Surrogate Modeling of Ion Acceleration in the Near-Critical Density Regime with Invertible Neural Networks

14.12.2021

Thomas Miethlinger

HZDR

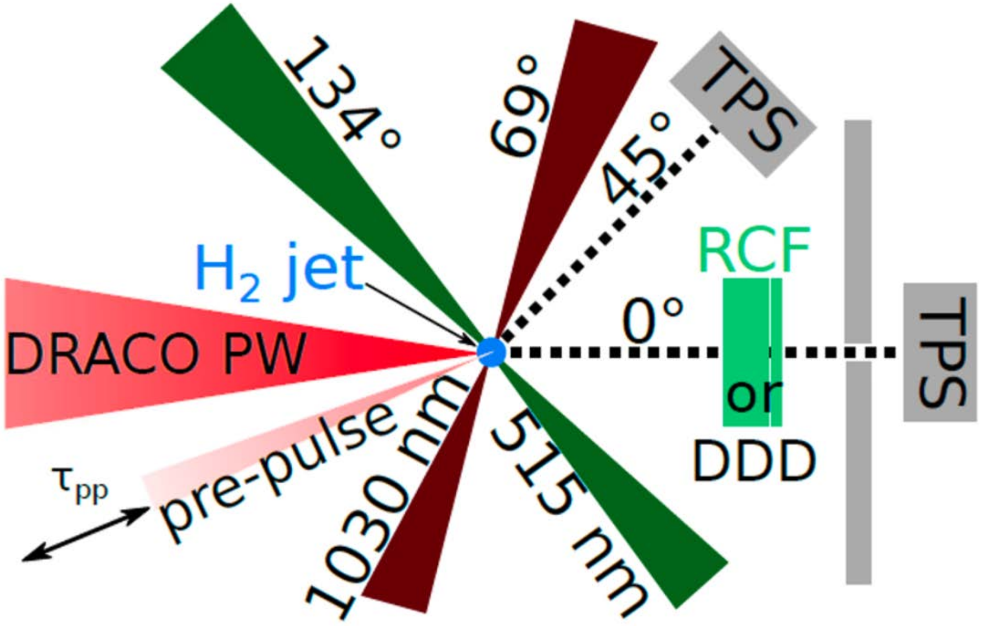
HELMHOLTZ ZENTRUM DRESDEN ROSSENDORF

Mitglied der Helmholtz-Gemeinschaft

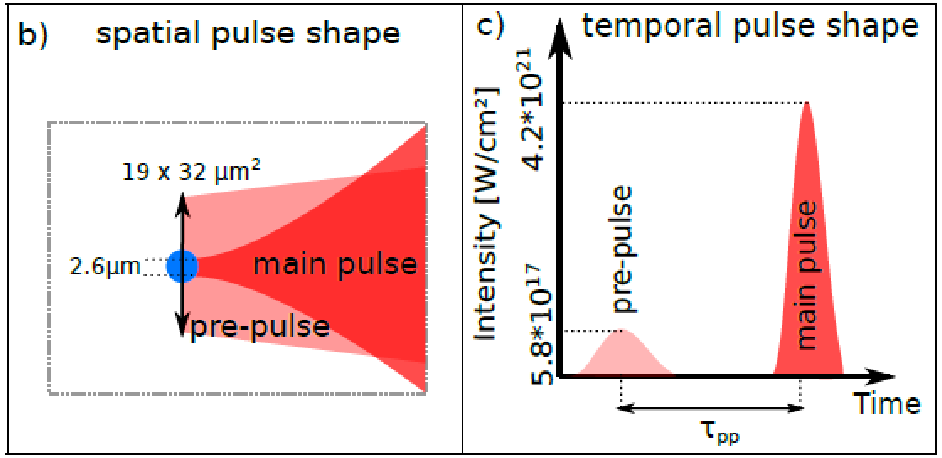
Experimental pre-pulse setup

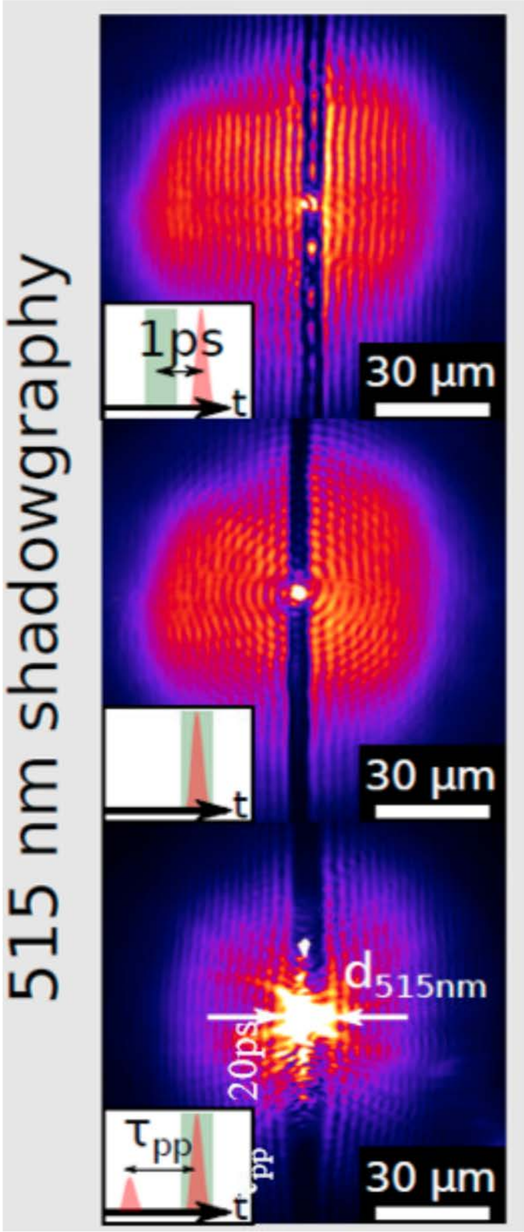
DRACO beam
18J, 30fs, 800nm
PM cleaned contrast
4 x 10^{21}W/cm^2

Timed artificial pre-pulse
5.8 x 10^{17}W/cm^2
(100mJ level, 19x32 µm spot size)
0…170 ps prior to the PW beam



- PM cleaned pulse contrast limits the premature target expansion to the last picosecond before the arrival of the high intensity peak
- Pre-expansion to several times the initial target diameter with pre-pulse

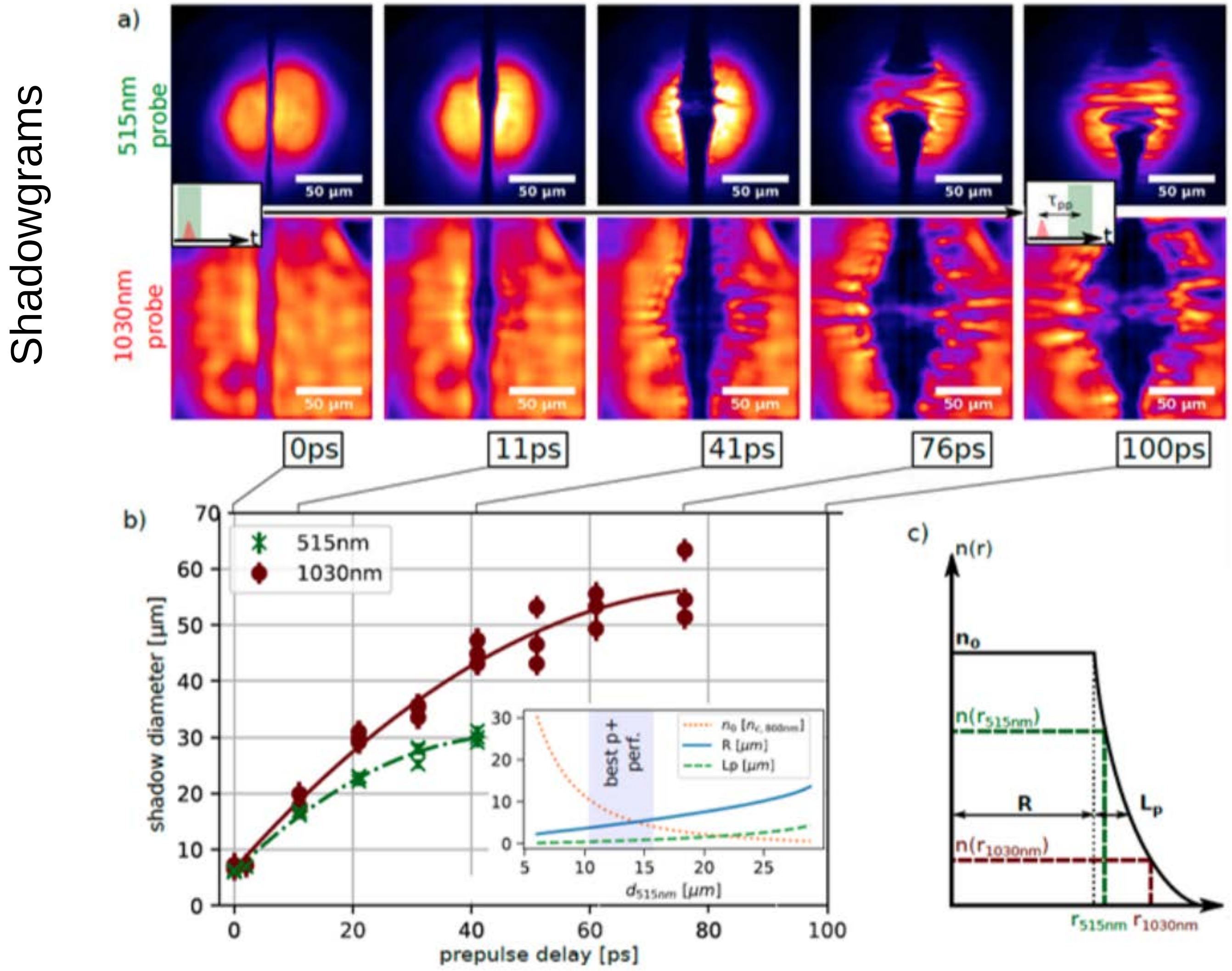




Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Target expansion study

- What change in the density profile was triggered by the pre-pulse?



Assumption:
"box-like" radial density distribution of pre-expanded H2 jet with:

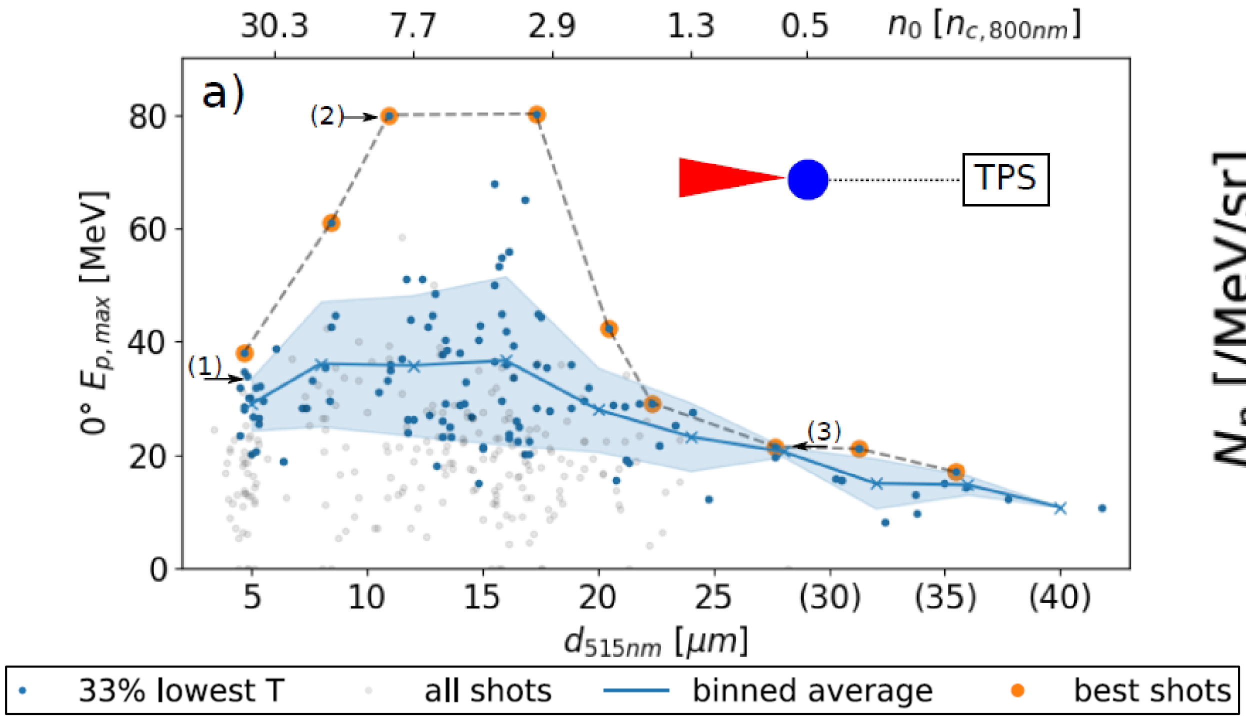
- R: core radius
- n_0: core density
- L_p plasma gradient

