ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-TUPOST051

USING DATA INTENSIVE SCIENCE FOR ACCELERATOR OPTIMIZATION*

C. P. Welsch†, University of Liverpool/Cockcroft Institute, UK

Abstract

Particle accelerators and light sources are some of the largest, most data intensive, and most complex scientific systems. The connections and relations between machine subsystems are complicated and often nonlinear with system dynamics involving large parameter spaces that evolve over multiple relevant time scales and accelerator systems. In 2017, the Liverpool Centre for Doctoral Training in Data Intensive Science (LIV.DAT) was established. With almost 40 PhD students, the centre is now established as an international hub for training PhD students in data intensive science. This paper presents results from studies carried out in LIV.DAT into novel high gradient accelerators with a focus on the data science techniques that were used. This includes studies into inverse-designed narrowband THz radiators for ultra-relativistic electrons, simulation of the transverse asymmetry and inhomogeneity on seeded self-modulation of beams in plasma, as well as studies into the physical aspects of collinear laser injection in Trojan Horse laser plasma experiments.

INTRODUCTION

In 2017 LIV.DAT, the Liverpool Centre for Doctoral Training in Data intensive science was established. LIV.DAT has quickly established itself as a center of excellence in data science, across a significant part of STFC research, including nuclear, particle and astrophysics, as well as accelerator science. The center also acts as a model for training cohorts of student in Data Intensive Science and currently trains 36 PhD students [1].

Recent years have witnessed a dramatic increase of data in many fields of science and engineering, due to the advancement of sensors, mobile devices, biotechnology, digital communication and internet applications. Very little targeted training is provided internationally to address a growing skills gap in this area. LIV.DAT provides a comprehensive training program to its students to close this skills gap through a cutting-edge research program.

SELECTED RESEARCH RESULTS

The focus of the LIV.DAT center is on addressing cutting-edge research challenges by using state-of-the-art data science techniques across three scientific work packages:

- Monte Carlo (MC) methods as tools to address a wide range of physics problems, from the dynamic behaviour of galaxies, cross sections in specific particle interactions to dose delivery planning in ion beam therapy;
- High Performance Computing (HPC) and Machine Learning (ML) using computing clusters to simulate cutting edge physics and engineering problems that cannot be dealt with on desktop computers;
- Data Analysis across the entire spectrum of physics research.

All three work packages are highly relevant for accelerator science R&D. In the following, selected research highlights from studies that are presented at this year's IPAC conference are given.

Inverse-Designed Narrowband THz Radiators for Ultra-relativistic Electrons

THz radiation sources are extremely useful for a wide range of applications and a number of methods [2-4] have been used to generate this radiation with lies between the microwave and infrared spectrum. Amongst these methods, the Smith Purcell effect has proven to provide a cost effective and compact solution [5].

Recent advances in computational design and optimization have paved the way for designing the underpinning optical structures algorithmically and thus allowing computer-guided structure optimization. LIV.DAT student Gyanendra Yadav has successfully carried out 3D-simulation studies into the generation of THz radiation from structures designed using the inverse function algorithm [6]. These have formed an important contribution to experiments demonstrating the capabilities of such radiators [7].

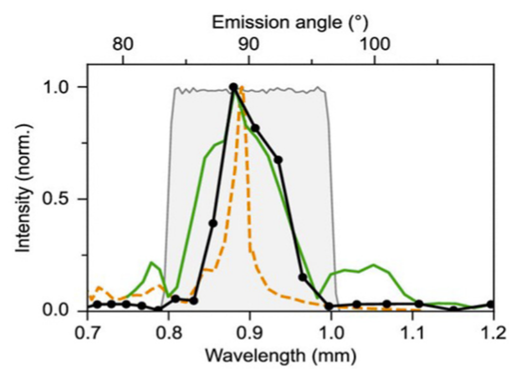

Figure 1: Measured and simulated emission spectrum. The black curve shows experimental data, whilst the orange and green curves show results from frequency domain and time domain simulations, respectively [7].

Full 3D time domain simulations for obtaining the resulting THz radiation spectrum were performed in the established simulation code CST Studio [8]. Figure 1 shows the resulting electromagnetic spectrum as obtained in CST Studio in direct comparison with experimental data and frequency domain simulations. It can be seen that an

* This work was supported by STFC under grant agreement ST/P006752/1.
† c.p.welsch@liverpool.ac.uk

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-TUPOST051

excellent agreement overall was found with a radiation peak at 90 degrees and at a frequency of 0.34 THz.

The method and results have excellent potential for application in future pump-probe experiments and in designing tuneable light sources. It shows a pathway towards narrowband and coherent THz radiation sources at arbitrary wavelength which would be extremely hard to generate through conventional sources.

