Max-Planck-Institut für Plasmaphysik

Deep learning for fast Bayesian inference of plasma diagnostic models

A. Pavone, J. Svensson, A. Langenberg, M. Brix, M. Krychowiak, U. Hoefel, S. Kwak, R.C. Wolf, the W7-X Team and JET Contributors

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

Content

- Modeling of plasma diagnostics measurements
- Bayesian modelling and inference within the Minerva framework
- Deep learning of Bayesian inference
- Proofs of concept:
  - Electron and ion temperature at the W7-X experiment
  - Edge electron density at the JET experiment
  - Joint probability distribution of a Z_{eff} - bremsstrahlung model at W7-X
- Conclusions and outlook

Modeling of nuclear fusion plasma diagnostics

- Few key plasma parameters:
  - Electron and ion temperature (T_e/T_i)
  - Particle density (n_e)
  - Impurity concentration: *effective charge* Z_{eff}

- Inferred from observations of *several different processes* with plasma diagnostics:
  - Thomson scattering → T_e, n_e
  - Interferometry → n_e
  - Electron cyclotron emission → T_e
  - Several types of spectroscopy (X-ray, visible, etc.) → T_e, T_i, Z_{eff} ...

Inverse problem

- Starting from data : one *inverse* function for each diagnostic observation.
  - $f_i^{-1}(d_i) \rightarrow p$

d: Thomson scattering, interferometer, spectroscopy
p: $T_e, n_e, T_i, \ldots$

- How to merge different observations of same *p* from different *d?*
- Uncertainties estimation required to compare different results
- Conventional statistics: use *estimators* to infer underlying distribution. Assumptions 'hidden' in the choice of the estimators.

Bayesian inference and modeling

- Definition of a model explaining the data.
  - Forward/generative models: predict diagnostic data from plasma parameter $f(p) \rightarrow d$
- Limited explanatory power.
  - Modelling uncertainties: probability distributions (prior, likelihood, posteriors)
- One single rule to infer the posterior -> estimate uncertainties: Bayes formula.

- *One* model of the plasma.

Bayesian inference and modeling

- Model *'m'* of a plasma process can predict observations *'d'* (data)
- *Probability distributions p:* uncertainties in model assumptions and predictions
- Bayes rule:

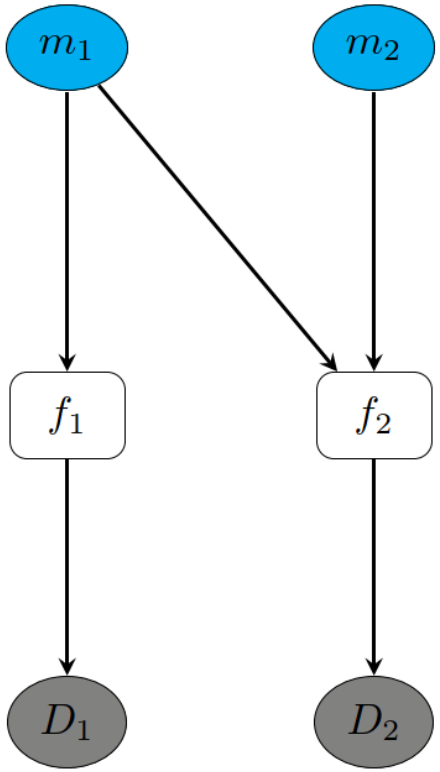
$$\underbrace{p(m|d)}_{\text{posterior}} = \frac{\overbrace{p(d|m)}^{\text{likelihood}}\overbrace{p(m)}^{\text{prior}}}{p(d)} = \frac{\overbrace{p(d,m)}^{\text{joint distribution}}}{p(d)} \propto p(d,m)$$

- Joint distribution $p(d,m)$: landscape of all possible assumptions and predictions

Reverend Thomas Bayes

The Minerva Bayesian modeling framework

- Common computational implementation across different plasma physics models
- *Graphical models* express probabilistic relations according to Bayes rule
- Generalization of inference algorithms

