Acceptance Rates of Invertible Neural Networks on Electron Spectra from Near-Critical Laser-Plasmas: A Comparison

03.02.2023

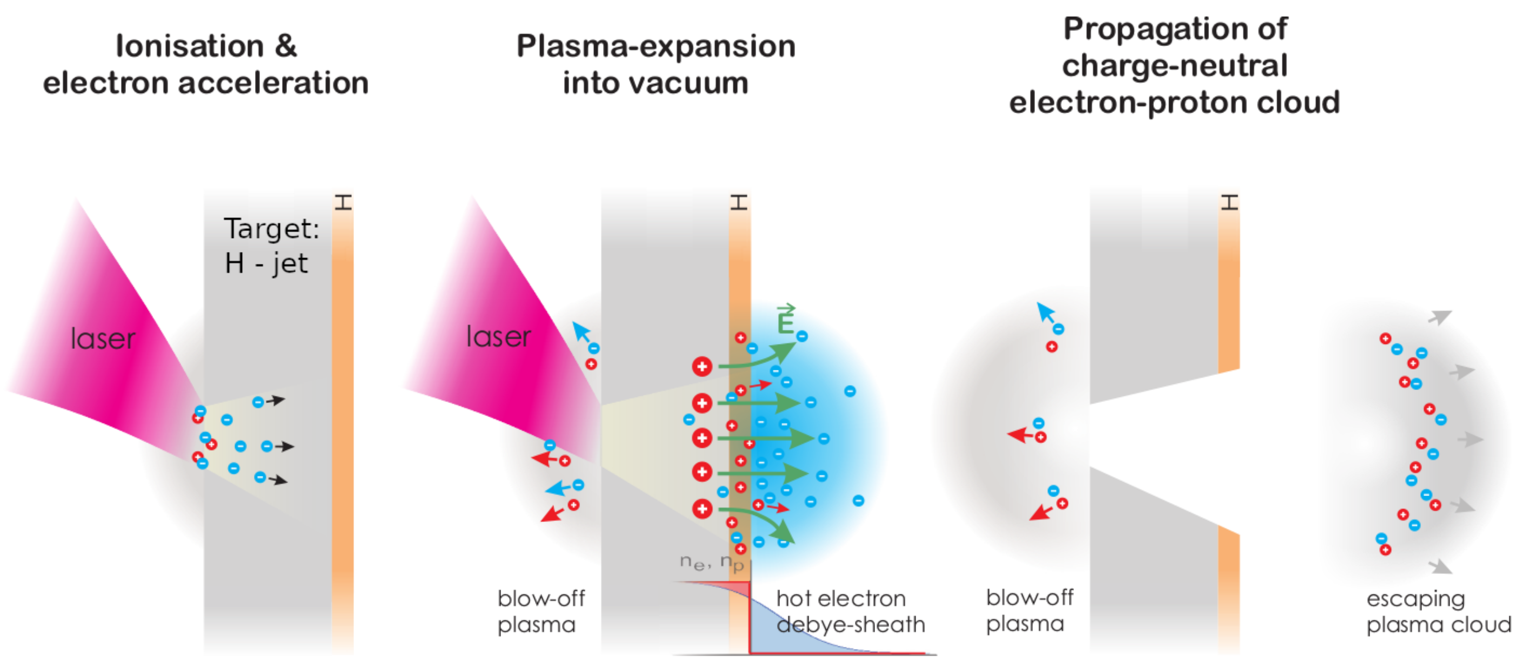
Thomas Miethlinger[1,2], Nico Hoffmann[1] and Thomas Kluge[1]

[1]Helmholtz-Zentrum Dresden-Rossendorf, 01328 Dresden, Germany
[2]Technische Universität Dresden, 01069 Dresden, Germany

HZDR
HELMHOLTZ ZENTRUM DRESDEN ROSSENDORF

Ion Acceleration from Laser-Plasmas

- Ultra-high intensity (10^{19}-10^{21} W/cm^2), ultra-short (20-50 fs) laser pulses
  - Energy: 1-30 J on target
  - E-fields: up to several TV/m
- Extreme conditions of matter: hot, dense plasma
  - Hot: $k_B T_e$ = 10−30 keV
  - Dense plasma: n_e = 5 – 300 n_c



- Goal: <u>reliably</u> produce protons with energies around 200 – 250 MeV
  - Current world record: ~100 MeV[1] (target: plastic foils)
  - Stable production at around 60 MeV[2]
- Prospective applications:
  - Radiation oncology
  - Material science
  - Plasma probing

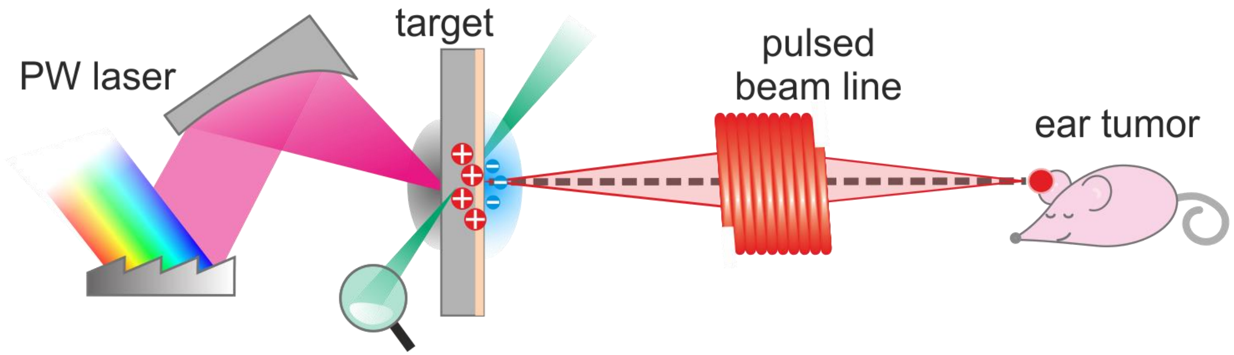


Illustration: Tumor irradiation in mice with laser-driven protons.[2]

[1]Ziegler: Sci Rep 11, 7338 (2021)
[2]Kroll: Nat. Phys. 18, 316–322 (2022)

DRESDEN concept HZDR

Physical Motivation and Problem Statement

- **Major goal** in laser-ion acceleration: increase $E_\mathrm{i}^\mathrm{max}$
  - Depends on **all** experimental/simulation parameters, $E_\mathrm{i}^\mathrm{max}(\mathbf{x})$, where $\mathbf{x}$:
    - Laser strength a_0
    - Laser pulse duration (approx. Gaussian-shaped pulse)
    - Target/plasma density n_0
    - …

→ Optimize parameter $\mathbf{x}$ in order to maximize $E_\mathrm{i}^\mathrm{max}$

