Generative and Invertible Networks for LHC Theory

Tilman Plehn

Universität Heidelberg

STRUCTURES 6/2021

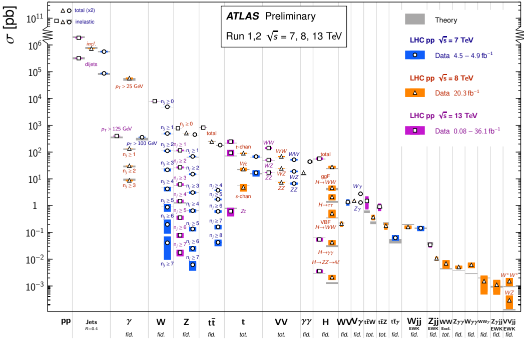
LHC goals

Fundamental questions

- particle nature of dark matter?
- origin of the Higgs mechanism? [hierarchy problem?]
- matter-antimatter asymmetry? [CP-symmetry]
- Standard Model all there is?

Impressive measurements

- many processes
- vastly different rates
- high precision
- predicted by theory



LHC goals

Fundamental questions

– particle nature of dark matter?

– origin of the Higgs mechanism? [hierarchy problem?]

– matter-antimatter asymmetry? [CP-symmetry]

– Standard Model all there is?

Impressive measurements

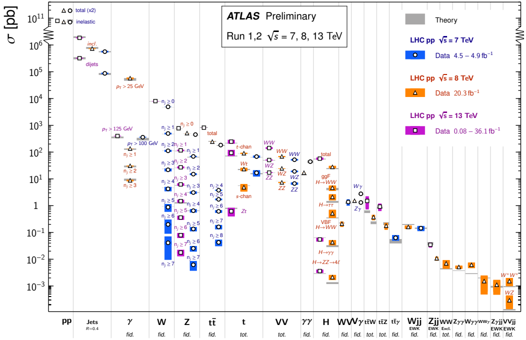
– many processes

– vastly different rates

– high precision

– predicted by theory

– but completely useless!



LHC goals

Fundamental questions

- particle nature of dark matter?
- origin of the Higgs mechanism? [hierarchy problem?]
- matter-antimatter asymmetry? [CP-symmetry]
- Standard Model all there is?

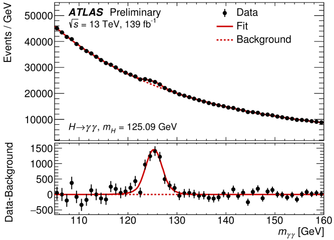
Impressive measurements

- many processes
- vastly different rates
- high precision
- predicted by theory
- but completely useless!

Rates not interesting

- energy distributions dropping
- new physics heavy

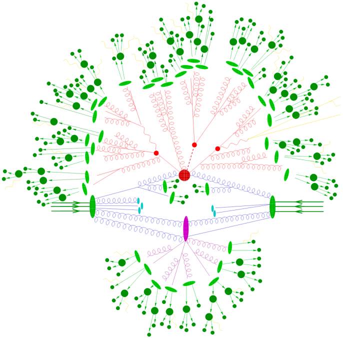
$\Rightarrow$ bumps, tails, kinematics instead

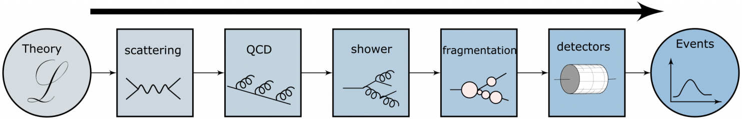


LHC simulations

Simulation-based inference [likelihood-free inference]

– start with Lagrangian
– calculate scattering in perturbative QFT
– simulate events [theory: Sherpa, Madgraph, Pythia]
– simulate detectors [experiment: ATLAS, CMS, Delphes]
$\Rightarrow$ LHC events in virtual worlds





LHC simulations

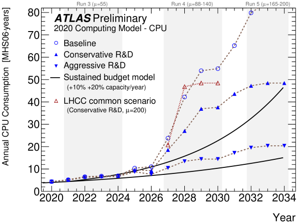
Simulation-based inference [likelihood-free inference]

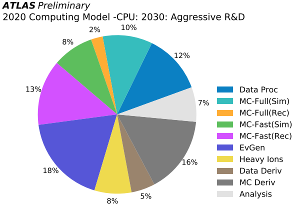
- start with Lagrangian
- calculate scattering in perturbative QFT
- simulate events [theory: Sherpa, Madgraph, Pythia]
- simulate detectors [experiment: ATLAS, CMS, Delphes]

$\Rightarrow$ LHC events in virtual worlds

HL-LHC: preparing for 25-fold data set

- simulated event numbers $\sim$ expected events
- statistics requiring 1%-2% uncertainty [NNLO/N^{3}LO]
- flexible signal hypotheses [time-dependent]
- low-rate high-multiplicity backgrounds