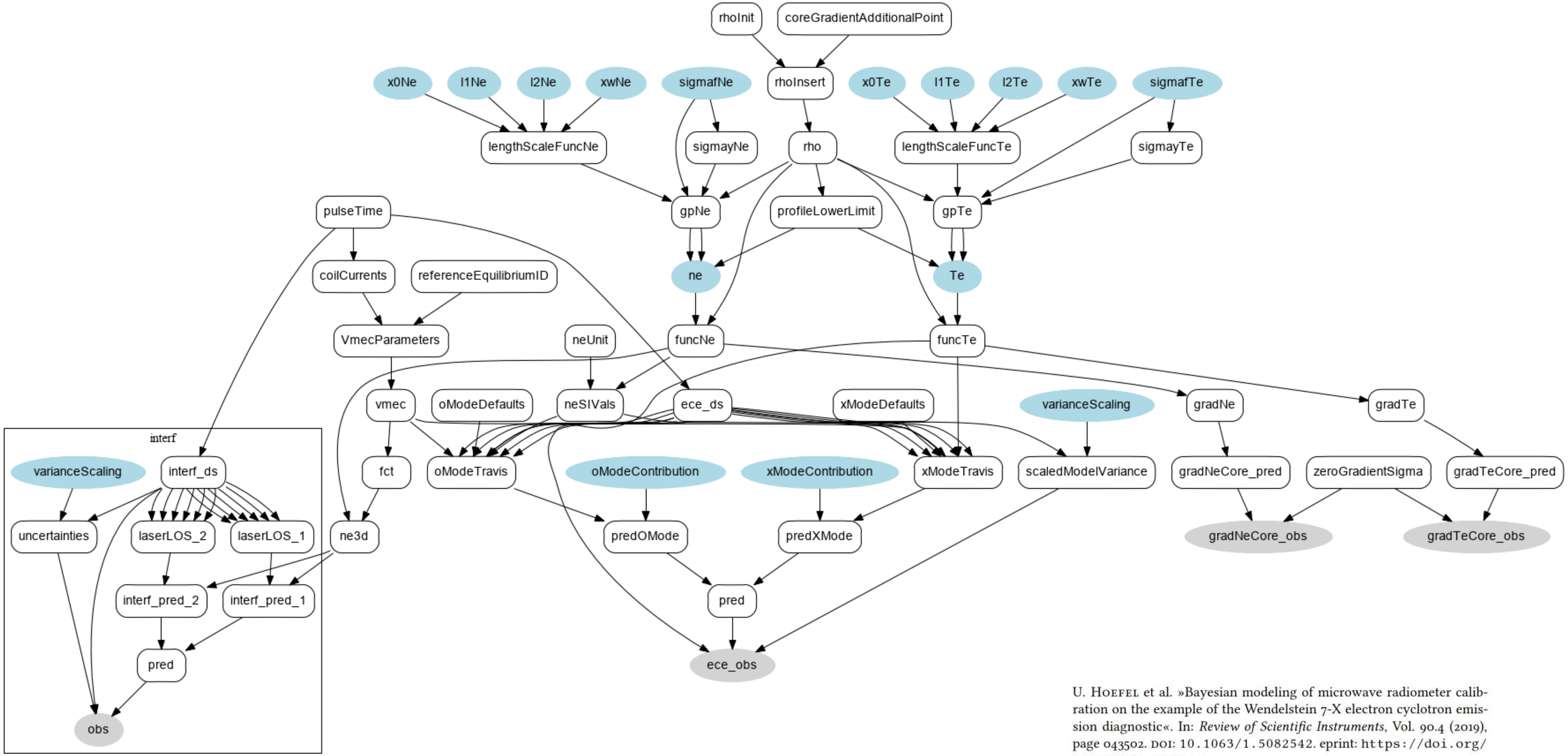


Simplified Bayesian graphical model

- Free parameters m_1 and m_2
- Observation sources D_1 and D_2

J. Svensson and A. Werner. Large Scale Bayesian Data Analysis for Nuclear Fusion Experiments. *IEEE International Symposium on Intelligent Signal Processing*, pages 1–6, 2007.

Graph of the electron cyclotron emission diagnostic at W7-X



U. Hoefel et al. »Bayesian modeling of microwave radiometer calibration on the example of the Wendelstein 7-X electron cyclotron emission diagnostic«. In: *Review of Scientific Instruments*, Vol. 90.4 (2019), page 043502. DOI: 10.1063/1.5082542. eprint: https://doi.org/10.1063/1.5082542.

Bayesian inference within the Minerva framework

- Complex models of multiple diagnostic processes $\longrightarrow$ computationally demanding inference
- Tens of minutes / hours for one single measurement (one measurement record)
- Posterior inference: hard to sample and/or optimize
- Historical approaches to Bayesian inference acceleration: variational Bayes, etc.

- Deep learning to approximate Bayesian inference?

Deep learning of Bayesian inference

- Bayesian models define statistical relations among quantities: $p(d, m)$
- Deep learning models learn statistical relations in the training data (d_t, m_t)
- Training data sampled from $(d_t, m_t) \sim p(d, m)$
- Deep learning of approximate conditionals induced by the Bayesian model
- *'Inverse model'*: $p(m_t|d_t) \sim p(m|d)$
- *'Generative model'*: $p(d_t|\, m_t) \sim p(d|m)$
- Joint probability distribution: $p(d_t, m_t) \sim p(d, m)$

One framework, multiple inference approaches

- Modeling: unified generative Bayesian models of plasma physics processes
- DL model is not about learning new physics (at this stage, at least):
- Physicist implements Bayesian physics model into the framework
- DL model learns to approximate inference under given modeling assumptions

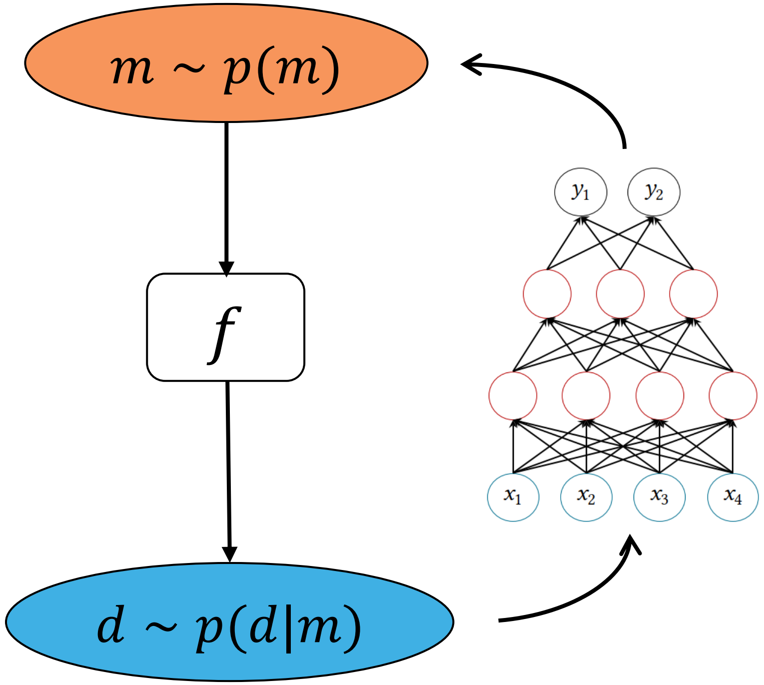
- Bayesian models can be used to generate training data to train DL models
- DL models can learn approximated version of Bayesian inference
- Bayesian inference can be made faster and scalable within one environment, when required

