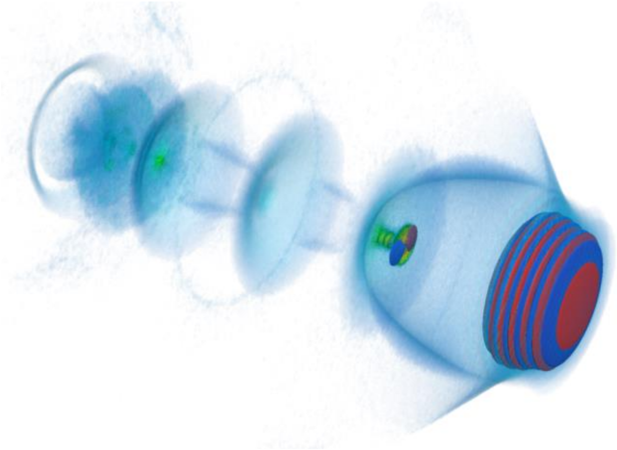
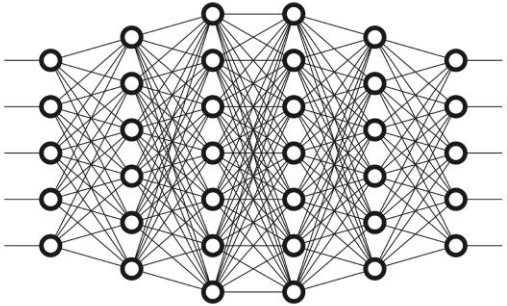
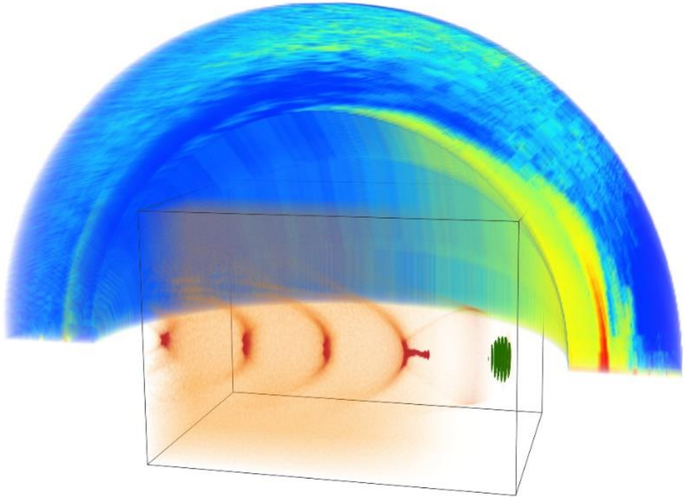
Machine-learning to better understand radiation emitted by laser-plasma interactions

ML@HZDR Symposium 2021 – Open data challenges



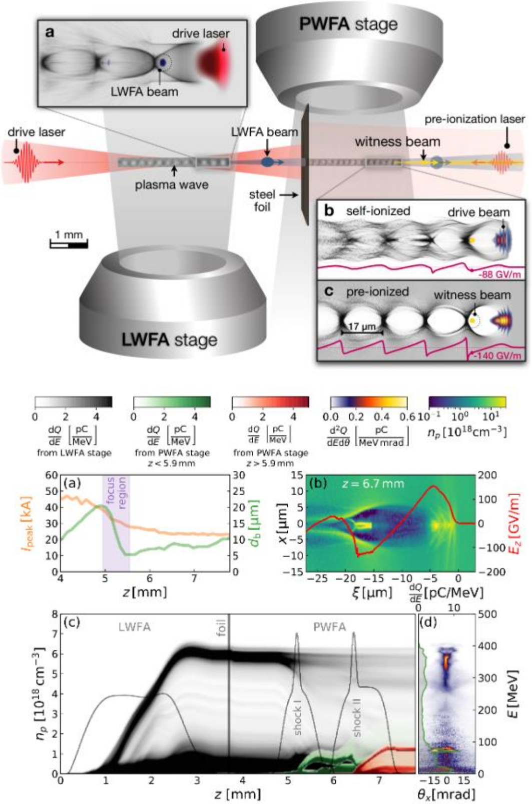




Alexander Debus, Richard Pausch, Matthew Leinhauser, Sunita Chandrasekaran,
Michael Bussmann, Friedrich Bethke, Anna Willmann, Jakob Dieckmann, Nico Hoffmann