LHC simulations

Simulation-based inference [likelihood-free inference]

- start with Lagrangian
- calculate scattering in perturbative QFT
- simulate events [theory: Sherpa, Madgraph, Pythia]
- simulate detectors [experiment: ATLAS, CMS, Delphes]

$\Rightarrow$ LHC events in virtual worlds

HL-LHC: preparing for 25-fold data set

- simulated event numbers $\sim$ expected events
- statistics requiring 1%-2% uncertainty [NNLO/N^{3}LO]
- flexible signal hypotheses [time-dependent]
- low-rate high-multiplicity backgrounds

Three ways to use ML

- improve current tools: iSherpa, ML-MadGraph...
- new ideas: fast ML-generator-networks...
- conceptual ideas for simulations and analyses

ML for LHC

LHC and Ben Nachman [EPS prize 2021]

– 1991: NN-based quark-gluon tagger
– 2005: TMVA in Root — analysis
– 2015: jet images — classification
– 2017: CaloGAN — jet generation

USING NEURAL NETWORKS TO IDENTIFY JETS

Leif LÖNNBLAD*, Carsten PETERSON** and Thorsteinn RÖGNVALDSSON***
Department of Theoretical Physics, University of Lund, Sölvegatan 14A, S-22362 Lund, Sweden

Received 29 June 1990

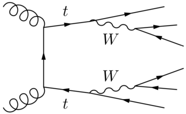
A neural network method for identifying the ancestor of a hadron jet is presented. The idea is to find an efficient mapping between certain observed hadronic kinematical variables and the quark-gluon identity. This is done with a neuronic expansion in terms of a network of sigmoidal functions using a gradient descent procedure, where the errors are back-propagated through the network. With this method we are able to separate gluon from quark jets originating from Monte Carlo generated e^+e^- events with ~85% approach. The result is independent of the MC model used. This approach for isolating the gluon jet is then used to study the so-called string effect.

In addition, heavy quarks (b and c) in e^+e^- reactions can be identified on the 50% level by just observing the hadrons. In particular we are able to separate b-quarks with an efficiency and purity, which is comparable with what is expected from vertex detectors. We also speculate on how the neural network method can be used to disentangle different hadronization schemes by compressing the dimensionality of the state space of hadrons.

Heidelberg history

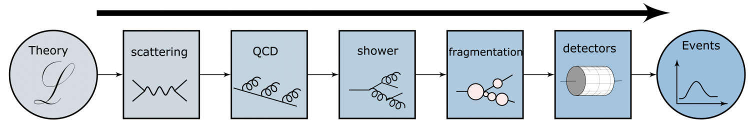
– 2017: top-jet tagger — classification
– 2018: jet autoencoder — unsupervised
– 2019: jet classification with uncertainties [w/ Manuel Haußmann]
– 2019: event GAN — generation [How to GAN #1-5]
– 2020: conditional INN — unfolding [w/ Ulli Köthe, Lynton Ardizzone]
– 2020: BayesFlow — inference [w/ Ulli Köthe, Stefan Radev]
– 2021: Bayesian INN — generation with uncertainties [w/ Manuel Haußmann]
– 2021: Dirichlet VAE — unsupervised
⇒ unsupervised, uncertainties, generation

INN-generating LHC events

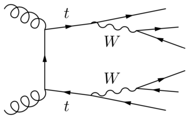


LHC scattering benchmarks

- n-particle phase space $\quad n \times 3$ d.o.f.
  training from event samples [optimal transport vs likelihoods]
  energy-momentum conservation learned
  no detector effects [smoother structures]
- top-quark pairs $\quad t\bar{t} \to 6$ jets
  $t/\bar{t}$ and $W^\pm$ on-shell [Breit-Wigner $\Gamma/m \sim \mathcal{O}(\%)$]
- n-jets/WZ+jets
  variable number of particles
  topology through jet-jet separation



INN-generating LHC events

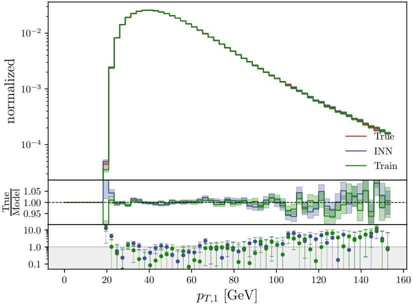


LHC scattering benchmarks

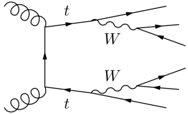
- n-particle phase space $\quad n \times 3$ d.o.f.
  training from event samples [optimal transport vs likelihoods]
  energy-momentum conservation learned
  no detector effects [smoother structures]
- top-quark pairs $\quad t\bar{t} \to 6$ jets
  $t/\bar{t}$ and $W^\pm$ on-shell [Breit-Wigner $\Gamma/m \sim \mathcal{O}(\%)$]
- n-jets/WZ+jets
  variable number of particles
  topology through jet-jet separation