  - Physical considerations:
    - Max. ion energies are proportional[1] with **mean kinetic energy of laser-driven electrons**, T_e:

$$E_\mathrm{i}^\mathrm{max} \propto T_\mathrm{e}$$

    - Definition of T_e: function of electron spectrum, $f_\mathrm{e}(E)$
    - Mostly depends[2] on the laser strength a_0

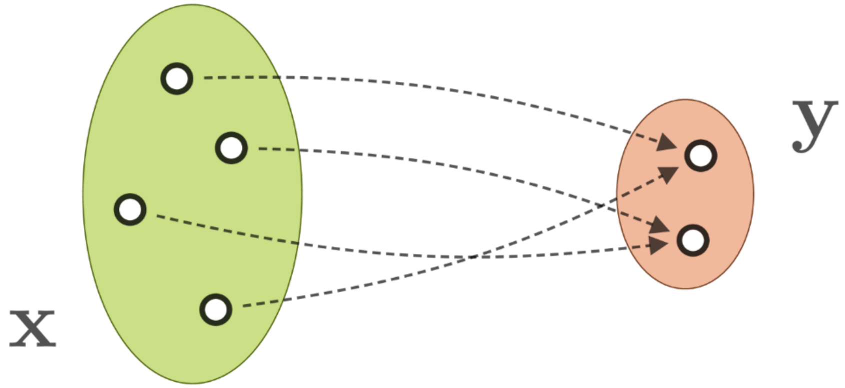
[1]Mora: Phys. Rev. Lett. 90.18, 185002 (2003)

[2]Kluge T., et al: Phys. Rev. Lett. 107.20, 205003 (2011)

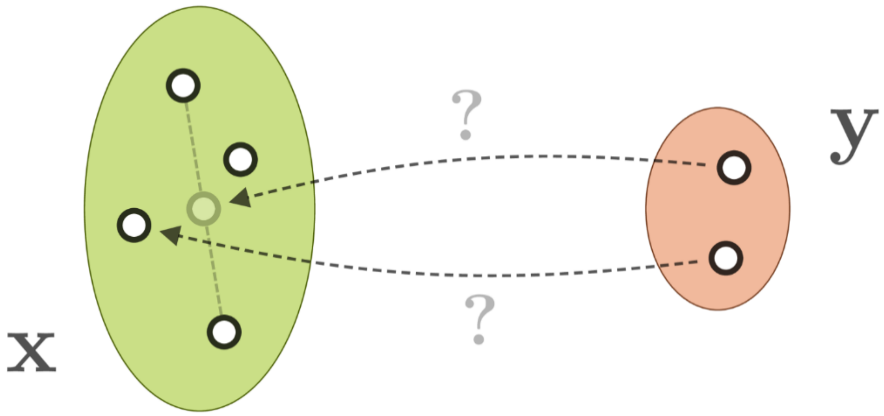
Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de

Ambiguous Inverse Problems in Natural Science

- **Forward problem:**
  - We have some system with internal state, → vector $\mathbf{x}$
  - Through some mechanism (experiment, simulation, …) $f(\mathbf{x})$, we measure observables $\mathbf{y}$
  - Each measurement is called an observation $\mathbf{y}^*$



- **Inverse Problem:**
  - Find possible $\mathbf{x}$ that would lead to an observation $\mathbf{y}^*$
  - However, information provided by $\mathbf{y}^*$ typically incomplete:
    - Information loss in forward process
    - Reduction of effective dimensionality → **multiple** $\mathbf{x}$ map onto same $\mathbf{y}^*$
    - Due to **ambiguity**, correct answer is not obvious



- **Example: X-ray diffraction**
  - State $\mathbf{x}$: Crystal structure / atomic positions
  - Measurement $\mathbf{y}^*$: Diffraction pattern

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

Posterior Distribution

- **Conditional posterior:**
  - Due to ambiguity, the full **conditional posterior**

$$p(\mathbf{x}\,|\,\mathbf{y}^*),$$

has to be determined.

  - Typically, we try to solve an inverse problem by finding a **representative sample**

$$\{\mathbf{x}\},$$

that approximates $p(\mathbf{x}|\mathbf{y}^*)$, and fulfills with $f(\{\mathbf{x}\}) = \mathbf{y}^*$.

  - *General problem* concerning many of our observables:
    - Proton spectrum $f_\mathrm{i}(E)$
    - **Electron spectrum** $f_\mathrm{e}(E)$ (relevant for **this** work)
    - …

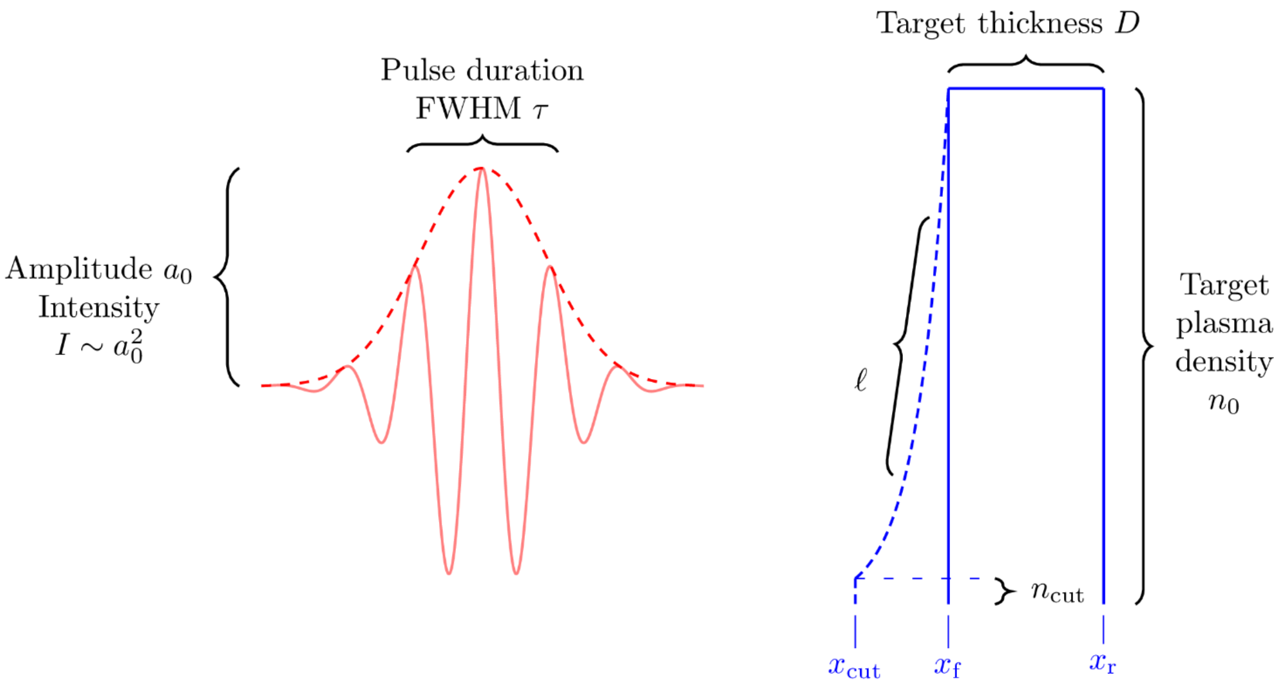
Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de