Deep learning of Bayesian model probability distributions

1. Given a (generative) Bayesian model
2. Use it to train the network to learn the *inverse* function $d \rightarrow m$
3. Sample model parameters m and predictions d from the joint distribution of the model:

$$d_t, m_t \sim p(d, m) = p(d|m)p(m)$$

4. Network learns an approximate MAP:
   - MSE error function: mean of Gaussian posterior
   - From training data distribution $p(m_t|d_t)$



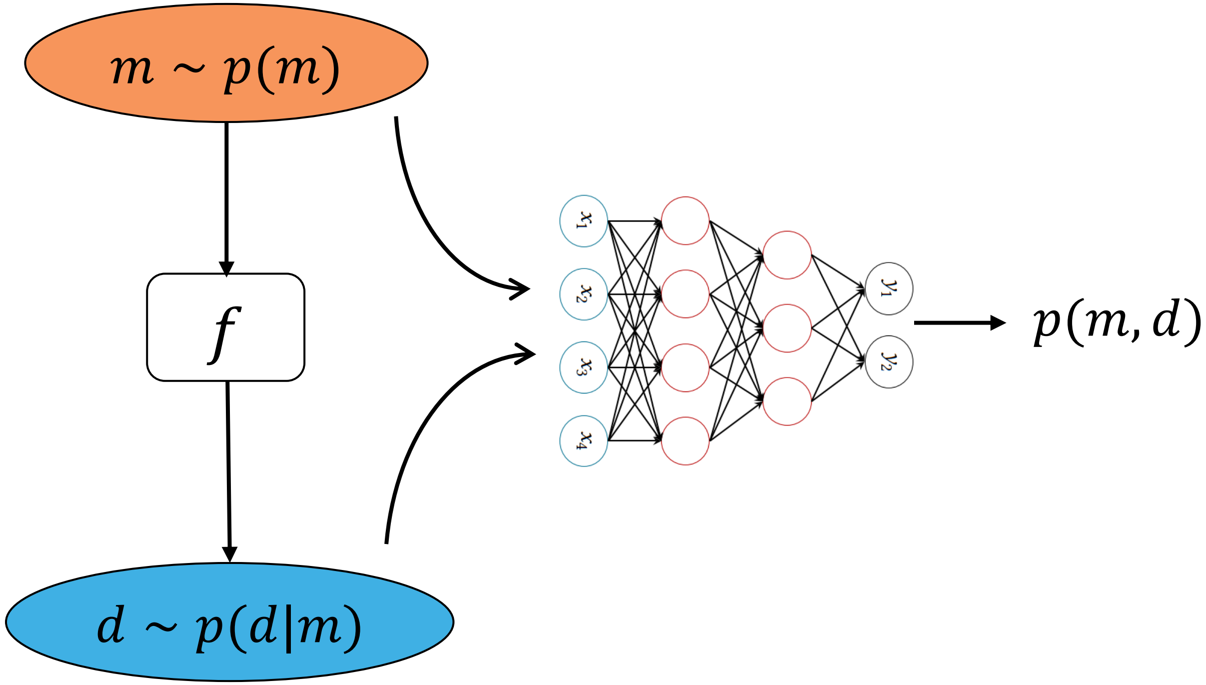
Artificial neural network (ANN) approximate Bayesian inference

1. Sample model parameters m and predictions d from the joint distribution of the model:

$$d_t, m_t \sim p(d,m) = p(d|m)p(m)$$

2. Network learns the joint probability distribution:

$$(m_t, d_t) \rightarrow p(m,d)$$



Artificial neural network (ANN) approximate Bayesian inference

- Training data: *only synthetic*, generated with the Bayesian model according to their distributions
- Evaluation on experimental measurements
- Inference acceleration can be significant:

  from 10 mins to 100 $\mu s \rightarrow 10^6$ acceleration

- Leveraging on a shared computational implementation (Minerva framework) we can automate:
  - Creation of training data
  - Training of machine learning model (ML libraries: e.g., TensorFlow)
  - Deployment and use of machine learning model for fast inference

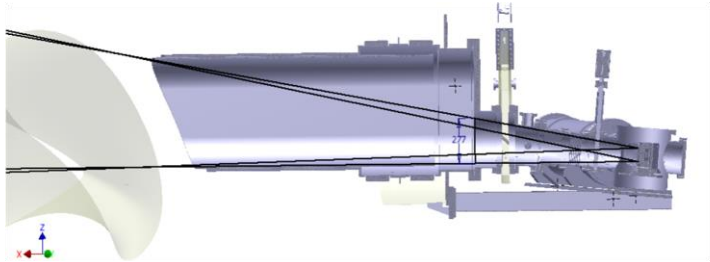
- How to choose machine learning model?
  - Regression problems: simple is better, and enough (MLP, CNN).
  - Time series: RNN.
  - Hyperparameters can be optimized: TensorFlow, AutoML, etc.