Transverse Asymmetry and Inhomogeneity on Seeded Self-Modulation of Beams in Plasma

The Advanced Wakefield Experiment (AWAKE) in 2016 became the first every proton-beam-driven plasma wakefield acceleration (PWFA) experiment [9-11]. Proton-beam-driven PWFA (PDPWFA) works by driving an electrostatic Langmuir wave in a plasma channel, using a relativistic beam of charged particles. The electromagnetic fields of the proton beam leads to transverse attraction of plasma electrons towards the beam's propagation axis, setting up an oscillation. The finite speed of the beam hence sets up a series of high and low electron density regions near the axis in its wake, at a characteristic 'plasma wavelength', between which the longitudinal electric field can reach up to 50 GV/m - three orders of magnitude higher than a conventional RF accelerator. A relativistic witness beam injected at the right position into this wakefield can, in principle, continuously gain energy from the plasma wakefield. Using the SPS proton beam as a driver as in AWAKE allows limitations on acceleration length due to energy depletion as with an electron beam driver to be overcome. The stability of particle bunches undergoing seeded selfmodulation (SSM) over tens or hundreds of meters is crucial to the generation of GV/m wakefields that can accelerate electron beams. LIV.DAT student Aravinda Perera used 3D particle-in-cell simulations in *QuickPIC* [12] and compared the results to an analytical model of seeded self-modulation (SSM) of elliptical beam envelopes using linear wakefield theory [13]. The extracted scaling law shows good agreement of power-law scaling, see Fig. 2. The coefficient of the fitted curve is slightly higher than in the model, likely due to simplifying assumptions made for the radial profile.

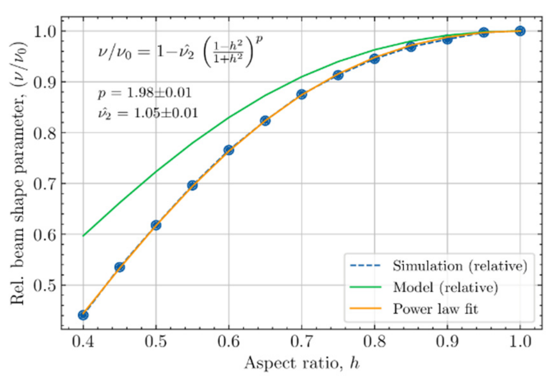

Figure 2: Scaling law for growth rate as a function of beam aspect ratio, h [13].

Collinear Laser Injection at the FACET-II E-310 Trojan Horse Experiment

The Facility for Advanced Accelerator Experimental Tests (FACET-II) is a test accelerator infrastructure at SLAC dedicated to the research and development of advanced accelerator technologies. LIV.DAT student Monika Yadav and co-workers performed simulations of electron beam driven wakefields, with collinear lasers used for ionization injection of electrons [14]. Using the fully relativistic 3D particle-in-cell (PIC) simulation code OSIRIS [15], they ran simulations of the E-310 experiment, using the parameters expected at FACET-II. The longitudinal electric field inside the bubble is shown in Fig. 3. The peak accelerating field in the bubble of 40 GV/m can clearly be seen. The magnitude of the radial fields are dependent on the beam and are highly sensitive to betatron oscillations and current fluctuations. The electric fields at the wake vertex and drive should be lower than the ionization threshold of the HIT gas to ensure that there is no uncontrolled injection which could degrade the beam quality.

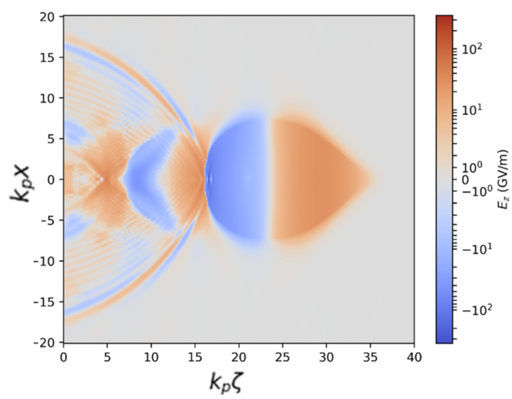

Figure 3: On-axis longitudinal wakefield evolution of the blowout when propagating through the pre-ionized plasma channel [14].

As the wavelength of the betatron radiation is significantly smaller than the scale of the PIC simulation, the resulting radiation was calculated by importing the particle trajectories from the PIC code into a Liénard–Wiechert code. Because most of the betatron radiation in this experiment will be generated by the drive beam, diagnosing the witness beam through betatron radiation presents a significant challenge and detailed simulation studies are key to a full understanding of the underpinning physics processes.

TRAINING

The training program in LIV.DAT builds on existing modules drawn from the University of Liverpool's MSc in Big Data Science. All students undertake 45 credits from this MSc program in their first year, including mandatory courses on data mining and data analysis.

In addition, the center is offering a wide-ranging training program in close collaboration with other major training initiatives. To date, this has included an international school on Monte Carlo Simulations [16], researcher skills

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-TUPOST051

training with researchers from the innovative training networks AVA [17] and OMA [18], an HPC training week hosted by Tech-X in spring 2019, as well as an STFC Summer School on Data Science in 2020 [19]. LIV.DAT has also co-organized a major Symposium on Accelerators for Science and Society in summer 2019 with contributions made by all of its students and talks now available via the event website [20].