Particle-In-Cell (PIC) Simulation Setup

- Generate data using PIC simulations:
  - Hydrogen target; initially cold (0K), fully ionized
  - Simulation box: 240 µm x 0.25 µm
  - PBC in transverse direction → (quasi-)1D simulations
  - Large parameter space, including both overdense and near-critical regime:

$$0.2 \leq n_{\mathrm{e}}/\gamma \leq 50$$

  - Numerics:
    - λ_{L} / Δx = 64
    - CFL = 0.99
    - 50 PPC



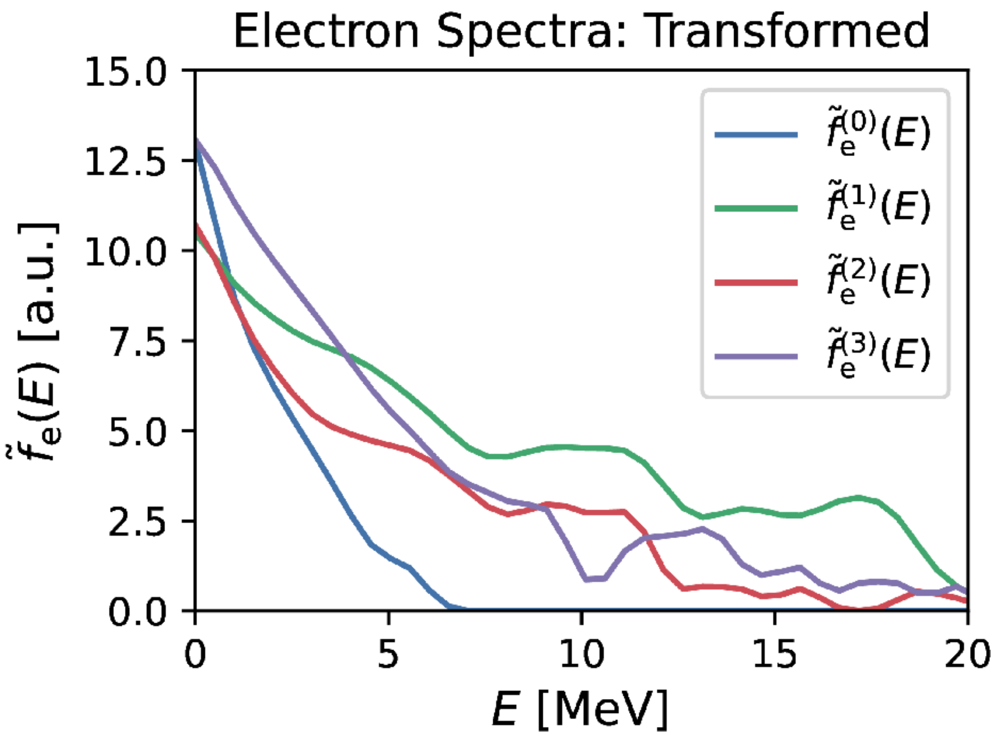
- Assume pre-plasma scale length ℓ at front → 5 simulation input parameter
  - Parameter chosen via quasi Monte-Carlo (QMC)
  - Budget: 5000 simulations in total (code: Smilei*)
  - → $(0.8 \cdot 5000)^{1/5}$ = 5.25 sims per dimension

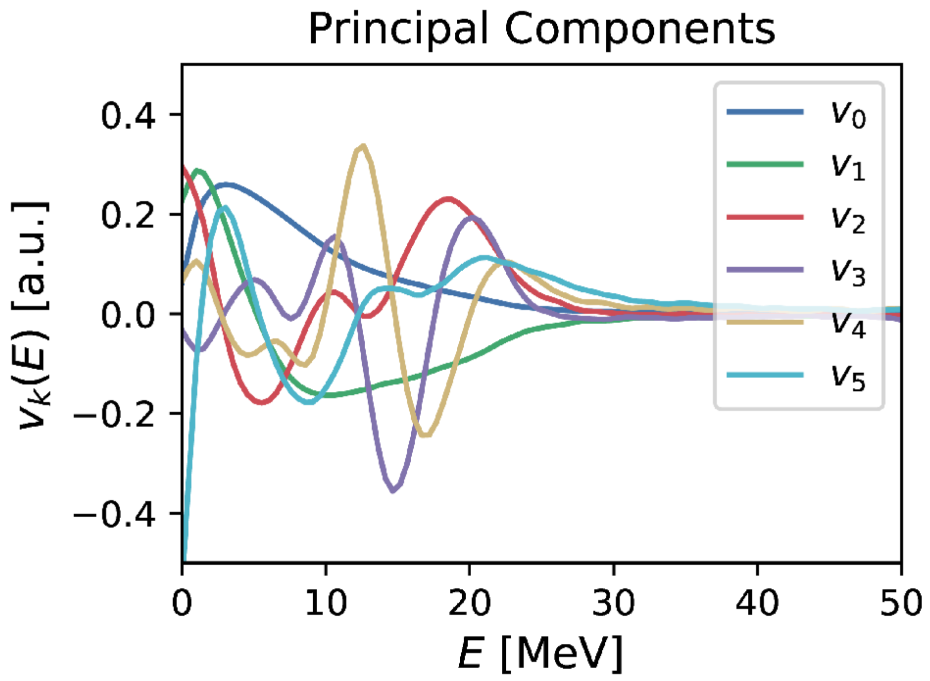
Table 1. Parameter space for PIC simulations.