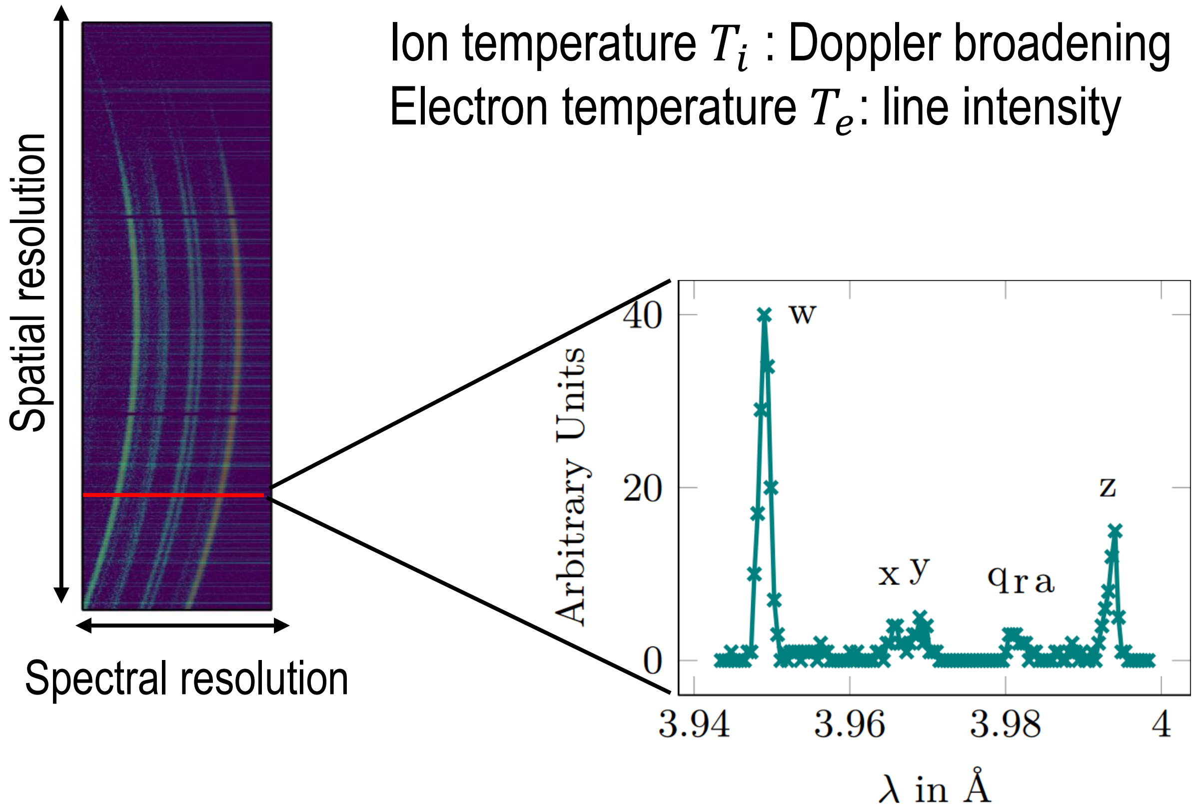
Inference of ion and electron temperature at W7-X

- X-ray imaging crystal spectrometer (XICS) diagnostic: X-rays are emitted in the interaction between injected Argon ions and plasma particles

Observations: 2D images of X-ray spectra across several lines of sight

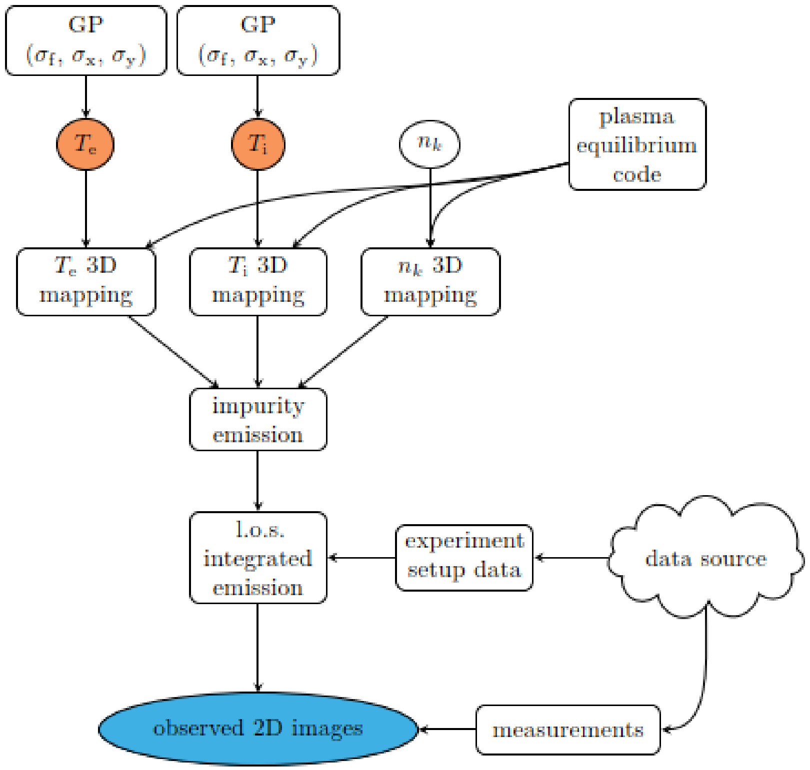
Ion temperature T_i : Doppler broadening
Electron temperature T_e: line intensity