- Can by solved for each pre-pulse delay with particle number conservation, the exponential scale length and the threshold density for the transition opaque to transparent
- Correlation between shadow diameter and target model parameters (R, n_0, L_p)

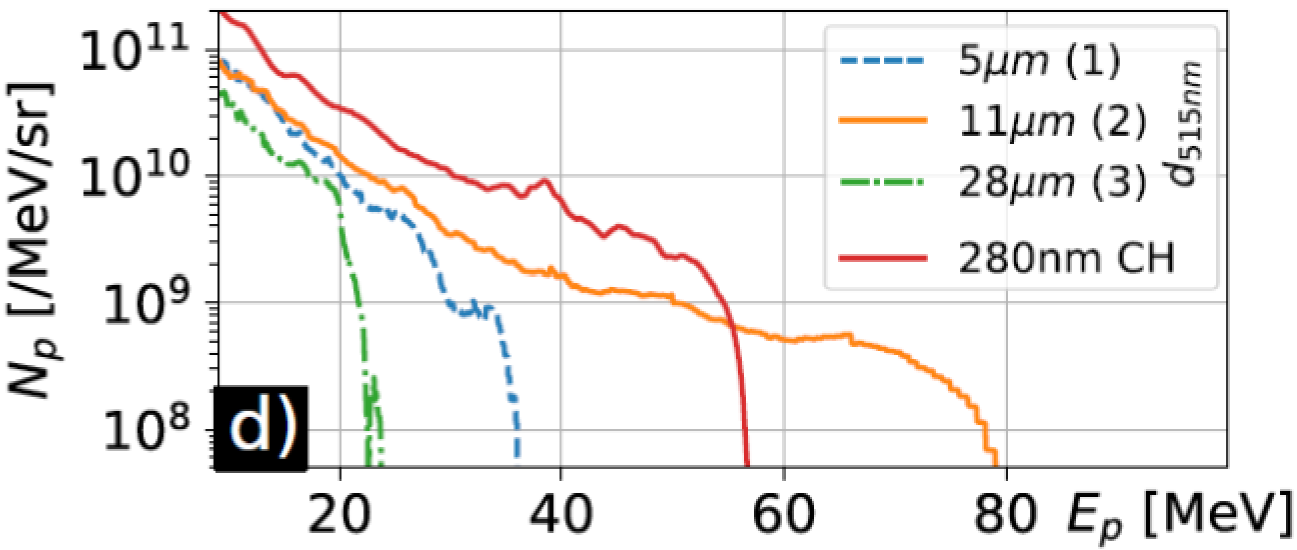
Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Changes in proton acceleration when applying a pre-pulse

Maximum proton energy improvement



Pre-expansion



(1) PM only
(2) „best“ pre-expansion
(3) Larger expansion

- Maximum proton energies of up to 80 MeV
- Increase in maximum proton energies for a certain pre-expansion (smaller and larger expansions result in decreased proton energies)

Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de

3D PIC simulations by Ilja Göthel

Scan of pre-expansion by assuming a constant density in the core, pre-/rear plasma scale length taken from experimental values
→ Radius scan is a scan of the initial density

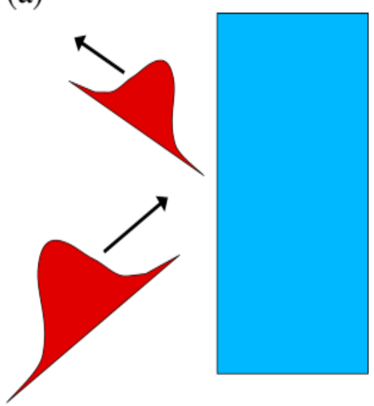
Simulation parameters:
LASERPROFILE=GaussianBeam
RES=24 cells/wavelength
particles per cell=12u
a0=33
FWHM=4.1 micron
TARGETPROFILE=flat with exponential pre- and rear plasma; scale-lengths matching experimental probing

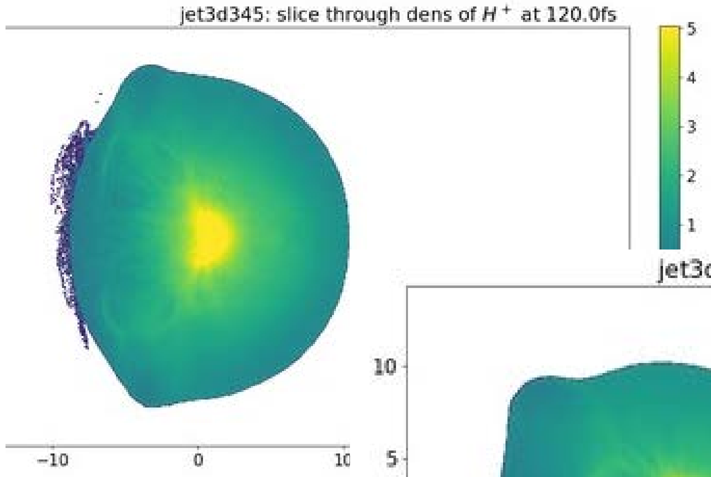
Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de

Real space proton distribution

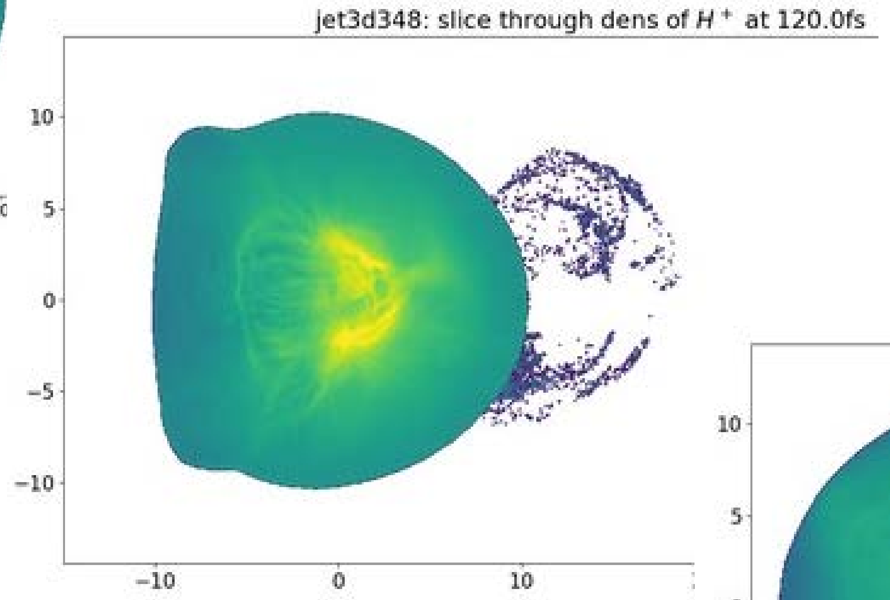
overcritical

(a)

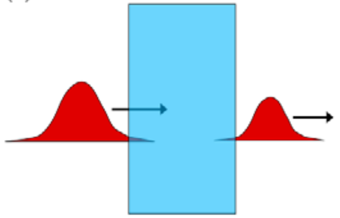




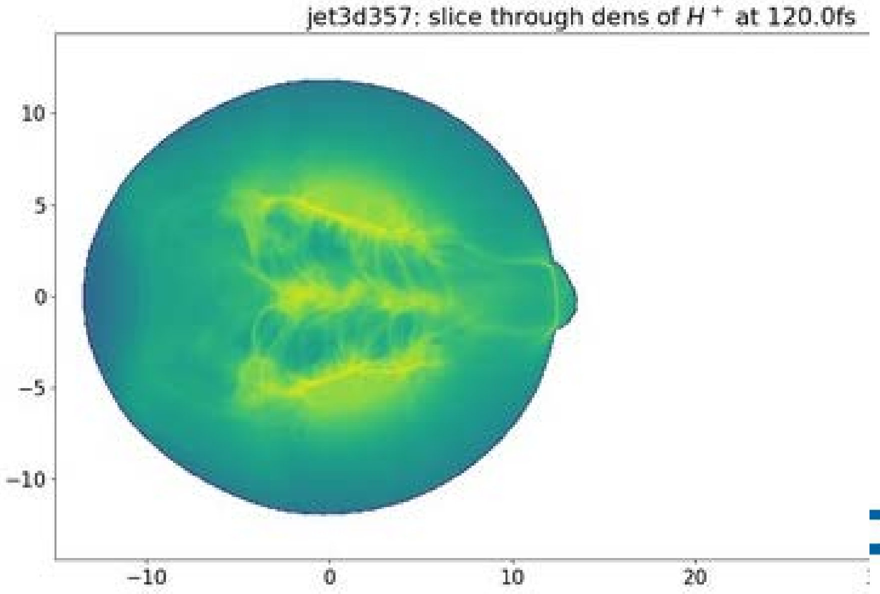
near critical



(b)