Quantity	Symbol	Unit	Min	Max	Scaling
Normalized vector potential	a_0	1	2	22	linear
Full width at half maximum	τ	fs	20	50	linear
Number density (bulk)	n_0	n_c	3	50	linear
Target thickness	D	µm	0.1	10	square
Pre-plasma scale length	ℓ	nm	1	1000	cubic

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

Synthetic Electron Spectra





- Measure electron spectra $f_e(E)$ 500 fs after laser-maximum reaches target front

- Many orders of magnitude present in electron spectra → *nonlinearly* transform spectra

- Represent transformed spectra via a *linear basis function* model from PCA:

$$\tilde{f}_e(E) \approx \sum_{k=1}^{6} c_k v_k(E)$$

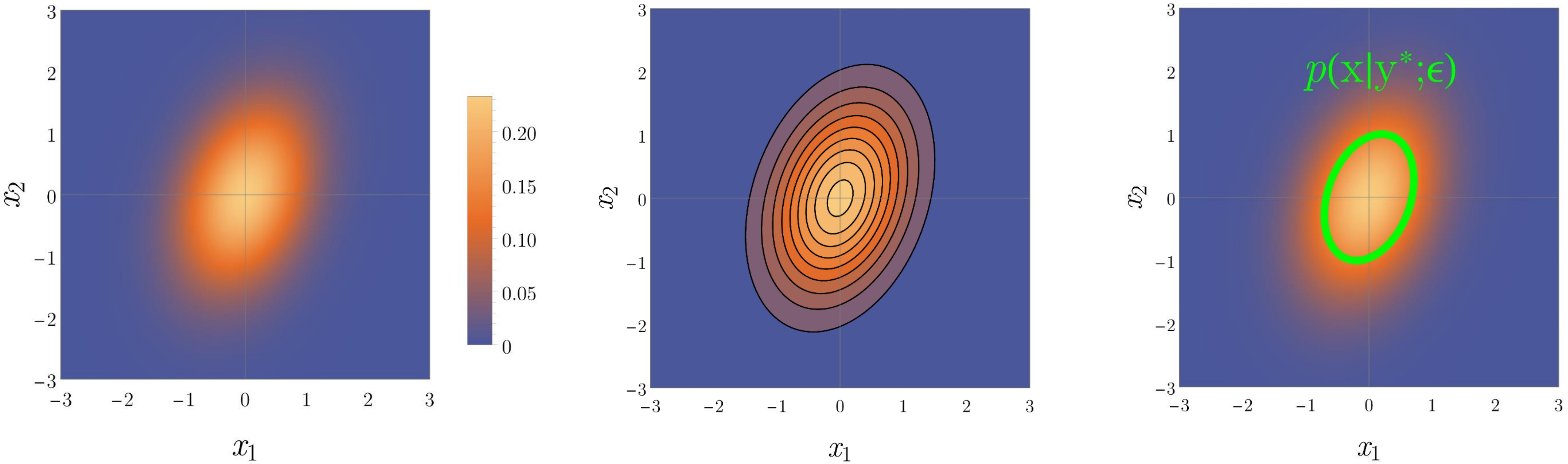
- Train a surrogate model (multilayer perceptron) to predict PCA coefficients

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

Acceptance Condition

- **Problem statement**:
  - Given an observation $\mathbf{y}^*$, return a representative sample $\{\mathbf{x}\}$ of $p(\mathbf{x}|\mathbf{y}^*)$ w.r.t. a **reference model** f for the forward process.
    - Define a distance / measure of (dis)similarity d, and threshold distance ε
    - Accept (i.e., add to solution set) proposed $\mathbf{x}$ if <u>*acceptance condition*</u> is fulfilled:

$$d(f(\mathbf{x}), \mathbf{y}^*) \leq \varepsilon$$



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

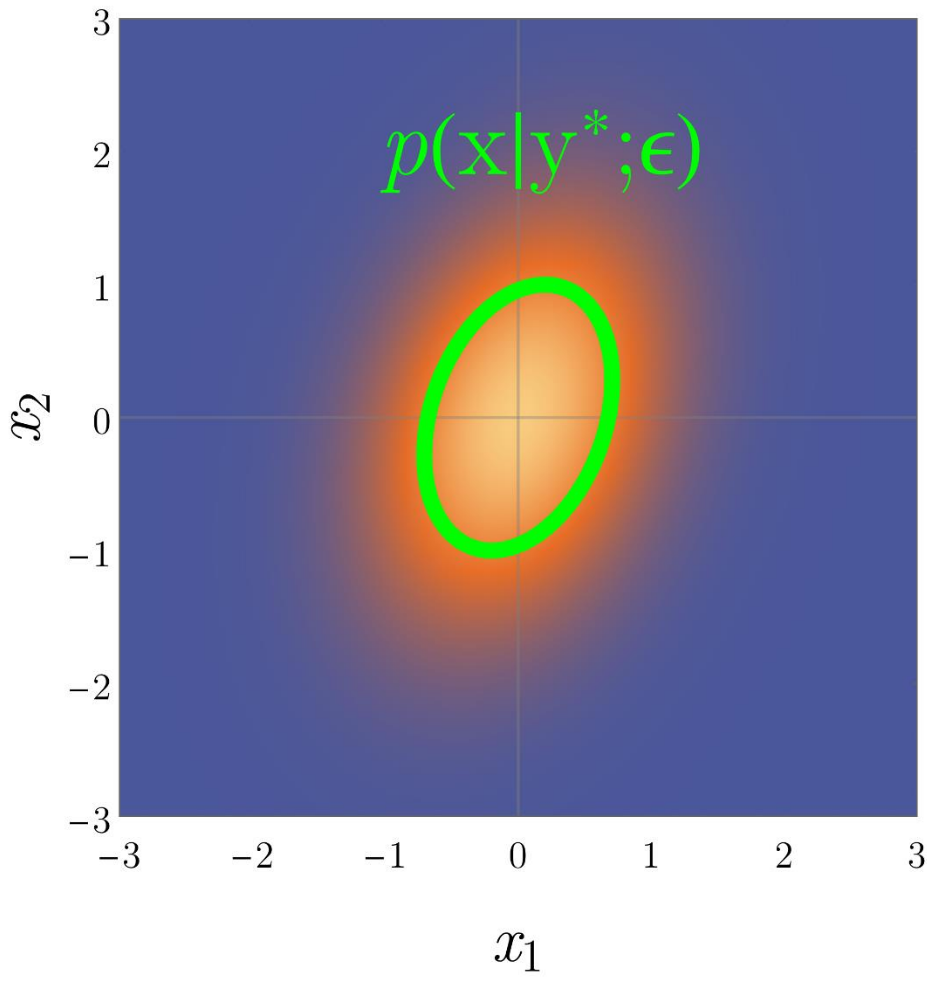
Solution Strategies for Inverse Problems

- **Approximate Bayesian Computation (ABC)**:
  - Propose **randomly drawn** parameter vector $\mathbf{x}$
  - Check $\mathbf{x}$ for acceptance criterion.