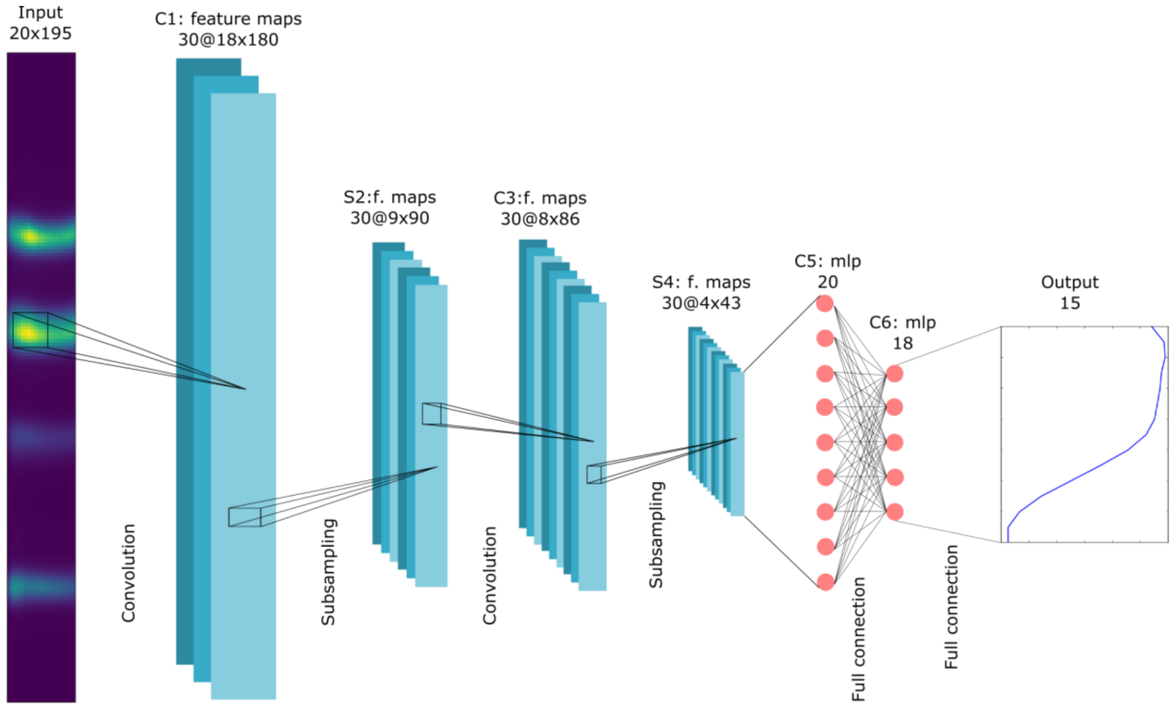




A. Langenberg et al. »Inference of temperature and density profiles via forward modeling of an x-ray imaging crystal spectrometer within the Minerva Bayesian analysis framework«. In: *Review of Scientific Instruments*, Vol. 90.6 (2019), page 063505. DOI: 10.1063/1.5086283. eprint: https://doi.org/10.1063/1.5086283.

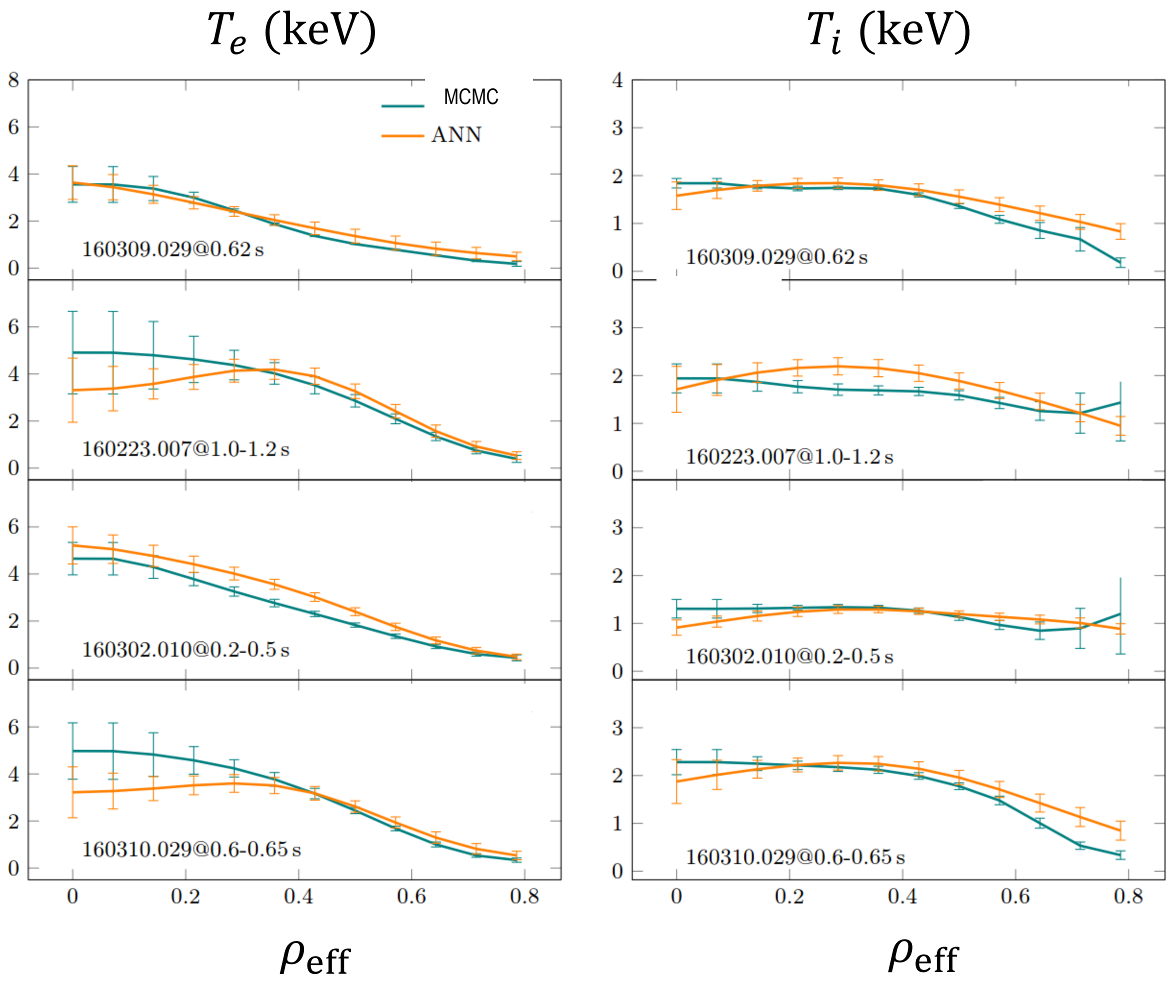
Convolutional neural network for 2D X-ray spectra





A. Pavone et al. »Neural network approximation of Bayesian models for the inference of ion and electron temperature profiles at W7-X«. In: *Plasma Physics and Controlled Fusion*, Vol. 61.7 (May 2019), page 075012. DOI: 10.1088/1361-6587/ab1d26.

