equilibrium stationary states of overdamped anharmonic stochastic oscillators driven by Lévy noise are typically multimodal. The very same situation is recorded for an underdamped Lévy noise-driven motion in single-well potentials with linear friction. Using computer simulations the underdamped motions in fixed, single-well potentials in the regime of nonlinear friction are studied. It is relatively easy to produce multimodality in the velocity distribution as it is determined by the friction itself and it is the same as the multimodality in the overdamped case with the analogous deterministic force. Contrary to the velocity marginal density, it is more difficult to induce multimodality in the position. Furthermore, modality of the non-equilibrium stationary states can be affected by the shape of the potential or stochastic resetting. Stochastic resetting is a motif that links diffusive motion and search strategies as starting anew can be used to optimize the time needed to find a target. The interplay between deterministic forces and stochastic restarting can further modify the shape of stationary states, e.g., it can produce additional or destroy existing modes.

Exploration of stochastic dynamics in fixed potentials was performed by methods of stochastic dynamics. Computer simulations have been performed using high-performance computing infrastructure PLGrid (HPC Centers: ACC Cyfronet AGH) within computational grants no. PLG/2018/011464, PLG/2019/012543, PLG/2020/013656, PLG/2021/014682 and PLG/2023/016175.

References

1. I. M. Sokolov, B. Dybiec and W. Ebeling, Harmonic oscillator under Lévy noise: Unexpected properties in the phase space, *Phys. Rev. E* 83, 041118 (2011).
2. K. Capała and B. Dybiec, Stationary states for underdamped anharmonic oscillators driven by Cauchy noise, *Chaos* 29, 093113 (2019).
3. M. Cieśla, K. Capała and B. Dybiec, Multimodal stationary states under Cauchy noise, *Phys. Rev. E* 99, 052118 (2019).
4. K. Capała, B. Dybiec and E. Gudowska-Nowak, Nonlinear friction in underdamped anharmonic stochastic oscillators, *Chaos* 30, 073140 (2020).
5. P. Pogorzelec and B. Dybiec, Resetting induced multimodality, *Chaos* 33, 063105 (2023).

Theoretical Studies of the Propylene Oxide Opening Reaction Mechanism in the Copolymerization of CO₂ with Epoxides Process Catalyzed by a Cobalt(III) Complex with Salen Ligand: Unraveling the Computational Complexity

Olga Żurowska^{1,2}, Artur Michalak¹

¹ Jagiellonian University, Quantum Chemistry Group Institution, Gronostajowa St 2, Kraków Poland

² Jagiellonian University, Doctoral School of Exact and Natural Sciences, Prof. St. Łojasiewicza St 11, Kraków, Poland

olga.zurowska@doctoral.uj.edu.pl

Keywords: bifunctional catalyst, epoxide/CO₂ copolymerization, effective reaction rate constant, reaction pathways, RPO/SPO isomers, catalytic activity

1. Introduction

Navigating the intricacies of predicting catalyst activity in computational studies poses significant challenges. Considerations extend beyond energy barriers to include the populations of alternative isomers, playing a pivotal role in shaping reaction pathways. Despite advancements in computational chemistry, modeling the catalytic processes of these complexes remains challenging in computational chemistry, given the multitude of isomers, various chain conformations, and intricate spatial arrangements, incurring significant computational costs.

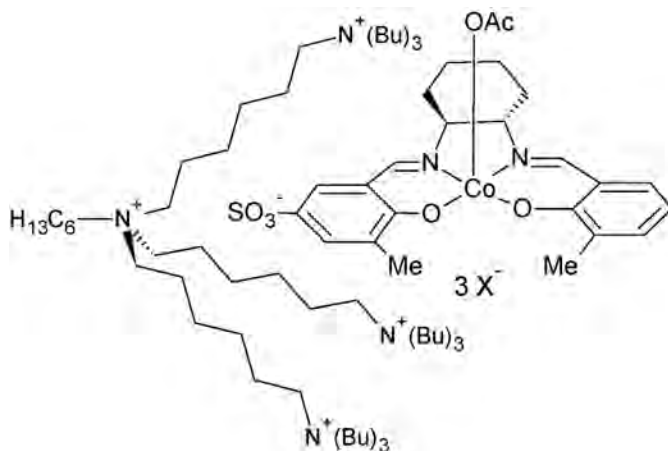


Fig. 1. Novel CO₂/PO copolymerization bifunctional catalyst with a sulfonic group.

2. Description of the problem

Synthesizing eco-friendly polypropylene carbonate through carbon dioxide and propylene oxide copolymerization shows promise. Cobalt(III)-salen catalysts, especially binary and bifunctional types, exhibit remarkable activity [1], [2]. Scalability challenges hinder large-scale production. This study explores a novel Co(III)-salen bifunctional catalyst with a sulfonic group (Fig 1.) [3], addressing scalability limitations. The challenge lies in optimizing efficiency, considering observed lower efficiency compared to systems with two and four N⁺ chains.

3. Related work

Prior studies [2] explored bifunctional catalysts (2-4 N+ chains) at Jagiellonian University. Investigations included propagation and cyclization competition, stereochemical factors in epoxide opening, and assessing Co(III)-salen catalysts with four N+ chains, revealing promising activity discrepancies between racemic and enantiomeric scenarios.

4. Solution of the problem

Developed a two-stage algorithm for computational studies:

1. Conformational space exploration and energy calculations for epoxy complexes.
 - Used semi-empirical molecular dynamics (MD) simulations with a 1 fs time step.
 - Applied PM7 MOPAC optimization for every 500th MD geometry, generating 500 new PM7 geometries.
 - Conducted DFT ADF optimization for every 10th PM7 geometry and MD geometry, resulting in 278 new DFT geometries.
2. Opening reaction analysis of epoxide.
 - Selected „prereactive complexes” based on O-C distance criteria ($<3.2 \text{ \AA}$), yielding 173 structures.
 - Performed DFT „linear transit” calculations, varying C-O distances linearly.
 - Populations of prereactive complexes were calculated, and contributions of reaction pathways to the effective rate constant were assessed.

ADF package (2017.103) for DFT calculations with ZORA relativistic effects was used.

5. Conclusions

The computational predictions, based on extensive geometric variations and numerous reaction pathways, qualitatively align with experimental data. The observed lower catalytic activity of the studied catalysts compared to those with two or four N+ chains highlights an avenue for improvement.

Acknowledgements. This research has been possible through computing allocation on the Ares system at the ACC Cyfronet AGH under the grants PLG/2022/015272, PLG/2022/015762, PLG/2023/016368.

References

1. Kim, T.-J.; et. al. *Macromolecules* 2009, 42, 6932; Kim, T.-J.; et. al. *Chemistry - A European Journal* 2010, 16, 5630–5644; Srebro, M., et. al. *Organometallics* 2010, 29, 5341–5352; Srebro, M., et. al. *Macromolecular Research* 2010, 18, 960–966.
2. Dyduch, K.; Srebro-Hooper, M.; Lee, B.Y.; Michalak, A. *J. Comp. Chem.* 2018 39, 1854; Roznowska, A.; Dyduch, K.; Lee, B.Y.; Michalak, A. *J. Mol. Model.* 2020, 26, 113; Dyduch, K.; Roznowska, A.; Srebro-Hooper, M.; Lee, B.Y.; Michalak, A. *Catalysts* 2021, 11, 328.
3. Y. H. Seo, Y. Bin Hyun, H. J. Lee, H. C. Lee, J. H. Lee, S. M. Jeong, B. Y. Lee, *Macromolecular Research*, 2021, 29(12), 855–863.

A Study of the Properties of Unsymmetrically Substituted 4,5-dialkoxy-2-nitroanilines by Spectroscopic and Theoretical Methods

Karina Synowiec, Marek Boczar, Jarosław Grolík

Jagiellonian University, Faculty of Chemistry, Gronostajowa 2, 30-387, Kraków, Poland

karina.synowiec@student.uj.edu.pl

Keywords: molecular dynamics, quantum mechanical calculations, molecular spectroscopy, hydrogen bonds

1. Introduction

The aim of this study is to synthesize and explore the physicochemical properties of two distinct constitutional isomers of 4,5-dialkoxy-2-nitroanilines, wherein each isomer is asymmetrically substituted with an ethoxy group and an isobutoxy group. This choice of substituents offers an intriguing platform for investigating the effects of asymmetric substitution on the structural, spectroscopic, and theoretical properties of the resulting compounds.



Fig. 1. Structures of 4-ethoxy-5-isobutoxy-2-nitroaniline (1a) and 4-isobutoxy-5-ethoxy-2-nitroaniline (1b).

2. Description of the problem

The study is particularly interested in how the different molecular arrangements of these constitutional isomers affect the spectroscopic properties of the molecules. By synthesizing the isomers and subjecting them to spectroscopic analysis, the study aims to explain the nature of electron and oscillatory transitions, providing insight into the complex relationship between molecular structure and spectroscopic properties.

3. Related work

In recent years, there has been only one significant publication on these compounds, which is based on their synthesis and basic characterization (X-ray, NMR) without explaining the differences between the isomers [1].

4. Solution of the problem

The crystal structures of the two systems were used for Born – Oppenheimer molecular dynamics calculations employing DFT in CP2K software [2]. BLYP function with dftd3 parameterization was applied. DZVP-GTH basis functions and GTH-BLYP-q pseudopotential were used with a GPW plane wave model (CUTOFF=450 Ry). Calculations were conducted in the NVT canonical ensemble at T=300 K with a time step of t=1 fs using Nose-Hoover thermostatting. In quantum mechanics calculation geometry optimizations were carried out using the DFT method (B3LYP) in Gaussian09 [3] software with the 6-311G+(2df,2pd) basis set. Vibrational analysis used the same parameters. Excited states were calculated via TD-DFT (B3LYP/6-311G+(2df,2pd)) considering solvent effects with the PCM model (tetrahydrofuran). Molecular orbitals were visualized using the B3LYP/STO-3G method.

There are no significant differences between the experimental and theoretically calculated infrared spectra. From their comparison, one can only conclude that hydrogen bonds are present in the crystal structure, as quantum mechanical calculations considered one molecule and molecular dynamics considered the elementary cell of the compounds studied. By calculating the energies of successive singlet states, analyzing the calculation results in detail, and visualizing the orbitals, it was possible to describe the HOMO orbitals from which the transitions come and the corresponding LUMO orbitals.

5. Conclusions

The geometries of the molecules in the crystal and averaged over the trajectory agree with each other. Analysis of the atomic power spectra has made it possible to precisely assign bands to the individual vibrations in the experimental spectrum. The theoretical UV - Vis spectra agree with the experimental data, which made it possible to explain the absorption bands by specific transitions between molecular orbitals.

Acknowledgements. The numerical experiment was possible through computing allocation on the Ares system at ACC Cyfronet AGH under the grants plghb11.

References

1. Grolik, J. et al. Regioselective Synthesis of the 4,5-Dialkoxy-2-Nitroanilines Bearing Two Different Alkoxy Substituents. *Tetrahedron Lett* 2022, 99, 153830.
2. Thomas D. Kühne, et al. CP2K: An electronic structure and molecular dynamics software package - Quickstep: Efficient and accurate electronic structure calculations. *The Journal of Chemical Physics* [Internet]. 2020 May;152(19):194103.
3. Gaussian 09, Revision A.02, M. J. Frisch, et al. Gaussian, Inc., Wallingford CT, 2016.

Investigating the Mechanical Characteristics of Al_2O_3 through Density Functional Theory and Molecular Dynamics

Mostafa Fathalian¹, Eligiusz Postek¹, Masoud Tahani^{2,1}, Tomasz Sadowski³

¹ Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawińskiego 5B, 02-106 Warsaw, Poland

² Department of Mechanical Engineering, Ferdowsi University of Mashhad, Mashhad 9177948974, Iran

³ Department of Solid Mechanics, Lublin University of Technology, 20-618 Lublin, Poland

mfath@ippt.pan.pl

Keywords: Al_2O_3 , fracture toughness, Density Functional Theory, Molecular Dynamics

1. Introduction

This exploration highlights the essential role of ceramics, notably aluminum oxide (Al_2O_3), in various technological applications due to its remarkable properties, including high mechanical strength and electrical insulation [1]. It underscores the transformative impact of computational approaches such as density functional theory (DFT) and molecular dynamics (MD) simulations in unraveling Al_2O_3 's mechanical characteristics. The focus is on key attributes like surface energy, Young's modulus, and fracture toughness, providing insights into the atomic-scale mechanisms governing these features. Through the application of DFT and MD simulations, a deeper understanding emerges regarding how cracks initiate, propagate, and influence overall fracture behavior, contributing to the advancement of enhanced materials for diverse applications [2,3].

2. Simulation methodology

The simulation process can be segmented into two components: initial DFT simulations and subsequent MD calculations. Both DFT and MD simulations focus on examining Alumina, characterized by a hexagonal crystal structure known as corundum.

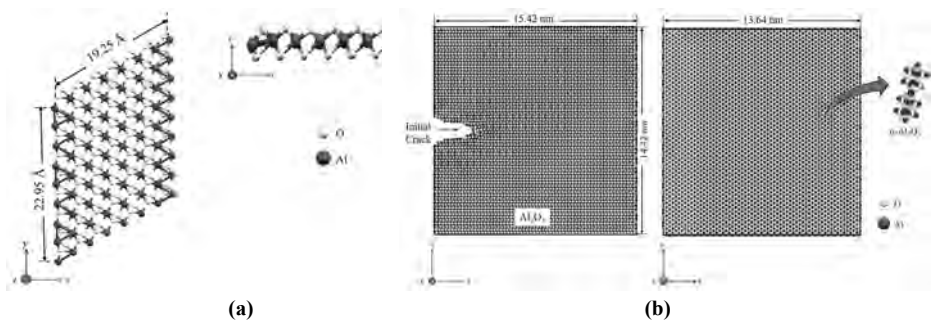


Fig. 1. (a) The defect-free structures of $\alpha\text{-Al}_2\text{O}_3$ configuration after relaxation. (b) Schematic representation of the $\alpha\text{-Al}_2\text{O}_3$ simulation model, dimensions, and coordinate system.

3. Discussion

The study delves into the effects of asymmetric oxygen (O) and aluminum (Al) vacancies on the surface of $\alpha\text{-Al}_2\text{O}_3$, examining their impact on mechanical and electronic properties. Vacancies, crucial in modeling with density functional theory (DFT), alter stress distribution

and local strain fields, influencing mechanical properties such as strength, ductility, and fracture toughness. They also introduce changes in electronic characteristics, affecting electrical conductivity and optical attributes. The investigation encompasses four models with distinct point vacancies and a defect-free supercell, exploring geometric parameters and calculating Young's modulus. The findings reveal varying degrees of structural alterations and disruptions in the crystal lattice associated with O and Al vacancies. A simulated tensile test explores the fracture processes of α -Al₂O₃, providing insights into stress–strain curves and ultimate strength. It is critical for understanding the material's structural limits and irreversible changes in its atomic arrangement. Crack growth in α -Al₂O₃ during z-direction tensile testing with MD provides vital insights into nanoscale material behavior. The observed events include initial crack appearance at zero strain, subsequent elongations triggering crack initiation and propagation, and eventual failure. The stress near the crack tip gradually increases until cracking begins and progresses steadily.

4. Software

In this study, the SIESTA package code was employed for conducting density functional theory (DFT) calculations, while MD simulations were executed using the LAMMPS software [4]. To conduct computations and simulations, high-performance computing (HPC) resources, specifically PLGrid, were utilized.

5. Conclusions

Aluminum vacancies demand higher energy than oxygen vacancies, making their occurrence less likely in natural conditions. Young's modulus significantly decreases with aluminum vacancies, impacting elongation in tensile tests. Fracture toughness diminishes notably with introduced aluminum vacancies. MD simulations yield a fracture toughness of 2.8 MPa·√m for α -Al₂O₃, aligning well with DFT calculations and experimental data.

Acknowledgements. The calculations were performed using PLGrid (plgdynacer04) facilities at the Academic Computer Centre Cyfronet, AGH University of Krakow, and at the Centre of Informatics Tricity Academic Supercomputer and Network in Gdansk, Poland. Grant from National Science Center, Poland; UMO 2019/33/B/ST8/01263.

References

1. P. Narottam, J. L. Bansal: Ceramic Matrix Composites: Materials, Modeling and Technology. Wiley, 2014; ISBN 978-1-118-23116-6.
2. M. Fathalian, E. Postek, T. Sadowski: Mechanical and Electronic Properties of Al(111)/6H-SiC Interfaces: A DFT Study. *Molecules* 2023, 28. doi: 10.3390/molecules28114345.
3. S. Blonski, S. H. Garofalini: Molecular Dynamics Simulations of α -Alumina and γ -Alumina Surfaces. *Surf. Sci.* 1993, 295, 263–274, doi: 10.1016/0039-6028(93)90202-U.
4. S. Plimpton: Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* 1995, 117, 1–19, doi: 10.1006/jcph.1995.1039.

Multi-scale Modeling of Texture Evolution in Deformed Metals

Bartosz Sułkowski

Department of Materials Science and Non-Ferrous Metals Engineering, Faculty of Non-Ferrous Metals,
AGH University of Krakow, 30 Mickiewicza Av., Kraków 30-059, Poland

bartosz.sulkowski@agh.edu.pl

Keywords: texture, modeling, vpsc, recrystallization

1. Introduction

Texture is one of the most important features which determines the properties of metals. It is very important to understand the texture evolution in deformed metals because of the strong anisotropy of mechanical properties. To learn relations between evolution of the texture and the mechanical properties, the simulations and modeling are very helpful. However, there are differences between experimental results and simulated textures especially at elevated temperatures. In the modeling of texture evolution in hot deformed metals, the recrystallization processes must be taken into account. In the presentation the model for simulations of the texture evolution in hot deformed metals is described. The new model is based on the visco-plastic self-consistent model (vpsc) combined together with the cellular automaton model (CA). The vpsc model is able to predict the deformation texture components while the CA model can trace the texture softening due to the activation of recrystallization processes. The model vpsc+CA was implemented in C++ with the use of the Message Passing Interface (MPI) library to increase the scalability of the calculations.

2. Description of the problem

The available models (vpsc or Taylor) for the texture evolution prediction are based only on the activation of the different deformation modes such as slip or twinning. The models are excellent to predict the deformation texture components at low temperatures [1-7]. At elevated temperatures, in hot deformed metals, recrystallization processes are activated, influencing the texture evolution. In the result, the softening and recrystallization texture components are observed. The vpsc neither Taylor models cannot predict these components.

3. Related work

Zecevic et al. [8] introduced a model for the texture evolution in hot-deformed metals. The model is a combination of the vpsc formulation and intragranular orientation gradients and strain energy fields. The strain energy is a driving force for grain boundary migration and is based on the total dislocation density stored in the grains. However, the Zacevic model is not distinguishing between the statistically stored (SSD) and geometrically necessary dislocations (GND). Thus, some discrepancies were observed between the experimental and simulated textures. The distinguishing between SSD and GND will allow better prediction of the simulated texture evolution and should be introduced to the basic formulations of vpsc or Taylor model.

4. Solution of the problem

In the vsc+CA model the structure of the deformed metals is divided into grains. Each grain is represented by its orientation and the total dislocation density. The orientation and dislocation density changes in the grains are calculated based on the vpsc model formulations. For the CA part of the model, each grain is divided into a certain number of subgrains whose orientation

and dislocation density differ a little from of the parent grain. Based on that differences in the subgrains, the rule of the subgrains state switch is introduced in the CA formulations. The rule contains the parameters describing the recrystallization mechanisms.

5. Conclusions

There are differences between experimental results and simulations of texture evolution in hot deformed metals. The vscp model can predict only the deformation texture components. The vscp+CA model is able to predict the texture softening due to the activation of the recrystallization processes. The support of the MPI library allowed to increase the scalability of the model and significantly reduced time needed for the calculations in the new vpssc+CA model.

Acknowledgements. The numerical experiment was possible through computing allocation on the Ares system at ACC Cyfronet AGH under the grants plgcyfsoftware2023.

References

1. D. Persaud-Sharma, A. McGoron, J. Biomim. Biomater. Tissue Eng. 3 (2012) 25–39.
2. A. Ojdanic, J. Horky, B. Mingler, M. Fanetti, S. Gardonio, M. Valant, B. Sułkowski, E. Schafner, D. Orlov, M. J. Zehetbauer, Metals 10 (8) (2020) 1064.
3. A. Mizelli-Ojdanic, J. Horky, B. Mingler, M. Fanetti, S. Gardonio, B. Sułkowski, E. Schafner, D. Orlov, M. J. Zehetbauer, Materials 14 (21) (2021) 6399.
4. J. F. Nie, K. S. Shin, Z. R. Zeng, Metall. Mater. Trans. A 51A (2020), December 2020–6045.
5. K. Kabirian, A. S. Khan, Int. J. Solids Struct. 67–68 (2015) 116–126.
6. H. Zhang, G. Huang, D. Kong, G. Sang, B. Song, J. Mater.Process. Technol. 211 (2011) 1575–1580, D. Raabe, Philos. Mag. A 79 (10) (1999) 2339–2358.
7. B. Sułkowski, R. Chulist, Materials Characterization 201 (2023) 112968.
8. M. Zecevic, R. A. Lebensohn, R. J. McCabe, M. Knezevic, Acta Mater. 164 (2019) 530–546.