The center has also developed a Massive Open Online Course (MOOC) on Data Science which took full advantage of Liverpool's virtual learning environment *Canvas*. It will form the basis of the core training program for PhD students in Liverpool's new STFC Center for Doctoral Training for Innovation in Data Intensive Science, LIV.INNO which will start later this year [21]. LIV.DAT also organizes a monthly seminar that covers wider R&D in Data Science and this is open to participants from around the world [22].

Finally, a particularly unique element of the center is that each student undertakes an industry placement for a minimum duration of six months, working on a topic outside of their core PhD project. This helps broaden their skills and expertise and boost their employability.

SUMMARY AND OUTLOOK

The LIV.DAT Center for Doctoral Training started in October 2017 and is currently training 36 PhD students in data intensive science techniques across a wide range of fundamental and applied research. The importance of these techniques for accelerator optimization was highlighted in this paper on the example of three accelerator science R&D projects which rely on high performance computing techniques. A brief overview of the broad training program that the center is offering was also given. The new CDT LIV.INNO will continue delivering cutting edge data science training to cohorts of students and the wider community for years to come.

REFERENCES

[1] LIV.DAT, `http://www.livdat.org`

[2] X. Zhang, Y. Jin, X.F. Ma, "Coherent measurement of THz optical rectification from electro-optic crystals", *Appl. Phys. Lett.* Vol. 61, pp. 2764–2766, 1992.

[3] E.A. Schneidmiller, *et al.,* "Tunable IR/THz source for pump probe experiments at the European XFEL" in Tschentscher, T. & Tiedtke, K. (eds.) Advances in X-ray Free-Electron Lasers II: Instrumentation, vol. 8778, pp. 151 – 156. International Society for Optics and Photonics (SPIE, 2013).

[4] J. Dai *et al.*, "High-power vacuum terahertz photomixer and integrated circuits based on microscale phototubes". *Opt. Express* vol. 29, 2021. `doi:10.1364/OE.409879`

[5] S.J. Smith *et al.*, "Visible light from localized surface charges moving across a grating", *Phys. Rev.* vol. 92, pp. 1069–1069, 1953.

[6] B. Hermann *et al.*, "Inverse-designed narrowband Thz radiator for ultrarelativistic electrons". *ACS Photonics* vol. 9, pp. 1143–1149, 2022. `doi:10.1021/acsphotonics.1c01932`

[7] G. Yadav *et al.*, "Simulation Study for an Inverse Designed Narrowband THz Radiator for Ultrarelativistic Electrons", presented at IPAC'22, Bangkok, Thailand, June 2022, paper TUPOST049, this conference.

[8] CST Studio, `www.3ds.com`

[9] E. Adli *et al.*, "Acceleration of electrons in the plasma wakefield of a proton bunch", *Nature* vol. 561, pp. 363–367, 2018.

[10] M. Turner *et al.*, "Experimental Observation of Plasma Wakefield Growth Driven by the Seeded Self-Modulation of a Proton Bunch", *Phys. Rev. Lett.* vol. 122, p. 054801, 2019.

[11] K. Rieger *et al*, "Experimental Observation of Proton Bunch Modulation in a Plasma at Varying Plasma Densities", *Phys. Rev. Lett.* vol. 122, p. 054802, 2019.

[12] A. Perera, O. Apsimon, J. Resta-López, and C. P. Welsch, "Modelling Seeded Self Modulation of Long Elliptical Bunches in Plasma", in *Proc. IPAC'21,* Campinas, Brazil, May 2021, pp. 4030-4033. `doi:10.18429/JACoW-IPAC2021-THPAB140`

[13] A. Perera *et al.*, "Simulations of Transverse Asymmetry and Inhomogeneity on Seeded Self-Modulation of Beams in Plasma", presented at IPAC'22, Bangkok, Thailand, June 2022, paper WEPOTK030, this conference.

[14] M. Yadav *et al.*, "Physical Aspects of Collinear Laser Injection at SLAC FACET-II's E-310: Trojan Horse Experiment", presented at IPAC'22, Bangkok, Thailand, June 2022, paper WEPOST041, this conference.

[15] R. A. Fonseca *et al.*, "Osiris: A three-dimensional, fully relativistic particle in cell code for modeling plasma based accelerators", In M.A. Peter et al., Computational Science — ICCS 2002, pages 342–351, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.

[16] Monte Carlo School, `indico.cern.ch/event/656336/`

[17] AVA project, `http://www.ava-project.eu`

[18] OMA project, `http://www.oma-project.eu`

[19] LIV.DAT School on Data Science, `http://indico.ph.liv.ac.uk/event/103/`

[20] Symposium on Accelerators for Science and Society, `indico.cern.ch/event/798052/`

[21] LIV.INNO IPAC

[22] LIV.DAT Data Science seminars, `www.liv.ac.uk/livdat/training-and-events/seminar_series/`