DL uncertainties: Bayesian neural networks



- Laplace approximation of weight posteriors:
  - Analytical solution of $p(w|t)$
  - Depending on the Hessian of the weight matrix
  - Sampling from committee of networks
  - Distribution of weights
  - Distribution of training local optima

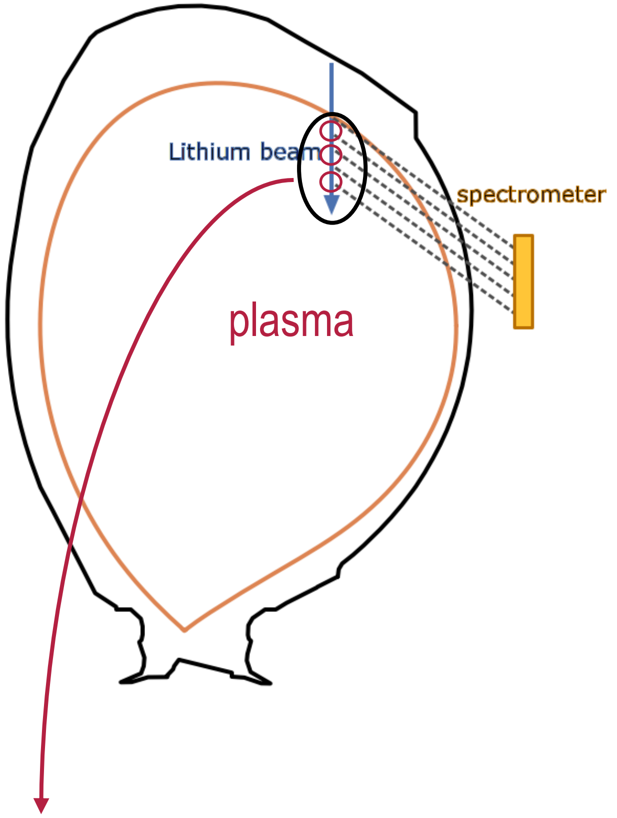
$$\boxed{\sigma_t^2 = \frac{1}{\beta} + \mathbf{g}^T\,\mathbf{A}^{-1}\mathbf{g} \qquad \mathbf{g} = \nabla_{\mathbf{w}} y\Big|_{MP} \qquad \mathbf{A} = \nabla\nabla_{\mathbf{w}} S\Big|_{MP}}$$

S: training loss function
y: network output

A. Pavone et al. »Bayesian uncertainty calculation in neural network inference of ion and electron temperature profiles at W7-X«. In: *Review of Scientific Instruments*, Vol. 89.10 (2018). DOI: 10.1063/1.5039286.

Inference of edge electron density at the JET tokamak

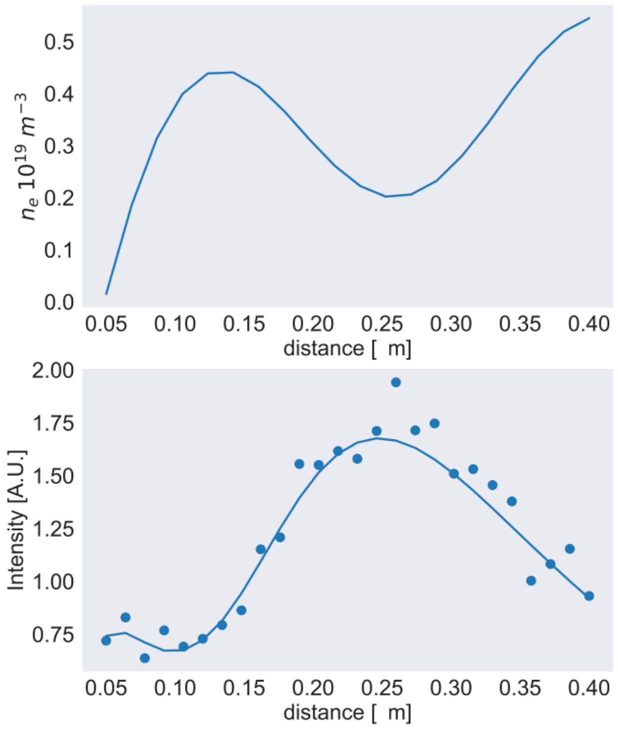
JET plasma cross section



Measuring Li line intensity along 40 cm distance from the top: n_e is inferred at these edge positions

M. Brix et al. »Recent improvements of the JET lithium beam diagnostic«. In: *Review of Scientific Instruments*, Vol. 83.10 (2012), page 10D533. doi: 10.1063/1.4739411. eprint: https://doi.org/10.1063/1.4739411.