Studies on the Interaction of Nanoplastics with Lipid Systems

Franciszek Włodek^{1,2}, Waldemar Kulig³, Jacek Korchowicz¹, Anna Stachowicz-Kuśnierz¹

¹Faculty of Chemistry, Jagiellonian University, Kraków, Poland

²Doctoral School of Exact and Natural Sciences, Jagiellonian University, Kraków, Poland

³Department of Physics University of Helsinki, Helsinki, Finland

franek.wlodek@doctoral.uj.edu.pl

Keywords: nanoplastics, phospholipid bilayers, molecular dynamics simulations

1. Introduction

Annual production of plastics is increasing and has now reached 400 million tonnes. Plastic is widely used in the machinery and clothing industries, as well as for food and water packaging. When plastic degrades, it releases nanoparticles (NP) and microparticles into the environment or food. When NP enters an organism, it can disrupt vital biological processes. Lipid membranes and monolayers control the penetration of NP into the cell centre or even into the organism, acting as critical gatekeepers. Plastic pollution and its interaction with lipid structures has attracted considerable research interest due to its potential impact on the environment and human health.

2. Description of the problem

Our work explores the interactions between small plastic nanoparticles and model lipid structures. Initially, we focused on the interactions of short polyethylene (PETH) chains with model membrane systems composed of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and dipalmitoylphosphatidylcholine (DPPC). PETH polymer was chosen because it is one of the most widely used plastics and is considered inert, what leads to conviction of being it safe for living organisms. On the other hand, we anticipate strong interactions between PETH chains and lipid tails. Exploring this topic can give us insight into the dangerous influence of inert, hydrophobic polymer chains on bilayers. In addition, we gain information on the ease of bilayer penetration by NP [1]. The next step in this research, which is still under preparation, involves studying the interactions of a four-component monolayer (made of POPC, DPPC, phosphatidylglycerol, and cholesterol) with different atactic polystyrene (STYR) chains at varying partial pressures. This approach, when compared with experimental results, will confirm our approach to the problem and allow us to explore the penetration of STYR chains into the monolayer. In addition, these simulations allow us to study the respiration process and make possible to determine in which pressure monolayer is most susceptible to penetration.

3. Solution of the problem

All simulations were performed using GROMACS 2022 software and all models were prepared via the CHARMM-GUI interface. The membrane models were composed of clean POPC and POPC-DPPC in a 1:1 ratio and were modelled in two sizes: 128 lipids per monolayer and 256 lipids per monolayer. To gain a deeper understanding of PETH-membrane interactions, we scanned through NP of different sizes defined by the number of repeat units (ranging from 10 to 100 with a step of 10 repeat units). Systems with PETH chains and bilayer models were simulated with a potential barrier placed 8 nm from the centre of the system, which prevented the NP from being placed too far from the bilayer. This approach allowed us to observe spontaneous penetration of bilayers. We performed controlled dynamics to pull large PETH chains into the bilayer. These simulations allowed us to determine the degree of perturbation of the bilayer

structure after penetration by calculating the order parameter and the area per lipid. In the next step of this research, we carried out a similar simulation of spontaneous penetration, but we did not introduce a potential barrier because penetration of the monolayer by STYR chains is easier.

4. Conclusions

In our initial research, we found a limiting size of PETH nanoparticles that can penetrate membranes. Larger PETH chains were found to be unable to penetrate model membranes and the limiting size was 40 repeat units. We observed that membranes built from 256 lipids per monolayer were less likely to be penetrated by NP. This fact is associated with larger amounts of water and system size. This research provides evidence for strong interactions of PETH nanoparticles with bilayer membranes and provides information about two mechanisms of penetration of the bilayer membrane, one faster for short chains and the other slower for long chains. PETH nanoparticles dissolve inside the bilayer and interpose themselves between the monolayers, interacting with the lipid tails [3] what leads to perturbation of the bilayer structure. In conclusion, absorbed chains have a negligible effect on the structure of the bilayers studied. In the second step of this research we concluded that STYR chains are absorbed on the surface of the monolayer. Throughout the simulation, STYR chains penetrate deeper into the monolayer and their radius increases, but they do not dissolve (untangle).

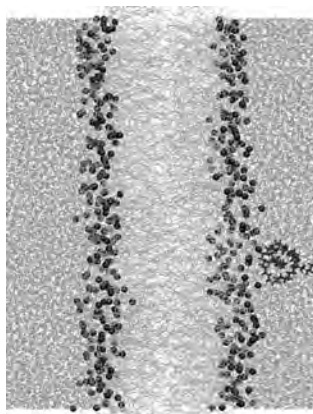


Fig. 1. Interactions of PETH with two-component membrane build from 256 lipids per monolayer.

Acknowledgements. This research has been supported by a grant from the Faculty of Chemistry under the Strategic Programme Excellence Initiative at Jagiellonian University. The numerical experiment was possible through computing allocation on the Ares and Athena systems at ACC Cyfronet AGH under the grants plgnpmembares.

References

1. Włodek F.; Kulig W.; Korchowicz J.; Stachowicz-Kuśnierz A. Insights into Short Chain Polyethylene Penetration of Phospholipid Bilayers via Atomistic Molecular Dynamics Simulations – submitted to editors.
2. Javanainen, M.; Lamberg, A.; Cwiklik, L.; Vattulainen, I.; Ollila, O. H. S. Atomistic Model for Nearly Quantitative Simulations of Langmuir Monolayers. *Langmuir* 2018, 34 (7), 2565–2572. <https://doi.org/10.1021/acs.langmuir.7b02855>.
3. Hollóczki, O.; Gehrke, S. Can Nanoplastics Alter Cell Membranes? *ChemPhysChem* 2020, 21 (1), 9–12. <https://doi.org/10.1002/cphc.201900481>.

Theoretical Study of Copper Complexes in Artificial Photosynthesis

Alicja Miklas¹, Mateusz Z. Brela²

¹ Doctoral School of Exact and Natural Sciences, Faculty of Chemistry, Jagiellonian University

² Faculty of Chemistry, Jagiellonian University

alicja.miklas@doctoral.uj.edu.pl, mateusz.brela@uj.edu.pl

Keywords: artificial photosynthesis, Cu(I) complexes, ETS, molecular modeling

1. Introduction

For the last decades, the main source of energy was fossil fuels. The combustion of which results in massive CO₂ production that is harmful to the atmosphere. Due to the damaging effects of CO₂, attempts to limit its production and to use CO₂ already present in the atmosphere are being made. The process that converts CO₂ into more environmentally friendly products is photosynthesis. Photosynthesis occurring naturally is a very complicated process and a much-simplified process that mimics natural photosynthesis is artificial photosynthesis (Fig. 1.). Artificial photosynthesis, in general, consists of the following processes: absorption of solar radiation, water oxidation, reduction of carbon dioxide or water. An artificial photosynthetic system consists of photosensitizer (PS), catalyst (CAT), and sacrificial electron donor (SED). There are many compounds studied for use in artificial photosynthesis and Cu(I) complexes can be included in this group [1]. Cu(I) complexes are attractive to be a photosensitizer because they often absorb relatively low energy radiation, exhibit high luminescence, strong reducing power, and appropriate lifetimes in excited states. This work presents a theoretical study of the Cu(I) complex, in particular interactions between Cu(I) core and various ligands.

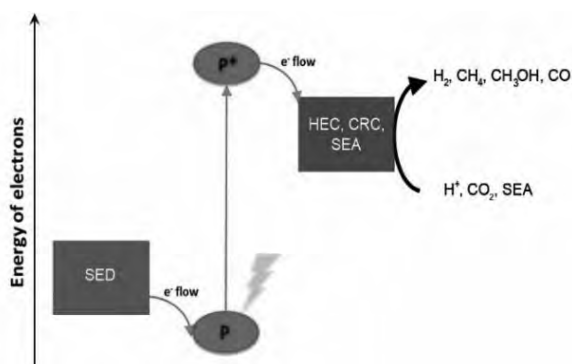


Fig. 1. Schematic representation of artificial photosynthesis.

2. Description of the problem

In the process of artificial photosynthesis, PS may either go a reductive quenching with SED, or oxidative quenching with CAT. After that CAT is reduced to activate CO₂. In analyzed Cu(I) complexes, Cu(I) has a +1 oxidation state and throughout the process undergoes reduction which is accompanied by its ligands. In many Cu(I) complexes there are redox-active (or so-called

non-innocent) ligands which often contain N atoms in sp^2 -hybridized heterocycles, employed to coordinate metal ions, forming PS and CAT. Usually, these ligands are called 'diimine' and are σ -donors and π -acceptors. Diimine ligands assist in the reduction of copper in Cu(I) complexes. Due to the fully filled d-shell of Cu(I), the lowest unoccupied molecular orbital (LUMO) is very often located on the diimine. Due to that diimine ligand may actively participate in the absorption of electromagnetic radiation on the lowest energetic level and this process formally is electron transfer from orbital centered on metal to the first free π^* diimine orbital. Due to that, when aiming at efficient Cu(I) based PS, properly tailored and selected non-innocent ligand is needed. In particular, the one that makes the LUMO level of the final complex low enough to make radiation absorption possible and at the same time enable convenient electron transfer to a catalyst [2].

3. Related work

A novel mononuclear Cu(I) complexes were synthesized by Prof. Claudia Bizzari [3] (collaborator from the University of Hohenheim) for example: via coordination with a benzoquinoxalin-2'-one-1,2,3-triazole chelating diimine and the bis[(2-diphenylphosphino)phenyl] ether (DPEPhos), to target a new and efficient photosensitizer for photocatalytic CO_2 reduction.

4. Solution of the problem

Theoretical calculations with Density Functional Theory (DFT) methodology with several correlation functionals were performed using the Amsterdam Density Functional (ADF) program. [4] A standard triple- ζ STO basis containing one set of polarization functions (TZP) was adopted for all atoms. The 1s electrons of C, N, O, and F as well as the 1s-2p electrons of Cu and Cl were treated as a frozen core. The bonding between model particle chains and the considered anions was analyzed by ETS-NOCV [5]. In this research, ETS calculations for several Cu(I) complexes were done. All of the analyzed complexes have Cu(I) cores but differ with ligands, however, all ligands contain either N, or P atoms. The goal is to analyze interactions between the Cu(I) core and ligands containing N atoms as well as between ligands containing P atoms.

5. Conclusions

Charge analysis and energy decomposition (ETS) of bonds between Cu(I) core and N-containing or P-containing ligands are analyzed. With the results of energy decomposition, bonding energy and interactions as well as bonds between Cu(I) core and ligands may be characterized. Such analysis may be helpful for future modeling of the photosensitizer.

Acknowledgements. The numerical experiment was possible through computing allocation on the Prometheus system at ACC Cyfronet AGH under the grant plghb11.

References

1. C. Steinlechner, A. F. Roesel, E. Oberem, A. Pöpcke, N. Rockstroh, F. Gloaguen, S. Lochbrunner, R. Ludwig, A. Spannenberg, H. Junge, R. Francke, M. Beller: "Selective Earth-Abundant System for CO_2 Reduction: Comparing Photo- and Electrocatalytic Processes", *ACS Catal.* 2019, 9, 2091-2100.
2. S. Berardi, S. Drouet, L. Francás, C. Gimbert-Suriñach, M. Guttentag, C. Richmond, T. Stoll, A. Llobet: "Molecular artificial photosynthesis", *Chem. Soc. Rev.*, 2014, 43, 7501.
3. L.-L. Gracia, P. Henkel, O. Fuhr, C. Bizzarri: "Selectivity control towards CO versus H_2 for photo-driven CO_2 reduction with a novel Co (II) catalyst", *Beilstein Journal of Organic Chemistry*, 2023, 19, 1766-1775.
4. G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders and T. Ziegler, *Chemistry with ADF*, *J. Comput. Chem.*, 2001, 22, 931.
5. Mitoraj, M., Michalak, A., Ziegler, T., *Journal of Chemical Theory and Computation*, 2009, 5, 962.

Hydrolysis as an Alternative Path of Nylon 6 and PHB Degradation: a Computational Study

Yuliia Didovets^{1,2}, Mateusz Z. Brela¹

¹ Molecular Spectroscopy Group, Department of Physical Chemistry and Electrochemistry, Faculty of Chemistry, Jagiellonian University, Gronostajowa St 2, Kraków, 30-387 Poland

² Doctoral School of Exact and Natural Sciences, Jagiellonian University, Prof. St. Łojasiewicza St 11, Kraków, 30-348 Poland

yuliia.didovets@doctoral.uj.edu.pl

Keywords: hydrolysis, degradation, nylon 6, PHB, AIMD

1. Introduction

Continuous research on the topic of polymer degradation highlights the need for detailed knowledge about the mechanisms of simple reaction types often applied during the process. Hydrolysis is one of the most popular paths of polymer degradation as it is especially successful for polymers with amide and ester functional groups. Polyamide 6 (nylon 6) and polyhydroxybutyrate (PHB), well-known representatives of polyamides and polyesters, both undergo acid and alkaline hydrolysis through the chain-scission reaction yielding oligomers of different molecular weights. The hydrolysis of nylon 6 and PHB can be additionally accelerated by increasing the temperature of the reaction environment [1,2].

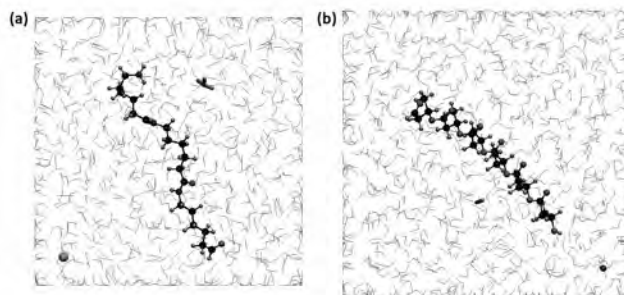


Fig. 1. Model systems of nylon 6 in acid (a) and PHB in alkaline (b) environments.

2. Description of the problem

As the experimental research works provide a great amount of data about the nylon 6 and PHB hydrolysis rates and mechanisms in acid and alkaline environments [1,2], such conclusions are not well explained on a molecular level. Further, a deeper characterization of the hydrolysis mechanism in different environments is lacking. The influence of temperature as well as pH on the thermodynamics of hydrolysis should be explained.

3. Related work

Previous research work was conducted for the estimation of the temperature and implicit solvation impact on the thermodynamics of nylon 6 and PHB thermal degradation processes [3].

In this research, dynamic calculations are applied to describe and compare the nylon 6 and PHB hydrolysis mechanisms and parameters (kinetics and thermodynamics) in acid and alkaline environments.

4. Solution of the problem

Ab initio molecular dynamics (AIMD) of model systems (Fig. 1) was run using the Quickstep approach in the CP2K package [4]. The DFT method was applied with the gradient functional: BLYP with Grimme dispersion correction (D3) [5]. The mixture of plane waves (cutoff: 450Ry) with the DZVP basis set was used. Periodic boundary conditions were applied to all directions, the timestep was set to 1 fs and the temperature to 300 K in the NVT ensemble. The ACC Cyfronet AGH resources were used, consuming in total ca. 2M of computational hours on the Ares supercomputer utilizing 8 nodes and 48 cores for each job.

The initial model systems' geometries contain a trimer structure of nylon 6 surrounded by 708 water molecules and one OH⁻ for alkaline hydrolysis and H₃O⁺ for acid hydrolysis. Similar molecular models were built for the PHB polymer. Obtained trajectories allow for the description of the initial structures' conformational stability. Additionally, several distance constraints were used for modelling the nylon 6 and PHB degradation reactions by molecular dynamics simulations.

5. Conclusions

Molecular dynamics results showed proton transfers, which affect preferable paths of nylon 6 and PHB hydrolysis. Further, the initial description of acid as well as alkaline hydrolysis mechanisms is discussed. The stability of intermediate ionic states has been elucidated.

Acknowledgements. This research has been supported by the program “Excellence Initiative – Research University” at the Jagiellonian University (Research Support Module application WSPR.WCh.1.9.2023.3(1)). The numerical experiment was possible through computing allocation on the Ares system at ACC Cyfronet AGH under the grants *plghb11-cpu*, *plgpbff2-cpu* and *plgsmshps-cpu*.

References

1. M. M. Brette, A. H. Holm, A. D. Drozdov, and J. d. C. Christiansen: “Pure Hydrolysis of Polyamides: A Comparative Study”, *Chemistry* 2024, 6, pp.13-50.
2. J. Yu, D. Plackett, and L. X. L. Chen: “Kinetics and mechanism of the monomeric products from abiotic hydrolysis of poly[(R)-3-hydroxybutyrate] under acidic and alkaline conditions”, *Polymer Degradation and Stability* 2005, 89, pp.289-299.
3. Y. Didovets and M. Z. Brela: “Theoretical Study on the Thermal Degradation Process of Nylon 6 and Polyhydroxybutyrate”, *Physchem*, 2022, 2, pp.334–346.
4. T. D. Kühne, M. Iannuzzi, *et al.*: “CP2K: An Electronic Structure and Molecular Dynamics Software Package - Quickstep: Efficient and Accurate Electronic Structure Calculations.” *J Chem Phys*, 2020, 152 (19), 194103.
5. S. Grimme, J. Antony, S. Ehrlich, H. Krieg: “A consistent and accurate ab initio parameterization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu” *J. Chem. Phys.* 2010, 132, 54104.

A Detail That Matters? Rotational Isomerism in Some *Ortho*-substituted Benzophenones

Andrzej J. Kałka, Mateusz Z. Brela, Andrzej M. Turek

Jagiellonian University in Kraków, Faculty of Chemistry
Gronostajowa 2, 30-856, Kraków, Poland

{andrzej.kalka, mateusz.brela}@uj.edu.pl
turek@chemia.uj.edu.pl

Keywords: Density Functional Theory (DFT), atropisomerism, conformational analysis

1. Introduction

Benzophenone (BP), known also under the name of diphenylmethanone (Ph_2CO), is a rather small organic compound which, despite its relatively simple structure, exhibits quite an interesting photophysical and photochemical behavior (phosphorescence, delayed fluorescence, photosensitization, etc.) [1]. Although over the decades it has been subjected to various types of chemical investigation, still, certain issues related to the aforesaid molecule – as well as its derivatives – tend to remain not fully explored.

2. Description of the problem

One of such unapprehended questions relates to atropisomerism (type of conformational isomerism) within the group of *ortho*-substituted benzophenones. To specify, by considering their spatial structure, it can be postulated that the aforesaid molecules should appear in the form of two atropisomers (stable rotational isomers), *syn* and *anti*, differing in orientation of the pendant group relative to the carbonyl fragment (see Fig. 1) [2]. While this detail may, at first sight, seem negligible, as the topology of bonds remains formally the same, yet, in fact, it may have considerable implications on the selected properties of the discussed compounds, since the two resulting rotamers may (slightly) differ in (photo)physical and (photo)chemical characteristics [3,4]. Interestingly, in spite of the fact that the above conjecture tends to be evidenced by several papers (see e.g. [3,4]), no comprehensive research has been reported so far to systematically investigate the impact (together with implications) of rotational isomerism onto the above-mentioned class of chemical compounds.

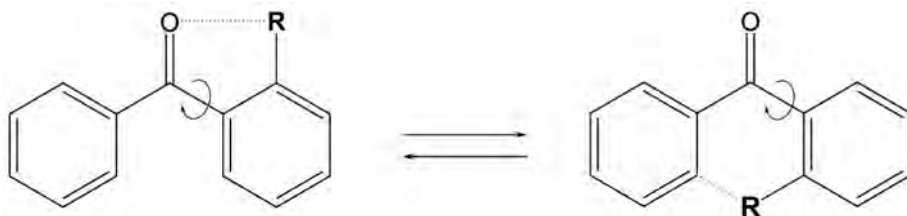


Fig. 1. Rotational isomerism within *ortho*-substituted benzophenones.

3. Solution of the problem

Going against the indicated gap, the Authors have made an attempt to start a dedicated research project, aimed at exploration of the rotamerism phenomenon within the group of *ortho*-substituted benzophenones.

The latter began with a preliminary DFT computational study, during which an initial set of 15 *ortho*-benzophenone molecules was analyzed with regard to the relative thermodynamic stability of the *syn* and *anti*-conformeric isomers. In the next step, for the compounds that were considered to exhibit the phenomenon of atropisomerism, some selected spectroscopic properties (IR, Raman, UV-Vis) were simulated and cross-checked between the corresponding rotamers.

All the computations were carried out with the Gaussian 16 software [5], making use of the resources provided by the HPC clusters Ares (1 node, 48 cores) and Eagle (1 node, 28 cores).

4. Related work

In parallel to the performed computational study, the outlined phenomenon is currently being investigated experimentally with a particular use of the UV-Vis and IR spectroscopy. The preliminary results of the measurements appear to be in line with the theoretical predictions.

5. Conclusions

The results of the performed DFT simulations clearly indicate that at least several of the investigated *ortho*-benzophenone molecules are expected to appear in the form of two thermodynamically stable rotamers. In addition, both vibrational and electronic spectra of the *syn* and *anti* isomers turned out to remain non-identical, which further implies the explored atropisomerism phenomenon to indeed have an observable impact on the considered class of compounds.

Acknowledgements. For their ongoing cooperation with regard to the experimental part of the conducted research, the Authors wish to sincerely thank A. Orlef and Prof. A. Kaczor from the Chiroptical Spectroscopy Group at the Jagiellonian University in Kraków.

The numerical experiments were possible through computing allocation on the Ares (grant plghb11) HPC cluster at ACC Cyfronet AGH.

