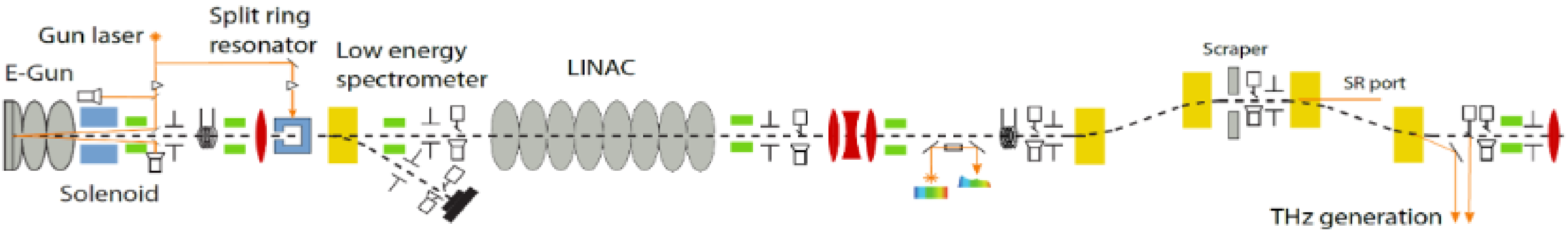
Karlsruhe Institute of Technology

Case Study: Surrogate Modelling for FLUTE

Chenran Xu, Andrea Santamaria Garcia

Artifact Preparation Workshop, 27-28 Nov. 2023

Test Facility FLUTE



The linear accelerator FLUTE at KIT serves as a test facility for various accelerator physics studies.

Wide parameter space:

- Charge: 1 pC - 1 nC
- Energy: 40 - 100 meV
- Rep. rate: up to 50 Hz

Multiple operation modes:

- Test for beam diagnostics (TDS)
- THz Generation (CSR, CTR, Undulator)
- Injector into cSTART storage ring

Surrogate Model for Reinforcement Learning

Reinforcement learning is a promising method to achieve autonomous accelerator operation.

- A. Eichler, First Steps Toward an Autonomous Accelerator, a Common Project Between DESY and KIT
- J. Kaiser, Learning to Do or Learning While Doing: Reinforcement Learning and Bayesian Optimisation for Online Continuous Tuning
- C. Xu, Beam Trajectory Control with Lattice-Agnostic Reinforcement Learning

However, (model-free) RL algorithms are very sample inefficient, requiring often $10^5 - 10^6$ interactions to train.

A **fast**, **accurate** surrogate model will greatly benefit the RL method development for accelerators.

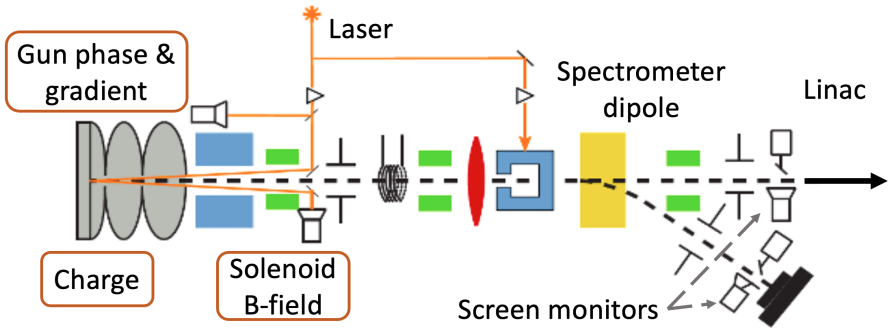
- Fast → reduced training iteration time
- Accurate → less challenging sim2real transfer

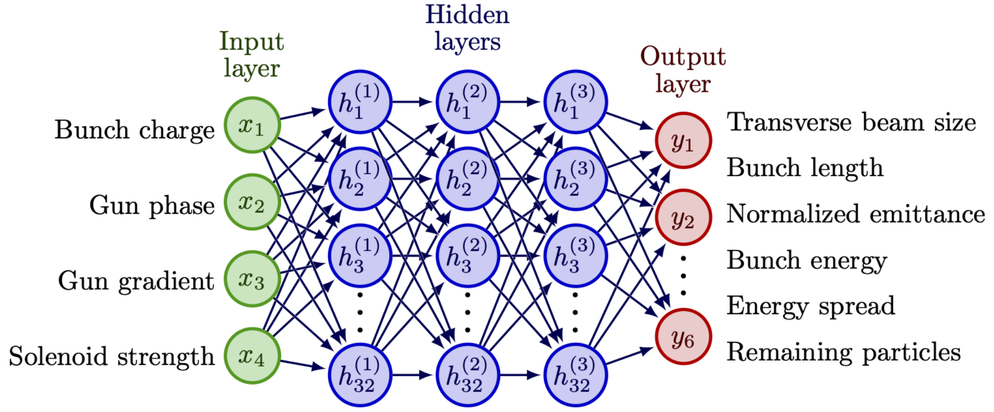
Previous Work: Low-Energy Section Modelling