- **Search Algorithms**
  - Start at some initial parameter vector $\mathbf{x}_0$
  - **Iteratively improve solution** via search algorithm

- **Invertible Neural Networks**
  - **Novel method**
  - Different architectures / approaches exists
  - Discussed later

- **Markov Chain Monte-Carlo (MCMC)**
  - Not discussed here



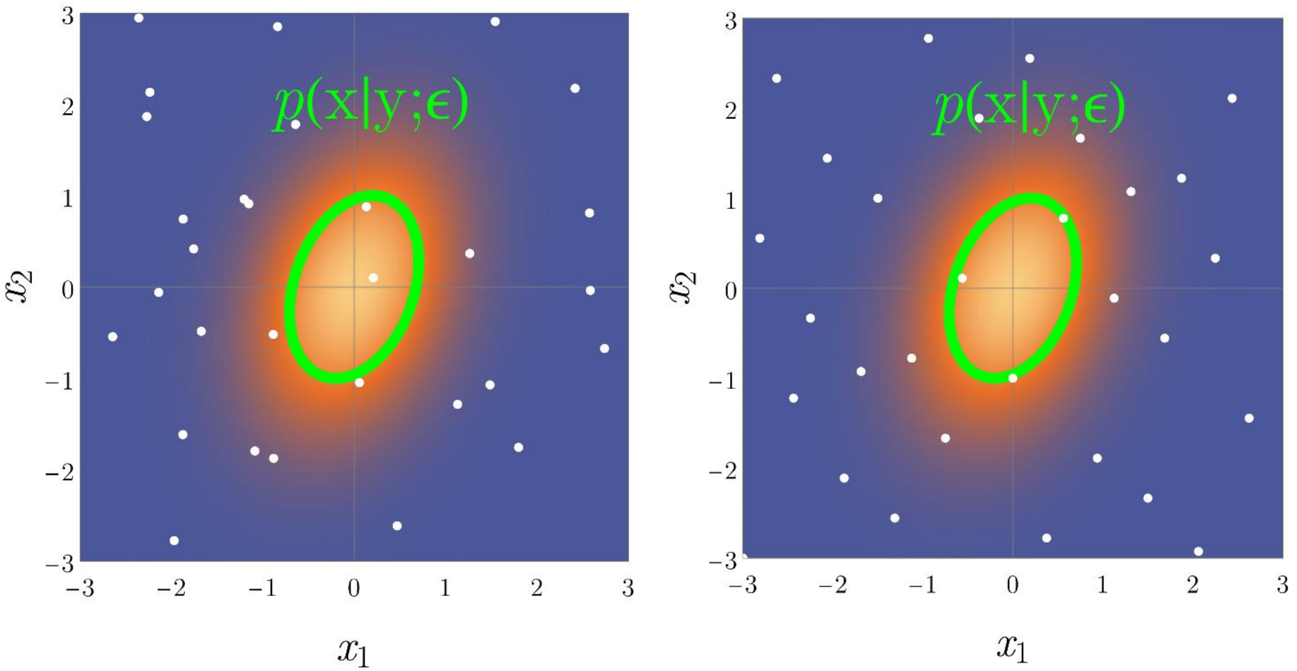
Thomas Miethlinger · Computational Radiation Physics · www.hzdr.de/crp
t.miethlinger@hzdr.de

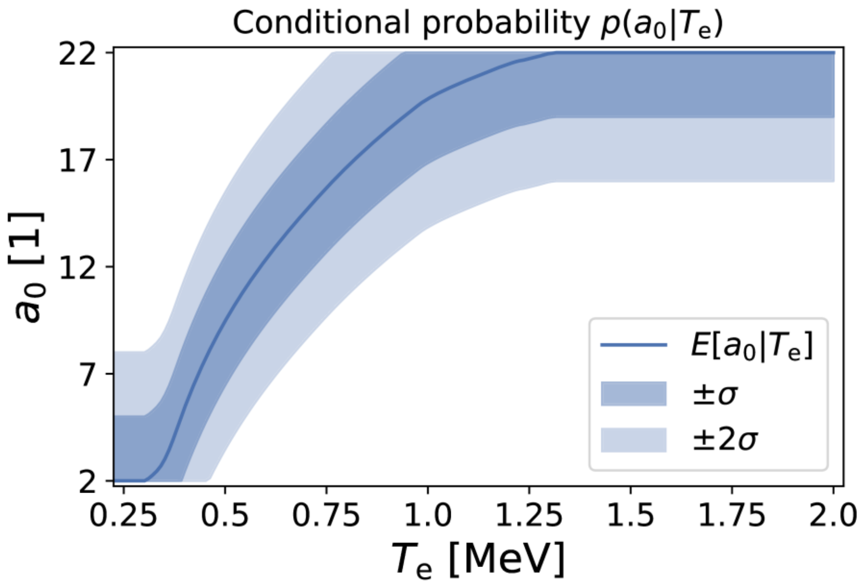
Approximate Bayesian Computation

- **(Standard) ABC**:
  - Propose **randomly drawn** parameter vector $\mathbf{x}$
  - Check $\mathbf{x}$ for acceptance criterion.

- **Quasi ABC**:
  - Construct **low-discrepancy** sequence
  - Propose parameter vector $\mathbf{x}$ according to sequence
  - Check $\mathbf{x}$ for acceptance criterion.

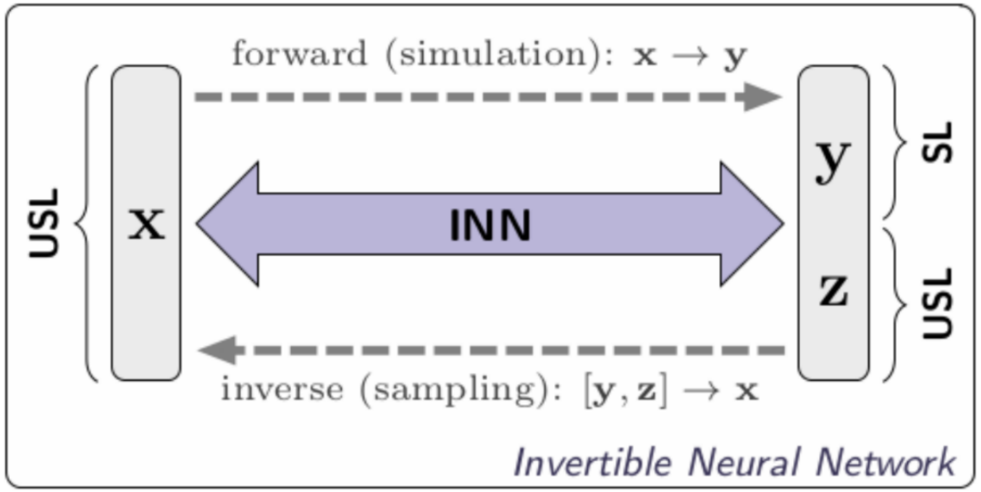
- **Informed ABC:**
  - Include **prior knowledge** to sampling procedure
  - In our example, due to physical considerations:
    - $T_\mathrm{e} = T_\mathrm{e}[f_\mathrm{e}(E)]$
    - but also: $T_\mathrm{e} = T_\mathrm{e}(a_0)$
    - sample $a_0 \sim p(a_0 | T_\mathrm{e}[f_\mathrm{e}(E)])$