References

1. A. M. Turek, G. Krishnamoorthy, K. Phipps, J. Saltiel: "Resolution of benzophenone delayed fluorescence and phosphorescence with compensation for thermal broadening" in *J. Phys. Chem. A* vol. 106, 2002, pp.6044-6052.
2. P. J. Cox, D. Kechagias, O. Kelly: "Conformations of substituted benzophenones" in *Acta Cryst. B* vol. 64, 2008, 206-216.
3. W. Li, J. Xue, S. C. Cheng, Y. Du, D. L. Phillips: "Influence of the chloro substituent position on the triplet reactivity of benzophenone derivatives: a time-resolved resonance Raman and density functional theory study" in *J. Raman Spectrosc.* vol. 43, 2012, pp. 774-780.
4. M. A. Strzemechny, S. G. Stepanian, D. I. Zloba et al.: "Scenario of temperature-related variation of phosphorescence spectra of *ortho*-bromobenzophenone crystal" in *Chem. Phys.* vol. 463, 2015, pp. 58-64.
5. Gaussian 16, Rev. A.03, Gaussian, Inc., Wallingford, CT, 2016.

The System for Automatic Skin Cancer Detection – Achievements and Challenges

Rafał Frączek², Michał Karwatowski², Jakub Grzeszczyk¹, Jakub Caputa¹, Piotr Pindel¹, Daria Łukasik¹, Maciej Wielgosz^{1,2}, Paweł Russek^{1,2}, Agnieszka Dąbrowska-Boruch^{1,2}, Ernest Jamro^{1,2}, Marcin Pietron², Sebastian Koryciak², Kazimierz Wiatr^{1,2}

¹ ACC Cyfronet AGH, Nawojki 11, 30-950 Kraków, Poland

² AGH University of Krakow al. Mickiewicza 30, 30-059 Kraków, Poland

{rafalfr, mkarwat, wielgosz, russek, adabrow, jamro, pietron, koryciak, wiatr}@agh.edu.pl, {d.lukasik, j.grzeszczyk, j.caputa, p.pindel}@cyfronet.pl

Keywords: image processing, veterinary, data annotation, deep learning, skin cancer

1. Introduction

Artificial intelligence (AI) has shown promising applications in the field of veterinary oncology. One notable use is the interpretation of microscopic images, where AI algorithms can analyse images from cytology samples to assist veterinarians in identifying potential cancerous lesions, which contributes to more accurate and timely diagnosis.

2. Description of the problem

Cytology images can vary widely in terms of quality, staining techniques, and sample preparation. Building robust AI models requires large and diverse datasets. In veterinary oncology, there might be limitations in the availability of annotated cytology images, hindering the development of highly accurate models.

3. Solution of the problem

The originally developed system for the semi-automatic cytology image annotation as well as automatic cancer cells detections has been presented in [1] and [2]. In this paper, the most important achievements and challenges in the development of this system are presented.

The most important component is a model based on a visual transformer neural network which is currently the most popular and successful method in computer vision. The model performs image segmentation, which identifies areas on the image that are characteristic for the different categories of cells. In the case of cytological images, these categories include such types as red blood cells as well as the most common cancer cells such as mast cells, histiocytes, melanocytes, plasma cells as well as lymphocytes. The model architecture core is the Swin-Transformer, which performs detection based on image features obtained using the CBNetV2 (Composite Backbone Network) [3] network, serving as the backbone for visual data input.

Another challenge in system development is to acquire a good training set. For this purpose, a special application was created that makes it possible for a specialist to quickly annotate a large number of images. To increase the number of labelled data, standard augmentation techniques (e.g. image scaling, rotations, colour transformation, histogram normalisation, etc.) were applied. As a result, the current training set consists of over 2,000 images, annotated by pathologists who identified around 100,000 objects, including cancer cells and other cells typically found in cytological skin smears. The images were captured using a microscope with a four hundred times magnification and also equipped with a high-resolution digital camera.

Another achievement is the creation of a functional web browser-based user interface that enables easy interaction with the cancer cell recognition model. The architecture of this subsystem consists of a service server which hosts services responsible for the images database, user authorization and queries, result report generation as well as mail distribution. From the software perspective, the central element is the MySQL database that stores information about users, examination tasks, and obtained results. The AI engine module communicates with the database, fetching orders and then sending them to the Ares supercomputer, where the AI model for pathological cell recognition is executed. Another module responsible for generating examination result reports also communicates with the database. This module retrieves results from the database and, based on them, generates reports that are sent back to a user and also saves them on the file server. Finally, another module is responsible for the user interface, which is displayed in the web browser.

4. Results

The training dataset consists of 9 distinct cancer types and an additional category for inflammations. The main task for the ML model is to perform the instance segmentation; while cancer cells are the most important, there are other, less numerous, types of cells that are also recognized. Model performance was measured using AP@IoU0.75 metric. Results varied depending on cell type, for main cancer cells values ranged from 0.84 to 0.96. All tests were run on samples from patients that were not used during training (excluding 3 types of cancer for which this information was not available). The number of cell instances in the dataset was sufficient to train the model; however, the limited number of patients (~40) poses potential problems with matching the distribution of data from other sources. The model training procedure was conducted on the dedicated resources for AI computations available at ACC Cyfronet AGH. The complete training process takes approximately six hours on a computing node with six Nvidia V100 GPGPU cards on the Ares server. For the learning process speedup, the transfer learning procedure was applied.

5. Conclusions

The paper summarises the most important achievements and challenges in the developed system for the automatic dogs' skin cancer detection. Compared to the previous system version, the new one can detect more cancer types with higher accuracy and provides a web user interface. The best-achieved results according to the AP@IoU0.75 metric reach 0.96. The web user interface provides new and upgraded features such as report generation in PDF format and user interaction in a forum.

Acknowledgements. PLGrid grant ID: plglaoui22/plglaoui23. All models were trained on the PLGrid infrastructure.

References

1. R. Frączek, M. Karwatowski, J. Grzeszczyk, J. Caputa, D. Łukasik, M. Wielgosz, P. Russek et al., A system for the fast and accurate cytology images annotation, KUKDM 2022: fourteenth ACC Cyfronet AGH HPC users' conference, 7-8 April 2022: proceedings.
2. R. Frączek, M. Karwatowski, J. Grzeszczyk, J. Caputa, P. Pindel, D. Łukasik, et al, Advances in the System for the Automatic Dogs' Skin Cancer Detection, KU KUDM 2023: fifteenth ACC Cyfronet AGH HPC users' conference: 19-21 April 2023: proceedings.
3. Liang T., Chu X., Liu Y., Wang Y., Tang Z., Chu W., Chen J., Ling H.: CB-NetV2: A Composite Backbone Network Architecture for Object Detection, 2021.

Transcriptomics Atlas Pipeline: HTC with Salmon

Piotr Kica^{1,2}, Sabina Licholai^{1,3}, Maciej Malawski^{1,2}

¹ Sano Centre for Computational Medicine, Czarnowiejska 36, 30-054 Kraków, Poland

² Faculty of Computer Science, AGH University of Krakow, al. Mickiewicza 30, 30-059 Kraków, Poland

³ ACC Cyfronet AGH, Nawojki 11, 30-950 Kraków, Poland

{p.kica, s.licholai, m.malawski}@sanoscience.org

Keywords: high-throughput computing, transcriptomics, array-job, AWS, Apptainer

1. Introduction

The transcriptome, which provides indirect information about gene activity under specific conditions, is defined as the complete set of RNAs of a cell at a given time. The aim of the Transcriptomics Atlas Project is to process publicly available data from the sequencing of RNA for a representative collection of human tissues and make it available in a convenient format for researchers. Salmon [1] is a tool that produces transcript-level quantification estimates from RNA sequences. It is the core part of the Salmon Pipeline which has been implemented in the project. Processing of thousands of such sequences is a problem that can be solved by high-throughput computing (HTC).

2. Problem – large processing pipelines

The preparation of Transcriptomics Atlas requires processing thousands of .sra files for selected human tissues using transcriptomics pipeline, called in the project as Salmon Pipeline. The pipeline consists of 4 steps:

- Downloading .sra file from NCBI NIH [2] using prefetch tool
- Converting .sra to .fastq format using fasterq-dump tool
- Performing quantification of reads using Salmon software
- Normalization of counts using DESeq2 package.

In this case, valid .sra files are 200MB to 30GB in size and the extracted .fastq files can be even 17 times larger. For the Transcriptomics Atlas to be comprehensive it is required to process over 8500 of such files (ideally over 200 files per tissue with a good mapping rate).

3. Previous work – cloud-based pipeline

Our most up-to-date prior work regarding Transcriptomics Atlas Project has been published in [2]. We present the early implementation of the Salmon Pipeline for preparation of the Transcriptomics Atlas. We identify resource requirements for each step of the pipeline and propose cloud architecture. We also perform small scale (99 files) comparison of performance between EC2 instances and Apptainer containers on the Ares cluster.

4. Solution: containerized HTC pipeline on the HPC cluster

The pipeline and required tools, dependencies have been containerized using Docker software. The Dockerfile uses Ubuntu:22 as a base image, installs python3 with boto3 library, SRA-Toolkit, Salmon, R with DESeq2 library. Alongside we containerize a precomputed Salmon Index (680MiB) created beforehand using human transcriptome. In order to create an Apptainer image on Ares, we first upload the Docker image to DockerHub, then we create an interactive job with MEMFS to quickly download and convert the image in memory using *apptainer pull* command. The result is an *image.sif* file of 1.3 GB in size.

In this scenario, the best solution is to utilize the array-job functionality of SLURM. Single sbatch script, which requests 500 array-jobs for 12 hours, each having 2 cores and 8 GB of RAM, is sufficient for this workload. Each job starts the container from a given image, mounts TMPDIR to the pipeline's WORKDIR, provides pipeline configuration using ENV variables, all the using *aptainer run* command with correct arguments.

The inputs are sent beforehand to an AWS SQS queue from which each container/worker polls from. After processing the given SRA_ID with the pipeline the resulting *normalized_counts.txt* is sent to S3 service for storage. The collected metadata of execution is uploaded to DynamoDB table. If there is nothing else in the queue the job is completed. Each worker has access to short-lived (12h) credentials with least-privilege permissions. Longer computations may require renewing the credentials.

5. Obtained results

The workload consumed 6646 CPU-hours with 62% average efficiency across all jobs. All intermediate files (.sra + .fastq) weighed 96.6 TiB in total. Because of the slurm queue load, 306 jobs (out of 500 requested) launched immediately and processed 99.6% of the workload in 12 hours.

6. Conclusions

Using array-jobs for this specific problem is the correct approach. Containerization of the pipeline made it easier to deploy and run on the cluster. By using an HPC cluster, each worker had enough storage space with good performance to process an .sra file of any size. Also, such an approach is beneficial for improving cluster utilization as such small jobs are easier to schedule. The workload has been successfully computed for the creation of the Transcriptomics Atlas.

Acknowledgements. This publication is supported by the European Union's Horizon 2020 research and innovation programme under grant agreement Sano No 857533 and was created within the project of the Minister of Science and Higher Education "Support for the activity of Centers of Excellence established in Poland under Horizon 2020" on the basis of the contract number MEiN/2023/DIR/3796. This publication is also partly supported by the European Union's Horizon Europe research and innovation programme under grant agreement NEARDATA No 101092644. This research was supported by the PLGrid infrastructure. Computations have been performed at the Ares supercomputer at ACC Cyfronet AGH (grant no. PLG/2023/016227).

References

1. Patro R., Duggal G., Love M. I., Irizarry R. A., Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 2017 Apr;14(4):417-419. doi: 10.1038/nmeth.4197. Epub 2017 Mar 6. PMID: 28263959; PMCID: PMC5600148.
2. National Library of Medicine, National Center for Biotechnology Information web site: <https://www.ncbi.nlm.nih.gov>.
3. Bader, Jonathan, et al. "Novel Approaches Toward Scalable Composable Workflows in Hyper-Heterogeneous Computing Environments." *Proceedings of the SC'23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis*. 2023.

Simulation of Neurotransmitter Flow in a Presynaptic Bouton of a Neuron with Cellular Automaton

Andrzej Bielecki, Maciej Gierdziewicz

AGH University of Krakow, Al. Adama Mickiewicza 30, 30-059 Kraków, Poland

{bielecki, gierdzma}@agh.edu.pl

Keywords: simulation, neurotransmitter, synapse, cellular automaton

1. Introduction

The process of reaction-diffusion of neurotransmitter (NT) inside the presynaptic bouton of the chemical synapse has been modeled by using cellular automaton. The environment in the chemical synapse is heterogeneous, so efficient tools are needed to model the flow of NT. Partial differential equations models work accurately but at a high cost due to their computational complexity. A novel approach has been proposed below, allowing us to improve the performance of the algorithm, presumably at the price of a slight decrease in precision.

2. Description of the problem

The transmission of nerve impulses in the presynaptic bouton may be split into: the appearance of synaptic vesicles, diffusion, exocytosis and NT reuptake from the cleft. The first three stages are considered in this paper. So far, they were modeled by using ordinary differential equations (ODE) [1] and partial differential equations (PDE) [2].

The analysis performed either by using PDE or a cellular automaton required that a three-dimensional mesh of the object should be constructed. The same mesh has been used in both cases; it contained about 40000 tetrahedrons and the active (exocytosis) zone on its surface – about 400 triangles. Since the simulation by using PDE requires solving large system of equations in each time step, the calculations take, of necessity, long time.

It seems that switching to cellular automata like in [3] precisely describes transport processes and the algorithm has lower complexity. The research described in this work verified the above hypothesis.

3. Related work

The model of the presynaptic bouton for solving PDE to simulate NT flow was designed in two dimensions [4] and in three dimensions [5] with mixed continuous and discrete mathematical models. In another work [6] the model contained a supply zone and a single release site. Later on, distorted symmetry and separate pools of NT vesicles were added.

4. Solution of the problem

The main program loop in PDE method iterated on tetrahedrons and, inside each tetrahedron, on pairs on its vertices. As mentioned above, that required solving the system of linear equations in each time step. If the cellular automaton was employed, each of the tetrahedrons was connected by common faces with 3 or 4 other elements (“neighbors”). The only loop structures in the program was the main time loop and the inside loop iterated over the tetrahedrons. The innermost loop over the neighbors had constant complexity. The equation used to model the reaction-diffusion (Equation 1) was a partial differential equation:

$$\partial_t q = D\nabla^2 q + R(q) \quad (1)$$

where q is the vector of unknowns – the density of NT, D is the diagonal matrix of diffusion coefficients (all of them were equal in this case), and R contains all local reactions (here –exocytosis). The simulated time was 0.1s and the modeled frequency of stimulation was 100Hz. The Ares supercomputer installed in Academic Computer Center Cyfronet AGH was used. The program was written in Python, primarily as single-threaded. The amount of time used for calculations was about 6 hours, and the memory was around 6 GB.

One of the resulting graphs (Fig. 1) is very similar to those present in the cited literature. The process of depleting the synapse of the neurotransmitter vesicles was captured.

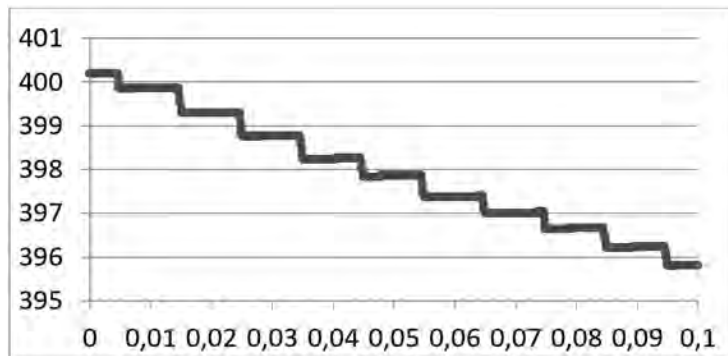


Fig. 1. The change in the amount of the neurotransmitter in the presynaptic bouton in time. Horizontal axis – time [ms]; vertical axis – equivalent of the number of synaptic vesicles.

5. Conclusions

We applied the cellular automaton methodology to simulate diffusion and reaction of neurotransmitter in the presynaptic bouton of the neuron. Time complexity of the algorithm is lower than that of partial differential equation, at the cost of a slight decrease in precision.

Acknowledgements. The numerical experiment was possible through computing allocation on the Ares system at ACC Cyfronet AGH under the grant *plgneuron2023*.

References

1. F. Aristizabal and M. I. Glavinovic: “Simulation and parameter estimation of dynamics of synaptic depression”, *Biological Cybernetics*, 90(1), 2004, pp. 3-18.
2. A. Bielecki and M. Gierdziewicz: „Simulation of neurotransmitter flow in three dimensional model of presynaptic bouton”, *Lecture Notes in Computer Science*, 12139, 2020, pp. 132-143.
3. A. Strader, K. E. Schubert, M. Quintana, E. Gomez, J. Curnutt and P. Boston: „Estimation, Modeling, and Simulation of Patterned Growth in Extreme Environments”, in Arabnia, H. and Tran, Q. N.(eds.): „Software Tools and Algorithms for Biological Systems. Advances in Experimental Medicine and Biology”, Springer, New York, USA, 696, 2011.
4. A. Bielecki, P. Kalita, M. Lewandowski, and B. Siwek: “Numerical simulation for a neurotransmitter transport model in the axon terminal of a presynaptic neuron”, *Biological Cybernetics*, 102, pp. 489-502, 2010.
5. M. M. Knoedel et al.: “Synaptic bouton properties are tuned to best fit the prevailing firing pattern”, *Frontiers in Computational Neuroscience*, 8:101, pp. 1-16 2014.
6. A. Bielecki, M. Gierdziewicz and P. Kalita: „Three-dimensional model of signal processing in the presynaptic bouton of the neuron”, *Lecture Notes in Artificial Intelligence* 10841:3-14, 2018.

Magnetic Resonance T_1 - and T_2 -weighted Data Synthesis from Heterogeneous Sources: a Federated Learning Approach

Jan Fiszler^{1,2}, Dominika Ciupek¹, Maciej Malawski^{1,2}, Tomasz Pięciak^{3,1}

¹ Sano Centre for Computational Medicine, Kraków, 30-054 Poland

² AGH University of Krakow, al. Mickiewicza 30, 30-059 Kraków, Poland

³ ETSI Telecomunicación, Universidad de Valladolid, Valladolid, 47011 Spain

{j.fiszler, d.ciupek, m.malawski}@sanoscience.org, tpieciak@tel.uva.es

Keywords: magnetic resonance imaging, data synthesis, federated learning, deep learning

1. Introduction

Data synthesis (DS) is a new field in medical imaging, particularly in magnetic resonance imaging (MRI), that reduces the acquisition time or reconstructs missing samples from sparsely acquired data (Fig. 1 (a)). DS is widely utilized for brain tumor classification and radiotherapy planning [1]. To use machine learning (ML) models efficiently in DS, it is critical to integrate heterogeneous data from multiple centers. Federated Learning (FL) is a promising approach to address this issue while preserving privacy concerns (Fig. 1 (b)). This study explores the feasibility of using FL for image-to-image DS between T_1 - and T_2 -weighted brain scans.

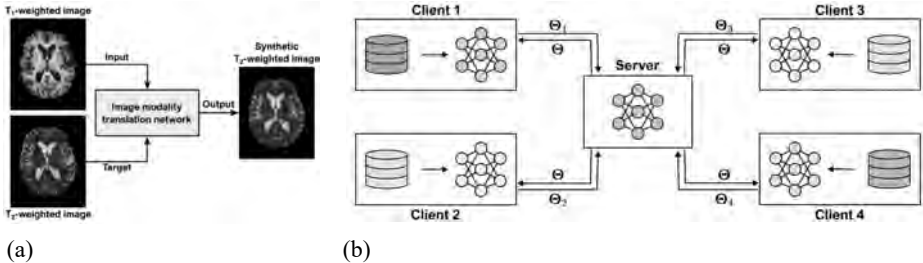


Fig. 1. (a) Visual representation of the translation process, (b) federated learning scheme involving four institutions (Θ_c – parameters of client c local model, Θ – parameters of the global model).

2. Federated learning

The FL is a decentralized method for training ML models without sharing the data and sensitive information. It is being used for translating medical images, with studies employing generative adversarial networks, basic aggregation algorithms, and more advanced techniques like split learning [2] and adversarial diffusion models [3]. Unlike traditional ML that gathers data from various sources onto a single server, FL only shares the features of locally trained models in client $c \in \{1, \dots, C\}$ to create a global model through aggregation [4]:

$$\Theta^{r+1} = \sum_{c=1}^C w_c \Theta_c^r \quad (1)$$

where: r – federated round, w_c – aggregation weight.

Recent studies suggest that the FL approach may perform comparably to those trained on centralized datasets and outperform models trained on individual institutional data. However, FL still requires significant computational resources and new technological solutions.

3. Results and conclusions

The study employed four publicly available datasets and tested three FL methods (FedAvg [4], FedAdam [5], and FedMRI [6]) with the U-Net model. Training a model on a personal computer takes ~ 2957 s/epoch, while on the Ares GPUs, it takes only ~ 10 s/epoch. The FL algorithm requires 128 epochs, which would take ~ 105 hours on a personal computer versus ~ 21 minutes on the supercomputer. Therefore, using HPC can accelerate the process up to 300 times.

Our study compared models trained on a single dataset versus those trained through the FL. We presented the relative error values in Tab. 1 and found that the FedMRI produced superior results across tested FL approaches and standalone ML models. Our analysis confirms the critical role of FL, as the simple FedAvg algorithm yields results as good as those of traditional ML algorithm.

Tab. 1. Relative error values (given in %) for every model trained on a single dataset and every aggregation method (T_1 -weighted to T_2 -weighted image translation). Bolded values indicate the lowest error for each dataset; underlined the error for training and testing the model on the same dataset.