Challenge: Understand µm- and fs-scale dynamics

Compact plasma wakefield accelerators

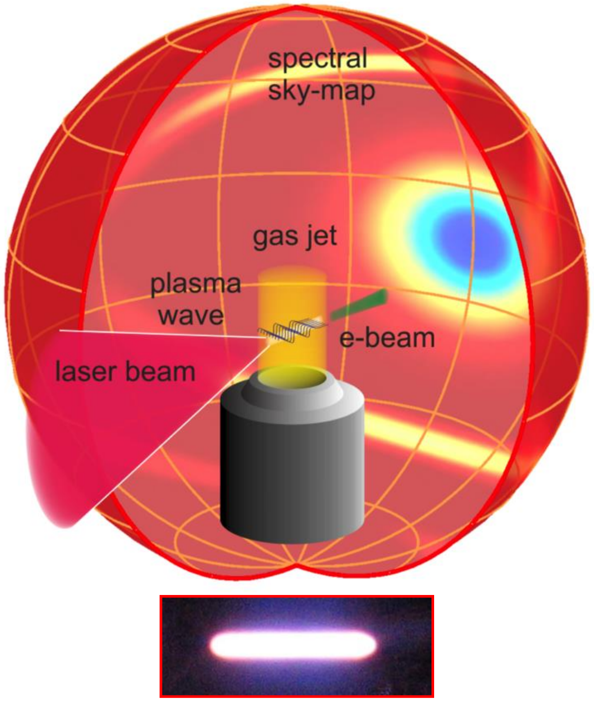


Demonstration of a compact plasma accelerator powered by laser-accelerated electron beams
T. Kurz et al., Nature Communication (2021)
Gas-dynamic density downramp injection in a beam-driven plasma wakefield accelerator
J. P. Couperus Cabadağ et al., Phys. Rev. Res. 3, L042005 (2021)

Simulating the radiation from Laser Plasma Interactions

shedding new light into the dynamics of laser-accelerated electrons

- Modeling complex plasma dynamics require **simulations**.
- The predicted **particle dynamics not directly accessible**.
- Synthetic (simulated) radiation **bridges the gap** between microscopic particle dynamics and observable radiation by predicting how the particles radiate.
- Radiation spectra give insight into the **momentum distribution.**
- **Radiation provides a direct window to the µm and fs-scale dynamics of laser-plasma interactions.**



Laser-plasmas radiate!

Simulating the radiation from Laser Plasma Interactions

shedding new light into the dynamics of laser-accelerated electrons

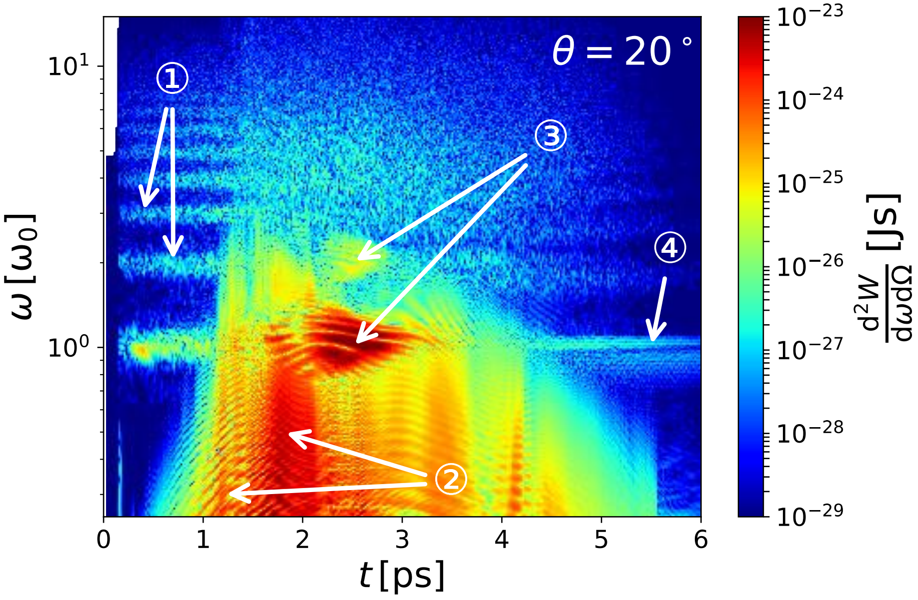
- **Radiation** from laser-plasma interactions **is ubiquitous and experimentally accessible** via imaging and spectroscopy.
- Yet, the data is hard to interpret and thus is often discarded.
- What do the radiation signatures mean? Are these robust and unambiguous?