Proof-of-principle surrogate modelling using fully-connected neural network

Restrict to 4 input and 6 output parameters

- **NN Structure**: $[32, 32, 32]$, $\tanh$ activation
- **Train data**: 10^4 ASTRA simulations with uniform randomly sampled parameters
- **Training**: 200 epochs

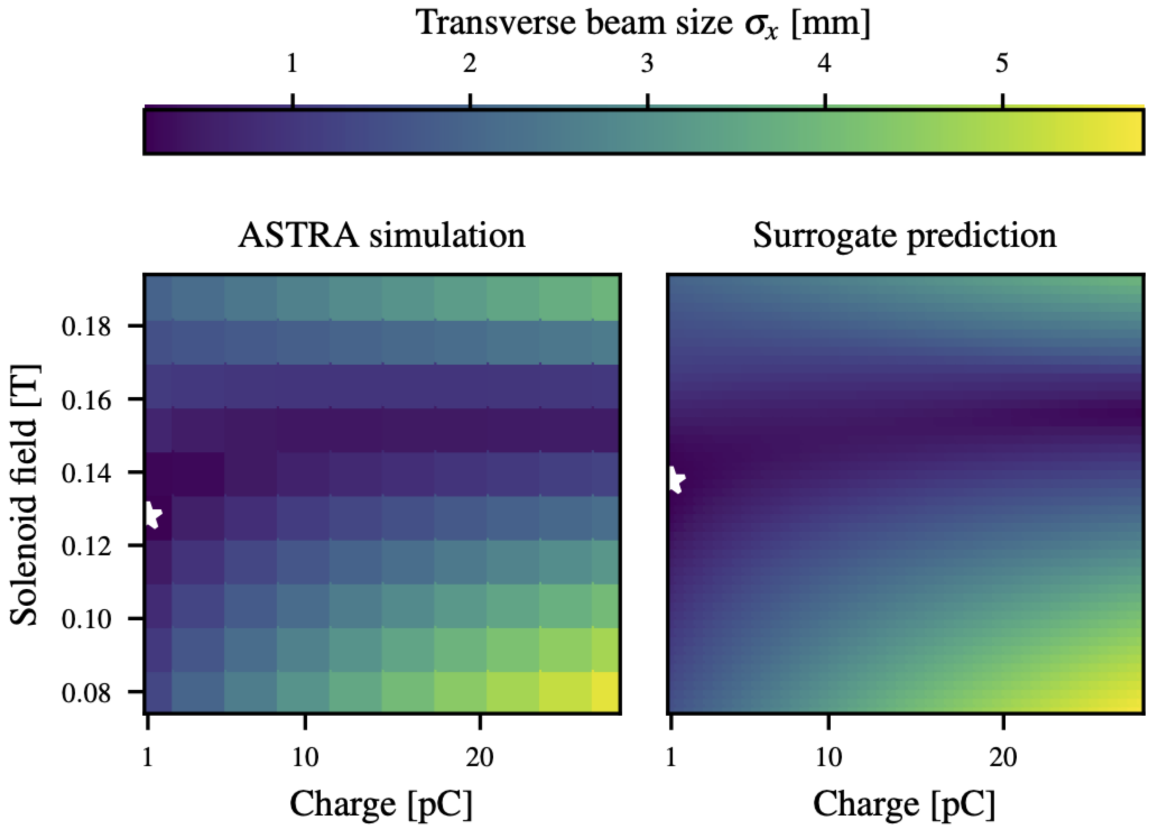




See IPAC22-TUPOPT070 for more details.