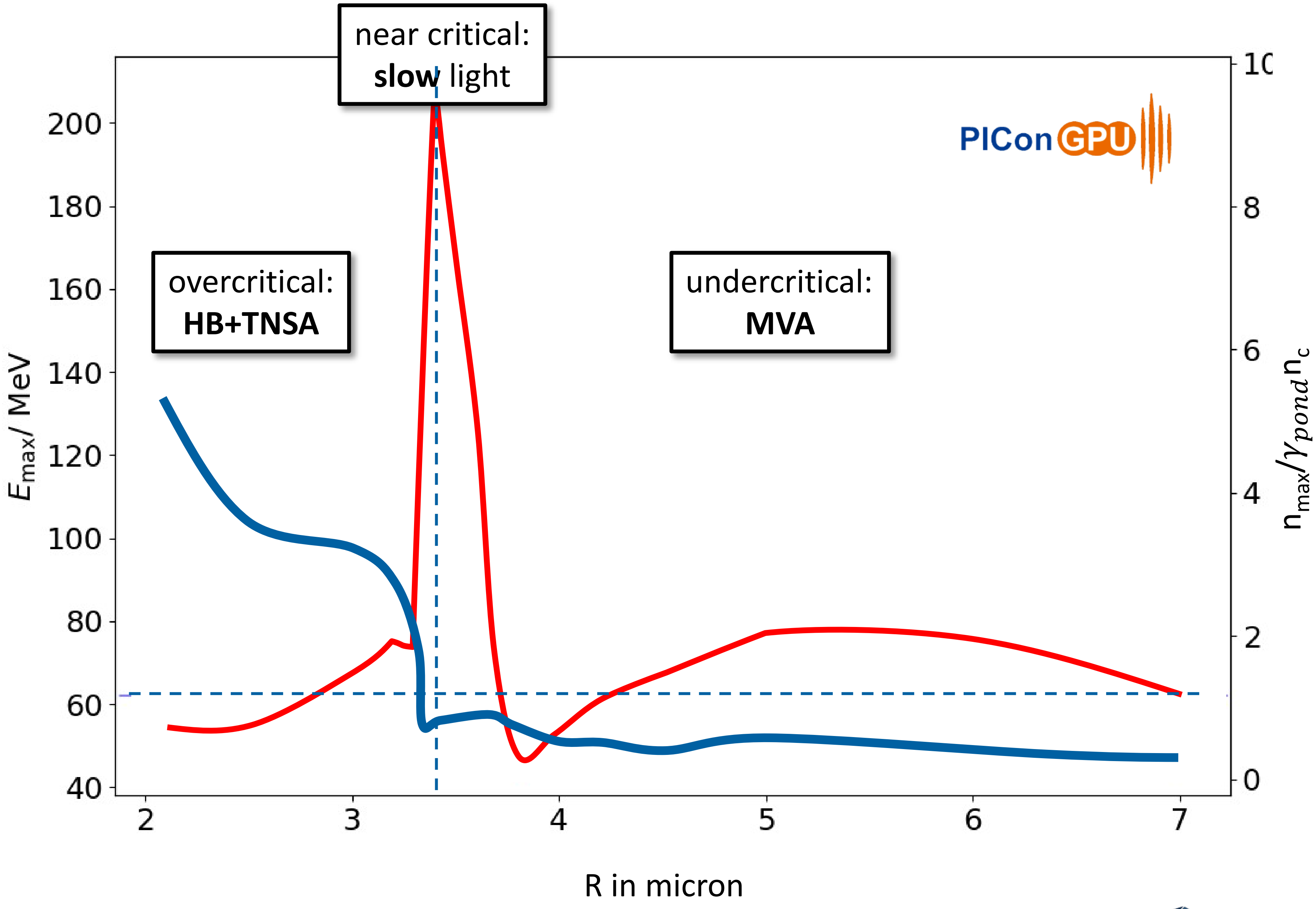


undercritical

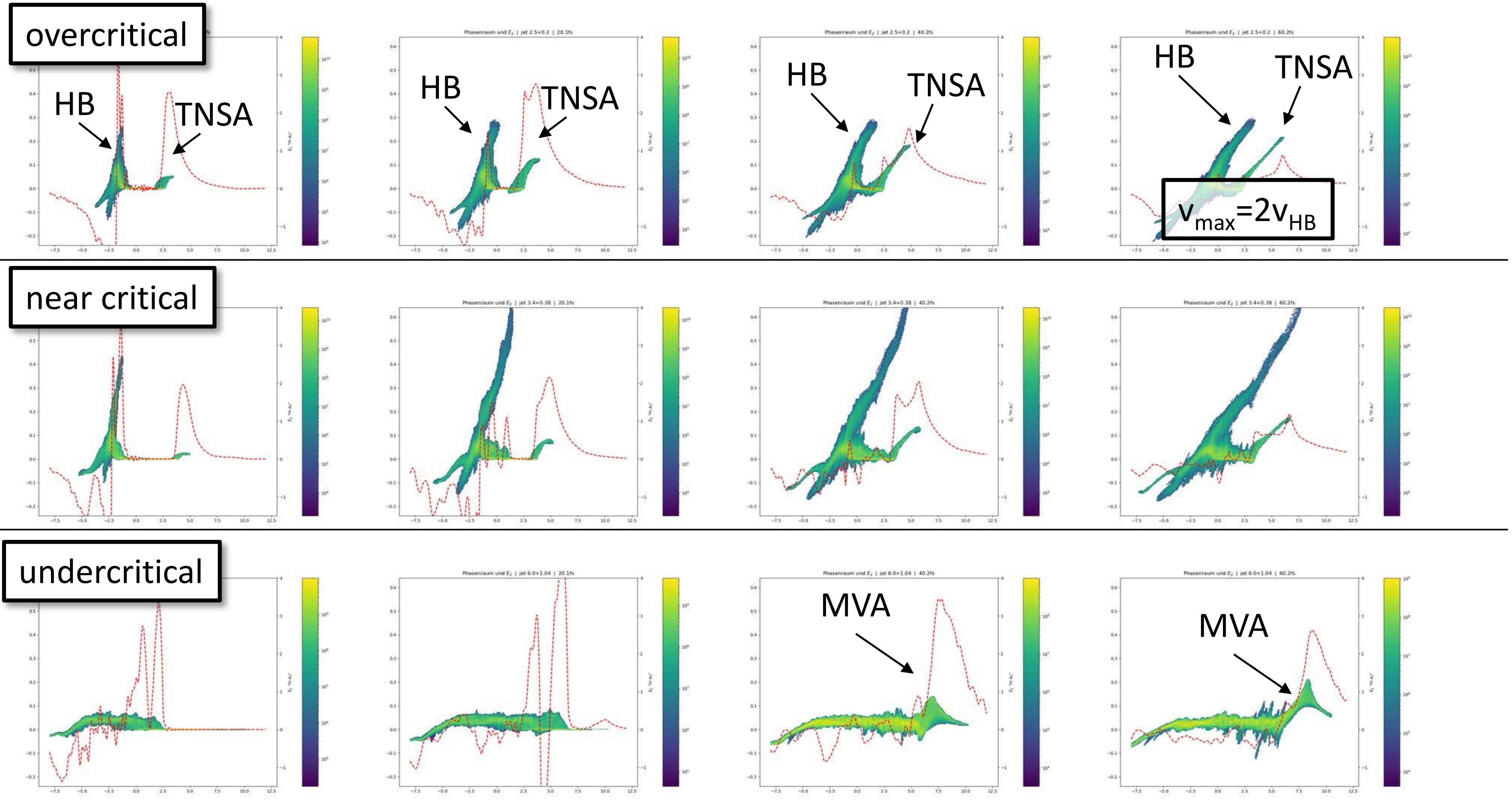


PIConGPU

Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de
Mitglied der Helmholtz-Gemeinschaft


near critical:
slow light
overcritical:
HB+TNSA
undercritical:
MVA
PIConGPU
E_{max}/ MeV
$n_{max}/\gamma_{pond}\, n_c$
R in micron
200
180
160
140
120
100
80
60
40
2
3
4
5
6
7
10
8
6
4
2
0

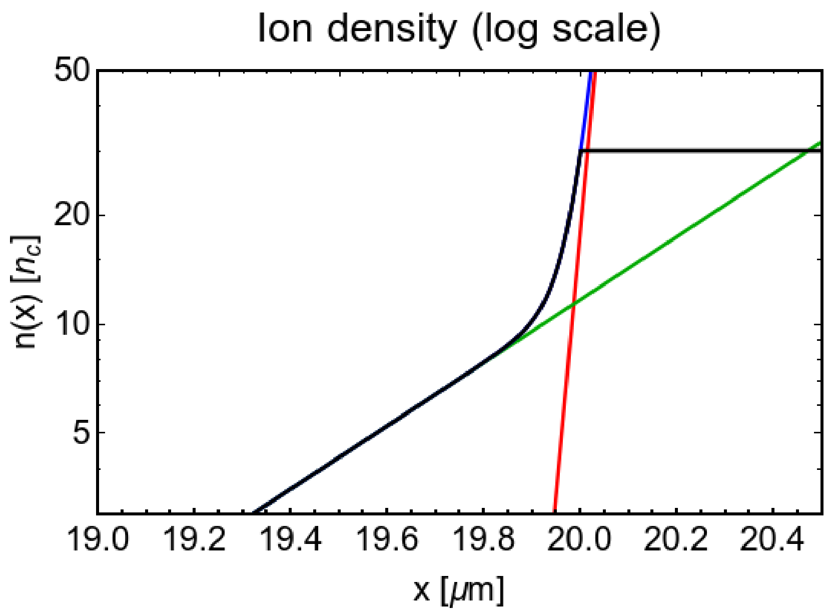
Phase space x - px



Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Surrogate modeling of ion acceleration from 1D simulations

- Generate data using PIC simulations:
  - Hydrogen; cold (0K) target
  - PBC in transverse direction → (quasi-)1D simulations
  - 2 simulation campaigns:
    - Overdense (opaque) targets: $(20 \leq n_0 \leq 50, \mathrm{T} \leq 1\%)$
    - Near critical (relativistic transparent) targets: $(8 \leq n_0 \leq 50, n_0/\gamma < 1)$



- Vary 6 input parameter $(\alpha_0, \tau_{\mathrm{FWHM}}, n_0, d, \ell_0, \ell_1)$
  - → $(0.8 \cdot 1367)^{1/6} = 3.2$ sims. per dim.
  - DOE via quasi Monte-Carlo
  - 1367 simulations in total (code: Smilei*)

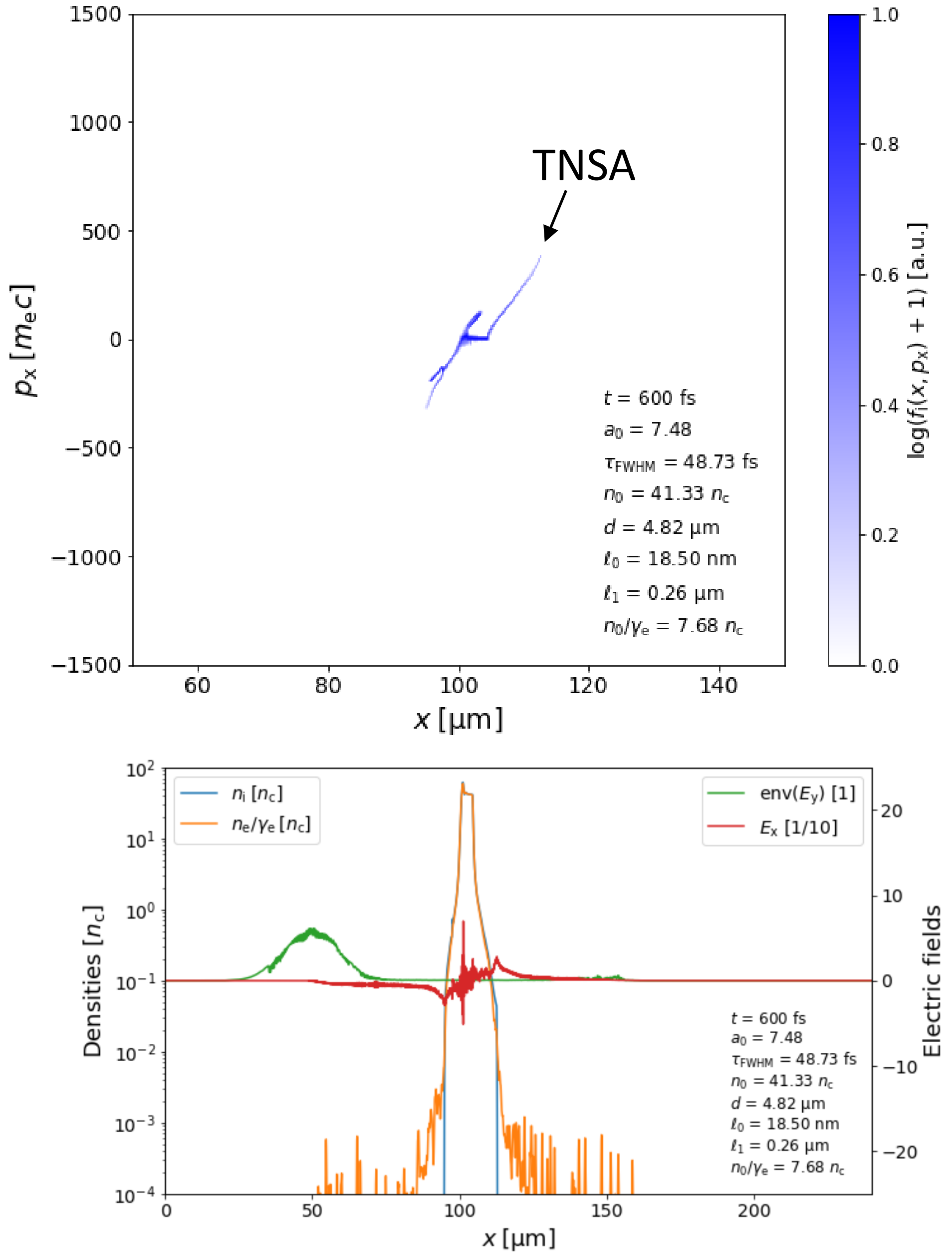
Quantity	Symbol	Unit	Min.	Max.	Scaling
Normalized vector potential	a_0	[1]	7	22	linear
Full width at half maximum	τ_{FWHM}	[fs]	15	50	linear
Number density (bulk)	n_0	$[n_c]$	20	50	linear
Target thickness	d	$[\mu\mathrm{m}]$	1	10	exponential
Short pre-plasma scale length	ℓ_0	[nm]	0.0	40.0	linear
Long pre-plasma scale length	ℓ_1	[nm]	$\ell_0 + 10\,\mathrm{nm}$	500.0	linear