- Injection of Lithium atoms in the plasma
- Excitation through collisions with plasma electrons
- Observations: Li emission spectra
- Plasma parameter: (edge) electron density



S. Kwak et al. »Bayesian modelling of the emission spectrum of the Joint European Torus Lithium Beam Emission Spectroscopy system«. In: *Review of Scientific Instruments*, Vol. 87.2 (2016), page 023501. doi: 10.1063/1.4940925. eprint: https://doi.org/10.1063/1.4940925.

n_e

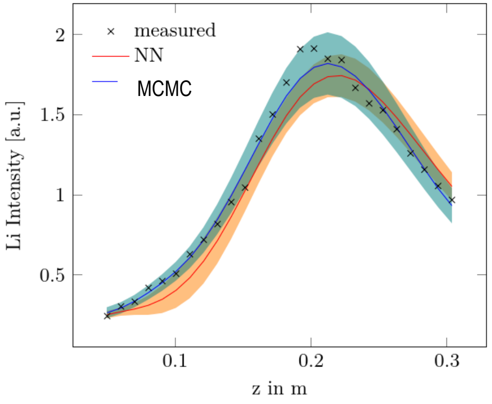
Multi-state collisional radiative model

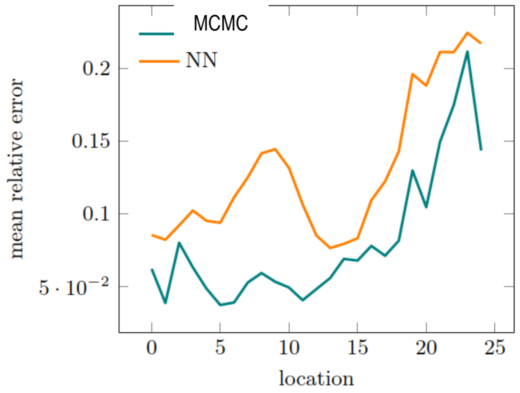
Li line intensity

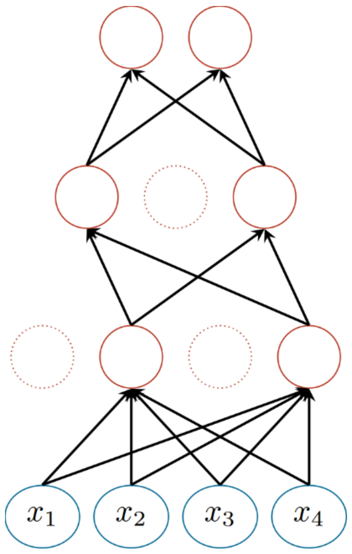
MC dropout training for online uncertainty estimation

- Reconstruction of observed line intensities
- Error (prediction p – measurements m) across several pulses and plasma conditions: $< 20\%$
- Uncertainties: MC dropout
- Variational inference of weight posterior
- Minimization of training loss function (mean squared error) = minimization of Kullback-Leibler divergence $KL(q|p)$

$$KL(q|p) = \int q(w) \log\left(\frac{p(w)}{q(w)}\right) dw$$





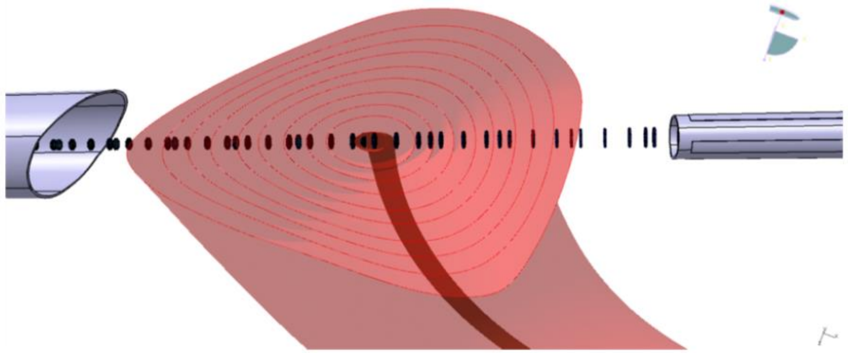


Units/weights dropped with probability q at training **and** inference

A. Pavone et al. »Neural network approximated Bayesian inference of edge electron density profiles at JET«. In: *Plasma Physics and Controlled Fusion*, Vol. 62.4 (Mar. 2020), page 045019. doi: 10.1088/1361-6587/ab7732.

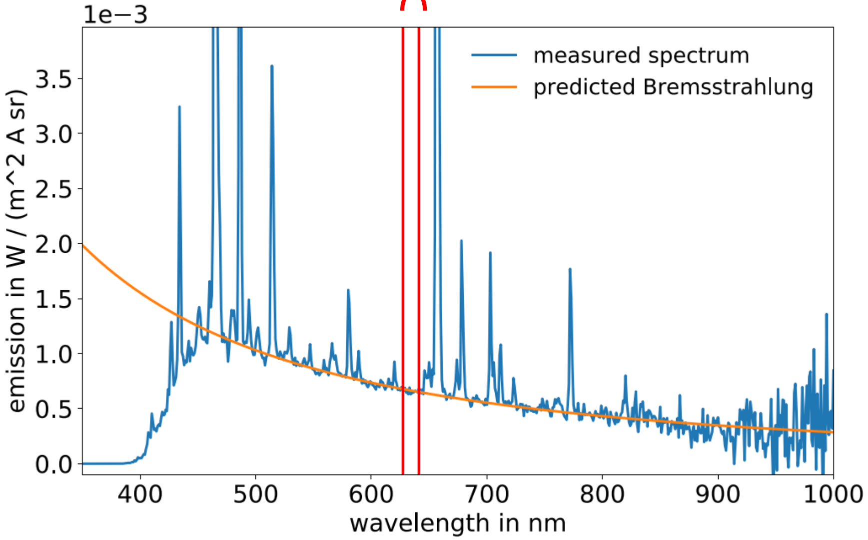
Inference of joint distribution of a Zeff-bremsstrahlung model at W7-X