	HCP MGH	HCP WU-Minn	OASIS-3	BraTS-HGG	BraTS-LGG	Average
HCP MGH	<u>17.0</u>	265.9	60.7	43.5	52.3	87.9
HCP WU-Minn	29.8	<u>50.9</u>	42.8	34.2	42.6	40.1
OASIS-3	39.9	122.9	<u>32.2</u>	46.5	54.5	59.2
BraTS-HGG	34.4	127.3	52.5	<u>28.1</u>	32.3	54.9
BraTS-LGG	35.1	158.1	52.4	26.1	<u>27.6</u>	59.9
FedAvg	28.8	96.3	38.1	26.3	30.5	44.0
FedAdam	29.4	76.6	37.1	24.1	27.7	39.0
FedMRI	19.0	54.3	35.9	23.4	32.1	32.9

Acknowledgements. The numerical experiments were possible through computing allocation on the Ares and Athena systems at ACC Cyfronet AGH under the grant PLG/2023/016117. Tomasz Pieciak acknowledges the Polish National Agency for Academic Exchange for grant PPN/BEK/2019/1/00421 under the Bekker programme and the Ministry of Science and Higher Education under the scholarship for outstanding young scientists (692/STY/13/2018). This work is supported by the European Union’s Horizon 2020 research and innovation program under grant agreement Sano No 857533 and the project of the Minister of Science and Higher Education (MEiN/2023/DIR/3796).

References

1. Y. H. Moshe, et al.: “Handling missing MRI data in brain tumors classification tasks: Usage of synthetic images vs. duplicate images and empty images”. *Journal of Magnetic Resonance Imaging*, 2023.
2. J. Wang, et al.: “FedMed-GAN: Federated domain translation on unsupervised cross-modality brain image synthesis”. *Neurocomputing*, 546, 2023.
3. M. Özbey, et al.: “Unsupervised medical image translation with adversarial diffusion models”. *IEEE Transactions on Medical Imaging*, 2023.
4. B. McMahan, et al.: “Communication-efficient learning of deep networks from decentralized data”. *Artificial Intelligence and Statistics*, PMLR, 2017, pp.1273-1282.
5. S. Reddi, et al.: “Adaptive federated optimization”. *arXiv preprint arXiv:2003.00295*, 2020.
6. C. M. Feng, et al.: “Specificity-preserving federated learning for MR image reconstruction”. *IEEE Transactions on Medical Imaging*, 42(7), 2023, pp.2010-2021.

Deep Learning Classification of Cytology Images with Uncertain Training Data

Jan Krupiński³, Szymon Mazurek¹, Krystian Strzałka¹, Maciej Wielgosz², Jakub Caputa¹, Rafał Frączek², Michał Karwatowski², Jakub Grzeszczyk¹, Daria Łukasik¹, Paweł Russek², Agnieszka Dąbrowska-Boruch², Ernest Jamro², Marcin Pietron², Sebastian Koryciak², Kazimierz Wiatr^{1,2}

¹ ACC Cyfronet AGH, Nawojki 11, 30-950 Kraków, Poland

² AGH University of Krakow al. Mickiewicza 30, 30-059 Kraków, Poland

³ Cracow University of Technology, Warszawska 24, 31-155 Kraków, Poland

{rafalfr, mkarwat, wielgosz, russek, adabrow, jamro, pietron, koryciak, wiatr}@agh.edu.pl, {j.krupinski, k.strzalaka, s.mazurek, d.lukasik, j.grzeszczyk, j.caputa}@cyfronet.pl

Keywords: cytology, classification, deep learning

1. Introduction

Training data is crucial for Deep Learning (DL). This also holds for the described system, which classifies cancer cells in cytological images. The purpose of the classification is to assist the veterinarian in making a cancer diagnosis. During tests of the system [1], we found out that the classification results performed by DL are moderate. One of the reasons for the low classification accuracy turned out to be an uncertain training set, especially non-diagnostic cells.

2. Description of the problem

Non-diagnostic cells are difficult to assign to a proper class. Examples of non-diagnostic cells include cells cut off at the ends of the image, damaged cells, cells that cannot be clearly assigned to a known class as e.g. cells are at the initial stage of cancer development. Moreover, as a result of the tests, it turned out that the pathologist in most cases was not able to correctly classify individual cells taken out of context (image of only a single cell obtained after the first step of classification [1]), even if he had previously made a similar correct classification of the same cells spotted on the whole image. Consequently, the pathologist gives in to suggestions: he looks globally but classifies locally. Moreover, even with the given image context, the same cells may be classified as e.g. cancerous or as non-diagnostic.

3. Related work

A similar problem arises in cervical cytology in humans, where anomalous cells can be classified into several classes based on the implied cancer risk [2]. The boundaries between those classes aren't precise, and to tackle this problem in DL a label smoothing algorithm has been used for classification [3].

4. Solution of the problem

Non-diagnostic cells are very often classified differently even by the same pathologist. A solution to this problem may be to ignore errors in the classification of non-diagnostic cells. If we swap a cancer cell with a non-diagnostic one during the inference process, we obtain a lot of non-diagnostic cells (false negative), which should cause concern for a vet doctor, but is unlikely to result in an incorrect diagnosis. Usually, there are many (same type of cancer) cells

in the image, so such a mistake is not harmful because different cancer cells may still be properly classified. Therefore, such a mistake is much more acceptable than misclassifying a non-cancerous cell as cancerous (false positive). A false positive classification error may be obtained when a non-diagnostic (or non-annotated) cell is classified as a cancer cell. This, however, does not mean misclassification but rather indicates that the pathologist was not certain during annotation process. Summing up, ignoring non-diagnostic misclassifications in most cases does not influence the inference process.

The results for the classification described above are given in Tab. 1. The accuracy of the classification, when ignoring non-diagnostic cells, is at the level of 99% and is much better than when counting non-diagnostic cells as a separate class.

In the web application where inference results are presented, it is possible to easily change the classification threshold. In the case of a high threshold, we want to have a lot of non-diagnostic cells, so the probability of false positives should be low (classification precision should be high). To achieve this, the ResNet18 (the second step of classification [1]) was retrained for additional 64 epochs (with early stopping) employing a scheme where the true cell label was changed to “non-diagnostic” if the model output did not match the true cancer class.

Tab. 1. Classification results with and without (ignoring) non-diagnostic cells.

	with non-diagnostic			ignore non-diagnostic		
model	accuracy	f1 macro	f1 weigh.	accuracy	f1 macro	f1 weigh.
segmentation	37.32%	43.09%	33.08%	62.91%	64.02%	67.46%
detection (th=0.75)	49.00%	39.02%	33.75%	99.15%	82.12%	99.40%
detection (thr=0.25)	25.24%	46.64%	16.08%	87.08%	55.46%	89.79%
CNN	23.41%	31.51%	14.92%	86.30%	64.42%	89.86%
ResNet18	20.18%	32.36%	14.53%	71.44%	33.22%	79.55%
ResNet18 extra training	20.58%	26.37%	12.21%	93.67%	78.37%	96.36%

5. Conclusions

Results presented in Tab. 1 show that the key factor for the classification accuracy are non-diagnostic cells, which in most cases are misclassified or mis-annotated. Ignoring these classification errors significantly improves the accuracy.

Acknowledgements. PLGrid grant ID: plglaosi22/plglaosi23. All models were trained on the PLGrid infrastructure, using the Athena supercomputer.

References

1. J. Krupiński et al. “Comparison of One and Two-Step Deep Learning Models for Cytology Image Classification”, KU KDM 2024.
2. D. Solomon, D. Davey, R. Kurman, A. Moriarty, D. O’Connor, M. Prey, et al.: „Bethesda System. Terminology for reporting cervical cytology.” JAMA 2002 Apr24;287(16):2114-9.
3. Y. Nambu, T. Mariya, S. Shinkai, M. Umemoto, H. Asanuma, I. Sato, Y. Hirohashi, T. Torigoe, Y. Fujino, T. Saito: „A screening assistance system for cervical cytology of squamous cell atypia based on a two-step combined CNN algorithm with label smoothing.” Cancer Med. 2022 Jan; 11(2):p. 520-529.

Comparison of One and Two-Step Deep Learning Models for Cytology Image Classification

Jan Krupiński³, Szymon Mazurek¹, Krystian Strzałka¹, Maciej Wielgosz², Jakub Caputa¹, Rafał Frączek², Michał Karwatowski², Jakub Grzeszczyk¹, Paweł Russek², Agnieszka Dąbrowska-Boruch², Ernest Jamro², Marcin Pietron², Sebastian Koryciak², Kazimierz Wiatr^{1,2}

¹ ACC Cyfronet AGH, Nawojki 11, 30-950 Kraków, Poland

² AGH University of Krakow, al. Mickiewicza 30, 30-059 Kraków, Poland

³ Cracow University of Technology, Warszawska 24, 31-155 Kraków, Poland

{rafalfr, mkarwat, wielgosz, russek, adabrow, jamro, pietron, koryciak, wiatr}@agh.edu.pl, {j.krupinski, k.strzalaka, s.mazurek, d.lukasik, j.grzeszczyk, j.caputa}@cyfronet.pl

Keywords: cytology, classification, deep learning

1. Introduction

This paper describes the improvements of the deep learning (DL) image classification system [1]. The system is employed to assist veterinary doctors in detection of dogs' skin cancer based on the cytology images. This work compares different DL methods: one-step and two-step. One of the most popular one-step methods is YOLO, for which cells detection (or segmentation) and classification are implemented simultaneously. For two-step methods, the first step finds locations of the cells in the image (same as for one-step method). Then, in the second step, each cell is classified separately by another DL module. Furthermore, in this paper, classification results for image segmentation and detections are compared.

2. Description of the problem

In the previous version of the system [1], only one-step neural network was employed. The initial results of this system were satisfactory when it was trained and tested on a set of photos taken by the only one veterinarian. However, the results turned out to be moderate in the case of foreign images. This was mostly due to the relatively small training set and e.g. different zoom / light conditions. The assumption of the previous system was the use of cell segmentation (see Fig. 1). In the case of segmentation, the boundaries of each cell are outlined both in the annotation (and training) process and in the inference process. Segmentation was used to avoid ambiguity of which cell was selected. For detection (a rectangular bounding box), in some cases there may be an overlap phenomenon, i.e. we are not able to indicate which cell is currently of interest.

3. Solution of the problem

The above-mentioned phenomenon of ambiguity is rare and, as practical tests have shown, without great significance. However, the choice between segmentation and detection is crucial for the accuracy of the classification process. This holds as the segmentation process is much more complex than detection, and consequently the neural network allocates much more resources to segmentation, which results in worse classification accuracy. It is worth emphasizing that at the initial stage of the project, it was not obvious which of these methods is better. For example, the exact cell shape (envelope) may be somehow used within the neural network in the classification process. However, the results shown Table 1 indicate that the segmentation process comes at the cost of worse classification accuracy.

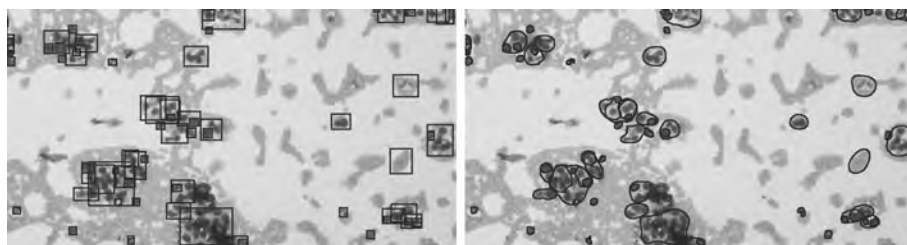


Fig. 1. Comparison of image detection (left) and segmentation (right).

Additionally, this article validates implementation results for two-step methods. As part of the neural networks of the second step, many different solutions were tested, of which two networks stand out: (Tab. 1, row 5) ResNet18 and (row 4) a small CNN with 130k parameters (2 convolutional layers with batch normalization and max pooling + 1 dense layer). For the segmentation (row 1) SwinTransformer [3] was employed as a backbone in the CBNetV2 framework [4]. For the detection (rows 2 and 3) YOLOv8 was employed. Slightly better results for smaller version of the dataset (and number of classes) are given in [5].

Tab. 1. Classification results for different methods and different thresholds (th).

#	model	accuracy	f1 macro	f1 weigh.
1	segmentation (th=0.75)	37.32%	43.09%	33.08%
2	detection (th=0.75)	49.00%	39.02%	33.75%
3	detection (th=0.25)	25.24%	46.64%	16.08%
4	2-nd step: CNN	23.41%	31.51%	14.92%
5	2-nd step: ResNet18	20.18%	32.36%	14.53%

5. Conclusions

The detection results are better than segmentation used in [1], at the expense of slight ambiguity as to which cell the classification applies to. The use of two-stage classification does not bring any additional benefits in classification accuracy, this is in contradiction with e.g. [2].

Acknowledgements. PLGrid grant ID: plglaosi22/plglaosi23. All models were trained on the PLGrid infrastructure, using the Athena supercomputer.

References

1. J. Grzeszczyk, et al.: "Segmentation of the veterinary cytological images for fast neoplastic tumors diagnosis", arXiv:2305.04332, 7 May 2023.
2. Y. Nambu, T. Mariya, S. Shinkai, M. Umemoto, H. Asanuma, I. Sato, Y. Hirohashi, T. Torigoe, Y. Fujino, T. Saito: "A screening assistance system for cervical cytology of squamous cell atypia based on a two-step combined CNN algorithm with label smoothing." *Cancer Med.* 2022 Jan; 11(2): pp. 520-529.
3. Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, B. Guo: "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows.", 2021.
4. T. Liang, X. Chu, Y. Liu, Y. Wang, Z. Tang, W. Chu, J. Chen, H. Ling: "CB-NetV2: A Composite Backbone Network Architecture for Object Detection.", 2021.
5. J. Caputa, D. Łukasik, M. Wielgosz, M. Karwatowski, R. Frączek, P. Russek, K. Wiatr: "Fast Pre-Diagnosis of Neoplastic Changes in Cytology Images Using Machine Learning.", *Applied Sciences*, 2021; 11(16):7181. <https://doi.org/10.3390/app11167181>.

Integrating Data Repositories with HPC Resources for Execution of Simulation Models

Taras Zhyhulin¹, Karol Zając¹, Piotr Nowakowski^{1,2}, Maciej Malawski^{1,2}, Marek Kasztelnik², Piotr Połec², Marian Bubak^{1,2}

¹ Sano Centre for Computational Medicine, Czarnowiejska 36, 30-054 Kraków, Poland

² ACC Cyfronet AGH, ul. Nawojki 11, 30-950 Kraków, Poland

{t.zhyhulin, k.zajac, p.nowakowski, m.malawski, m.bubak}@sanoscience.org,
{m.kasztelnik, p.polec}@cyfronet.pl

Keywords: data repositories, high performance computing, medical simulations

1. Introduction

The digital twin concept is gaining traction in research, demanding substantial computational power for simulations and a lot of memory resources for storing models, input and output data. Sano Centre for Computational Medicine, in collaboration with ACC Cyfronet AGH, is actively developing tools to optimize high performance computing (HPC) resources and facilitate easy sharing and publication of the data. The architecture of the platform we are developing, is described in [1]. Our focus is on providing scientists with a user-friendly toolkit for seamless model execution. This paper introduces the integration of the Model Execution Environment platform with data repositories, streamlining data management for researchers.

2. Description of the problem

Harnessing HPC resources necessitates specific expertise and extensive data management, posing challenges for researchers. Additionally, sharing processed data and research results among teams demands adherence to fair involvement rules, involving external services and consuming valuable time. Moreover, storing data locally utilizes a significant amount of storage. Our aim is to alleviate these challenges by providing a comprehensive platform for efficient data management.

3. Related work

An example of a possible solution to this problem is described in [2]. The pySUMMA model API integrates the HydroShare data repository with the SUMMA model to provide it with necessary input data and enables user to configure the model conveniently. It is executed on the Jupyter Notebook, which allows the utilization of HPC resources for execution; however, it does not provide support for other model execution environments or languages.

4. Solution to the problem

Our platform streamlines medical simulation execution, integrating with Dataverse and Zenodo. The Model Execution Environment requires configuration of data repository instance credentials to address API calls. This data allows the HPC infrastructure to seamlessly communicate with the data repository via API. It eliminates the role of intermediaries, enhancing efficiency and collaboration. Customizable rules in data repositories ensure fair sharing, safeguarding confidentiality during research and enabling straightforward publication.

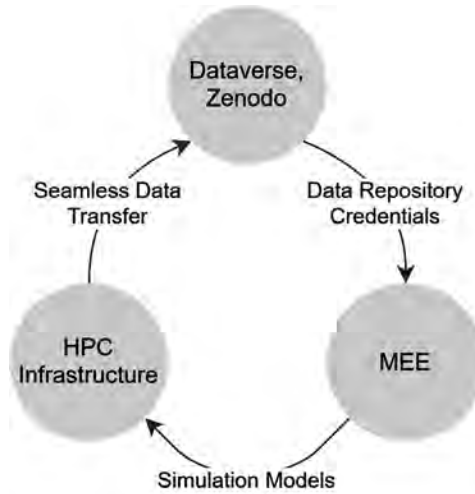


Fig. 1. Process of Data Transfer.

5. Conclusions

Our research has resulted in a sophisticated toolkit optimizing the efficiency of medical research. The advantage of our solution is the possibility of integration with multiple data repository types, no restrictions on executed model type, and straightforward but multifunctional configuration. Moreover, the provided integration mechanism allows two-way communication with the data repository, facilitating seamless upload of the results of the work. Future plans include broader integration, utilizing Digital Object Identifiers for data retrieval, eliminating complex configuration requirements. We are committed to enhancing the accessibility and ease of use of our toolkit.

Acknowledgements. This publication is supported by the European Union’s Horizon 2020 research and innovation programme under grant agreement Sano No 857533. This publication is supported by Sano project carried out within the International Research Agendas programme of the Foundation for Polish Science, co-financed by the European Union under the European Regional Development Fund. This publication is (partly) supported by the European Union’s Horizon 2020 research and innovation programme under grant agreement ISW No 101016503. We gratefully acknowledge Polish high-performance computing infrastructure PLGrid for providing computer facilities and support within computational grant no. PLG/2023/016227.

References

1. Bubak M., Czechowicz K., Gubała T., Hose D. R., Kasztelnik M., Malawski M., Meizner J., Nowakowski P. and Wood S. The EurValve model execution environment, *Interface Focus* 11:20200006. 20200006. <http://doi.org/10.1098/rsfs.2020.0006>.
2. Young-Don Choi, Jonathan L. Goodall, Jeffrey M. Sadler, Anthony M. Castronova, Andrew Bennett, Zhiyu Li, Bart Nijssen, Shaowen Wang, Martyn P. Clark, Daniel P. Ames, Jeffery S. Horsburgh, Hong Yi, Christina Bandaragoda, Martin Seul, Richard Hooper, David G. Tarboton: Toward open and reproducible environmental modeling by integrating online data repositories, computational environments, and model Application Programming Interfaces, *Environmental Modelling & Software*, Volume 135, 2021, 104888, ISSN 1364-8152, <https://doi.org/10.1016/j.envsoft.2020.104888>, <https://www.sciencedirect.com/science/article/pii/S1364815220309452>.

Research towards Virtual Human Twins

Marian Bubak^{1,2}, Marek Kasztelnik¹, Maciej Malawski^{1,2}, Jan Meizner^{1,2},
Piotr Nowakowski^{1,2}, Piotr Połec¹

¹ACC Cyfronet AGH, ul. Nawojki 11, 30-950 Kraków, Poland

²Sano Centre for Computational Medicine, Czarnowiejska 36, 30-054 Kraków, Poland

{m.bubak, m.malawski, j.meizner, p.nowakowski}@sanoscience.org,
{m.kasztelnik, p.polec}@cyfronet.pl

Keywords: high performance computing, medical simulations, personalized medicine, digital twin, virtual human twin

1. Introduction

According to a common vision of the future medicine, health systems should enable personalized, optimized health promotion, disease prevention, diagnosis, and finally efficient treatment for the benefit of patients [1]. Moreover, health care providers should act accordingly to elaborated standards, in a conscious, empowered, committed and responsible way. To realize this idea, we need good mathematical models of human physiology implemented in a form of computer simulation modules as well as availability of health data. As the result we could expect more profits from research and optimized health treatment processes, informed, empowered, engaged and responsible citizens, and finally, more economically organized the whole health service system.

2. Description of the problem

Currently, the way to personalized medicine leads through elaboration of so called virtual human twin (VHT), which is a specific implementation of the idea of digital twin - a digital representation of a physical object, person, or process, placed in context in a digital version of its environment. VHT is an integrated, multi-scale, multi-temporal and multi-disciplinary representation of quantitative human physiology and pathology. To put the idea of VHT in practice, according to expert of the EU project EDITH [2], we require an inclusive ecosystem of digital twins in healthcare, implementation a federated cloud-based repository, gathering human digital twin resources such as models, data sets, algorithms, practices, and designing the architecture of a simulation platform to facilitate the transition towards personalized medicine.

3. Related work

A series of research in this direction started by Peter Hunter with the Physiome Project [3] and was continued in the framework of the Virtual Physiological Human (VHP), sometimes equated with „in silico medicine” - a field involving the use of individualized computer simulations based on physiology in aspects of prevention, diagnosis, prognostic assessment and treatment of disease and biomedical product development [4].