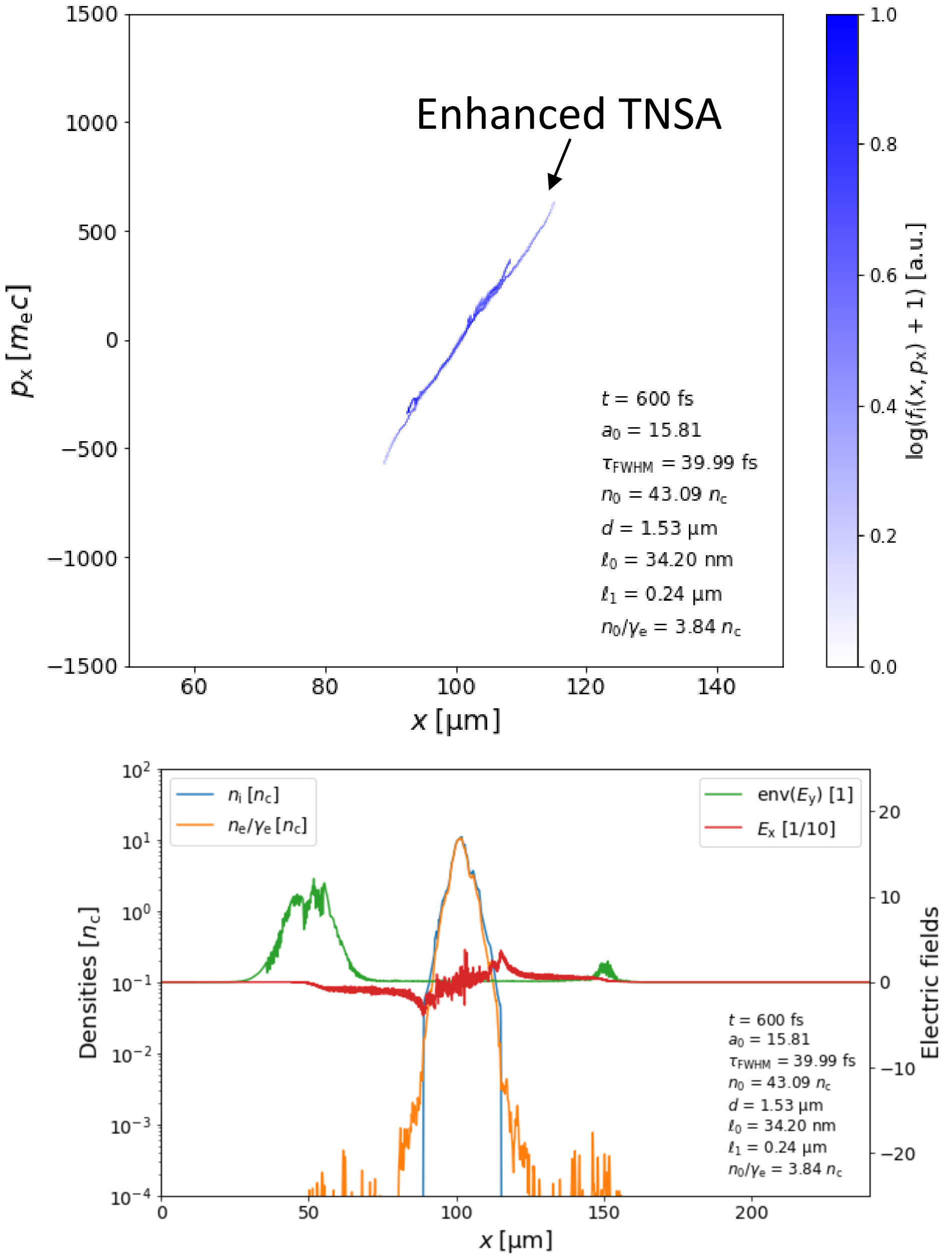
Table 1: Overview of simulation parameter and design space for overdense case.

*Derouillat: Comput. Phys. Commun. 222, 351-373 (2018)

Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

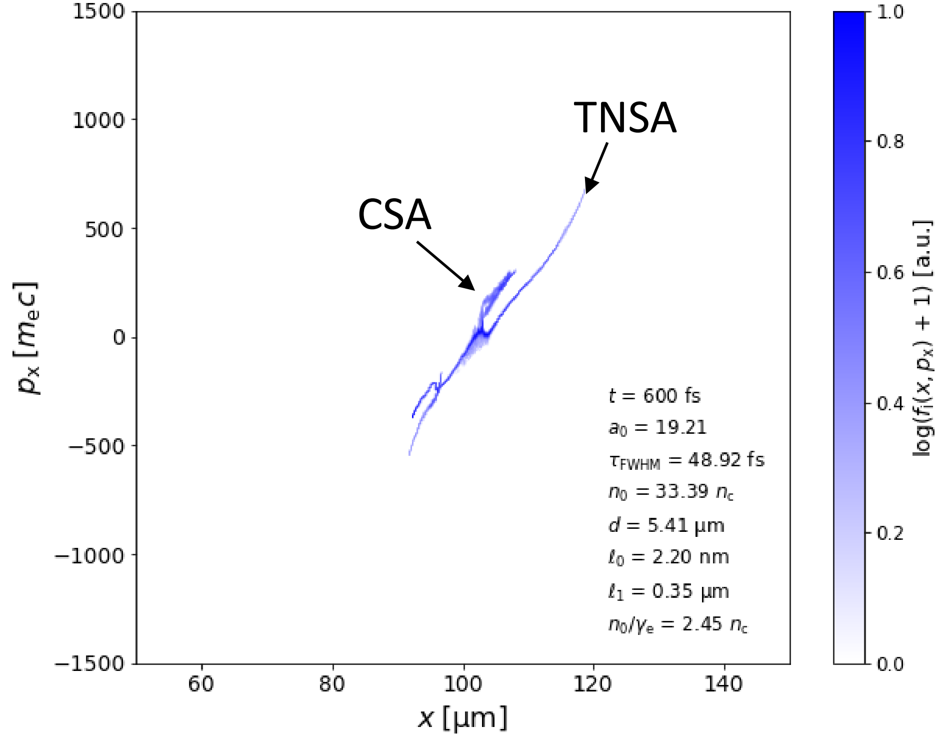
Acceleration mechanisms

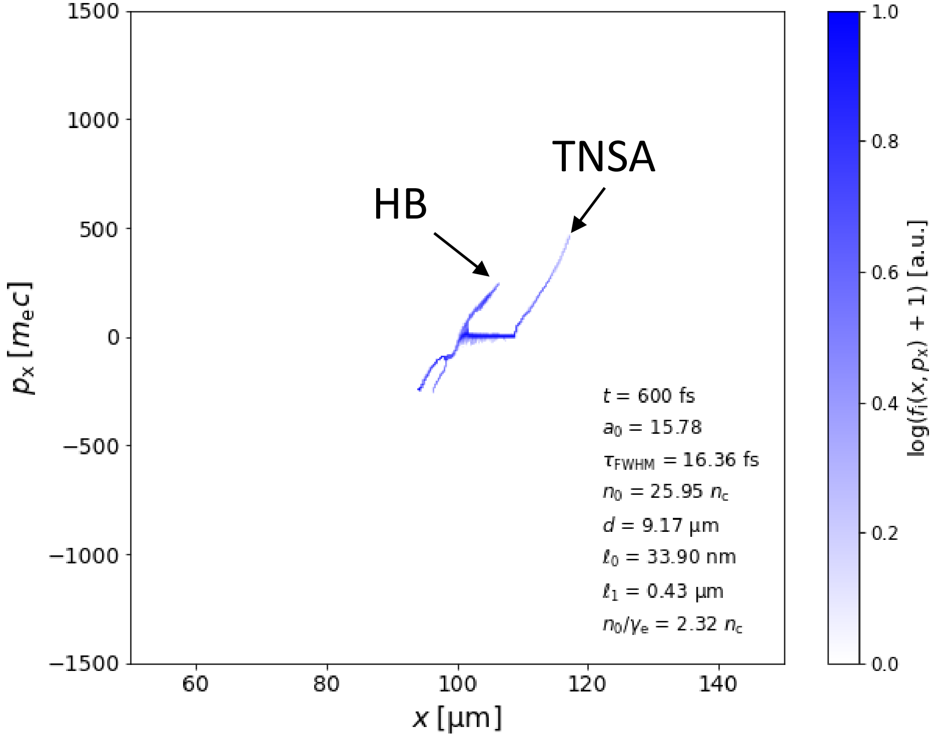


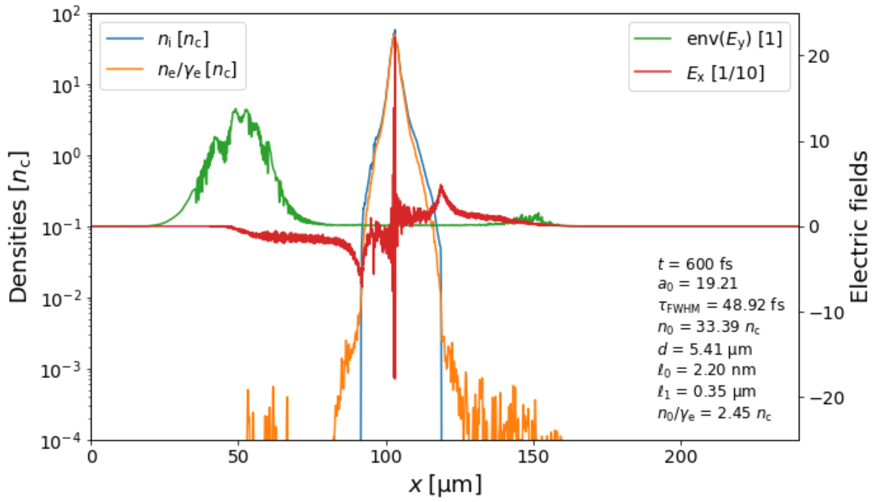


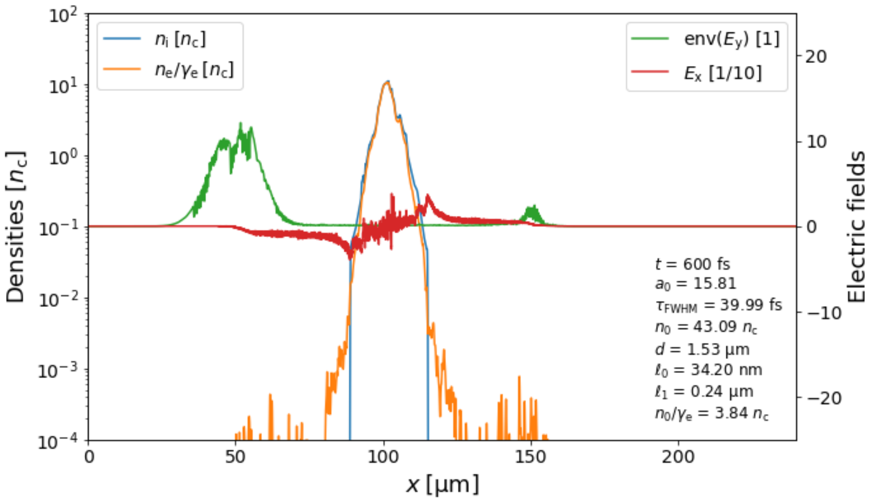
Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Acceleration mechanisms









Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Modeling with neural networks

Learn quantities from simulations with neural networks:

- From input to output quantities
  - $E_i^{max}, T_e, f_i(p_x), f_e(p_x)$
  - Campaign 2: R, T
- Different NNs:
  - Standard Multilayer Perceptrons (MLPs)
  - Invertible Neural Networks (INNs)
  - Autoencoders (AEs)

