

# SCALABLE BAYESIAN INFERENCE WITH MODEL BASED MACHINE LEARNING

A. Pavone  
Max-Planck Institute for Plasma Physics  
Greifswald, Germany  
Email: [andrea.pavone@ipp.mpg.de](mailto:andrea.pavone@ipp.mpg.de)

J. Svensson  
Max-Planck Institute for Plasma Physics  
Greifswald, Germany

A. Langenberg  
Max-Planck Institute for Plasma Physics  
Greifswald, Germany

M. Krychowiak  
Max-Planck Institute for Plasma Physics  
Greifswald, Germany

S. Kwak  
Max-Planck Institute for Plasma Physics  
Greifswald, Germany

R.C. Wolf  
Max-Planck Institute for Plasma Physics  
Greifswald, Germany

W7-X Team  
Max-Planck Institute for Plasma Physics  
Greifswald, Germany

## Abstract

At large nuclear fusion experiments, plasma discharges are assessed through measurements collected with several diagnostic devices. Each instrument collects data generated in different physics processes: a consistent and efficient exploitation of the information contained in each different data source regarding a few common plasma parameters can be achieved with Bayesian inference. In Bayesian inference, a model of the physics processes generating the observations is prescribed. Such a model consists of physical assumptions and the quantification of the uncertainties with probability distributions regarding the model parameters prior to any observation and the predictive process. The former are described with a *prior* distribution, whereas the latter with a *likelihood* function. The product of these two distributions is known as the *joint* probability distribution of the model, which provides a comprehensive view of the distribution of the parameters and the observations which can be generated with the Bayesian model. The Minerva Bayesian modelling framework offers a modular environment where physics models of measuring devices and corresponding uncertainties can be implemented computationally as Bayesian models such that inference can be carried out on the available experimental observations. Conventional Bayesian inference algorithms such as Markov Chain Monte Carlo for sampling from the *posterior* distribution of the model parameters and Maximum A Posteriori for point estimates of the most likely values are typically robust but slow, especially for computationally expensive forward models. Ideally, it would be desirable to have the soundness and integrity of Bayesian modelling and inference with the efficiency of fast machine learning algorithms, which modern artificial neural networks have demonstrated in different domains of science. This can be achieved by training neural networks on data sampled from a Bayesian model joint probability distribution. Networks trained on such data can learn to approximate different aspects of Bayesian inference: they can learn to reconstruct the model parameters given the observations, the observations given the model parameters and the posterior probability distribution values. Each one of these tasks has its own applicability in the acceleration of the Bayesian inference process. MAP point estimates, MCMC sampling and direct posterior sampling can be estimated with a properly trained neural network model in time scales of hundreds to thousands of microseconds. In principle, this allows to develop “real time” tools based on reliable Bayesian inference results. Moreover, the neural network predictions can also be adorned with uncertainties, calibrated in a Bayesian fashion, without introducing significant time delays. Finally, by leveraging on a common computational implementation of different Bayesian models within the Minerva modelling framework, it is possible to produce fast and approximate machine learning based representation of any model in an automatic fashion, independently of its specific features and implementation details. This method could therefore be thought as a general approach to make Bayesian inference a fast, scalable, and versatile modelling tool.