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

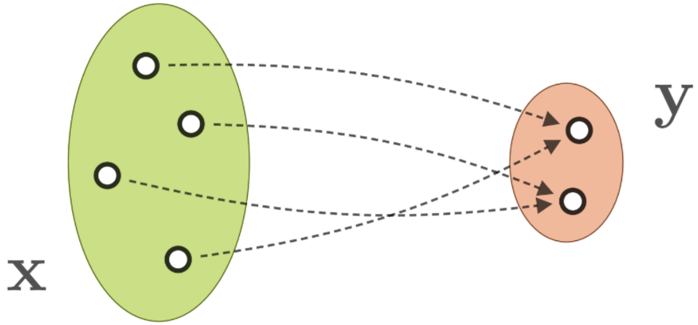
Invertible Neural Networks

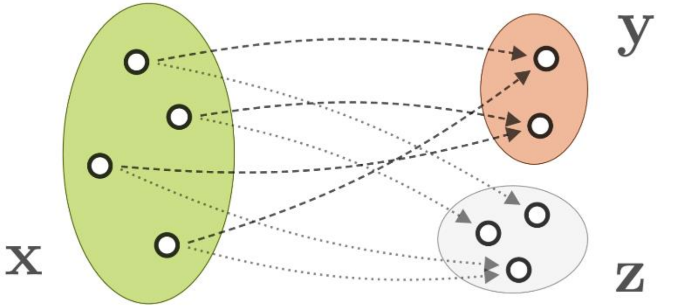


Mode of action of an INN.[1]

- Mapping from in- to output is bijective: inverse (= INN^{-1}) exists
  - Introduction of latent space $\mathbf{z}$
- Train forward process $\mathbf{x} \rightarrow [\mathbf{y},\ \mathbf{z}]$ jointly with inverse process $\mathbf{x} \leftarrow [\mathbf{y},\ \mathbf{z}]$
- Generate set of "input vectors":

$$\{\mathbf{x}\} = \mathrm{INN}^{-1}(\mathbf{y}, \mathbf{z} \sim \mathcal{N}(0,1))$$





Invertibility by introduction of latent vector $\mathbf{z}$:

$$\dim(\mathbf{x}) = \dim(\mathbf{y}) + \dim(\mathbf{z})$$

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

Benchmark Experiment

1. Prepare ground-truth value $\mathbf{y}^*=g(E)$

- Select/specify ground-truth spectrum, $f_{\mathrm{e}}(E)$
- Nonlinearly transform spectrum
  - Filter, take logarithm, etc.
  - Project onto PCA space → $g(E) = \mathbf{y}^*$

2. Prepare proposal spectra $\{f(E)\}$:

- Use inverse solver to generate $\{\mathbf{x}\}$
- Compute $\mathrm{MLP}(\{\mathbf{x}\})=\{\mathbf{c}\}$, the set of proposal PCA coefficient vectors
- Compute proposal spectra from $\{\mathbf{c}\}$ → $\{f(E)\}$, the set of proposal spectra

3. Acceptance criterion:

- Check if relative error:

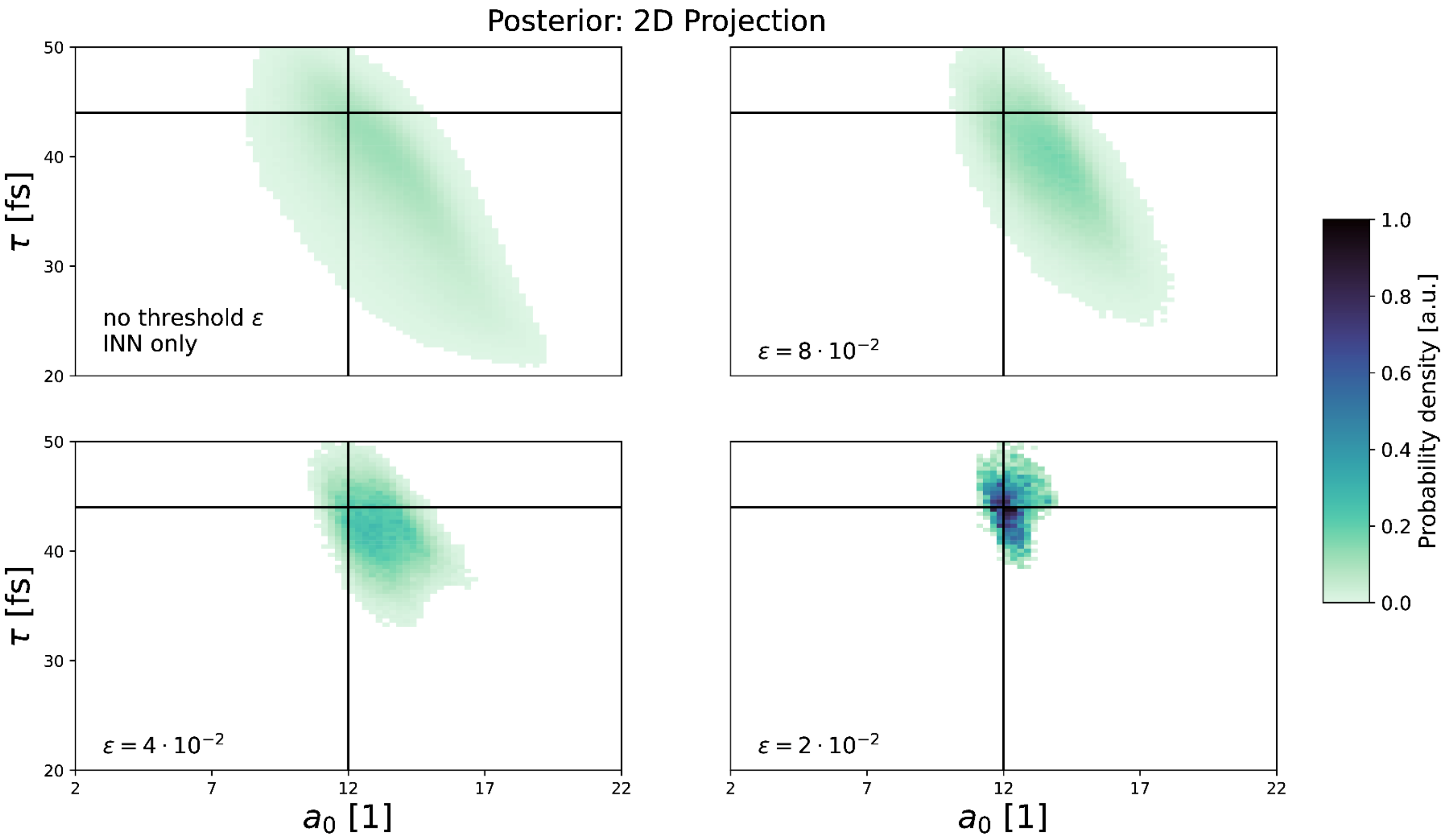
$$d[f(E), g(E)] = \frac{||f(E) - g(E)||_2}{||g(E)||_2} = \frac{\sqrt{\int (f(E) - g(E))^2 \, \mathrm{d}E}}{\sqrt{\int g^2(E) \, \mathrm{d}E}} \leq \varepsilon$$

is fulfilled (relative L2 distance).