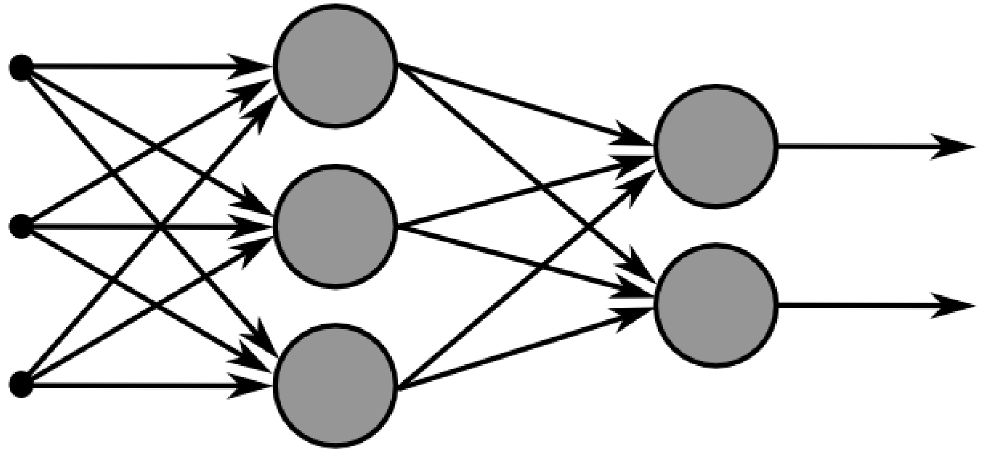
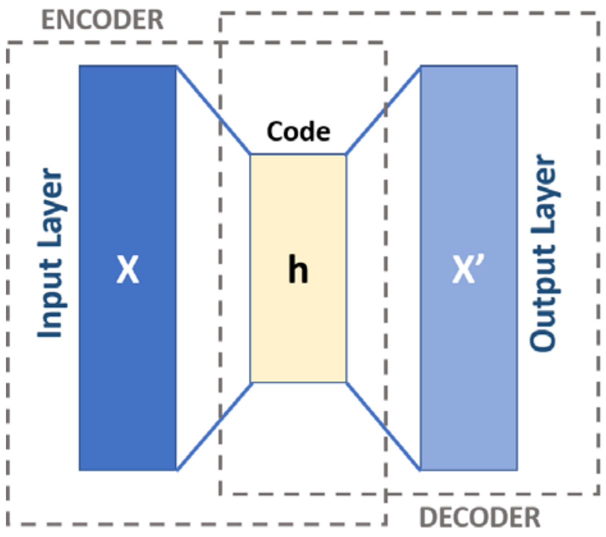


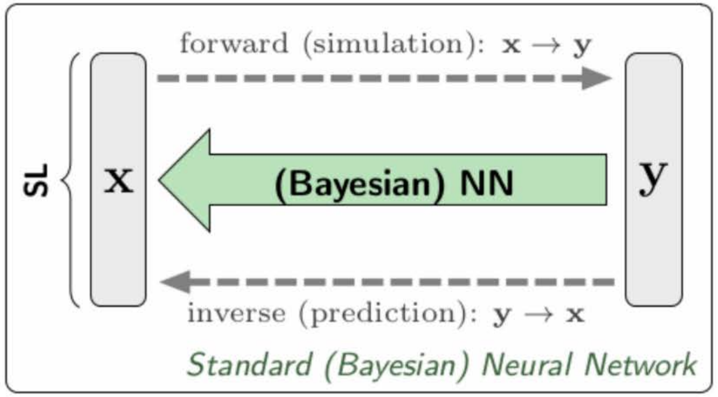
Diagram of a multi-layer feedforward artificial neural network.*



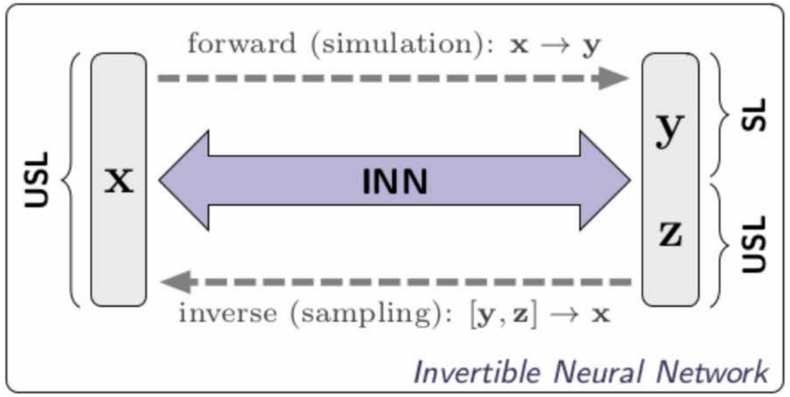
Autoencoder schema.*

*Chrislb, CC BY-SA 3.0 <http://creativecommons.org/licenses/by-sa/3.0/>, via Wikimedia Commons
*Michela Massi, CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0>, via Wikimedia Commons

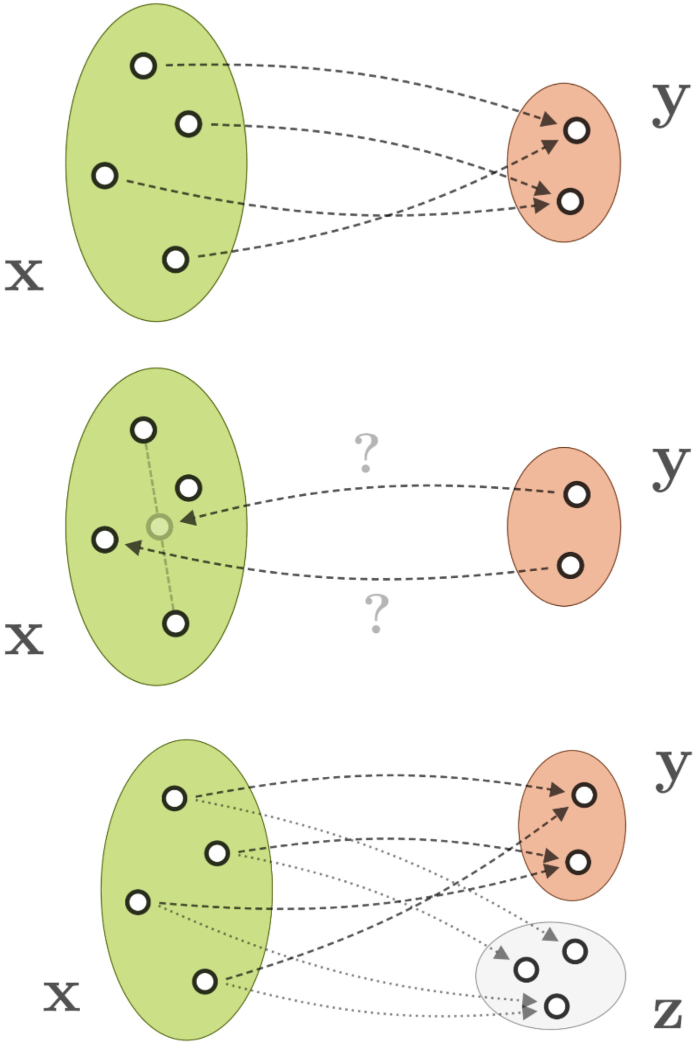
Invertible Neural Networks



Solving inverse problems with BNNs.



Mode of action of an INN.*



Introduction of latent vector resolves ambiguity of inverse problem.*

Invertible Neural Networks:

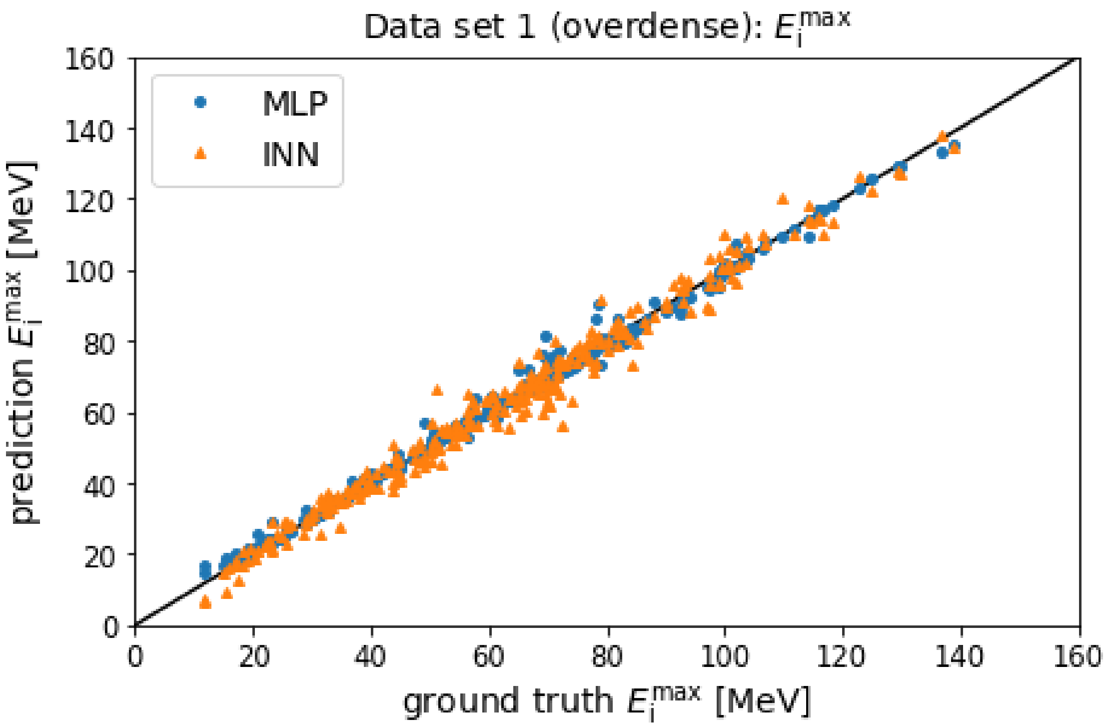
- Mapping from $\mathbf{x}$ to $\mathbf{y}$ is bijective, i.e., inverse $(= \mathrm{INN}^{-1})$ exists
  - Introduction of latent space
- Learn inverse process jointly with forward process
- Obtain conditional posterior probabilities via sampling of $\mathbf{z} \sim \mathrm{N}(0, 1)$.

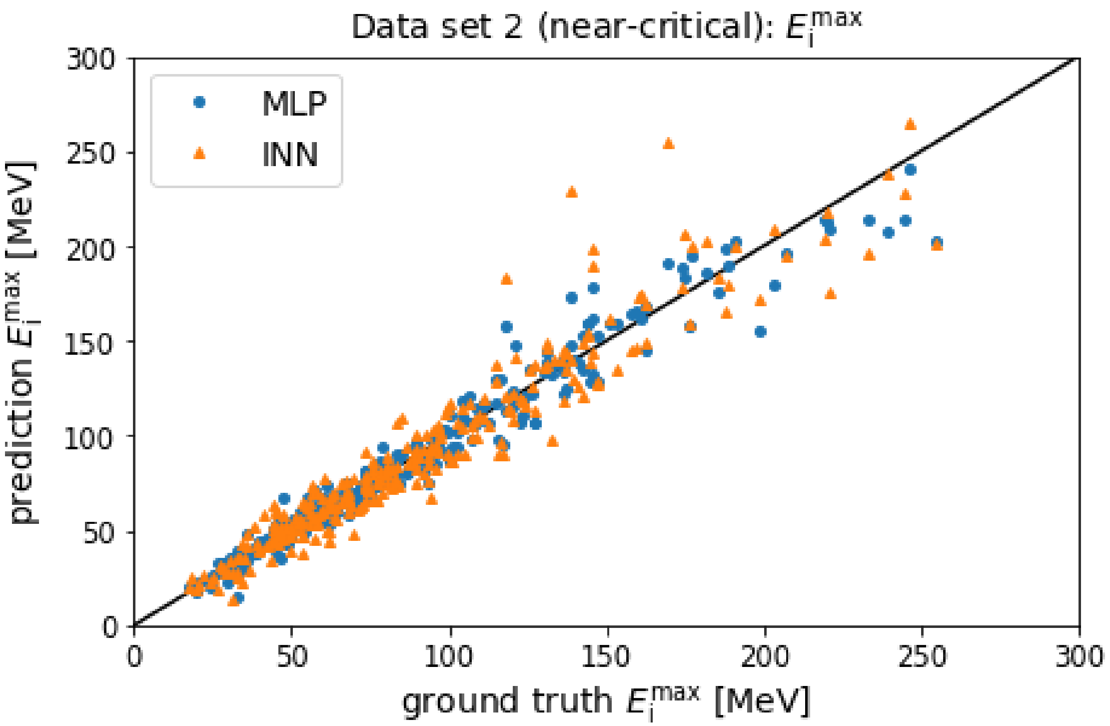
*L. Ardizzone, et. al: Analyzing Inverse Problems with Invertible Neural Networks (ICLR, 2019)

Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de
DRESDEN concept
HZDR
Mitglied der Helmholtz-Gemeinschaft