(1) ?

(2) ?

(3) ?

(4) ?



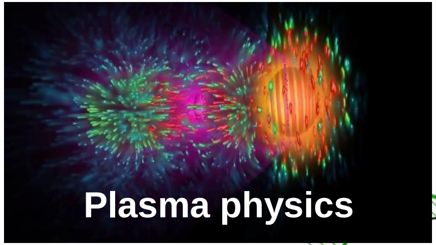
So what is the challenge in radiation from plasmas?

It is computationally very expensive!

Compute radiation within particle-in-cell simulation (**in-situ**)

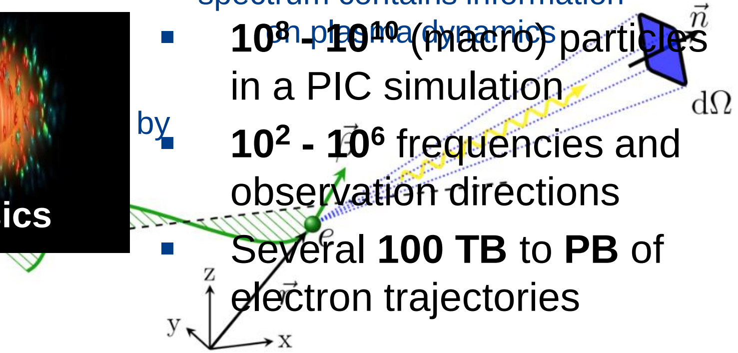
Requirements

- spectra range from **IR to x-ray**
- resolve **coherent** and **incoherent** radiation **simultaneously**
- include **polarization** properties
- resolve **temporal evolution** of spectra
- compute radiation of all **billions of particles** in PIC simulation → **Processing must happen online during PIC (in-situ).**



spectrum contains information on plasma dynamics

by



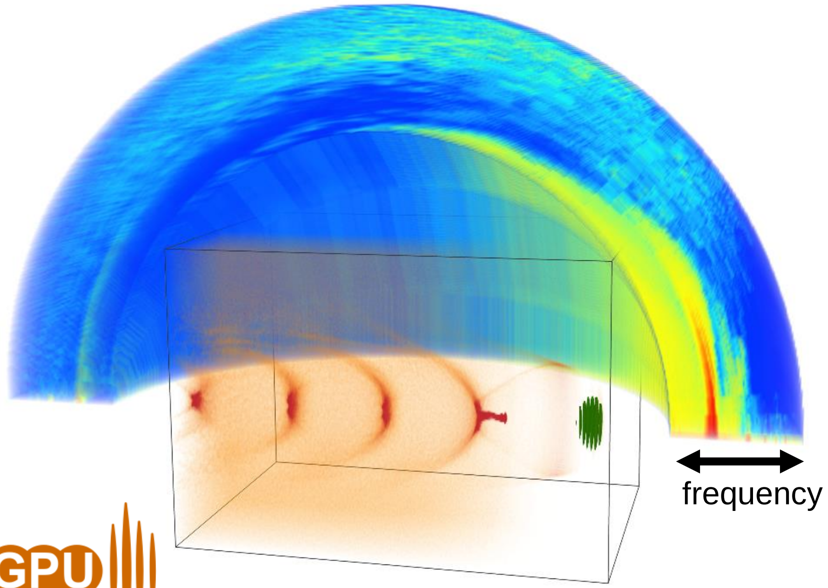
- **10^8 - 10^{10}** (macro) particles in a PIC simulation
- **10^2 - 10^6** frequencies and observation directions
- Several **100 TB** to **PB** of electron trajectories

$$\frac{\mathrm{d}^2 W}{\mathrm{d}\Omega\,\mathrm{d}\omega} = \frac{1}{16\pi^3\varepsilon_0 c}\left|\sum_{k=1}^{N_p}\int_{-\infty}^{+\infty} q_k \cdot \frac{\vec{n}\times\left[\left(\vec{n}-\vec{\beta}_k\right)\times\dot{\vec{\beta}}_k\right]}{\left(1-\vec{\beta}_k\cdot\vec{n}\right)^2}\cdot \mathrm{e}^{\mathrm{i}\omega(t-\vec{n}\cdot\vec{r}_k(t)/c)}\,\mathrm{d}\,t\right|^2$$

So what is the challenge in radiation from plasmas?

