Contribution ID: **0** Type: **not specified**

# Learning the latent structure of collider events

*Thursday 16 July 2020 18:00 (20 minutes)*

We describe a technique to learn the underlying structure of collider events directly from the data, without having a particular theoretical model in mind. It allows to infer aspects of the theoretical model that may have given rise to this structure, and can be used to cluster or classify the events for analysis purposes. The unsupervised machine-learning technique is based on the probabilistic (Bayesian) generative model of Latent Dirichlet Allocation. We pair the model with an approximate inference algorithm called Variational Inference, which we then use to extract the latent probability distributions describing the learned underlying structure of collider events. We provide a detailed systematic study of the technique using two example scenarios to learn the latent structure of di-jet event samples made up of QCD background events and either a pair-produced top-quark signal or a new physics W' signal. We also present results from using this technique to infer new physics from the LHC Olympics datasets.

## Summary

We describe an unsupervised method based on a probabilistic generative model called Latent Dirichlet Allocation, used in conjunction with a stochastic variational inference technique, to uncover new physics in di-jet events at high-energy colliders. We provide a number of examples, and will present results obtained from running this technique on the LHC Olympics datasets.

**Primary authors:** Dr DILLON, Barry (Jozef Stefan Institute); Dr FAROUGHY, Darius (Zurich University); Prof. KAMENIK, Jernej (IJS); Mr SZEWC, Manuel (International Center for Advanced Studies, Argentina)

**Presenter:** Dr DILLON, Barry (Jozef Stefan Institute)

**Session Classification:** Anomaly Detection