Example Predictions: Maximum proton energies

Overdense and Near-critical data sets





Density range: $20 \leq n_0 \leq 50$
Opaque, transmission $\mathrm{T} \leq 1\%$

Density range: $8 \leq n_0 \leq 50$
Opaque, transmission $\mathrm{T} \leq 50\%$

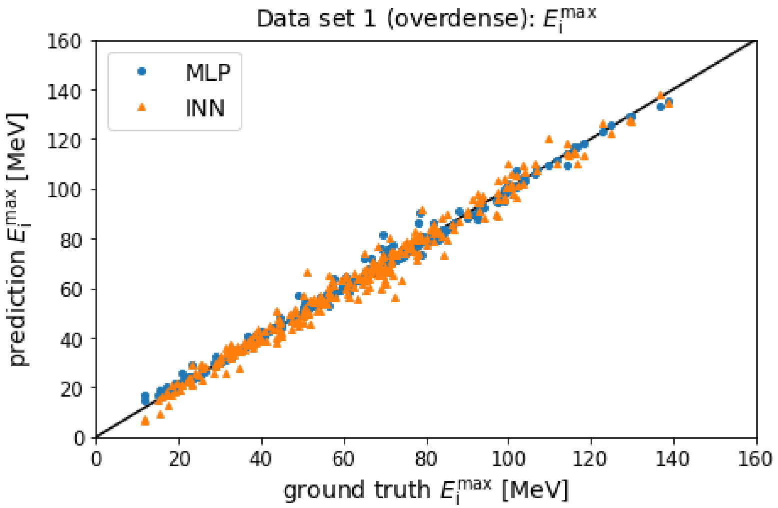
Rest of talk concerned with data set 1

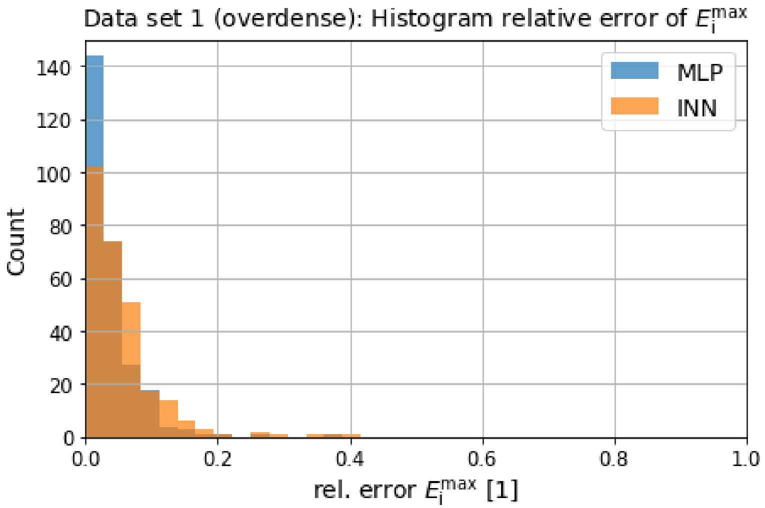
Example Predictions: Scalars

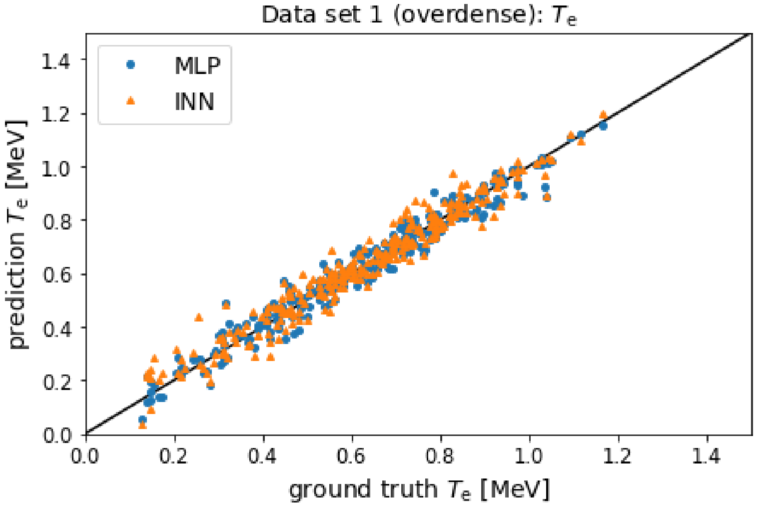
Maximum proton energies, electron temperatures

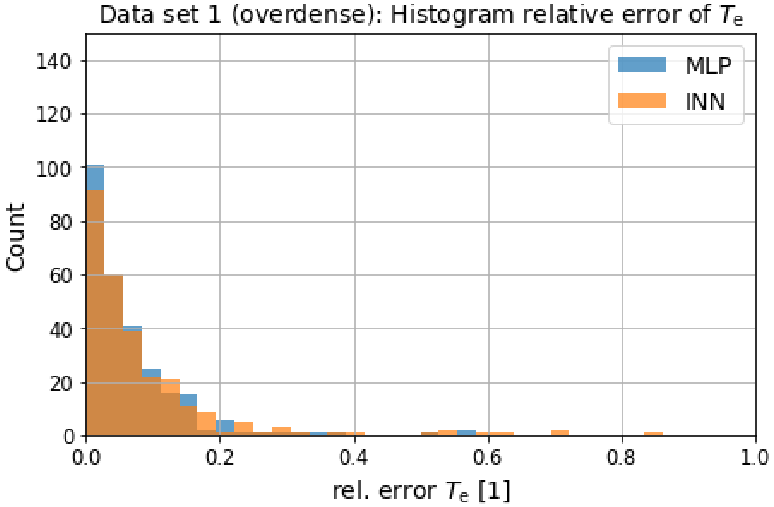
If accuracy of forward pass of INN is ok, then what about speed/performance?

Check against ABC! →









	mean		median	
	MLP	INN	MLP	INN
E_p^{max}	0.040	0.050	0.026	0.035
T_e	0.060	0.079	0.044	0.052

Table 2: Mean and median relative errors of the MLP and the INN models for E_p^{max} and T_e.

INN: Performance test

Maximum proton energies, electron temperatures

What about accuracy of backward pass? →

	ABC				INN-ABC			
	min	max	median	mean	min	max	median	mean
Comp. time [s]	0.14	122.3	0.77	4.60	0.008	3.06	0.05	0.19
Accept. rate [1]	$1.4 \cdot 10^{-6}$	$1.5 \cdot 10^{-3}$	$2 \cdot 10^{-4}$	$4 \cdot 10^{-4}$	$3 \cdot 10^{-4}$	$1.2 \cdot 10^{-1}$	$1.6 \cdot 10^{-2}$	$2.4 \cdot 10^{-2}$

Table 3: Benchmark comparison of ABC and INN-ABC for scalar observables $\mathbf{y} = (E_\mathrm{p}^\mathrm{max}, T_\mathrm{e})^\mathsf{T}$. Total run times are for ABC and INN-ABC are 1260.18 s and 50.89 s, respectively.

Factor 25 speedup!

ABC algorithm

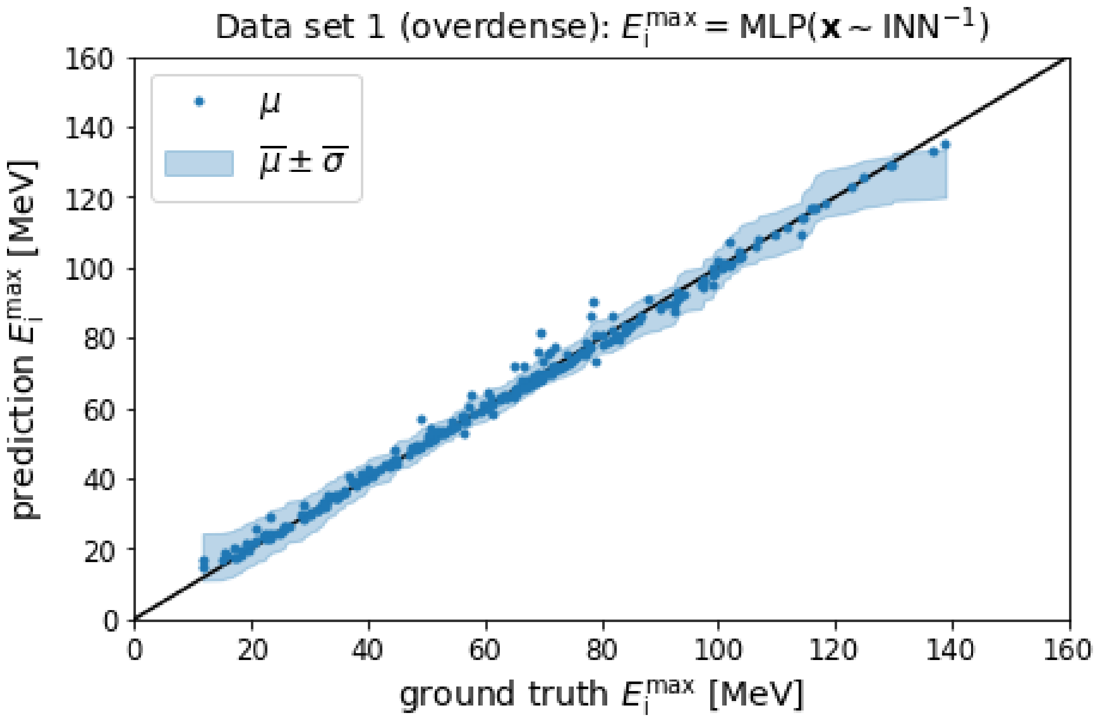
0. Define distance function d, acceptance threshold ε and minimum number of accepted samples $N_\mathrm{acc}^\mathrm{min}$.
1. Generate N_rnd random input vectors $\{\mathbf{x}_k\}$ from parameter space.
2. Compute vectors in output space $\{\mathbf{y}_k\} = \mathrm{MLP}(\{\mathbf{x}_k\})$
3. If $d(\mathbf{y}^{(i)}, \mathbf{y}_k) \leq \varepsilon$, then accept $\mathbf{y}_k$ and update count of total number of accepted vectors $N_\mathrm{acc}^\mathrm{tot}$.
4. If $N_\mathrm{acc}^\mathrm{tot} < N_\mathrm{acc}^\mathrm{min}$, goto 1. Else, return set of accepted vectors.

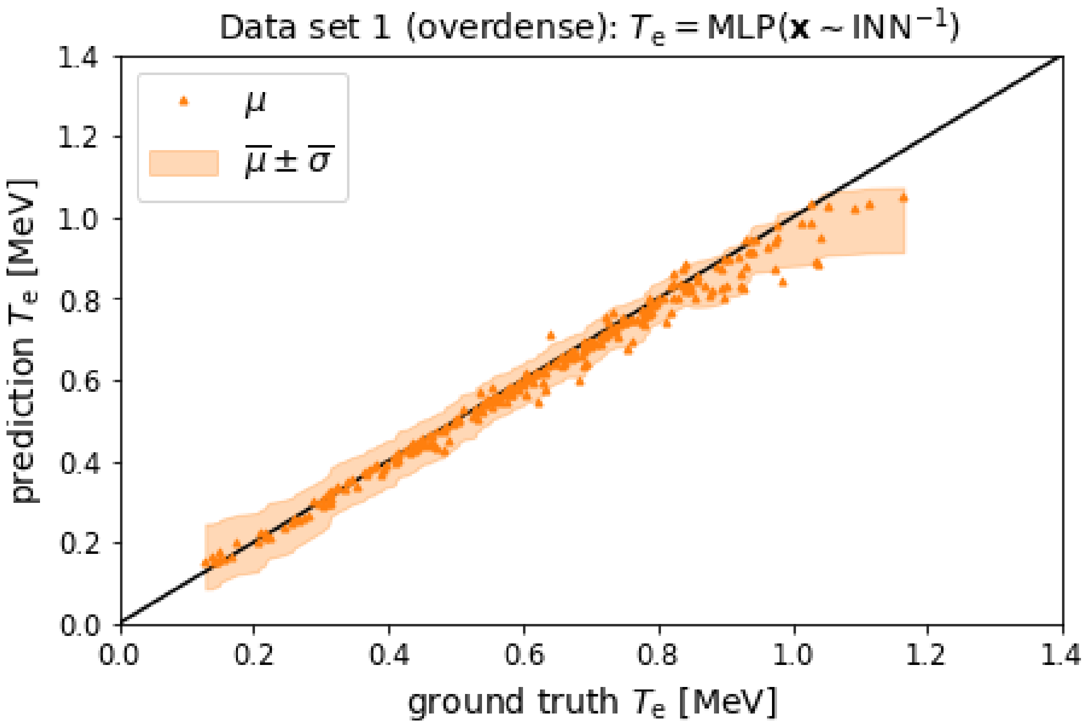
INN performance test routine

0. Define distance function d, acceptance threshold ε and minimum number of accepted samples $N_\mathrm{acc}^\mathrm{min}$.
1. Generate N_rnd latent vectors $\{\mathbf{z}_k\}$ with $\mathbf{z} \sim \mathcal{N}(\mathbf{z}; \mathbf{0}, \mathbf{I}_{\mathrm{dim}_\mathbf{z}})$.
2. Compute approximation of conditional posterior distribution $\{\mathbf{x}_k^{(i)}\} = \mathrm{INN}^{-1}(\mathbf{y}^{(i)}, \{\mathbf{z}_k\})$.
3. Compute vectors in output space $\{\mathbf{y}_k^{(i)}\} = \mathrm{MLP}(\{\mathbf{x}_k^{(i)}\})$
4. If $d(\mathbf{y}^{(i)}, \mathbf{y}_k^{(i)}) \leq \varepsilon$, then accept $\mathbf{y}_k^{(i)}$ and update count of total number of accepted vectors $N_\mathrm{acc}^\mathrm{tot}$.
5. If $N_\mathrm{acc}^\mathrm{tot} < N_\mathrm{acc}^\mathrm{min}$, goto 1. Else, return set of accepted vectors.