4. Solution to the problem

As our contribution to development of the VHT idea, at first, we have analyzed internal structure and functional requirements of typical applications simulating human physiology gathered by partners of the EDITH project [2]. This formed a basis for elaboration of a demonstrator of the execution subsystem of the VHT ecosystem. It is a software system agnostic to

the supported classes and formats of data items, easy to support a comprehensive data repository where various data items may be queried, retrieved, and fed into the computational models which constitute the simulation workflow, to run on classical HPC resources for scale-out studies which involve processing large amounts of data and “parameter study” types of computations. It enables model versioning: previous versions of the model are stored and may be referred to if needed, as well as reproducibility of computer simulations. The system also enables execution of computational models controlled by a set of scripts with a versioning system enabling collaborative editing and tagging specific versions that may be later selected to suit the researchers’ need. It provides a straightforward way to display, download and analyze simulation results. The functionality of the demonstrator was successfully validated with a set of typical VHT modules on the ACC Cyfronet HPC resources.

5. Conclusions and future work

Our research has resulted in elaboration of an demonstrator which enables running on HPC resources modules of VHT and may be integrated with a repository of models and data. We consider it as the first step towards elaboration of the whole VHT ecosystem.

We expect to demonstrate the value of VHT paradigm by evaluating model selection in the treatment of two types of strokes in a multicenter clinical trial that will compare treatment and patient outcomes in situations with and without digital twin in healthcare (DTH) availability. It will be achieved in the framework of the EU project GEMINI [5] in which our team is contributing to elaboration of good practices to produce complex medical computer models considering their ability to run on HPC resources. Based on the elaborated practices, virtual twin models will be created and then be run for a large cohort of patients on the ACC Cyfronet AGH infrastructure. We also expect that Artificial Intelligence solution will considerably enhance virtual human twins. Our team just started research in the EU project Meetween [6] which will develop new solutions for language-free collaboration between people using comprehensive, integrated algorithmic capabilities offered by fundamental models and self-supervised training on large data sets. This in turn, will enable a real-time machine learning-powered translation, summarization, and virtual assistant services for online meetings. We will create tools to test the efficiency and correctness of AI models as well as software environment for experiments with the machine learning processes on a large scale using the HPC infrastructure.

Acknowledgements. We acknowledge the support of EU under grants EDITH No 101083771 (Digital Europe), GEMINI No 101136438 (HORIZON), Meetween No 101135798 (HORIZON), and Teaming Sano No 857533, as well as ACC Cyfronet AGH for providing computer facilities and support within computational grant no. PLG/2023/016723.

References

1. B Vicente, A. M., Ballensiefen, W. & Jönsson, J.-I. How personalised medicine will transform healthcare by 2030: the ICPeMed vision. *J. Transl. Med.* 18, 180 (2020), <https://doi.org/10.1186/s12967-020-02316-w>.
2. EDITH – European Virtual Digital Twin, EU Project, Digital Europe, <https://www.edith-csa.eu/>.
3. Physiome Project - <https://www.auckland.ac.nz/en/abi/our-research/research-groups-themes/physiome-project.html>.
4. VPH Institute - <https://www.vph-institute.org/>.
5. GEMINI - A Generation of Multi-scale Digital Twins of Ischaemic and Haemorrhagic Stroke Patients, EU Project, HORIZON, <https://dth-gemini.eu/>.
6. Meetween - My Personal AI Mediator for Virtual MEETings BETWEEN People, EU Project, HORIZON.

Numerical Simulations of Short and Long Gamma-Ray Bursts on Ares Supercomputer

Gerardo Urrutia, Agnieszka Janiuk

Centre for Theoretical Physics, Polish Academy of Sciences, Al. Lotnikow 32/46, 02-668 Warsaw, Poland

{gurrutia, agnes}@cft.edu.pl

Keywords: accretion, jets, Gamma Ray Bursts, computer science

1. Introduction

Gamma-ray bursts (GRBs) are the most energetic explosions after the big bang. It emits very luminous pulses of γ -radiation whose energy per second is twenty orders of magnitude greater than the sun [1]. These cosmic explosions are located at cosmological distances and every year are detected at least 300 GRBs [2,3].

The GRB emission is produced by the energy dissipation of powerful relativistic jets. These jets are mainly arising from massive stars before their collapse, or launched after binary neutron star mergers. The origin and propagation of GRB involve complicated physical systems whose interpretation requires the solution of magneto hydrodynamics equations, nuclear network evolution and radiative processes. Numerical simulations are useful tools to solve the most realistic scenarios. Notwithstanding, these simulations are computationally expensive. For this reason, we need to allocate our numerical experiments in supercomputers.

2. Description of the problem

The GRB problem, have been become more challenging due to the technological improvement of the new generation of observatories that provide high-resolution information with some times change the paradigm of our understanding. For example, in August 2017, was detected the first gamma-ray burst (GRB) associated with gravitational wave (GW) emission [4]. The GRB modelling has been improved in the last few years. However, more scenarios should be explored to improve our understanding. In addition, the GRBs are a multiscale problem, each part of the evolution should be solved with a different approach. One problem that I will describe in the talk is the self-consistency between simulations at large and small scales.

Typically, the evolution of GRB jet is solved separately in two regions, at small or large scales. Small-scale simulations are frequently focused on the physical conditions for jet launching, and the jet evolution is followed during milliseconds [5]. However, the implications at large scales are not solved. The jet crosses a post merger outflow, but its description is often simplified, then the consistency with the previous evolution of jet/outflows at small-scales is washed out. Here, the main challenge is to connect the results of small-scale simulations to a numerical setup that solves for large-scale.

3. Related work

Our implementation shows that the dynamics of the jet suffers substantial modifications during its propagation in a more realistic environment (Fig. 1).

4. Solution of the problem

For the dynamics of short GRB, we have to connect results of small-scale simulations to large scale. We perform the accretion dynamics by General Relativistic Magnetohydrodynamics

simulations, performed with HARM_Cool [9]. We follow the wind evolution by tracers which store the evolution of temperature and electron fraction. Before the injection in the large-scale simulation, we reprocessed the tracers into a nuclear network code SkyNet, and we recovered the gas pressure after the r-process with the inversion of the Helmholtz equation. The final distributions of material were imposed in the SRHD code mezcald [10] that solves the hydrodynamics equations employing an adaptive mesh refinement methods. For the simulations, we use mainly cpp, gfortran and open mpi.

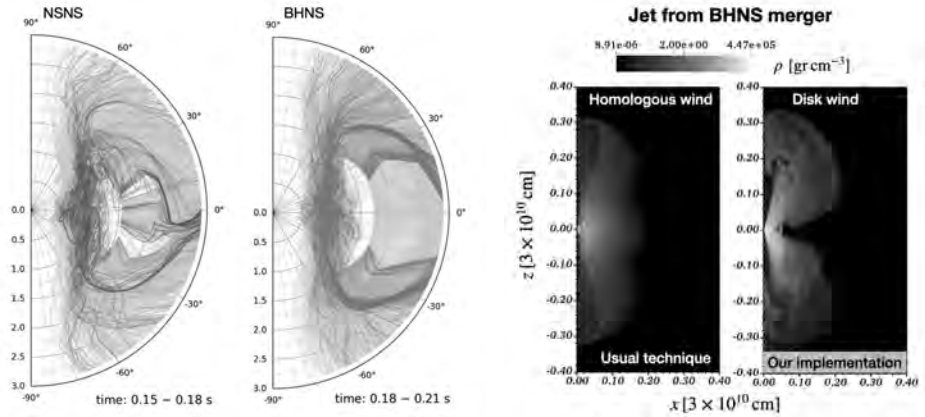


Fig. 1. Left: Trajectories of tracers particles arising from accretion disc of two configuration of GRB progenitors (NSNS merger and BHNS merger). Right: Comparison between usual methodology and our new implementation. See for example [8].

5. Conclusions

We improve our studies of jet evolution by connecting small-scale simulations with large-scale evolution, by importing tracers which include the evolution of nuclear r-process.

Acknowledgements. This work/research has been supported by 2019/35/B/ST9/04000 from Polish National Science Center. The numerical experiment was possible through computing allocation on the Ares supercomputer system at Cyfronet under the grant PLG/2024/016972.

References

1. Kumar P., Zhang B., 2015, PhysRep, 561, 1.
2. Ghirlanda, G., Salafia, O. S., Paragi, Z., et al. 2019, Science, 363, 968.
3. Mészáros P., Gehrels N., 2012, Research in Astronomy and Astrophysics, 12, 1139.
4. Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, PhRvL, 119, 161101.
5. Janiuk, A., James, B., A&A 668, A66 (2022).
6. Gottlieb O., Nagakura H., Tchekhovskoy A., Natarajan P., Ramirez-Ruiz E., Banagiri S., Jacquemin-Ide J., et al., 2023, ApJL, 951, L30.
7. Soker N., 2023 Res. Astron. Astrophys. 23 121001.
8. Urrutia G., Janiuk A., Nouri F. & James B., 2023, submitted to APJ.
9. Janiuk, A., Proga, D., ApJ 675, 1, 519 (2008).
10. De Colle F., Granot J., Lopez-Camara D., Ramirez-Ruiz E., 2012, ApJ, 746,122.

Viscous Torque in Turbulent Magnetized AGN Accretion Disks and Its Effects on EMRI's Gravitational Waves

Agnieszka Janiuk, Fatemeh Hossein Nouri

Center for Theoretical Physics, Al. Lotnikow 32/46, 02-668 Warsaw, Poland

agnes@cft.edu.pl

Keywords: computational astrophysics, numerical relativity, magnetohydrodynamics

1. Introduction

The mergers of supermassive black holes (SMBHs) produce MHz gravitational waves (GW), potentially detectable by future Laser Interferometer Space Antenna (LISA). Such binary systems are usually embedded in an accretion disk at the centre of the active galaxy. Plasma environment imposes measurable imprints on the GW signal if the mass ratio of the binary is around $q \sim 10^{-4} - 10^{-3}$. The effect of the gaseous environment on the GW signal is strongly dependent on the disk parameters, so future low-frequency GW detections will provide information about the physics of accretion disks. We investigate this effect by modelling magnetized disks around primary massive black hole. Using GRMHD HARM-COOL code, we perform 2D and 3D simulations. In our numerical analysis, we study the angular momentum transport and turbulence generated by the magnetorotational instability (MRI). We quantify the disk's effective viscosity and its evolution over time, and compute the relativistic viscous torque on a low-mass secondary black hole. Hence we estimate the GW phase shift due to the gas environment.

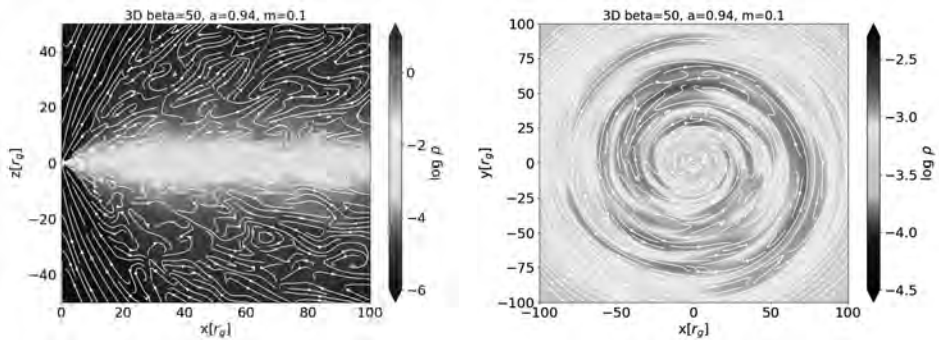


Fig. 1. The density profile and magnetic field lines for the meridional slice at $\phi = 0$ [left], and the equatorial slice of the same quantities for our 3D simulation [right] at the final snapshot.

2. Description of the problem

We use our version of the GR MHD code HARM described in [1] and [2], which uses numerical algorithms developed initially by [3]. HARM uses a conservative shock-capturing scheme, with fluxes calculated using classical Harten-Lax-van Leer method. The constrained transport maintains divergence free magnetic field. The background spacetime is fixed by the Kerr metric of the black hole with constant mass and angular momentum. The hydro equations are evolved in the modified spherical Kerr-Schild coordinates that are non-singular on the horizon. The radial and angular maps from Kerr-Schild (KS) coordinates to modified Kerr-

Schild (MKS) are used to increase the resolution close to the black hole and the equatorial plane, respectively, to resolve the thin disk accurately.

3. Related work

Several groups have studied the interaction of the gaseous circumbinary disk with the binary system with the 2D and 3D non-relativistic hydrodynamic simulations. They include the viscous gas described by the Navier-Stokes fluid equations, interacting with an equal mass circular binary, assuming the binary orbit to remain fixed (Moody et al. 2019; Tiede et al. 2020; Mahesh et al. 2023) or evolving with time (Franchini et al. 2022, 2023). Depending on the assumptions for the binary orbit, grid resolution and circumbinary disk parameters, they found that the exchange of angular momentum between the disk and the binary may lead to binary expansion or shrinkage. Our work is one of a very few, to best of our knowledge, general relativistic MHD simulations related to this topic. Only very recently, a new fully relativistic formalism has been introduced by Cardoso et al. (2022) to study GW emission by extreme mass-ratio inspirals in non-vacuum curved spacetimes.

4. Solution of the problem

Our initial density configuration is based on prescription for thin disks, proposed in [4]:

$$\rho(r, \theta) = \rho_e \exp(-\alpha_{\text{disk}}^2 z^2 / H^2); \quad z = r \cos(\theta). \quad (1)$$

The initial poloidal magnetic field configuration is with the nonzero azimuthal component of the magnetic vector potential defined as

$$A\phi = r^{3/4} m^{5/4} / (m^2 + \cos(2\theta))^{5/8}. \quad (2)$$

We performed set of four 2-dimensional simulations with varying parameters for magnetic field inclination, initial gas-to-magnetic pressure ratio, and black hole spin. The standard resolution was 1056*528 grid points in radial and polar direction. To investigate the resolution and the third dimension effects we perform a high-resolution 2D simulation with 2016*1056 grid points, and a 3D simulation with a moderate resolution of 288*256*96.

5. Conclusions

We observed the disks with well-resolved MRI have an average α viscosity that varies around 0.1 – 0.25 during the second part of the evolution. The MRI is suppressed at the inner part of the disk, close to the primary BH, so the value of the α viscosity is negligible in this region. However, time-averaged α reaches $\approx 0.1 - 0.2$ at larger radii where the fluid is turbulent and the MRI fastest growing mode is resolvable. We applied the numerical results from the GR MHD simulations to estimate the viscous torque using the GR-Hybrid approach for the general relativistic one-dimensional thin disk. We found that the time-averaged viscous torque can be as large as $\sim 1\%$ of the GW torque for a mass ratio of $q = 10-3$ at radii around $r \sim 100 r_g$, where α is maximal. This extra torque from the environment appears as faster shrinkage of the binary's orbit and phase shift in the GW signal.

Acknowledgements. This research has been supported by grant No. 2019/35/B/ST9/04000 from the Polish National Science Center. The numerical experiment was possible through computing allocation on the Ares system at ACC Cyfronet AGH under the grants plg-grb6.

References

1. A. Janiuk, 2017, ApJ, 837, 39.
2. K. Sapountzis, K. and A. Janiuk, 2019, ApJ, 873, 12.
3. C. Gammie, McKinney, J. C., & Tóth, G. 2003, ApJ, 589, 444.
4. I. K. Dhiingia, B. Vaidya, C. Fendt, 2021, MNRAS, 505, 3596.

Analysis of Accelerated Particles in PIC Simulations by Using Neural Networks

Gabriel Torralba Paz¹, Artem Bohdan², Jacek Niemiec¹

¹ Institute of Nuclear Physics Polish Academy of Sciences, PL-31342, Kraków, Poland

² Max-Planck-Institut für Plasmaphysik, Boltzmannstr. 2, DE-85748 Garching, Germany

gtorralba@ifj.edu.pl, artem.bohdan@ipp.mpg.de, jacek.niemiec@ifj.edu.pl

Keywords: neural networks, kinetic plasma simulations, particle acceleration

1. Introduction

Particle-In-Cell (PIC) codes are used for modelling collisionless plasma shocks at kinetic scales, resolving both the ion and electron dynamics. Within these shocks, particles can undergo acceleration by different mechanisms. Particle tracing, a tool implemented in many PIC codes, allows us to track individual plasma particles and gain a detailed comprehension of their energisation processes. However, pinpointing the exact mechanisms which affect particles proves difficult as there is a high level of uncertainty if one is to inspect multiple particle trajectories by eye. In a bid to enhance the accuracy and precision of particle tracing analysis, we propose for the first time to use neural networks (NNs). We propose two methods to analyse the particle tracing data: regression and anomaly detection by using convolutional NN. This methodology may prove useful in the analysis of PIC and hybrid-kinetic simulations.

2. Description of the problem

In high Mach number quasi-perpendicular shocks, the Buneman instability can be excited in the shock foot. This instability generates high-amplitude electrostatic waves, which can pre-accelerate electrons through the shock surfing acceleration process. To analyze this process, we use a dataset from our recent PIC simulation [1] and trace the momentum of 210,000 electrons over a span of 1,200 time steps, during which they travel through the region affected by the Buneman instability. Besides momentum, each particle is associated with a label representing the maximum kinetic energy that the particle attained along its trajectory. The majority of electrons are slightly heated by the electrostatic waves and only a few percent are pre-accelerated. This poses a challenge to analyze these data by conventional means.

3. Solution of the problem

We perform regression (supervised learning) and anomaly detection (unsupervised learning) by using convolutional NNs to analyse our large dataset. Regression compares the label computed in the dataset and the predicted label to train the NN. On the other hand, anomaly detection compares the momenta time series with a reconstruction of the same series created by the NN. If the reconstructed time series closely matches the original, the NN classifies the particle as a non-anomaly. Conversely, if there are significant differences, the particle is classified as an anomaly.

To build these NNs, we use the Python library Keras, which provides an easy-to-use API for the development of neural networks. To train and test our NNs, we use the Athena cluster installed at ACC Cyfronet AGH. A single node containing 8 Tesla V100-SXM2 GPUs is used for the calculations. The size of the dataset is ~20 GB and the training may take around 1 hour for regression and 2 hours for anomaly detection.

4. Results

Regression: Due to the imbalanced nature of the dataset, we use sample weighting and log scale transformation of the dataset to help the neural network during its training. The results are presented in Figure 1. We obtain a linear regression $y = 1.0015 (\pm 0.0004)x + 0.0080 (\pm 0.0006)$ with an $R^2=0.9886$, which is an excellent result for our imbalanced dataset.

Anomaly detection: To train the autoencoder used for anomaly detection, we only use non-energetic particles, that is, particles with a kinetic energy lower than $\gamma-1=0.07$. Table 1 shows the results for anomaly detection. The more diagonal the table, the better the results. Energetic particles are clearly distinguished as anomalies, as indicated by the ratio 1069/4. However, the NN occasionally struggles with some non-energetic particles, predicting them as anomalies, as shown by the ratio 1069/313. Nevertheless, the results seem promising, especially considering that anomaly detection uses no labels for NN training.

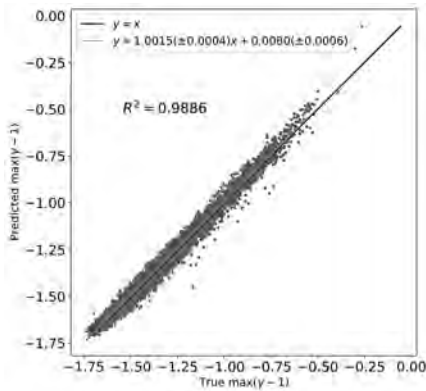


Fig. 1. Result for regression.

Tab. 1. Result for anomaly detection.

	Non-energetic	Energetic
Predicted Anomaly	313	1069
Predicted Non-anomaly	61614	4

5. Conclusions

Neural networks are used to analyse a dataset of traced particles obtained from recent PIC simulations. We investigate the use of regression and anomaly detection for the analysis of pre-accelerated electrons. Regression provides excellent results with a very high correlation between true and predicted labels, achieving an $R^2=0.9886$. The anomaly detection effectively identifies energetic particles with a success rate of 0.77.

Acknowledgements. This work has been supported by the National Science Centre through research project 2019/33/B/ST9/02569. The numerical experiment was possible through computing allocation on the Athena system at ACC Cyfronet AGH under the grant “plgastrogpu23”.

References

1. A. Bohdan, J. Niemiec, M. Pohl, Y. Matsumoto, T. Amano, M. Hoshino: *Astrophysical J.*, 2019, 878:5.

Particle-In-Cell Simulations of Merger Shocks in Galaxy Clusters

Oleh Kobzar¹, Jacek Niemiec², Takanobu Amano³, Masahiro Hoshino³,
Shuichi Matsukiyo⁴, Yosuke Matsumoto⁵, Martin Pohl^{6,7}

¹Astronomical Observatory, Jagiellonian University, ul. Orła 171, 30-244 Kraków, Poland

²Institute of Nuclear Physics, PAS, ul. Radzikowskiego 152, 31-342 Kraków, Poland

³Department of Earth and Planetary Science, University of Tokyo, 7-3-1 Hongo, Tokyo, Japan

⁴Faculty of Engineering Sciences, Kyushu University, 6-1 Kasuga-Koen, Kasuga, Fukuoka, Japan

⁵Department of Physics, Chiba University, 1-33 Yayoi, Inage-ku, Chiba, Japan

⁶Institute of Physics and Astronomy, University of Potsdam, 14476 Potsdam-Golm, Germany

⁷DESY, Platanenallee 6, 15738 Zeuthen, Germany

oleh.kobzar@uj.edu.pl

Keywords: shock waves, electron acceleration, magnetic waves, PIC simulations

1. Introduction

Shock waves in cosmic plasmas can be found on various scales. They are typically considered as places of the electromagnetic turbulence generation and acceleration of particles. Large scale merger shocks in galaxy clusters propagate with low Mach numbers ($M \ll 10$) in hot plasma with $\beta \gg 1$. X-ray and radio observations indicate the efficient electron acceleration at such shocks, although the mechanism of this process still requires further investigations.