If yes → add to solution set

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

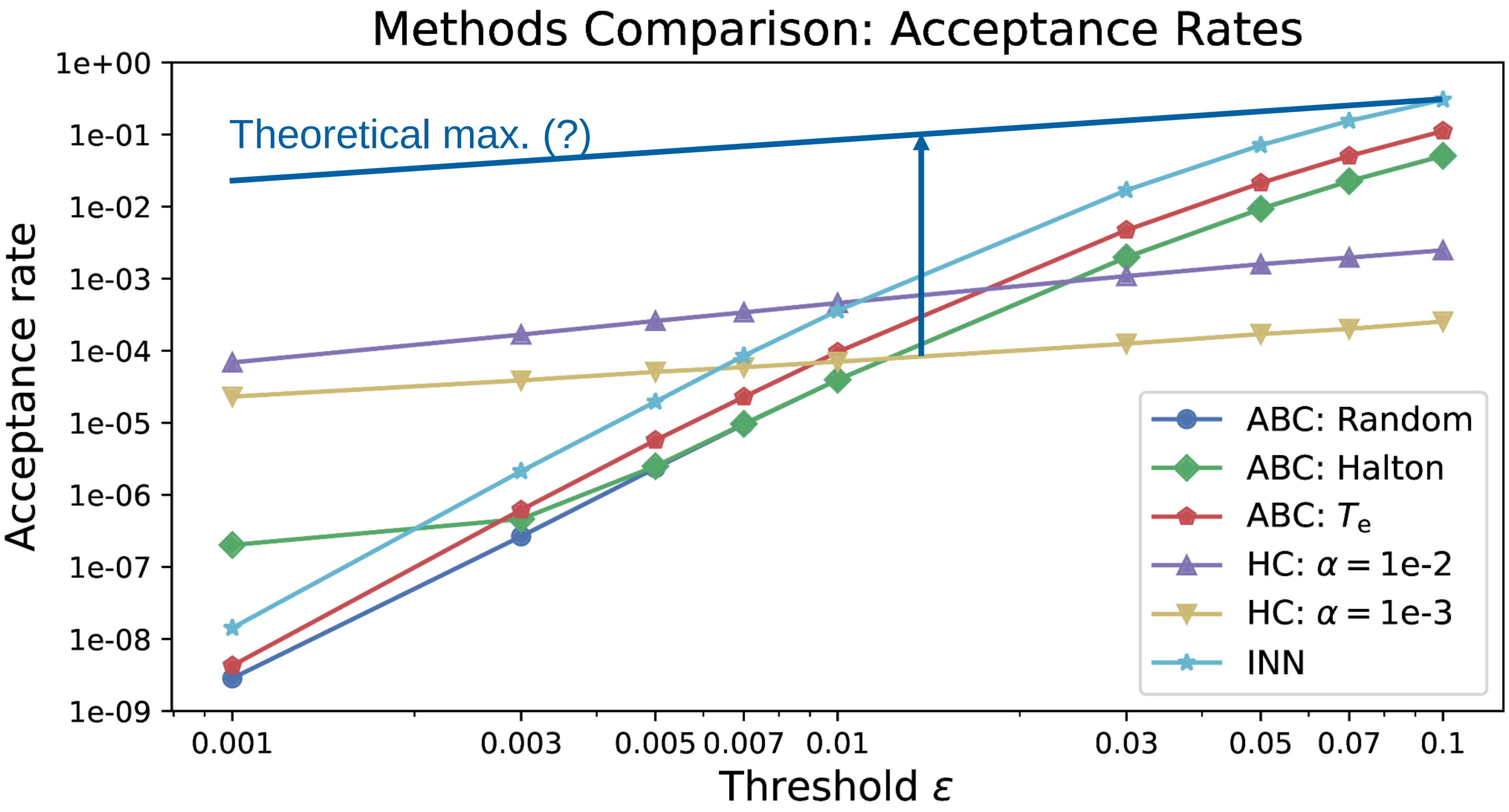
Conditional Posterior: Revisited



- **Typical:** High-dimensional problem → 2D projections
- Reasonable to (initially) assume multivariate Gaussian
  - Compute correlations/covariances
- Smaller ε → probability "localizes"