It is computationally very expensive!

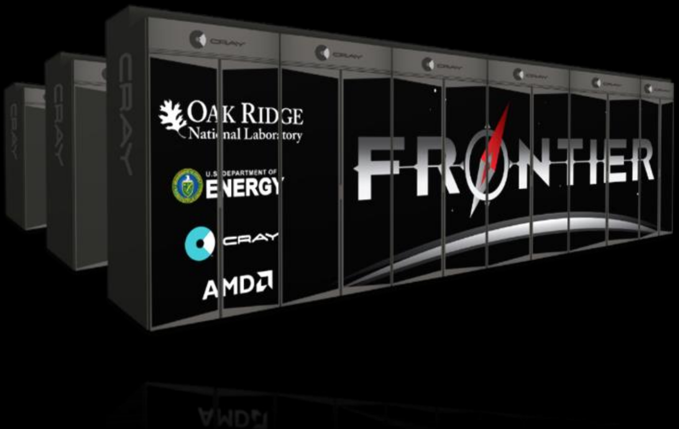
- For exploring radiation signatures in a first laser wakefield acceleration simulation
  Total simulation time: over 9 months



Simulation setup:

- 512 frequencies IR-VIS-UV
- 250 virtual observers on half-dome
- Computed for ~10^9 particles at each time step
- **Over 20x more compute intensive than simulation w/o radiation**

Road to Exascale for PIConGPU



Juwels Booster (Nov 2020), JSC at FZ Juelich

73 PetaFLOPS (7th Top 500, 3rd Green 500)

936 nodes, 3744 NVIDIA Volta A100s

AMD EPYC Rome CPU

Summit, ORNL (2018)

200 PetaFLOPS

4,608 nodes

27,648 NVIDIA Volta V100s

IBM POWER9 CPU

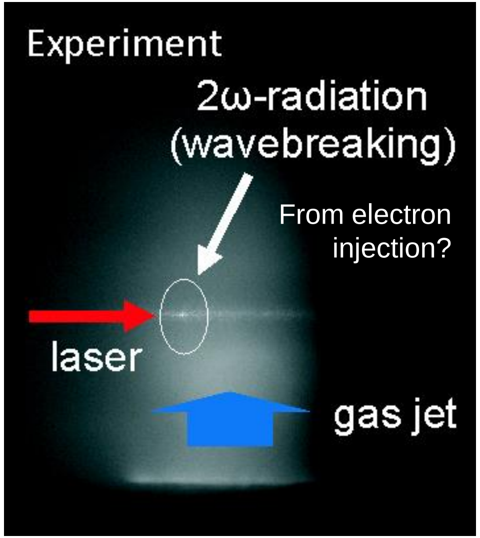
Frontier, ORNL (2022)

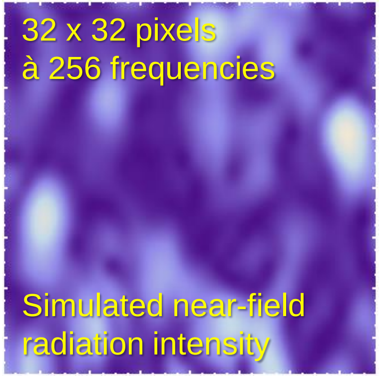
> 1.5 ExaFLOPS

AMD GPU hardware

Cray architecture / compilers

Knowledge extraction challenge:
Where are the sources of the emitted radiation?