2. Particle-In-Cell simulations as a key solution method

Since shock waves are strongly nonlinear, they can be studied with numerical methods, requiring extremely large computational resources utilized with high-performance codes. We used Particle-In-Cell (PIC) simulations, that is a self-consistent first-principle method for collisionless plasma. It follows the individual particles and solves Maxwell's equations for the electromagnetic fields on a spatial grid.

3. Related work

Recent investigations with PIC-simulations confirmed the shock-drift acceleration (SDA) as a mechanism of the electron energization [1 – 3], and indicate the importance of the electron scattering by multi-scale wave structures for the reaching of highest energies [4]. The latter requires a large system size to be resolved in a simulation.

4. Simulation setup and results

The numerical experiment has been realized with use of advanced MPI-based parallel relativistic code TRISTAN [5]. The code has been rewritten in FORTRAN90 and modified to use the HDF5 file format libraries and improved algorithm of the particle sorting [6].

In the framework of the present study we performed series of the large-scale 2D-3V PIC simulations, parameters of which are listed in Table 1. Run A is a reference run with parameters used in [4] ($\beta = 5$ and $\theta_{\text{bn}} = 75^\circ$). We increased plasma beta to $\beta = 20$ in Run B-20, and changed the shock obliquity to $\theta_{\text{bn}} = 60^\circ$ and $\theta_{\text{bn}} = 78^\circ$ in Run T-60 and Run T-78, correspondingly.

Figure 1 demonstrates the evolution of the upstream electron energy distributions observed in different simulation runs. These spectra are formed by the populations of the reflected electrons after their energization via stochastic shock drift acceleration (SSDA).

Tab. 1. List of the main simulation parameters.

Parameter	Run A	Run B-20	Run T-60	Run T-78
CPU-hours used	~5 mln	~5 mln	~2 mln	~5 mln
Disk space used	~50 TB	~40 TB	~20 TB	~30 TB
CPUs/nodes (max.)	6240/260	6000/250	5400/225	6000/250
Shock obliquity, θ_{Bn}	75°	75°	60°	78°
Plasma beta, β	5	20	5	5

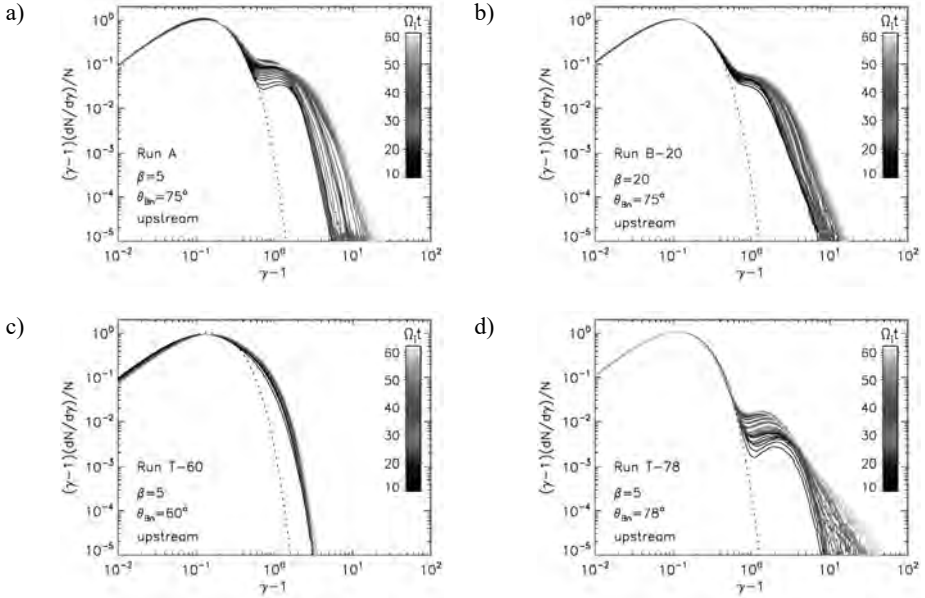


Fig. 1. Evolution of electron energy spectra in Run A (a), Run B-20 (b), Run T-60 (c) and Run T-78 (d).

5. Conclusions

Maximum electron energy in the course of SSSA process does not depend significantly on plasma beta, while the effect of the shock obliquity is much more essential.

Acknowledgements. This work has been supported by the Polish National Science Centre through projects UMO-2016/22/E/ST9/00061 and 2019/33/B/ST9/02569. The numerical experiment was possible through computing allocation on the Prometheus system at ACC Cyfronet AGH under the grants “astropic22” and “astropic23”.

References

1. X. R. Guo, L. Sironi, R. Narayan, 2014, *Astrophys. J.*, 794, 153.
2. X. R. Guo, L. Sironi, R. Narayan, 2014, *Astrophys. J.*, 797, 47.
3. S. Matsukiyo, Y. Matsumoto, 2015, *J. of Phys.*, 642, 012017.
4. O. Kobzar, J. Niemiec, T. Amano et.al. 2021, *Astrophys. J.*, 919, 97.
5. J. Niemiec, M. Pohl, T. Stroman, K.-I. Nishikawa, 2008, *Astrophys. J.*, 684, 1174.
6. A. Dorobisz, M. Kotwica, J. Niemiec, O. Kobzar, A. Bohdan, K. Wiatr, 2018, *LNSC*, 10777, 156.

Performance of LLAMA Large Language Model in Argument Detection

Michał Karwatowski¹, Marcin Pietroń¹, Kamil Faber², Dominik Żurek²

¹ Institute of Electronics, AGH University of Krakow, al. Mickiewicza 30, 30-059 Kraków, Poland

² Institute of Computer Science, AGH University of Krakow, al. Mickiewicza 30, 30-059 Kraków, Poland

{mkarwat, pietron, kfaber, dzurek}@agh.edu.pl

Keywords: Large Language Model, Transformers, deep learning, NLP

1. Introduction

The theory of argumentation and the use of logical reasoning to justify claims and conclusions is an extensively studied field, but the application of data science methods to automate these processes is a relatively recent development. Argument Mining is defined as the task of automatically identifying and extracting argumentative components (e.g., premises, claims, etc.) and detecting the existing relations among them (i.e., support, attack, rephrase, no relation). In the last few years automated argument mining has become a more popular task [3,4]. Deep learning models enable the analysis of arguments more efficiently and extract their semantics. Recently, large language models (LLM) have been achieving excellent results in many natural language processing tasks. These models have huge computational and memory complexity. This work presents the performance of the OpenLLAMA model using natural language benchmarks.

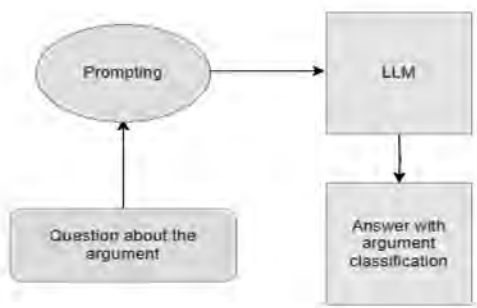


Fig. 1.The main framework for LLM argument classification.

2. Description of the problem

Argument Mining focuses on automatically identifying, extracting, and analyzing argumentative structures within natural language texts, which includes recognizing core components of arguments, such as claims and evidence. It can be decomposed into sub-tasks. First, the identification of argument components consists in distinguishing argumentative propositions from non-argumentative propositions. This allows the user to segment the input text into arguments. Second, the identification of clausal properties is the part of AM that focuses on finding premises among the argumentative propositions. Third, the last sub-task is the identification of relational properties. Two different argumentative propositions are considered at a time, and the main objective is to identify which type of relation links both propositions.

3. Related work

In recent years, we have seen the development of new methods and new corpora in the field of argument mining [1]. Initially, these were quite simple machine and deep learning models [2]. The Transformer-based architectures and their subsequent pretraining approaches became a turning point in the NLP research area [3, 4]. Then, the emergence of Chat-GPT4, a generative pre-trained transformer language model developed by OpenAI, has presented new opportunities for argument mining. This model possesses the ability to comprehend context, produce coherent and contextually appropriate responses, and participate in meaningful dialogues with users. It also gave rise to many open-source language models like OpenLLAMA.

4. Solution of the problem

An open model LLAMA was used to examine the capabilities of LLM in argument detection and classification. The LLAMA made it possible to automate the process. The three datasets Argsme, UKP and US2016 were taken as a benchmark. The question to the model was formed by taking the claim and the argument and using it in the question to the model by asking if the argument is for, against, or if there is no relation with the topic.

Due to the size of the large language models, it is not possible to run them on a single GPU. Therefore, the task was performed on multiple GPUs. The scalability and efficiency were measured using 3 size versions of the model: 7B, 13B, and 70B. The biggest version was not able to run on a smaller number of GPUs. The average accuracy of argument classification for OpenLLAMA models in the case of the Argsme dataset was between 60% and 75%, for UKP between 50% and 60%, and for US2016 between 40% and 65%. The simulations were run using NVIDIA A100 GPU. The work is the first stage of building a prompt-based algorithm that will improve the argument detection task. The error analysis will help to define prompt engineering rules.

Tab. 1. Model pipeline inference on multiple GPUs [sec/sample].

model size	GPUs			
	1	2	4	8
7B	6,74	6,09	7,20	6,35
13B	113,68	13,24	13,10	12,15
70B	---	---	---	47,37

Acknowledgements. PLGrid grant ID: plglaois23/plglaois24. All calculations were made on the PLGrid infrastructure, using the Athena supercomputer.

References

1. R. Schaefer and M. Stede: GerCCT: An Annotated Corpus for Mining Arguments in German Tweets on Climate Change. In Proceedings of the Thirteenth Language Resources and Evaluation Conference, pp. 6121–6130, Marseille, France, 2022.
2. A. Galassi, M. Lippi and P. Torroni, P.: Argumentative link prediction using residual networks and multi-objective learning. In Proc. of the 5th Workshop on Argument Mining, pp.1–10, 2018.
3. P. Srivastava, P. Bhatnagar and A. Goel: Argument Mining using BERT and Self-Attention based Embeddings. ICAC3N, 2022.
4. R. Ruiz-Dolz, J. Alemany, S. Heras and A. Garcia-Fornes. Transformer-Based Models for Automatic Identification of Argument Relations: A Cross-Domain Evaluation. IEEE Intelligent Systems, 2021.

Performance of Benchmarking Continual Curriculum Learning

Dominik Żurek¹, Marcin Pietroni², Kamil Faber¹, Michal Karwatowski²

¹ Institute of Computer Science, AGH University of Krakow, al. Mickiewicza 30, 30-059 Kraków, Poland

² Institute of Electronics, AGH University of Krakow, al. Mickiewicza 30, 30-059 Kraków, Poland

{dzurek, pietron, kfaber, mkarwat}@agh.edu.pl

Keywords: continual learning, deep learning, lifelong learning, curriculum learning

1. Introduction

Continual learning (CL) is one of the most promising trends in recent machine learning research. Its goal is to go beyond classical assumptions in machine learning and develop learning strategies that present high robustness in dynamic environments and avoid forgetting past knowledge [4]. The continual learning research is fragmented into several learning evaluation protocols, comprising different learning tasks, datasets, and evaluation metrics. The benchmarks adopted so far are still distant from the complexity of real-world scenarios.

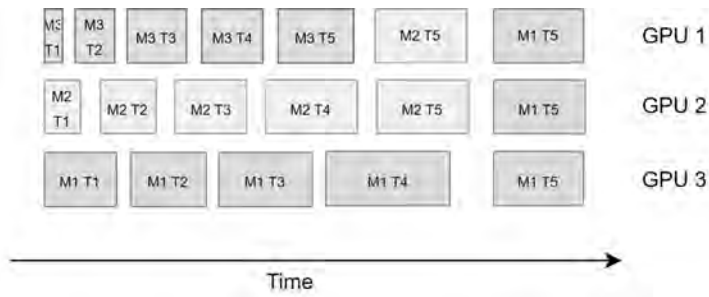


Fig. 1. Pipeline of the CL benchmark.

It is hard to clearly and objectively assess models and strategies using existing benchmarks. In this work, we fill this gap for CL on image data by introducing two novel CL benchmarks that involve multiple heterogeneous tasks from six image datasets (starting from MNIST and ending with ImageNet and vice versa, M2I and I2M), with varying levels of complexity and quality. Our aim was to fairly evaluate current state-of-the-art CL strategies on a ground that is closer to complex real-world scenarios. The new scenarios were evaluated on a set of different CL strategies in a parallel environment (Fig. 1).

2. Description of the problem

The majority of attention in CL has been devoted to computer vision, where the most common task is to learn models that can classify different kinds of images while preventing catastrophic forgetting. Every new model has been evaluated in a slightly different setting - using a different dataset, evaluation metrics, and learning protocols. The benchmark panorama of CL in computer vision is quite fragmented, and therefore it has become tougher to measure domain adaptation in a fair and homogeneous way. Therefore, curriculum-based M2I and I2M scenarios were designed [3]. Tests on many CL strategies were required to show the validity of the developed benchmark.

3. Related work

CL strategies belong to three main groups: using regularization, dynamic architectures, and rehearsal (also known as experience replay). Regularization strategies influence the model weights adjustment process that takes place during model training to preserve the knowledge of previously learned tasks (e.g., EWC, LwF, SI) [1,2]. Rehearsal strategies store samples for all classes from previous tasks in the scenario and use them in the next tasks (e.g., GDumb, Replay, Gradient Episodic Memory (GEM)). The dynamic architecture strategies modify the model topology for each scenario task.

4. Solution of the problem

Common CL benchmarks focus on homogeneous tasks, such as different classes of handwritten digits (e.g. as in splitMNIST). To deal with multiple heterogeneous tasks, our benchmark leverages 6 largely varying image classification datasets: MNIST (handwritten digits), Omniglot (alphabets), Fashion MNIST (clothing items), SVHN (street view house numbers), CIFAR10 (small real-world images), and ImageNet (multidomain large-scale real-world images).

Tab. 1. The time efficiency of benchmarking new CL scenarios with and without pipeline mode.

	1 GPU	2 GPUs	4 GPUs	8 GPUs
w/o pipeline	1079 [min]	614 [min]	423 [min]	360 [min]
pipeline	1079 [min]	500 [min]	290 [min]	220 [min]

It is important to devise scenarios with varying quality and task complexity since an ideal model should present generalization capabilities dealing with easy, moderate, and difficult tasks at the same time, as found in the real world. The proposed benchmark meets these requirements. The benchmark was tested on a wide spectrum of CL strategies (CWRStar, SI, LwF, GDumb, EWC, PNN, Cumulative, Naïve, Replay, GEM, MAS). Each dataset in a scenario was trained with 50 epochs. This generates high computational complexity. Therefore, the calculations were parallelized. The first mode involved trivial parallelization (each strategy was performed on a separate computing unit). The second version is the pipeline mode (Figure 1, example for 3 CL methods (M) and 5 tasks (T)). Each strategy was parallelized if additional computing resources were available at a given moment. In Table 1 the time efficiency is presented for different number of GPUs. The simulations were run for Resnet50 model using NVIDIA V100 GPU.

Acknowledgements. PLGrid grant ID: plglaois23/plglaois24. All calculations were made on the PLGrid infrastructure, using the Athena supercomputer.

References

1. G. I. Parisi, R. Kemker, J. L. Part, C. Kanan and S. Wermter: „Continual lifelong learning with neural networks: A review”. *Neural Networks* 113, pp. 54-71, 2019.
2. M. De Lange, R. Aljundi, M. Masana, S. Parisot, X. Jia, A. Leonardis, G. Slabaugh and T. Tuytelaars: „A continual learning survey: Defying forgetting in classification tasks. *IEEE transactions on pattern analysis and machine intelligence*”, 44(7), pp.3366-3385, 2021.
3. Y. Bengio, J. Louradour, R. Collobert and J. Weston: „Curriculum Learning”, in *Proc. ICML*, 2009, pp.41-48.
4. Z. Li, D. Hoiem: „Learning without forgetting”, *IEEE transactions on pattern analysis and machine intelligence*, 40(12), 2017, pp.2935-2947.

Comparison of GPU and CPU Implementations of New Variants of SDLS Algorithms for LABS Problem

Dominik Żurek¹, Marcin Pietron², Kamil Faber¹, Michał Karwatowski², Kamil Pięta¹

¹AGH University of Krakow, Faculty of Computer Science, Kraków, Poland

²AGH University of Krakow Institute of Electronics, Kraków, Poland

{dzurek, pietron, kfaber, mkarwat, kpietak}@agh.edu.pl

Keywords: LABS, steepest-descent local search, parallel computing, GPGPU, OpenMP

1. Introduction

Low autocorrelation binary sequence (LABS) remains an open hard optimization problem. One of the promising directions for solving the problem is designing advanced solvers based on local search heuristics. This paper compares the time efficiency of the two new variants of SDLS algorithms (2), SDLS-2 and SDLS-DT (1) which explore the wider neighborhoods. The implementations utilize parallel computing with CPU and GPU.

2. Description of the problem

LABS is an NP-hard combinatorial problem with a simple formulation. It consists of finding a binary sequence $S = \{s_0, s_1, \dots, s_{L-1}$ with length L , where $s_i \in \{-1, 1\}$ which minimizes the energy function $E(S)$:

$$C_k(S) = \sum_{i=0}^{L-k-1} s_i s_{i+k} \text{ and } E(S) = \sum_{k=1}^{L-1} C_k^2(S)$$

3. Solution of the problem

The new approach to solving the LABS problem based on SDLS algorithms relies upon increasing the search area during each iteration (1). In the GPGPU implementation of SDLS, SDLS-2 and SDLS-DT it was possible to achieve parallelism on number of the solution level (number of thread's blocks) and during performing SDLS calculation where L threads are calculating L energies at the same time (3). In the CPU's implementation, it is possible to use only one-level parallelism. As it turns out, the most effective way is to introduce a parallel on the number solution level which is equivalent to the K thread block in the GPGPU implementation. Consequently, K' solutions are calculating by K' threads at the same time, where K' means the number of available cores of the CPU (in our case this number equals 128). The experiments were performed on an AMD EPYC 7742 64-Core Processor, and NVIDIA A100.

Tab. 1. Comparison of time results [s].

Method	GPU TIME		CPU TIME	
	L=128	L=256	L=128	L=256
SDLS	0.03	0.08	0.12	0.37
SDLS-2	1.51	10.56	25.81	123.56
SDLS-DT	0.34	1.67	6.15	28.47

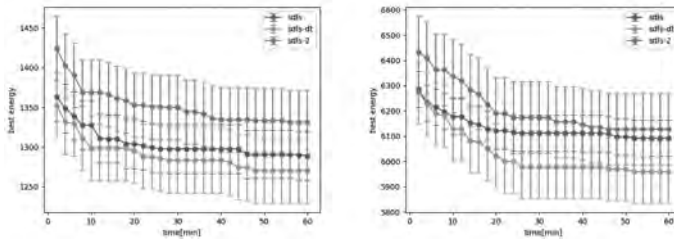


Fig. 1. Energies achieved on GPGPU by SDLS, SDLS-2 and SDLS-DT algorithms (1).

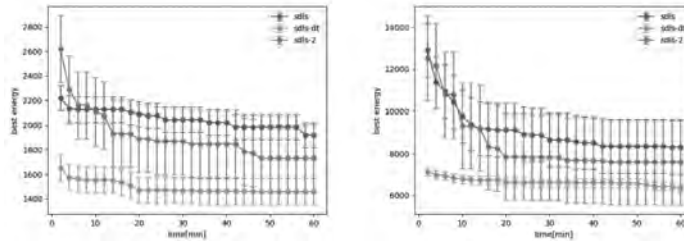


Fig. 2. Energies achieved on CPU by SDLS, SDLS-2 and SDLS-DT algorithms (1).

4. Conclusions

Table 1 contains the execution time of three versions of SDLS algorithms that were running through 512 iterations for two values of L , that is, 128 and 256. In each case, the most effective times were achieved for GPU implementation. Fig. 1 and Fig. 2. demonstrate the efficacy of the algorithms, where each of them was seeking the optimal solution through one hour for two different input length. As could be observed for each size of the problem, the best solution was obtained by the SDLS-DT algorithm.

Acknowledgements. This research was partially supported by the PLGrid infrastructure (plg-cyberneuroevol23).

References

1. Żurek, D., PiętaK, K., Pietroń, M., Kisiel-Dorohinicki, M. (2021). New Variants of SDLS Algorithm for LABS Problem Dedicated to GPGPU Architectures. In: Paszynski, M., Kranzlmüller, D., Krzhizhanovskaya, V. V., Dongarra, J. J., Sloot, P. M. A. (eds) Computational Science – ICCS 2021. ICCS 2021. Lecture Notes in Computer Science, vol 12742. Springer, Cham. https://doi.org/10.1007/978-3-030-77961-0_18.
2. Bartholomew–Biggs, M. (2008). The Steepest Descent Method. In: Nonlinear Optimization with Engineering Applications. Springer Optimization and Its Applications, vol 19. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-78723-7_7.
3. Toward hybrid platform for evolutionary computations of hard discrete problems, D. Żurek, K. PiętaK, M. Pietroń, M. Kisiel-Dorohinicki. Procedia Computer Science 108, 877-886.