- Single line-of-sight spectrometer collecting plasma emission



$$\mathrm{I}(\lambda, \boldsymbol{x}) \propto \frac{Z_{\mathrm{eff}} n_e^2}{\sqrt{T_e}}$$

line emission-free region

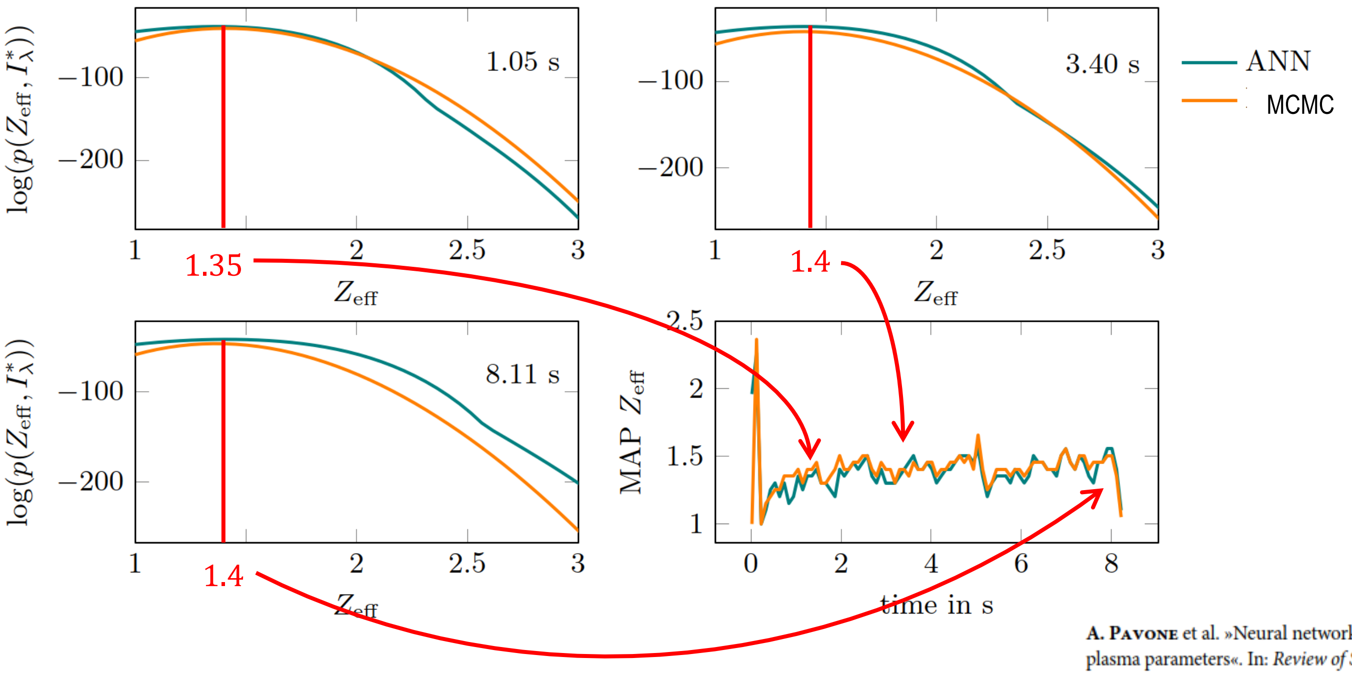


- bremsstrahlung background radiation ($\approx 630-640\ \mathrm{nm}$) estimated from a line emission-free region of the observed spectrum

A. Pavone et al. »Measurements of visible bremsstrahlung and automatic Bayesian inference of the effective plasma charge Zeff at W7-X«. In: *Journal of Instrumentation*, Vol. 14.10 (Oct. 2019), pages C10003–C10003. DOI: 10.1088/1748-0221/14/10/c10003.

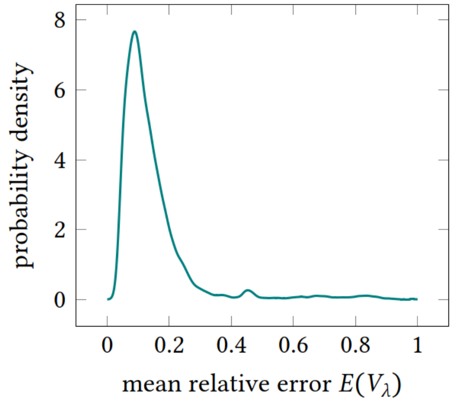
Inference of joint distribution of a Zeff-bremsstrahlung model at W7-X

- Training input: $Z_{\text{eff}}, I_\lambda$
- Training target: $p(Z_{\text{eff}}, I_\lambda) = p(Z_{\text{eff}})p(I_\lambda | Z_{\text{eff}}; n_e, T_e)$
- Reconstruction of posterior $p(Z_{\text{eff}} | I_\lambda^*)$ from experimental measurement of I_λ^*



plasma shot # 20180807.015

- Mean relative error below 0.2 for most cases



A. Pavone et al. »Neural network surrogates of Bayesian diagnostic models for fast inference of plasma parameters«. In: *Review of Scientific Instruments*, Vol. 92.3 (2021), p. 033531. DOI: 10.1063/5.0043772. eprint: https://doi.org/10.1063/5.0043772.

Conclusions

Bayes theory, NN approximate inference and shared computational modeling framework:

- Unified modeling of plasma physics processes to predict observations
- Training on **virtual** Bayesian models -> possible **generalization** to different devices/systems before experimental data are available
- Computationally sustainable and scalable inference independently of model complexity (≈100 μs)
- Deep learning model uncertainties computable at inference time

- Real time applications possible
- DL based fast approximate inference immediately generalized for any integrated model

- Acceleration of posterior sampling (MCMC/DL-based variational inference)
- Physics constraints into DL model (with forward model decoding)