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

Acceptance Rates

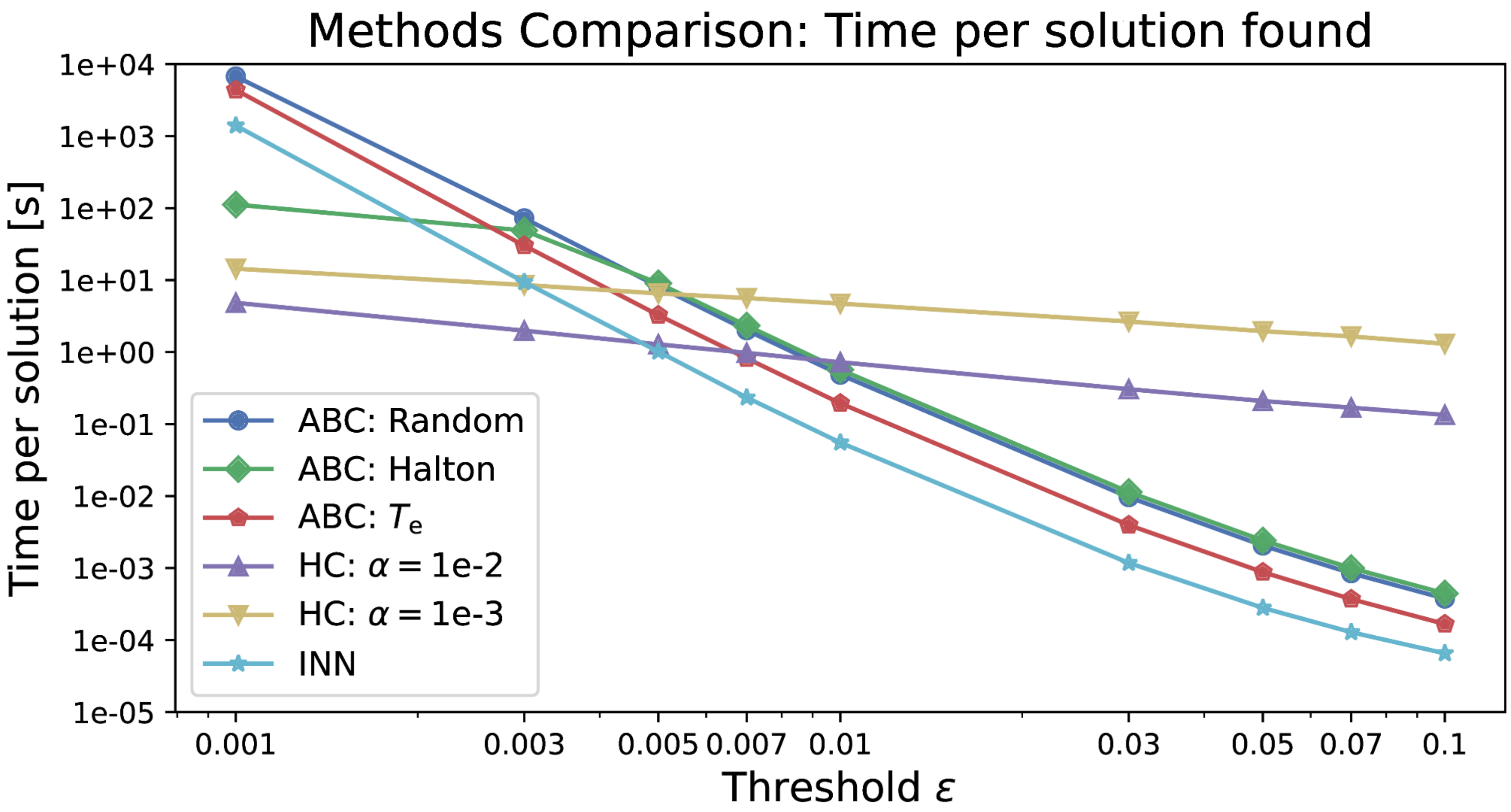


$$\text{acceptance rate} = \frac{\#\text{accepted}}{\#\text{total}}$$

Fig.: Acceptance rates of different approaches in dependence of acceptance threshold ε, averaged over 800 different electron spectra.

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

Runtimes



$$\text{Time per solution} = \frac{\text{sampling time}}{\text{acceptance rate}}$$

Fig.: Time per solution (acceptance) for different approaches in dependence of acceptance threshold ε, averaged over 800 different electron spectra.

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

Conclusion

- INNs need thorough hyperparameter optimization for hard problems
- INNs strongly benefit from dimensionality-reduction techniques
- Sampling time of INNs similar to ABC

- Acceptance rates:
  - ABC shows the worst performance for small ε
    - Partial improvement via quasi-random numbers
    - Informed prior increases acceptance rates by about $\times 2$.
  - HC has the worst performance for high ε, but the highest performance for small ε.
  - INN outperforms ABC by $\times 10$, but has difficulties for small ε.

- General tradeoff: accuracy $\leftrightarrow$ speed

  → Use composite algorithms (INN + “refinement” via, e.g., HC)

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

Acceptance Rates of Invertible Neural Networks on Electron Spectra from Near-Critical Laser-Plasmas: A Comparison

03.02.2023

Thomas Miethlinger[1,2], Nico Hoffmann[1] and Thomas Kluge[1]

[1]Helmholtz-Zentrum Dresden-Rossendorf, 01328 Dresden, Germany
[2]Technische Universität Dresden, 01069 Dresden, Germany

Optimized Inverse Solver: INN-HC

Algorithm 1 Inverse solver optimized for low acceptance thresholds ϵ.

Assuming: $\epsilon \ll \epsilon_{\text{bound}}$

procedure $\textsc{InverseSolver}(\mathbf{y}^\star,\ m,\ f,\ \text{inn},\ d,\ \epsilon,\ \alpha)$
- $\mathbf{Y}^\star \leftarrow \text{vstack}(\mathbf{y}^\star, m)$ $\triangleright$ Vertically stack $\mathbf{y}^\star$, i.e. $\mathbf{Y}^\star \in \mathbb{R}^{m \times n_{\mathbf{y}}}$
- $\mathbf{Z} \leftarrow \text{rand}(\mathcal{N}(0,1), (m, n_{\mathbf{z}}))$ $\triangleright$ $\mathbf{Z} \in \mathbb{R}^{m \times n_{\mathbf{z}}}$
- $\mathbf{X} \leftarrow \text{inn}^{-1}([\mathbf{Y}^\star, \mathbf{Z}])$
- **for** i in $1, \ldots, m$ **do**
  - $\mathbf{x} \leftarrow \mathbf{X}_{i,\cdot}$
  - **if** $d(\mathbf{y}^\star, f(\mathbf{x})) > \epsilon$ **then**
    - $\mathbf{x} \leftarrow \text{FirstChoiceHillClimbing}(\mathbf{y}^\star, f, d, \epsilon, \mathbf{x}, \alpha)$
    - $\mathbf{X}_{i,\cdot} \leftarrow \mathbf{x}$
- **return** $\mathbf{X}$

procedure $\textsc{FirstChoiceHillClimbing}(\mathbf{y}^\star,\ f,\ d,\ \epsilon,\ \mathbf{x}_0,\ \alpha)$
- $\mathbf{x} \leftarrow \mathbf{x}_0$
- **while** $d(\mathbf{y}^\star, f(\mathbf{x})) > \epsilon$ **do**
  - $\boldsymbol{\xi} \leftarrow \text{rand}(\mathcal{U}([-1,1]), (n_{\mathbf{x}}))$ $\triangleright$ Generate vector with random direction
  - $\boldsymbol{\xi} \leftarrow \boldsymbol{\xi} / |\boldsymbol{\xi}|$ $\triangleright$ Normalize to unit length
  - $\tilde{\mathbf{x}} \leftarrow \mathbf{x} + \alpha \boldsymbol{\xi}$
  - **if** $d(\mathbf{y}^\star, f(\tilde{\mathbf{x}})) \leq d(\mathbf{y}^\star, f(\mathbf{x}))$ **then**
    - $\mathbf{x} \leftarrow \tilde{\mathbf{x}}$.
- **return** $\mathbf{x}$

Search Algorithms

Instead of trying uncorrelated solutions randomly, we may try to improve the solution incrementally.

- **First-Choice Hill-Climbing**:
  - Start at random position in parameter space
  - Propose new position

$$\tilde{\mathbf{x}} = \mathbf{x} + \alpha\boldsymbol{\xi},$$

  where α is the "learning rate", and $\boldsymbol{\xi}$ is a unit vector with random direction
  - Check $\tilde{\mathbf{x}}$ for acceptance criterion
    - If accepted, start at new random position to restrict ourselves on uncorrelated samples (same as in ABC)

- **Gradient-based methods / optimizer**:
  - More involved, faster
  - Preferred if expressions for gradients are available
  - → Doesn't work with black-box models
  - Example: Adam, SGD

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