Accuracy Inverse Predictions:

Maximum proton energies, electron temperatures

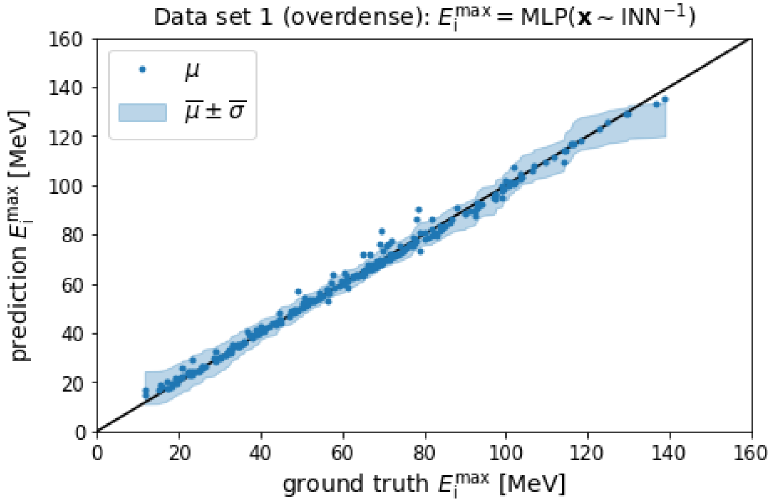


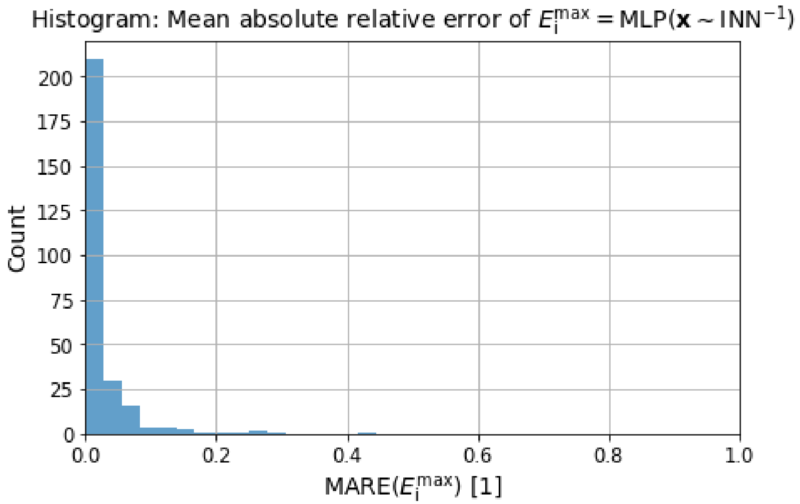


Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

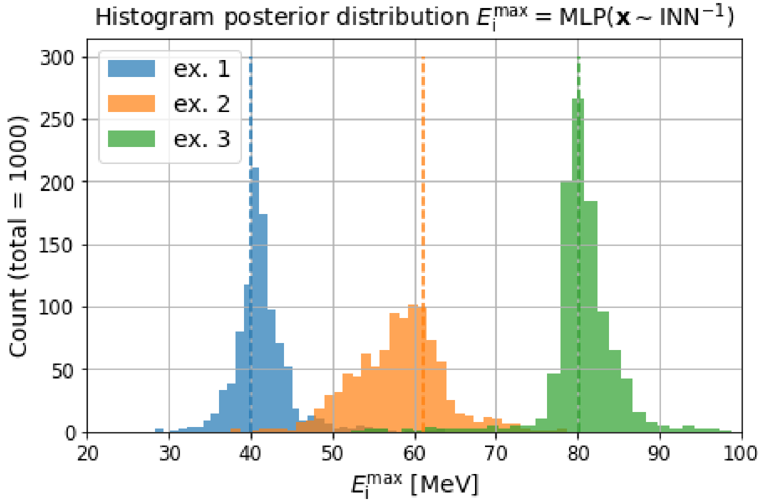
Accuracy Inverse Predictions:

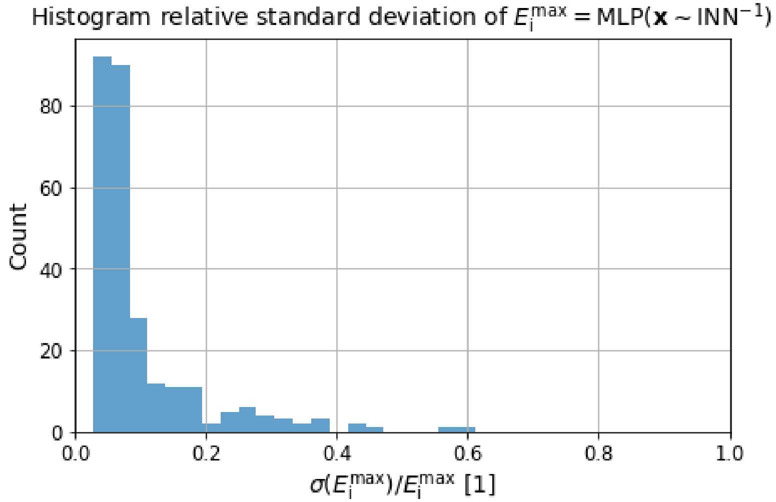
Maximum proton energies





$$\mathbf{MARE}^{(i)} = \frac{1}{N_{\mathrm{rnd}}} \sum_{k=1}^{N_{\mathrm{rnd}}} \left| \frac{\mathbf{y}_k^{(i)} - \mathbf{y}^{(i)}}{\mathbf{y}^{(i)}} \right|$$





$$\frac{\boldsymbol{\sigma}^{(i)}}{\mathbf{y}^{(i)}} = \frac{1}{\mathbf{y}^{(i)}} \sqrt{\frac{1}{N_{\mathrm{rnd}} - 1} \sum_{k=1}^{N_{\mathrm{rnd}}} \left(\mathbf{y}_k^{(i)} - \bar{\mathbf{y}}^{(i)} \right)^2}$$

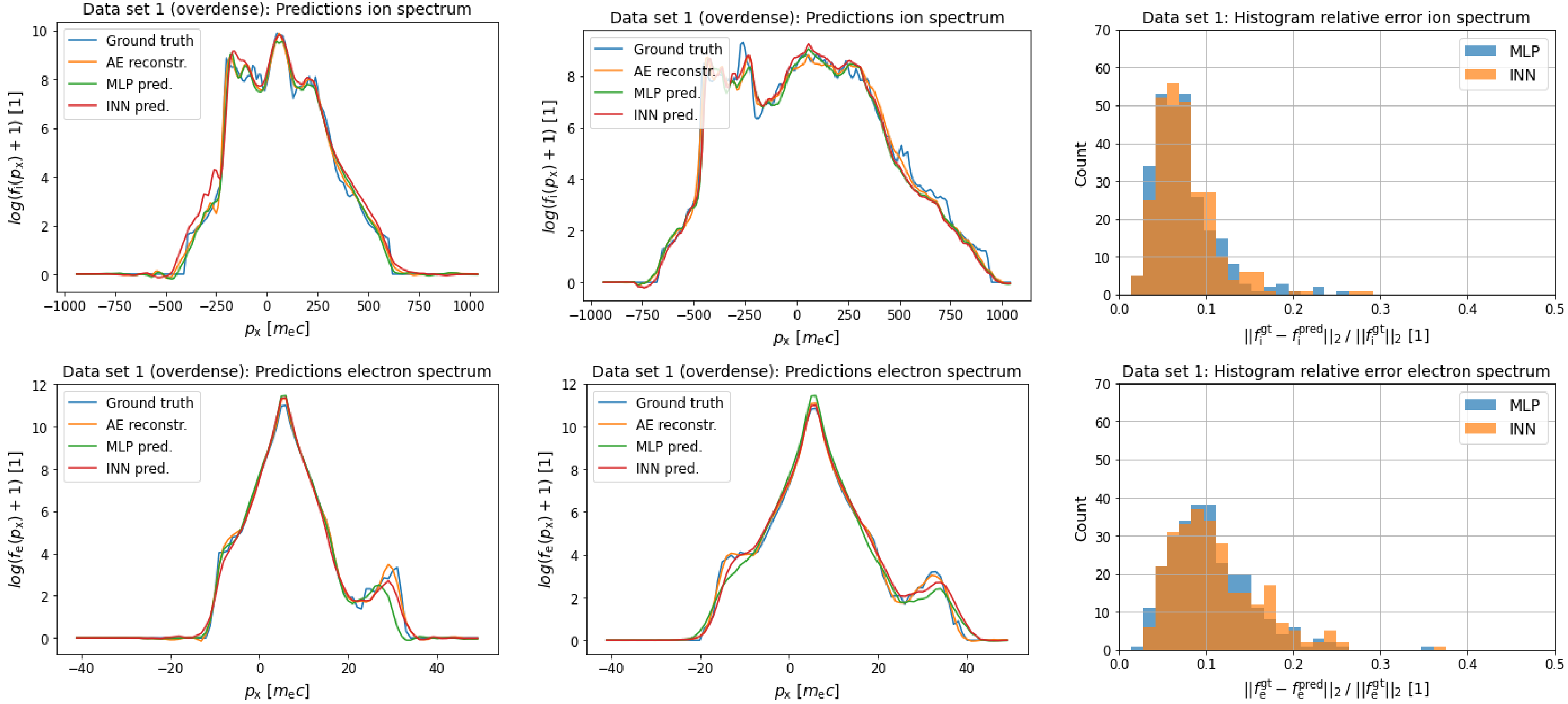
DRESDEN concept

HZDR

Example Pred. : Proton, electron (momentum) spectra (1D Functional Data)