Ab Initio Molecular Dynamics Simulations of Aqueous LiTFSI Solutions – Structure, Hydrogen Bonding, and IR Spectra

Andrzej Eilmes, Piotr Wróbel, Piotr Kubisiak

Jagiellonian University, Faculty of Chemistry, Gronostajowa 2, 30-387 Kraków, Poland

{eilmes, kubisiak}@chemia.uj.edu.pl, piotr.wrobel@doctoral.uj.edu.pl

Keywords: water-in-salt electrolytes, vibrational spectra, molecular dynamics, hydrogen bonds

1. Introduction

Highly concentrated salt solutions in water, water-in-salt (WiS) electrolytes are an emerging class of electrolytes with expanded electrochemical stability window [1]. A typical example of a WiS electrolyte is lithium bis(trifluoromethanesulfonyl)imide (LiTFSI) solution in water for which concentrations up to 21 mol/kg (21 m) can be achieved. Multiple experimental and computational works focus on the structure and dynamics of WiS systems. Vibrational spectroscopies (IR or Raman) are common tools to investigate interactions in such electrolytes.

2. Description of the problem

Analysis of the infrared (IR) spectra of LiTFSI solutions in water provides information on the changes in local environments of water molecules and the structure of hydrogen bonds (HBs) formed in the electrolyte. The structure of the solution can be simulated by means of classical, force field-based, molecular dynamics (MD). However, the ab initio molecular dynamics (AIMD) is necessary for the calculations of the IR spectrum (obtained as a Fourier transform of the dipole moment of the system). In this work, we wanted to investigate the properties of LiTFSI WiS using AIMD, in particular the relation of the hydrogen bonds structure to the IR spectra.

3. Related work

Experimental IR spectra of LiTFSI/H₂O electrolytes were measured at different concentrations and abundances of different environments of water molecules were extracted from the spectra or calculated from the classical MD [2,3]. AIMD was applied to calculate IR spectra of LiTFSI solutions at two concentrations [4], but the detailed analysis of the HBs network and its relation to the changes in the spectrum has not been performed.

4. Solution of the problem

Ab initio molecular dynamic simulations employing the density functional (DFT) methodology with PBE functional, the empirical dispersion correction D3, Goedecker's pseudopotentials and a molecularly optimized DZVP basis set were performed in the CP2K program [5]. Four LiTFSI concentrations (1, 5, 10, and 20 m) were examined; the number of atoms was between 718 and 741. Two independent replicas of each system were simulated for 40 ps with a time step of 1 fs in the NVT ensemble at T = 298 K using the Nosé-Hoover thermostat.

The IR spectra were obtained from the recorded AIMD trajectories as the Fourier transforms (FTs) of the autocorrelation function of the dipole moment. In order to analyze the effect of hydrogen bonding on the vibrational frequencies, FTs of all O-H bond lengths were calculated, yielding the power spectra of the O-H vibrations.

5. Results

The analysis of the structure of electrolytes revealed that Li⁺ ions interact mainly with water molecules and cation-anion interactions become important only at the highest concentrations. In WiS solutions, salt ions form aggregates, while water molecules are rather isolated or aggregate into small clusters. We correlated the shifts in the IR spectra to the local configurations of the HBs formed by water molecules and the change from the symmetric water-water binding at low concentrations, through the asymmetric water-anion configuration at the intermediate concentrations to the symmetric anion-anion environment in WiS electrolytes [6].

6. Conclusions

The IR spectra simulated from the AIMD trajectories are in a very good agreement with the experimental data [2,3]. The abundances of local environments extracted from our computational results [6] agree well with the decomposition of measured IR spectra [2,3], confirming the applicability of the AIMD methodology to WiS solutions.

Acknowledgements. The numerical experiment was possible through computing allocation on the Ares system at ACC Cyfronet AGH under the grant plgaenael7. The AIMD simulations were running on 4-5 nodes (192-240 CPU cores) and consumed about 1 M walltime hours.

References

1. L. Suo, O. Borodin, T. Gao, M. Olguin, J. Ho, X. Fan, C. Luo, C. Wang, K. Xu: "Water-in-Salt" Electrolyte Enables High-Voltage Aqueous Lithium-Ion Chemistries, in *Science* 2015, 350, 938-943.
2. Y. Zhang, N. H. C. Lewis, J. Mars, G. Wan, N. J. Weadock, C. J. Takacs, M. R. Lukatskaya, H. G. Steinrück, M. F. Toney, A. Tokmakoff, E. J. Maginn: Water-in-Salt Aqueous Electrolytes. 1. Liquid Structure from Combined Molecular Dynamics Simulation and Experimental Studies, in *J. Phys. Chem. B* 2021, 125, 4501-4513.
3. N. H. C. Lewis, B. Dereka, Y. Zhang, E. J. Maginn, A. Tokmakoff: From Networked to Isolated: Observing Water Hydrogen Bonds in Concentrated Electrolytes with Two-Dimensional Infrared Spectroscopy, in *J. Phys. Chem. B* 2022, 126, 5305-5319.
4. T. Malaspina, G. Colherinhas, S. E. Weitzner, B. C. Wood, E. E. Fileti: Unraveling Local Structures of Salt-in-Water and Water-in-Salt Electrolytes via Ab Initio Molecular Dynamics, in *J. Mol. Liquids* 2023, 383, 122097.
5. J. Hutter, M. Iannuzzi, F. Schiffmann, J. VandeVondele: CP2K: Atomistic Simulations of Condensed Matter Systems, in *Wiley Interdiscip. Rev.: Comput. Mol. Sci* 2014, 4, 15-25.
6. P. Wróbel, P. Kubisiak, A. Eilmes: Ab Initio Molecular Dynamics Simulations of Aqueous LiTFSI Solutions – Structure, Hydrogen Bonding, and IR Spectra, in *J. Phys. Chem. B* 2024, 128, 1001-1011.

Interaction Spheres for Molecular Interactions – Application in Drug Design

Rafał Kurczab

Maj Institute of Pharmacology, Polish Academy of Sciences, Smętna 12 Street, Kraków, Poland

kurczab@if-pan.krakow.pl

Keywords: molecular interaction, drug design, salt bridge, halogen bond, hydrogen bonds

1. Introduction

Understanding the interactions that bind a drug molecule to a receptor makes it possible to properly assess its mechanism of action and more rationally design new chemical compounds with the desired biological activity. Recently, a lot of molecular interactions have been used in the design strategy of new drugs, namely halogen bonds, fluorine bonds, hydrogen bonds, sulfur interactions, and aromatic interactions. Nevertheless, there is a lack of parameterization of some of these interactions in force fields and their lack of consideration in molecular modeling software.

2. Description of the problem

Mapping the interaction spheres on characteristic molecular fragments located at the binding site in the receptor will allow to estimate of modifications in the molecule to increase its biological activity.

3. Related work

Spherical scan was used to study several types of molecular interactions, namely halogen bonds [1], salt bridge [2], and interaction with sulphur [3]. The interaction spheres obtained for salt bridges have also been successfully applied in optimizing the structure of 5-HT₆ serotonin receptor ligands [4].

4. Solution of the problem

For selected types of molecular interactions (i.e. salt bridge, and fluorine bond), it defined the set of simple model systems. Each model system was divided into two fragments between which the molecular interaction takes place. In the next step, successive alignments of the two fragments with respect to each other are generated by changing two geometrical parameters of the interaction - distance and angle. Each of the generated geometries was next used to calculate the interaction energy by quantum mechanical methods in the GAUSSIAN program. The following settings were used:

- geometry optimization was carried out at B3LYP-D3/6-31G(d,p) level,
- spherical scan was performed using in-house script,
- single point energy was performed at B3LYP-D3/6-31G(d,p) level and with PCM model (solvent=water),
- an in-house script for visualization of interaction energy distribution (angle scan was performed with the step of 10°, whereas the distance scan was performed with the step of 0.05 Å).

The theoretical study of dependency between geometry and energy of X(O, N, C)-H \cdots F-C type of interaction (fluorine bond) was developed and used. It was tested for the optimized geometry of dimethylamine and fluorobenzene moiety (Fig. 1A, B). For fixed H \cdots F distance (taken from the output of a dimer geometry optimization), the angle between monomers was systematically changed in a way to cover a sphere around the hydrogen atom of dimethylamine. Single-point energy calculations were then performed (B3LYP-D3/6-31G(d,p)). The results of surface scan indicated (Fig. 1.) that the energy of this type of interaction depends on the donor-acceptor angle. In addition, analysis of the sphere indicates forbidden areas - i.e., such an orientation between donor and acceptor that destabilizes the complex (purple areas/dots in the spheres).

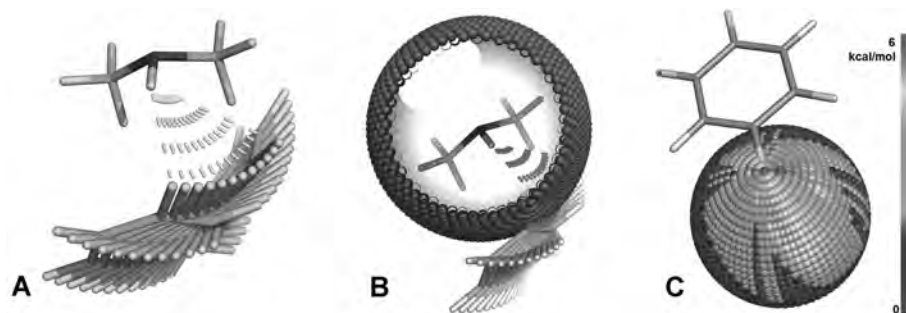


Fig. 1. Spherical scan approach performed for dimethylamine–fluorobenzene system as one of the simple model systems mimicking fluorine interaction.

5. Conclusions

The resulting interaction spheres simply show the areas (arrangements of fragments) in which the complex concerned obtains increased stabilization or destabilization. This allows the rational design of new derivatives with better parameters of biological activity. The spheres are created once, and not dynamically for a given system by which obtaining the necessary information is immediate.

Acknowledgements. This work/research has been supported by the Maj Institute of Pharmacology Polish Academy of Sciences. The numerical experiment was possible through computing allocation on the Prometheus system at ACC Cyfronet AGH under the grants (fluorine2, hbonds2, hbonds).

References

1. R. Wilcken, M. O. Zimmermann, A. Lange, S. Zahn, F. M. Boeckler: Using halogen bonds to address the protein backbone: a systematic evaluation, *J. Comput. Aided Mol. Des.* 2012, 26 (8), 935-945.
2. R. Kurczab, P. Śliwa, K. Rataj, R. Kafel, A. J. Bojarski: Salt Bridge in Ligand–Protein Complexes–Systematic Theoretical and Statistical Investigations, *J. Chem. Inf. Model.* 2018, 58, 11, 2224–2238.
3. X. Zhang, Z. Gong, J. Li, T. Lu: Intermolecular Sulfur \cdots Oxygen Interactions: Theoretical and Statistical Investigations, *J. Chem. Inf. Model.* 2015, 55, 10, 2138–2153.
4. K. Grychowska, S. Chaumont-Dubel, R. Kurczab, et. al.: Dual 5-HT6 and D3 Receptor Antagonists in a Group of 1H-Pyrrolo[3,2-c]quinolines with Neuroprotective and Procognitive Activity, *ACS Chem. Neurosci.* 2019, 10, 3183–3196.

Dynamic Disulfide Bonds in Molecular Dynamics Simulation

Pamela Smardz, Paweł Krupa

Institute of Physics Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warsaw, Poland

psmardz@ifpan.edu.pl, pkrupa@ifpan.edu.pl

Keywords: molecular dynamics, all-atom, disulfide bonds

1. Introduction

Proteins are crucial for all living organisms, and their structure is vital for their functioning. The structure of proteins is intricately tied to the composition of amino acids in the polypeptide chains, interactions between them, and the presence of disulfide bonds, along with external conditions such as the environment or stress factors. Therefore, the study of protein conformation and dynamics is essential in molecular biology, pharmacology, and immunology to unravel underlying mechanisms.

2. Description of the problem

About 24% of proteins contain disulfide bonds, covalent bonds between cysteine side-chains, the presence or absence of which may serve various roles in proteins. Computational simulations are often used to study biological systems, yet they are costly, particularly at the quantum level. Therefore, simplifications are necessary, resulting in all-atom molecular dynamics (MD). However, conventional MD (cMD) simulations cannot capture chemical reactions, including disulfide bond formation and breaking.

3. Related work

Previous studies used static disulfide bond models, which did not allow for formation or breaking, or Gō-like models to explore the effects of disulfide bond breaking on the folding process [1] and build-in-house coarse-grained model to examine the folding mechanisms of disulfide-bonded proteins [2].

4. Solution of the problem

To enable disulfide bonds to break and form during simulations, we introduced a pseudo-potential based on finite distance restraints between sulfur atoms of selected cysteines. This method was coupled with all-atom cMD simulations in Amber to capture equilibrium properties (GPUs on Athena) and Steered MD to impose mechanical tension and study degradation pathways (CPUs on Ares) (Fig. 1.).

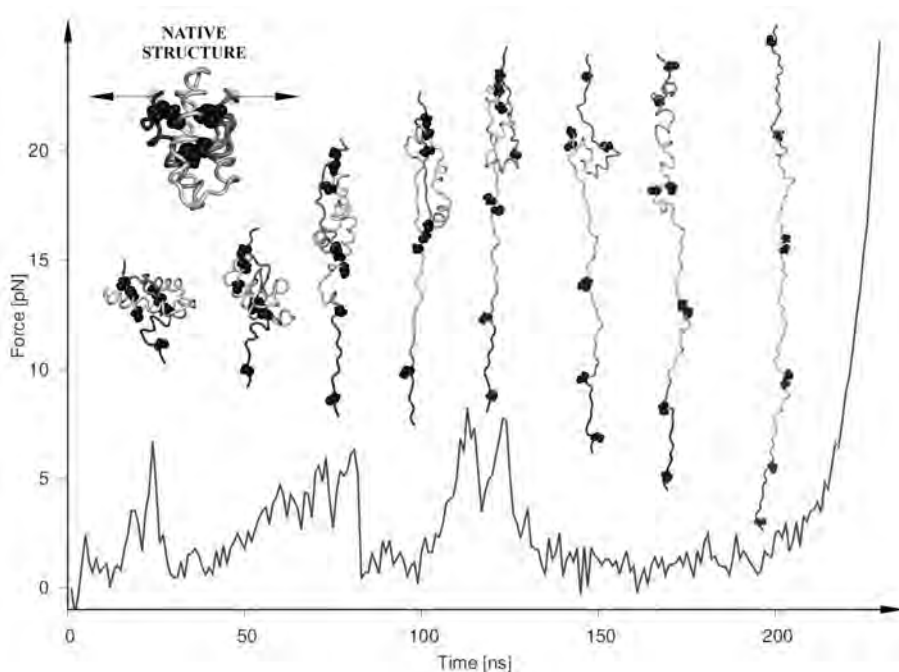


Fig. 1. Plot of force in function of extension with cartoon representation of conformations during stretching in SMD simulation (equivalent of AFM pulling) of nsLTP with cartoon representation of native structure (PDB code: 2n81) with arrows indicating external stretching force applied during simulation (top left corner).

5. Conclusions

Our method effectively simulates disulfide bond dynamics across diverse conditions, showcasing its utility in molecular studies. Performed cMD and SMD simulations in various conditions successfully explained differences in conformational dynamics between allergenic and non-allergenic nsLTPs.

Acknowledgements. This work was supported by the National Science Center (Poland) Sonata grant 2019/35/D/ST4/03156, and the PLGrid ACC Cyfronet AGH computational grant PLG/2023/016624.

References

1. M. Qin, J. Zhang, W. Wang, *Biophys. J.* 2006, 90, 272.
2. M. Qin, W. Wang, D. Thirumalai, *Proc. Natl. Acad. Sci. U. S. A.* 2015, 112, 11241.

Application of Fragment Molecular Orbital Method to Investigate G Protein Coupled Receptors

Paweł Śliwa

Faculty of Chemical Engineering and Technology, Cracow University of Technology
24 Warszawska St., Kraków 31-155, Poland

pawel.sliwa@pk.edu.pl

Keywords: FMO/PIEDA, GPCR, Glutamine, Dopamine, Serotonin

1. Introduction

The utilization of quantum chemical methods for studying biological systems is typically computationally demanding. The fragment molecular orbital method (FMO) [1-3] serves as a valuable approach for assessing the energy of large systems at the *ab initio* level. The applied pair interaction energy decomposition analysis method (PIEDA) provides the electrostatic, exchange, charge transfer, and dispersion contributions to the total interaction energies, which is particularly beneficial for studying protein–ligand complexes.

2. Description of the problem

In this study, the FMO methodology was employed to characterize the energy of ligand receptor (GPCR) complexes. This approach was utilized to evaluate several aspects, including:

- Allosteric modulation of the glutamatergic receptor
- Repulsive and attractive regions within the binding sites of selected serotonin receptors
- The similarities and differences in the active sites of three types of dopamine receptors
- Differences between the active and inactive states of ligand–D4 receptor complexes.

3. Related work

The FMO methodology has been effectively employed in analyzing various large biological systems, particularly in retrospectively studying binding sites and utilized as a tool to support drug design. The FMO scheme could be a valuable tool for structure-based drug design by estimating the chemical character of the protein's binding site [4,5].

4. Solution of the problem

FMO input files were prepared using Facio software. The fragmentation of complexes was generated mostly automatically, however, some bond detachments had to be done manually. The PIE was calculated between ligand and binding site amino acids. All computations were performed with GAMESS using the second order Möller-Plesser perturbation (MP2) level of theory with the 6-31G* basis set and polarizable continuum model. To evaluate nature of interactions, based on PIEDA results, the percentage of the sum of absolute values of charge transfer and electrostatic contributions in the total attraction energy was calculated. The above methodology was used to study four cases:

- the crystal structure of metabotropic glutamate receptor 1 complexed with a negative allosteric modulator and its twelve close structural analogs with a broad spectrum of affinities,

- the three non-crystallized serotonin receptors (5-HT1A, 5-HT2A and 5-HT7) complexed by eight of active ligands,
- the crystal structure of dopamine D2, D3, and D4 receptors with the co-crystallized ligands,
- the D4R's characteristic upon known seven agonists or ten antagonists binding.

5. Conclusions

Based on the results, several general conclusions can be made:

- The study of a series of negative allosteric modulators of mGluR1 showed that the allosteric binding site has mostly hydrophobic character.
- The obtained energetic map of the binding site showed that the orthosteric binding sites of selected serotonin receptors have both attractive and repulsive regions.
- The obtained pair interaction energy decomposition analysis provided valuable insights for binding sites of D2-Risperidone, D3-Eticlopride, and D4-Nemonapride complexes.
- The energy calculations using the FMO/PIEDA method indicate that most of the interactions observed during ligand binding by the dopamine D4 receptor were attractive.

Acknowledgements. The numerical experiment was possible through computing allocation on the Prometheus and Ares systems at ACC Cyfronet AGH under the grants: acconiom2015, bqmmm2015, micela2018gpu, plgpsliwamodel2022.

References

1. Fedorov D. G., Kitaura K. (2004) The importance of three-body terms in the fragment molecular orbital method. *J. Chem Phys* 120:6832–6840.
2. Fedorov D. G., Kitaura K. (2007) Pair interaction energy decomposition analysis. *J. Comput Chem* 28:222–237.
3. Sekino H., Sengoku Y., Sugiki S., Kurita N. (2003) Molecular orbital analysis based on fragment molecular orbital scheme. *Chem Phys Lett* 378:589–597.
4. Śliwa P., Kurczab R., Bojarski A. J. (2018) ONIOM and FMO-EDA study of metabotropic glutamate receptor 1: quantum insights into the allosteric binding site. *Int J. Quantum Chem* 118:e25617.
5. Heifetz A., Chudyk E., Gleave L. et al (2016) The fragment molecular orbital method reveals new insight into the chemical nature of GPCR-ligand interactions. *J. Chem InfModel* 56:159–172.

Theoretical Description of Peptide Bonds in Selected Systems Based on the ETS-NOCV Method

Maria Rózga, Artur Michalak

Department of Theoretical Chemistry, Faculty of Chemistry,
Jagiellonian University, Gronostajowa 2, 30-387 Kraków, Poland

maria.rozga@student.uj.edu.pl, michalak@chemia.uj.edu.pl

Keywords: ETS-NOCV analysis, dipeptides, peptide bond, bond strength model

1. Introduction

The aim of this work was to analyze qualitative features of dipeptides in order to quantitatively assess their impact on the peptide bond strength. By utilizing the ETS-NOCV approach a prediction model of peptide bond strength between all pairs of proteinogenic amino acids was constructed.

2. Description of the problem

The energy associated with the formation of the peptide bond is known to depend on the type of amino acid, yet no bond energy analysis, that includes this dependence, can be found in literature. The ETS-NOCV approach allows for a comprehensive analysis of chemical bonds within one theoretical framework by combining the NOCV method of deformation-density decomposition with the ETS bond-energy decomposition scheme.

3. Related work

The ETS-NOCV approach has been utilized for a vast array of chemical structures and bonds [1-3]. Extensive conformational search [4] support the hypothesis that conformational preferences of amino acids are transferable to poly-peptides. As of yet no search for transferability of energetic parameters has been performed.

4. Solution of the problem

Six model dipeptides were used in order to choose the most suitable partitioning scheme for NOCV description. The systematic analysis was performed on two dipeptide series (Ala-X and X-Ala, where X denotes one of 20 amino acids). Due to a complexity of conformational space of the studied systems, for all the considered dipeptides their respective energy minima were found by a computational protocol, combining the molecular dynamics at the (PM7)- level [5] with DFT-based geometry optimizations via ADF package (2017.104) [6].