Low-Energy Section Model Validation

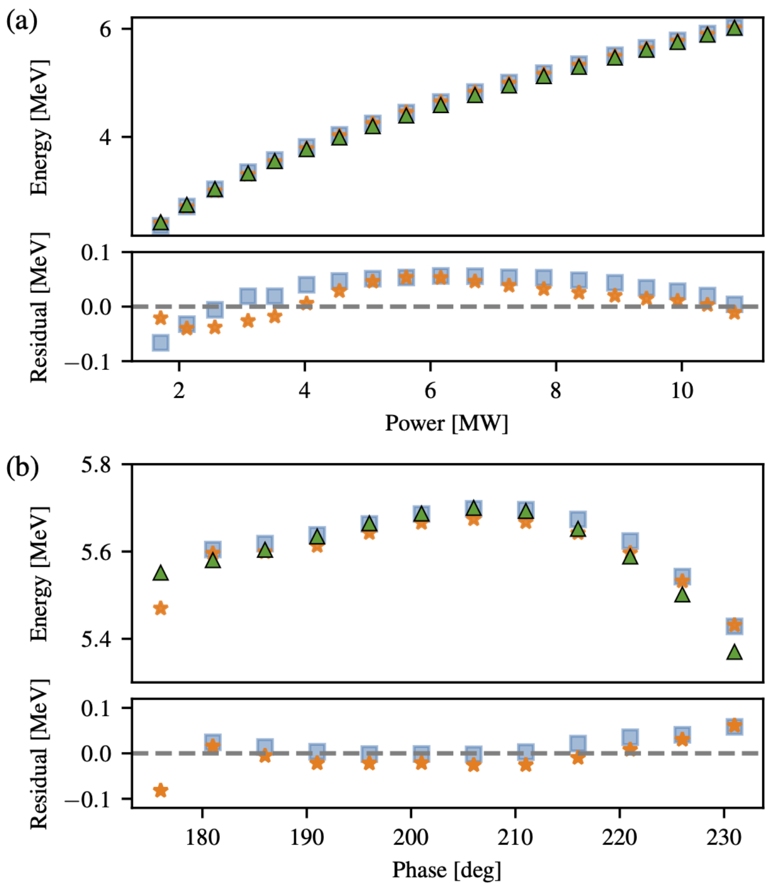
The NN model (with < 1 ms inference time) has overall good agreement with ASTRA results (≥ 1 min).



Comparison with Real Measurement

The train model could also provide accurate prediction for real-world measurement.

Possibility to use the surrogate as **virtual diagnostics** in case of destructive measurements.



Outlook 1: Active Learning

For random or grid sampling, $N_{\mathrm{Samples}} \propto \exp\left(D_{\mathrm{Input}}\right)$.

However, a large percentage of the parameter combinations lead to useless results.

Active learning can be used to reduce the required samples, for example

- Bayesian active learning (c.f. Gal2017)
- Surrogate for subsystems to validate/reject parameter regions, i.e. for DA estimation

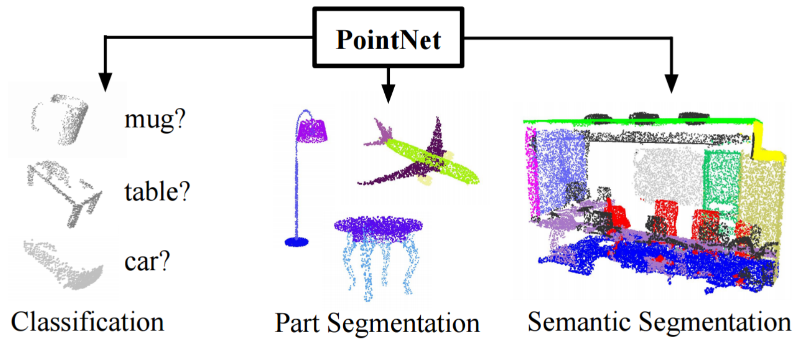
Outlook 2: Point Cloud Representation

In order to predict more complex beam dynamics (micro-bunches, sub-structure), we should move from scalar values to full 6-D phase space prediction.

Auto-encoders and NNs have been successfully applied to predict 2D phase space information, but are harder to be extended to 6D.

The **Point cloud** seems to be an ideal representation of the **macro-particles** used in tracking simulations.

Inspired by Hayg, see also the later talk on **LinacNet**.



Reference: PointNet

Outlook 3: Combine with Differentiable Simulation

Use a surrogate model only when necessary.

A complete (start-to-end) surrogate model could **over-complicate** the problem if the beam dynamics can be accurately modeled, which is actually the case in many simple sections of the accelerator.

Recent developments in **differentiable beam dynamics** simulations enables fast tracking. For example:

- Cheetah: GitHub repository
- BMAD-X: description ; use case phase space reconstruction,

Thanks for your attention

Backup