Protons

Electrons



Example A

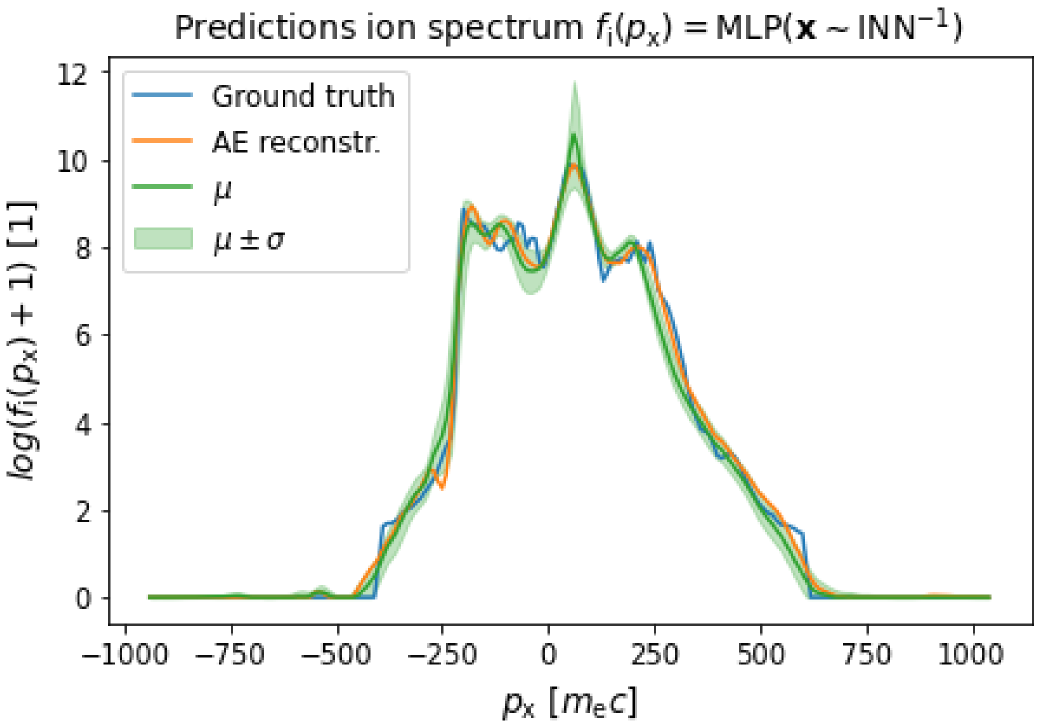
Example B

Relative error: spectra

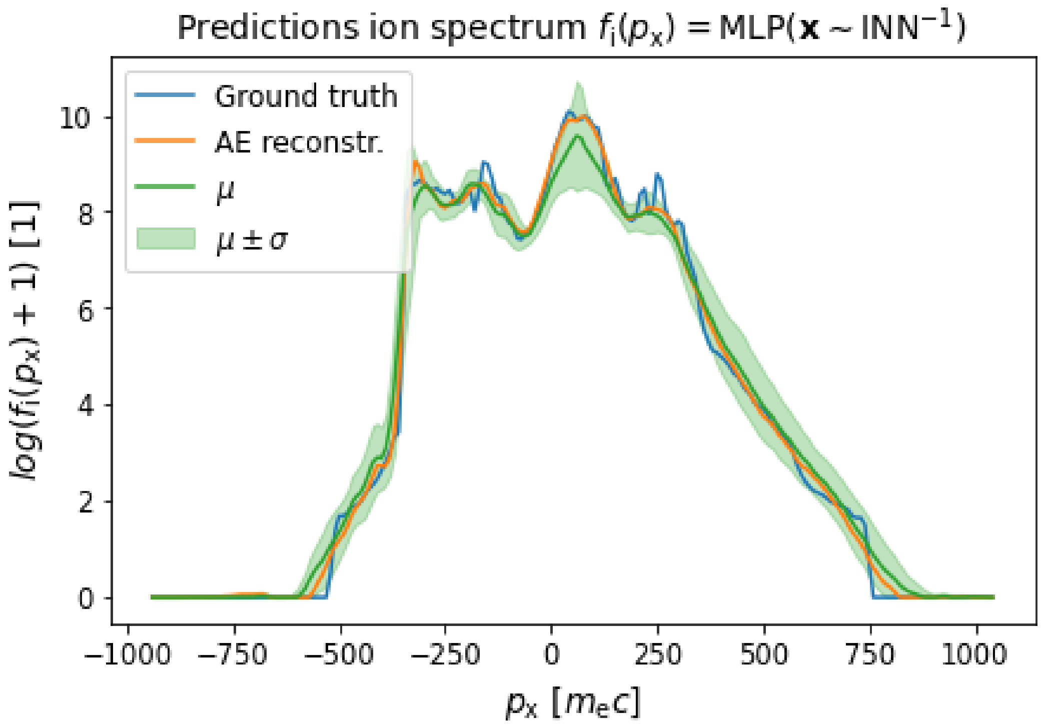
Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Accuracy Inverse Predictions: Proton momentum spectra

Protons



Example A

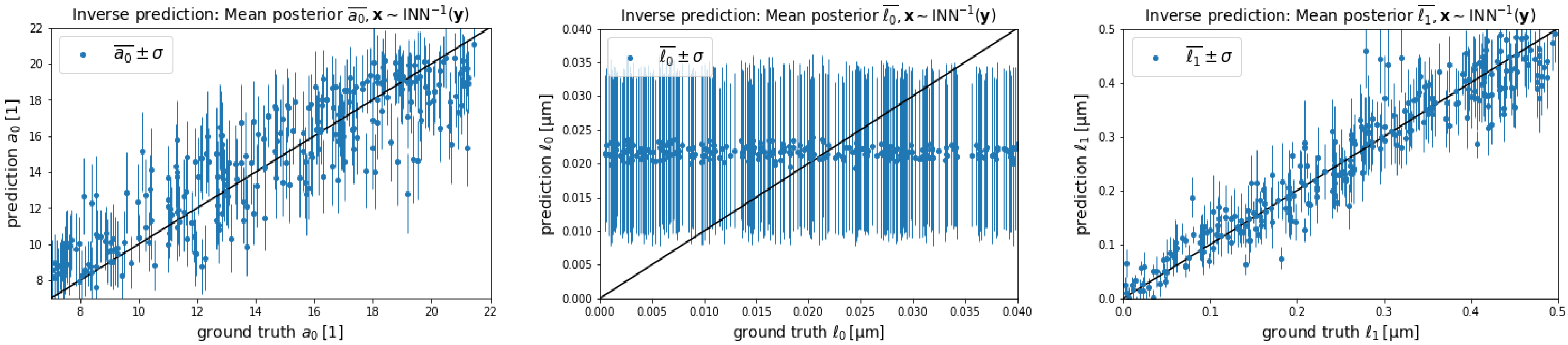


Example B

- INNs can learn both the forward as well as inverse process for different types (scalars, functional, images, …) of data
- Given a model for the forward process, the accuracy of the INN can be interpreted in y-space.
  - Furthermore, we can compute uncertainty by mapping the sampled posterior onto y-space.

Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Inverse Predictions: (Un)Ambiguity of parameters

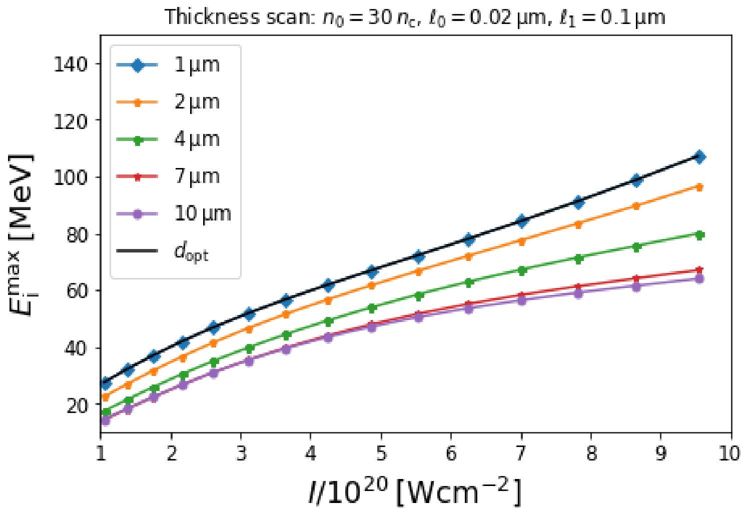


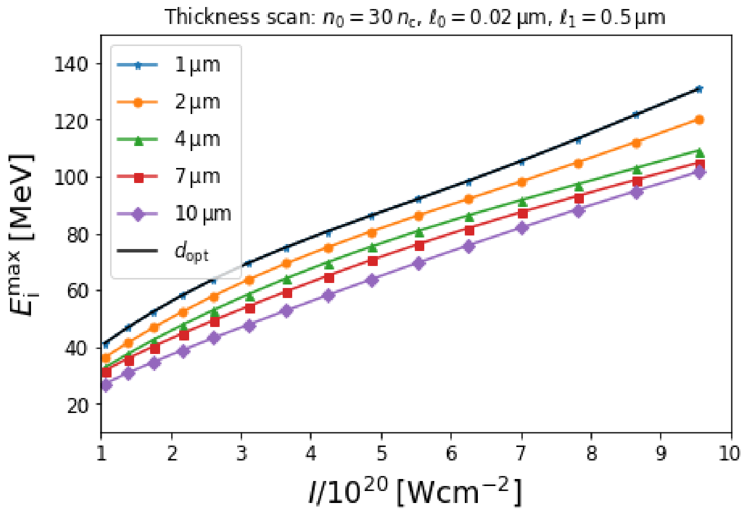
- Test if we can learn/recover from experimental observables f_i, f_e to (simulation) parameter $\mathbf{x}$
  - Of course, this is in general a distribution $p(\mathbf{x}|\mathbf{y})$
  - Nevertheless, integrating over all but one variable might already contain useful information

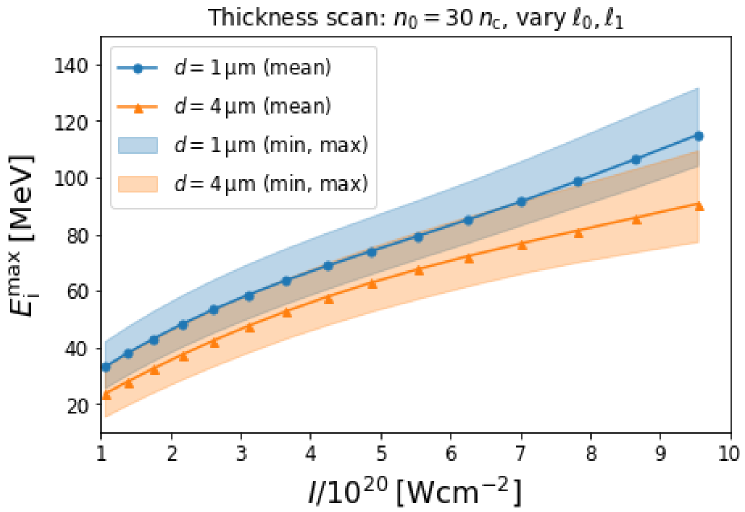
→ Means of posterior might (somewhat) recover parameter value.
→ ℓ_1 (long pre-plasma scale length) leaves a characteristic signature in spectra
→ In principle, models (trained on simulations) can also be applied to experimental data
  – Need corresponding & appropriate data set.

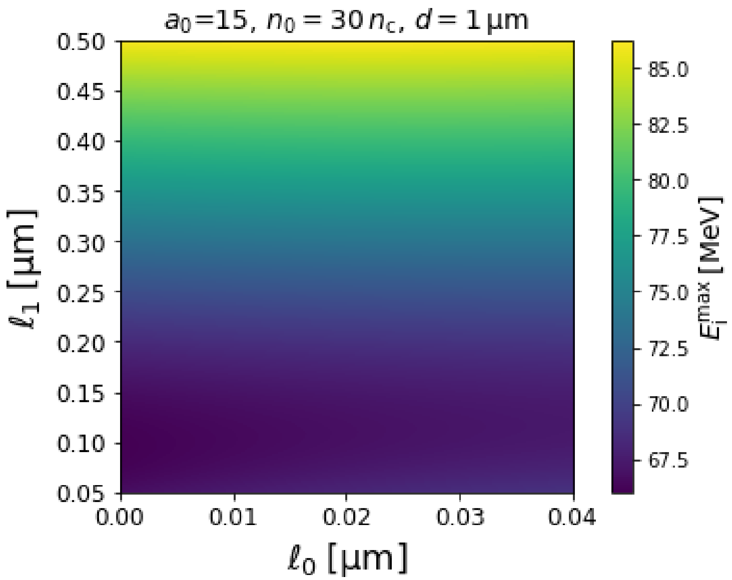
Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Model exploration









Surrogate models can help to increase rate of scientific discoveries.
E.g:

- Thickness d should be chosen smaller than 1 micron (at least for TNSA and according to (quasi)1d model)
  - Larger optimal thicknesses are expected for higher laser intensities (HB) and lower densities (MVA)
- Strong dependency on long pre-plasma scale length
  - which is in contrast to very overdense targets, say $n_0 > 100$
  - However, it seems $\nabla_{\ell_0} \approx \nabla_{\ell_1}$

DRESDEN concept

Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de

Conclusion and Outlook

Conclusion

- Trained machine learning models from performed simulation campaign (1367 simulations, 6 input parameter) with predictive accuracy.
- Both MLPs and INNs are useful and play an important role for modeling.
  - MLPs show strong predictive capabilities in the forward direction.
  - INNs show strong predictive capabilities in forward AND backward direction. Furthermore, in the backward direction INNs are much more efficient than standard method (they strongly outperform against ABC).
- Functional data, having several hundred dimensions, can be effectively learned with autoencoders (more robust AEs currently developed in Nico Hoffmann's group).

Outlook

- From observables / experimentally available outputs to "hidden" variables (e.g. phase space data).
- Cluster data to identify ranges of physical regimes.
- Check validity of theory in corresponding cluster.
- Do 2D/3D simulations, learn physics relevant to 2D/3D, e.g. instabilities.
- Use experimental data on models trained on simulations in the backward direction.

DRESDEN concept HZDR

Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de