10nm bandwidth at
400nm center wavelength

Synthetic diagnostics

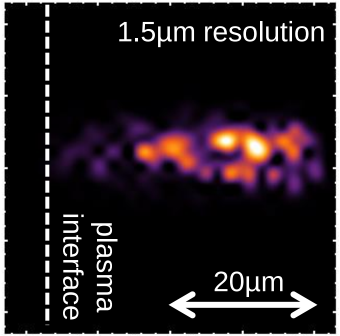
f/3.7 imaging

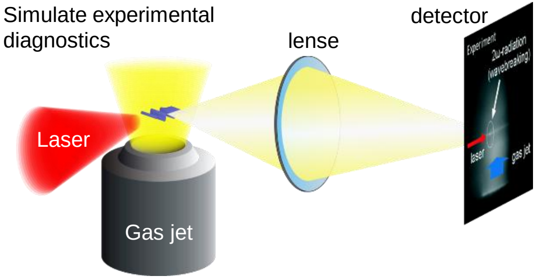
Huygen-Fresnel

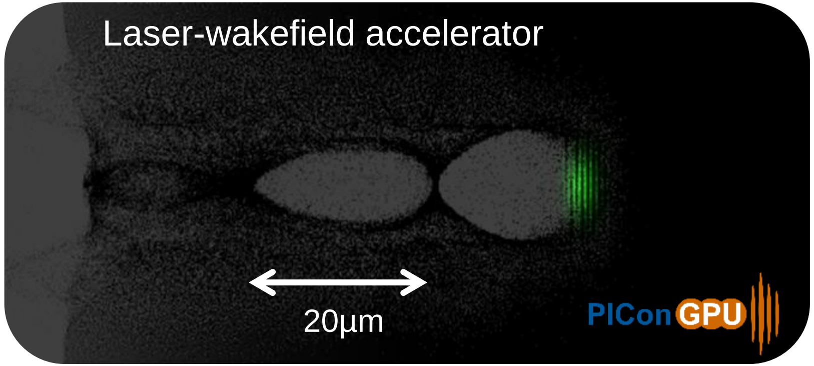
(15° x 15° solid angle)

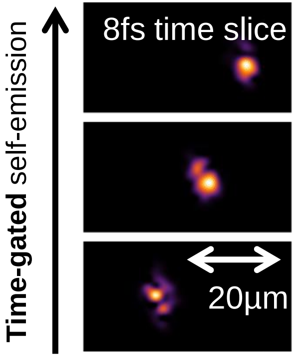
46.3 µm x 46.3 µm
field of view

Time-Integrated self-emission

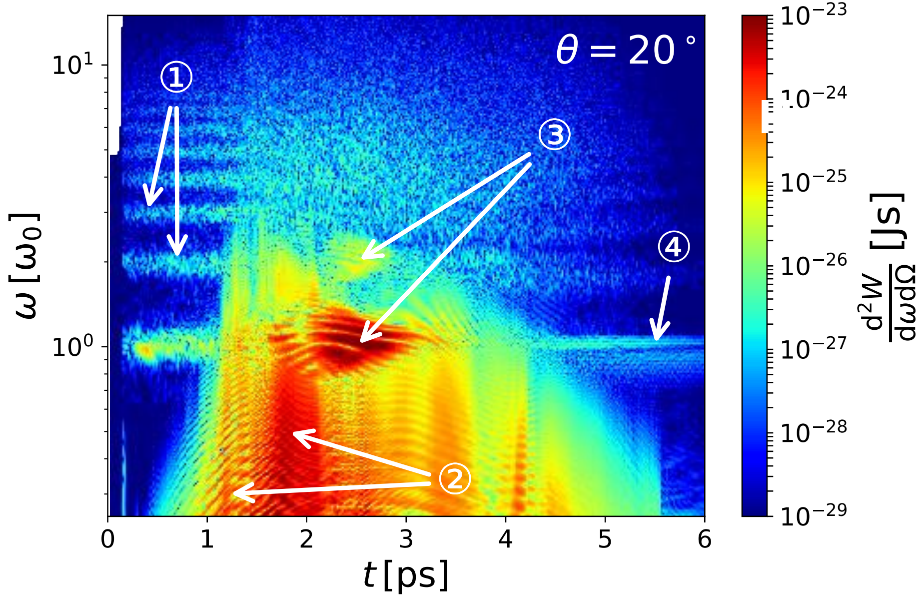








Knowledge extraction challenge:
What is the physics behind the emitted radiation?



(1) Higher harmonics

(2) Wave-breaking signature

(3) Blowout signal

(4) Red-shifted scattering from the depleted laser pulse

Current state-of-the art of knowledge extraction:
PhD student brain power and running lots of reduced-model simulations.