The NOCV results for two dipeptide series in the chosen cation-anion partitioning were found to exhibit a clear σ/π donation/back-donation picture. This, in turn, allowed to construct the prediction model of the peptide bond strength between any amino acid pair. To verify model's predictions His-Thr and Tyr-Thr, predicted to exhibit the strongest, and Gly-Val, Trp-Val – the weakest bonds were analyzed.

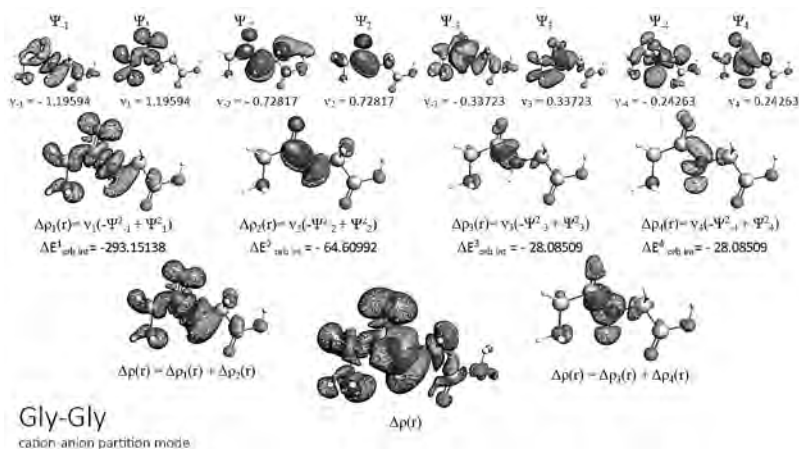


Fig 1. Four most significant NOCV pairs, their respective differential densities, and their sums of the Gly-Gly dipeptide in cation-anion partition.

5. Conclusions

Among other qualitative features, the ETS-NOCV method allowed to recognize a characteristic of the Dewar-Chatt-Duncanson model donation/back-donation picture on the basis of which the prediction model was derived. The results show an agreement of the qualitative trend between predicted and computed values, but an understandable quantitative discrepancy can be seen.

Acknowledgements. The numerical experiment was possible through computing allocation on the Prometheus and Ares systems at ACC Cyfronet AGH under the grant plgmichalak7.

References

1. M. Mitoraj, A. Michalak, *J. Mol. Model.*, 2007, 13 (2), 347–355.
2. A. Michalak, M. Mitoraj, T. Ziegler, *J. Phys. Chem.*, 2008, 112 (9), 1933–1939.
3. M. Mitoraj, A. Michalak, T. Ziegler, *J. Chem. Theory Comput.*, 2009, 5 (4), 962–975.
4. W. Yu, X. Xu, H. Li, R. Pang, K. Fang, Z. Lin, *Comput Chem.*, 2009, 30, 2105–2121.
5. MOPAC2016; Stewart, J.J.P. Stewart Computational Chemistry, Colorado Springs, CO, USA.
6. ADF2017.1, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, <http://www.scm.com>.

Analytical Inversion Regional Modeling System for Estimation of National-Scale CO₂ Fluxes – Results of the CoCO2 Project

Michał Gałkowski^{1,2}, Piotr Sekuła^{2,3}, Mirosław Zimnoch²

¹ MPI for Biogeochemistry, Department Biogeochemical Signals, Jena, Germany

² AGH University of Krakow, Faculty of Physics and Applied Computer Science, Kraków, Poland

³ Institute of Meteorology and Water Management, National Research Institute, Kraków, Poland

michal.galkowski@bgc-jena.mpg.de, piotr.sekula@imgw.pl, zimnoch@agh.edu.pl

Keywords: anthropogenic emissions, carbon dioxide, inversion modeling system, WRF model

1. Introduction

Monitoring anthropogenic emissions into the atmosphere is a key element in mitigation of climate change. Besides the widely used bottom-up methods, the top-down approach based on atmospheric observations and numerical models is becoming widely used. Bayesian inversion systems using state-of-the-art atmospheric transport models allow for estimating CO₂ emissions based on atmospheric observations of its mole fractions. Here, we present a system developed and tested for Poland within the CoCO2 Horizon 2020 project carried out in 2021-2023.

2. Description of the system

The inversion modeling system is based on the WRF-GHG model [1]. It has three embedded domains of 5 km, 1 km, and 200 m resolutions (Fig. 1). The modelled passive tracers, representing anthropogenic emission sectors and biogenic fluxes, were used in the inversion framework to assimilate surface data from stations across Europe and calculate a set of scaling factors (and their uncertainty) to correct the a-priori emission fields taken from the TNO inventory for anthropogenic emissions [2] and the offline Vegetation Photosynthesis and Respiration Model model for the biosphere [3].



Fig. 1. Extent and topography of the WRF domain for the parent domain (5 km) and location of two subdomains (1 km and 200 m) marked with rectangles.

3. Results

The simulations presented here have been performed using parallel WRF code run on up to 21 nodes of Ares cluster in ACC Cyfronet AGH, with approximately 500k computational hours and 44 TB of storage. Calculations based on two summer and two winter months (2018

and 2021) showed a good agreement between measured and simulated CO₂ mole fractions for most of the observation sites (e.g., Fig. 2), and correct trend of the signal with dominating anthropogenic emissions for winter season and photosynthetic sink for summer. It has also been demonstrated that, based on current measurement network in Poland (two sites), it is possible to improve the biogenic fluxes; however, the system is not sensitive to anthropogenic emissions. To overcome this issue, an expansion of the GHG observing network is necessary. Alternatively, satellite observations from the upcoming ESA missions might be assimilated. However, these will require development of the inversion system.

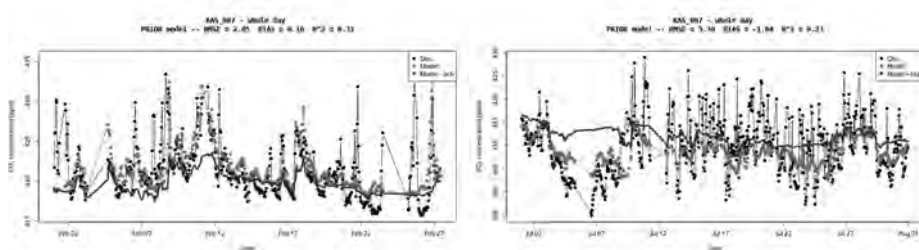


Fig. 2. Comparison of CO₂ molar fractions simulated and measured at Kasprzy Wierch station during summer and winter 2021.

4. Conclusions

- Based on the current results, we demonstrate that the system is able to make corrections of biogenic flux components on a national scale (Poland, Germany).
- Due to deficiencies in the GHG observation network, data-driven updates of anthropogenic fluxes in Poland were not possible - expansion of GHG observation is critically needed for that purpose.

Acknowledgements. This project has been partially supported by the European Union’s Horizon 2020 research and innovation programme under grant agreements No. 958927, and the subsidy of the Ministry of Education and Science. The research results presented in this paper have been developed with the use of equipment financed from the funds of the “Excellence Initiative – Research University” programme at the AGH University of Krakow. The authors gratefully acknowledge the Polish high-performance computing infrastructure PLGrid (HPC Centres: ACC Cyfronet AGH) for providing computer facilities and support within computational grant no. PLG/2022/015860.

References

1. Beck, V. et al. (2013): WRF-Chem simulations in the Amazon region during wet and dry season transitions: evaluation of methane models and wetland inundation maps, *Atmos. Chem. Phys.*, 13, 7961–7982, <https://doi.org/10.5194/acp-13-7961-2013>.
2. Kuenen, J. et al. (2022). CAMS-REG-v4: a state-of-the-art high-resolution European emission inventory for air quality modelling. *Earth System Science Data* 14(2), 491-515. 10.5194/essd-14-491-2022.
3. Pathmathevan, M. et al. (2008). A Satellite-based Biosphere Parameterization for Net Ecosystem CO₂ Exchange: Vegetation Photosynthesis and Respiration Model (VPRM). *Global Biogeochemical Cycles*. 22(2). 10.1029/2006GB002735.

Big Data in Proton Beam Monitoring: LGAD Detectors Study

Leszek Grzanka^{1,2}, Nicola Minafra⁴, Ronan McNulty³, Tomasz Nowak², Jan Swakoń²,
Piotr Rzeźnik¹

¹ AGH University of Krakow, Kraków, Poland

² Institute of Nuclear Physics, Polish Academy of Sciences, Kraków, Poland

³ School of Physics, University College Dublin, Belfield, Dublin 4, Ireland

⁴ Department of Physics & Astronomy, University of Kansas, Lawrence, KS 66045, USA

grzanka@agh.edu.pl

Keywords: proton beam, silicon detectors, radiotherapy

1. Introduction

The radiotherapy of cancer with proton beams is a novel method which provides very good clinical outcome and sub-millimeter precision in shaping the therapeutic beam. Proton therapy requires an accurate dosimetric method to capture the spatial dose distribution and ensure that the treatment planning goals are met. The additional quantities describing proton beams are useful in the research areas: kinetic energy of protons measured in single particle counting approach can be correlated with radiobiological effect and provide additional insight into treatment accuracy.

2. Description of the problem

In our work we focus on the application of fast silicon detectors (in particular Low Gain Avalanche Detectors, LGAD) as a prototype dosimeter sensor for proton therapy. Such detectors have sensitive areas with dimensions less than a millimeter, have linear response and are stable after large radiation doses. Its fast response makes them suitable for time-of-flight analysis, which can be used to estimate the kinetic energy spectrum of the beam.

3. Related work

The LGAD sensors were already tested in medical linac accelerator as radiation monitor and prototype dosimeter. In the work described in [1] the LGAD were exposed to the electron beams proved to have promising spatial (sub-mm) and temporal (sub-ns) resolution. The study reported in [1] doesn't present, however, a detailed comparison of the beam current measured by ionization chambers and derived from particle-counting. Another approach was reported in [2] where the LGAD sensors were used to perform proton radiography, exploiting their excellent time-of-flight capability. Our approach goes a step further, as we are exploiting the full waveform shape collected in a large dataset, which enables us to gain more insight into the therapeutic beam.

4. Solution of the problem

The LGAD detectors were used to characterize the beam delivered by the AIC-144 cyclotron in the Institute of Nuclear Physics in Krakow. The beam used in the irradiation was used in the routine treatment of patients with ocular tumors in the years 2011-15. It has a very special time structure which can be exploited by fast detectors. The detector readout board was connected to the LeCroy oscilloscope which saved the signal digitized at a 10 GHz rate. The time series with data of 4 LGAD channels were saved in the internal memory of the oscilloscope

and periodically dumped into the fast external SSD drive. With this mode we've been capturing continuously 500ms of the beam time, with another couple of seconds needed to flush the buffers. We collected 3 TB of the time series data, saved for further processing in the S3 file system available in ACC Cyfronet. After the first data reduction, based on peak selection, we ended up with around 500 GB of data. Cloud server storage S3 facilitates data exchange and allows data exploration for many GBs sets using JSROOT libraries and a simple web browser client. As the data doesn't fit into memory, the full data analysis was performed using various approaches on the Ares supercomputer: starting from simple batch processing, executing analysis workflows using the HyperQueue tool, deploying Jupyter notebooks on the full Ares node and distributing the data with Dask library.

5. Conclusions

The sensor readings in single particle counting mode were compared with standard methods where ionization chambers, integrating collected charge in 0.1s windows are applied. The LGADs were able to correctly estimate the proton fluence. On top of that the information on proton energy could be deduced using time-of-flight methods. The LGAD sensors revealed as well the internal time structure of the beam macropulses and its frequency distribution.

Acknowledgements. This work/research has been supported by the Eurolabs Transnational Access Program Grant number EURO-LABS-ALTO-001 and by the Polish Ministry of Education and Science, project 2022/WK/14. The numerical experiment was possible through computing allocation on the Prometheus system at ACC Cyfronet AGH under the grants plglgad2023 and plglgad2023raw.

References

1. Isidori, Tommaso, et al. „Performance of a low gain avalanche detector in a medical linac and characterisation of the beam profile.” *Physics in Medicine & Biology* 66.13 (2021): 135002.
2. Ulrich-Pur, Felix, et al. „First experimental time-of-flight-based proton radiography using low gain avalanche diodes.” arXiv preprint arXiv:2312.15027 (2023).

Pedestrian Crossing Intention Detection with Spiking Neural Networks in Challenging Road Conditions

Szymon Mazurek^{1,2}, Mustafa Sakhai², Maciej Wielgosz^{1,2}

¹ ACC Cyfronet AGH, Nawojki 11, 30-950 Kraków, Poland

² AGH University of Krakow, al. Mickiewicza 30, 30-059 Kraków, Poland

s.mazurek@cyfronet.pl {szmazurek, msakhai, wielgosz}@agh.edu.pl

Keywords: spiking neural networks, pedestrian intention, pedestrian detection, autonomous driving

1. Introduction

Recent breakthroughs in machine learning and artificial intelligence have led to ground-breaking applications in various domains, including autonomous driving. In this paper, we investigate the performance of the spiking neural networks (SNN) in the detection of pedestrians crossing the street. We use simulation datasets for this task that provide data from standard cameras and dynamic vision sensors in standard and challenging weather conditions. To train the network, we utilize the surrogate gradient method.

2. Description of the problem

The task is to detect if a pedestrian is crossing the street in the given frame of a recording. For this purpose, two subsets with sceneries in either good or bad weather conditions were generated using the CARLA simulator [2]. Each clip included the recording with DVS alongside a standard RGB camera. DVS data was converted to standard image format and resized to 256x450 resolution alongside the RGB images from the base 600x1600 resolution. The good weather subset consisted of 118 videos, the bad one of 82. Each frame in the clips was associated with a label, indicating if in the frame the pedestrian was crossing the street or not.

3. Related work

SNNs, due to their low latency and energy consumption, are especially promising as implementations of intelligent systems in autonomous driving. The concept of using such networks in edge devices has received increased interest in recent years. Cordone et al. [5] used surrogate gradient-trained networks for automotive object detection and classification adopting well-known deep network architectures to spiking versions. These models have reached high performance, comparable with the original dense models. Su et al. [3] introduced a spiking YOLO [1] variant, once again nearly matching the performance of a non-spiking one with a large reduction in computational costs.

4. Solution of the problem

Implementation details. For initial experiments, we conducted frame-level detection using a spiking version of ResNet18 [4] with ImageNet weights, trained with surrogate gradients. For that purpose, we chose to approximate the nonlinearity with simple leaky-integrate-and-fire neuron models, with a decay constant of 2, a firing threshold of 1 V, and a hard reset of membrane potential after emitting the spike. The neuron's response was approximated with sigmoid as a surrogate function, with the smoothing factor of 5.

The dataset was split into training, validation, and testing subsets, with the testing subset comprising 15% of the total dataset. As a loss function, we chose binary cross entropy with associated weights for the positive class to counter the present class imbalance, as the ratio of negative to positive frames was approximately 4:1. We used AdamW optimizer with a learning rate of 0.001, and a weight decay of 0.1 and trained for a maximum of 100 epochs. The batch size varied between 8, 16, and 32, depending on the GPU RAM capabilities for a given experiment. To measure the performance on the test subset, we used F1-score and AUROC. The experiments were implemented in Python 3.10 using PyTorch 2.0.1 and SpikingJelly 0.0.0.0.14. The training was performed on an A100 GPU with 16 CPU cores in the Athena supercomputer.

Results. The experiment results are shown in Tab. 1. The initial classification performance for the model trained using only a single sample shows poor performance, with most of the AUROC scores close to 0.5, therefore close to random choice. F1-scores for these runs were also unsatisfactory, failing to surpass the 40% mark. These results were however predictable, as spiking networks are inherently temporal, and their response is incremental, based on the series of events. We, therefore, chose to repeat the input frame 10 times to simulate the temporal properties. This change induced a profound increase in the model's performance. For every subset and modality, except the RGB data from the bad weather subset, were close to 1, indicating near-perfect class separability. Interestingly, for the good weather subset, the model performed equally well on RGB and DVS data. On the other hand, for the bad weather subset, the DVS modality performed much better, with AUROC of 0.9868 and an F1-score of 83.13% compared to the 0.8699 and 53.82% for the RGB modality for the same metrics respectively.

Tab. 1. Table shows the test AUROC and F1-score on the test subset for a spiking neural network trained with different amounts of sample repeats in both data subsets and modalities. Bold numbers indicate the best results for a given subset.

Data subset	Good weather				Bad weather			
Data modality	RGB		DVS		RGB		DVS	
Metrics	AUROC	F1 [%]	AUROC	F1 [%]	AUROC	F1 [%]	AUROC	F1 [%]
Single frame	0.5812	27.74	0.6069	35.46	0.7317	39.34	0.548	21.82
10 frames	0.9924	90.33	0.9895	90.59	0.8699	53.82	0.9868	83.13

5. Conclusions

Detecting pedestrians while driving autonomously is crucial but challenging. Spiking neural networks are becoming a promising solution, offering competitive performance with less latency and energy consumption compared to deep learning models. In the experiments, we have shown that a simple adaptation of a deep neural network to the spiking version trained with surrogate gradients can reach high performance in the pedestrian street crossing detection task. Our results also highlight the seminal role of the temporal aspect of the input data for the network, showing that even a simple repetition of input frames in a single sample can result in a dramatic increase in the model's performance. Also, it can be concluded that the data obtained with DVS can also be used for such classification, being equally good as RGB data, or even greatly superior when the weather conditions are bad. These findings may guide future research with the usage of such vision sensors and SNNs. However, this work can and should be expanded further, as some possible ways of improving the network performance, such as different neuron models or surrogate, functions were not explored.

Acknowledgements. This research has been supported by PLGrid. The numerical experiment was possible through computing allocation on the Athena system at ACC Cyfronet AGH under the grants plgdyplomanci5.

References

1. P. Jiang, D. Ergu, F. Liu, Y. Cai, B. Ma: A Review of Yolo Algorithm Developments, in: *Procedia Computer Science*, Volume 199, 2022, Pages 1066-1073.
2. M. Wielgosz, A. M. López, M. N. Riaz: CARLA-BSP: a simulated dataset with pedestrians, in: *Arxiv cv.CV*, eprint: 2305.00204, 2023.
3. Q. Su, Y. Chou, Y. Hu, J. Li, S. Mei, Z. Zhang, G. Li: Deep Directly-Trained Spiking Neural Networks for Object Detection, in: *Proceedings of IEEE International Conference on Computer Vision (ICCV) 2023*.
4. K. He, X. Zhang, S. Ren, J. Sun: Deep Residual Learning for Image Recognition, in: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
5. L. Cordone, B. Miramond, P. Thierion: Object Detection with Spiking Neural Networks on Automotive Event Data, in: *Proceedings of International Joint Conference on Neural Networks (IJCNN) 2022*.
6. E. Neftci, H. Mostafa, Friedemann Zenke: Surrogate Gradient Learning in Spiking Neural Networks: Bringing the Power of Gradient-based optimization to spiking neural networks, in: *IEEE Signal Processing Magazine*, Volume 36, 2019, pages 51-63.

Author Index

- Amano T. 51
- Bielecki A.** 33
Boczar M. 15
Bohdan A. 49
Brela M. Z. 23, 25, 27
Bubak M. 41, 43
- Caputa J. 29, 37, 39
Ciupek D. 35
- Dąbrowska-Boruch A.** 29, 37, 39
Didovets Y. 25
Družbicki K. 11
Dybiec B. 12
- Eilmes A. 59
- Faber K.** 53, 55, 57
Fathalian M. 17
Fischer J. 35
Frączek R. 29, 37, 39
- Gałkowski M.** 69
Gierdziewicz M. 33
Grolik J. 15
Grzanka L. 71
Grzeszczyk J. 29, 37, 39
- Hoshino M. 51
- Jamro E.** 29, 37, 39
Janiuk A. 45, 47
- Kalka A. J.** 27
Karwatowski M. 29, 37, 39, 53, 55, 57
Kasztelnik M. 41, 43
Kica P. 31
Kobzar O. 51
Korchowiec J. 21
Koryciak S. 29, 37, 39
Krupa P. 63
Krupiński J. 37, 39
Kubisiak P. 59
Kulig W. 21
Kurczab R. 61
- Lichołai S. 31
- Łukasik D.** 29, 37
- Malawski M.** 31, 35, 41, 43
Matsukiyo S. 51
Matsumoto Y. 51
Mazurek S. 37, 39, 73
McNulty R. 71
Meizner J. 43
Michalak A. 13, 67
Miklas A. 23
Minafra N. 71
- Niemiec J. 49, 51
Nouri F. H. 47
Nowak T. 71
Nowakowski P. 41, 43
- Pietroń M.** 29, 37, 39, 53, 55, 57
Pięciak T. 35
Piętak K. 57
Pindel P. 29
Pohl M. 51
Poleć P. 41, 43
Postek E. 17
- Rózga M.** 67
Russek P. 29, 37, 39
Rzeźnik P. 71
- Sadowski T.** 17
Sakhai M. 73
Sekula P. 69
Smardz P. 63
Stachowicz-Kuśnierz A. 21
Strzałka K. 37, 39
Sułkowski B. 19
Swakoń J. 71
Synowiec K. 15
- Śliwa P.** 65
- Tahani M.** 17
Torralba Paz G. 49
Turek A. M. 27

Urrutia G. 45

Wiatr K. 29, 37, 39

Wielgosz M. 29, 37, 39, 73

Włodek F. 21

Wróbel P. 59

Zajac K. 41

Zhyhulin T. 41

Zimnoch M. 69

Żurek D. 53, 55, 57

Żurowska O. 13

Published by

ACC Cyfronet AGH
ul. Nawojki 11
30-950 Kraków
www.cyfronet.pl



Sponsors



ISBN 978-83-61433-47-7