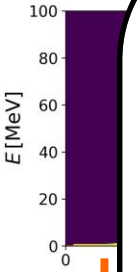
PhD thesis: Richard Pausch

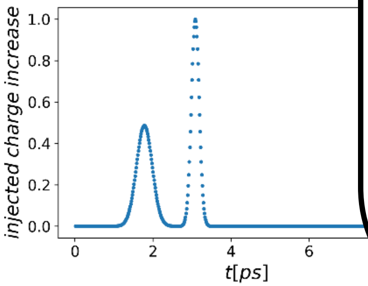
Initial results: Reconstructing plasma dynamics by invertible neural networks

From PIC simulation over synthetic data to normalized input for the neural network

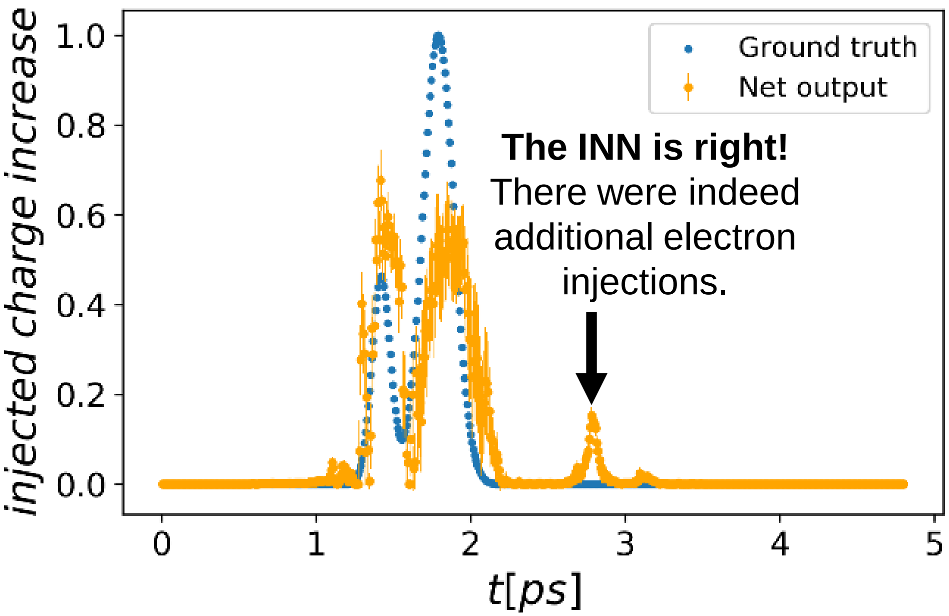
Time

electron

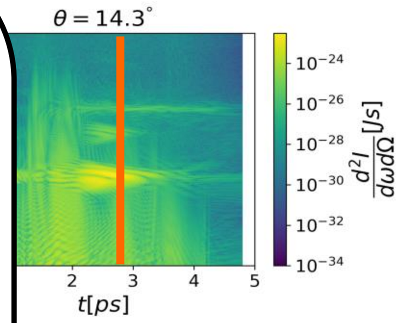


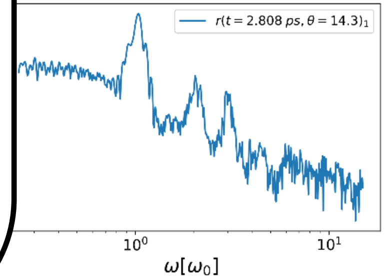


Comparison INN output to ground truth



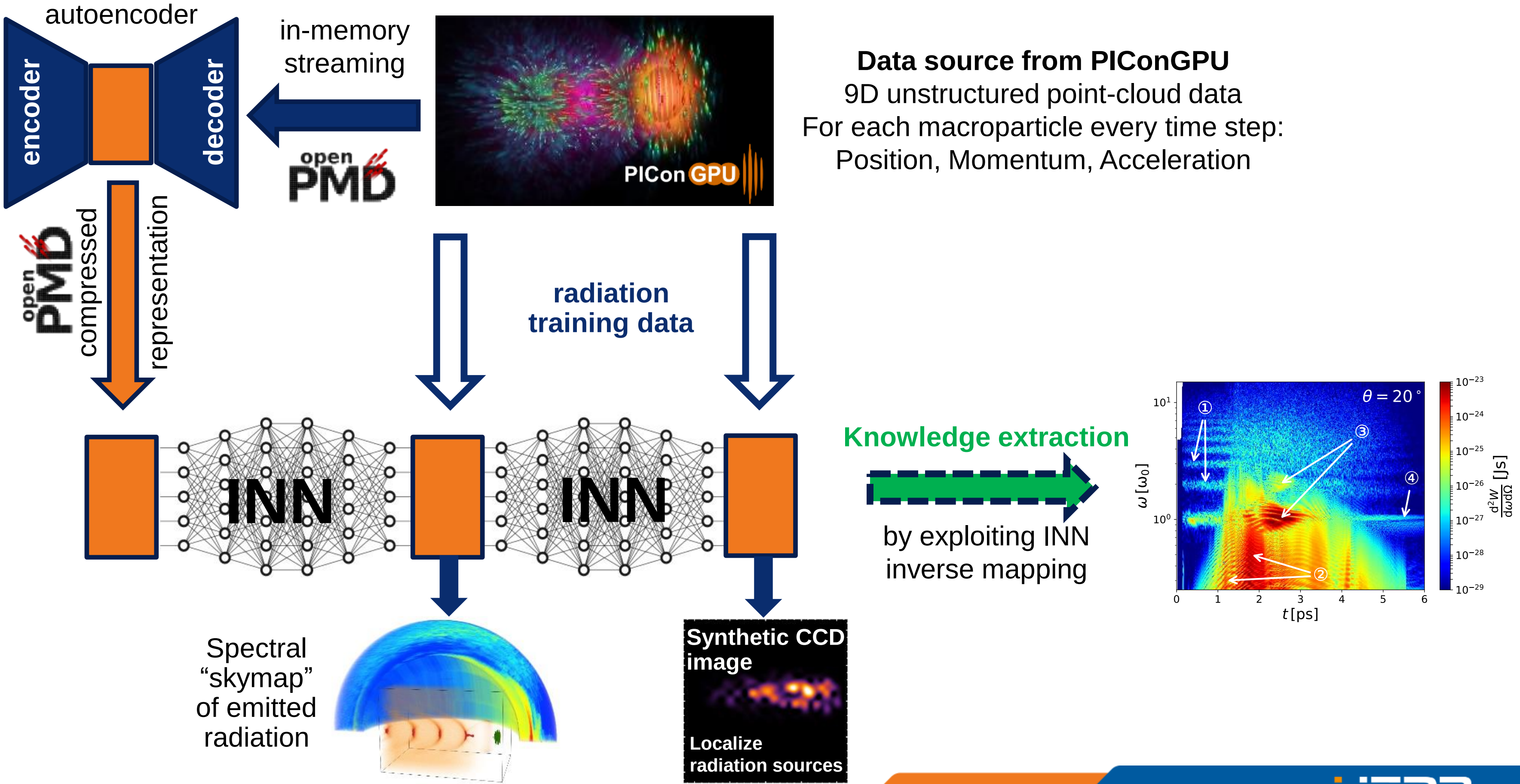
Radiation spectrum





Bachelor thesis: Friedrich Bethke

Data Challenge: Large-scale distributed ML



Data source from PIConGPU
9D unstructured point-cloud data
For each macroparticle every time step:
Position, Momentum, Acceleration

Conclusions

- **Electromagnetic radiation from laser-plasma interactions is key to resolving plasma dynamics at the fs and μm scale.**
- **Physics challenge:** What do radiation signatures in spectra and images mean? Which of these are robust, unambiguous and quantitative?
- **Performance challange:** Calculating plasma radiation for synthetic spectra and images requires high-performance computing resources at the exascale.
- **Knowledge extraction challenge:** Deducing from spectra and images the physics dynamics
is an inverse problem involving PBytes of data.
- **Required steps for machine learning**
  - Massively distributed ML: Integrate an autoencoder approach based on point-cloud data (macroparticle positions, momenta and acceleration) into PIConGPU.
  - Learning compressed representation mapped to spectral and imaging radiation diagnostic via invertible neural network.

„White light generation“ of DRACO laser pulse in air

Thank you for your attention!

Alexander Debus