INN-generator [Butter, Heimel, TP,... (prelim)]

- normalizing flow to phase space
  bijective, stable mapping
  Jacobian known
- training on 2M n-jet events
- goal 1% precision



INN-generating LHC events

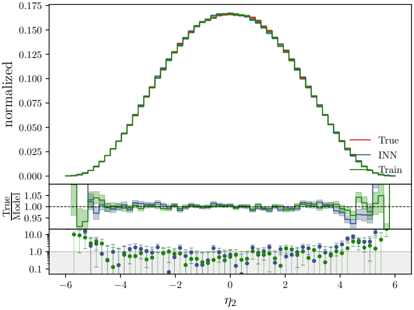


LHC scattering benchmarks

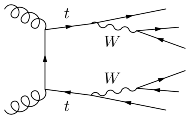
- n-particle phase space $\quad n \times 3$ d.o.f.
  training from event samples [optimal transport vs likelihoods]
  energy-momentum conservation learned
  no detector effects [smoother structures]
- top-quark pairs $\quad t\bar{t} \to 6$ jets
  $t/\bar{t}$ and $W^\pm$ on-shell [Breit-Wigner $\Gamma/m \sim \mathcal{O}(\%)$]
- n-jets/WZ+jets
  variable number of particles
  topology through jet-jet separation

INN-generator [Butter, Heimel, TP,... (prelim)]

- normalizing flow to phase space
  bijective, stable mapping
  Jacobian known
- training on 2M n-jet events
- goal $\quad$ 1% precision



INN-generating LHC events

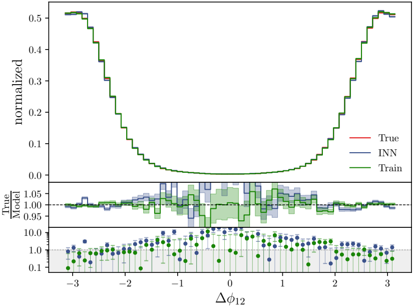


LHC scattering benchmarks

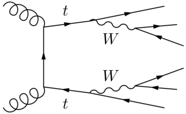
- n-particle phase space $\quad n \times 3$ d.o.f.
  training from event samples [optimal transport vs likelihoods]
  energy-momentum conservation learned
  no detector effects [smoother structures]
- top-quark pairs $\quad t\bar{t} \to 6$ jets
  $t/\bar{t}$ and $W^\pm$ on-shell [Breit-Wigner $\Gamma/m \sim \mathcal{O}(\%)$]
- n-jets/WZ+jets
  variable number of particles
  topology through jet-jet separation

INN-generator [Butter, Heimel, TP,... (prelim)]

- normalizing flow to phase space
  bijective, stable mapping
  Jacobian known
- training on 2M n-jet events
- goal 1% precision



INN-generating LHC events



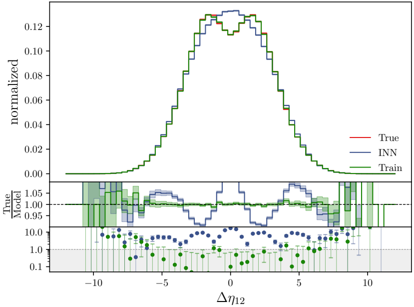
LHC scattering benchmarks

- n-particle phase space $\quad n \times 3$ d.o.f.
  training from event samples [optimal transport vs likelihoods]
  energy-momentum conservation learned
  no detector effects [smoother structures]
- top-quark pairs $\quad t\bar{t} \to 6$ jets
  $t/\bar{t}$ and $W^\pm$ on-shell [Breit-Wigner $\Gamma/m \sim \mathcal{O}(\%)$]
- n-jets/WZ+jets
  variable number of particles
  topology through jet-jet separation

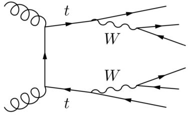
INN-generator [Butter, Heimel, TP,... (prelim)]

- normalizing flow to phase space
  bijective, stable mapping
  Jacobian known
- training on 2M n-jet events
- goal 1% precision

$\Rightarrow$ challenges remaining...



INN-generating LHC events



LHC scattering benchmarks

– n-particle phase space $\quad n \times 3$ d.o.f.
  training from event samples [optimal transport vs likelihoods]
  energy-momentum conservation learned
  no detector effects [smoother structures]

– top-quark pairs $\quad t\bar{t} \to 6$ jets
  $t/\bar{t}$ and $W^{\pm}$ on-shell [Breit-Wigner $\Gamma/m \sim \mathcal{O}(\%)$]

– n-jets/WZ+jets
  variable number of particles
  topology through jet-jet separation

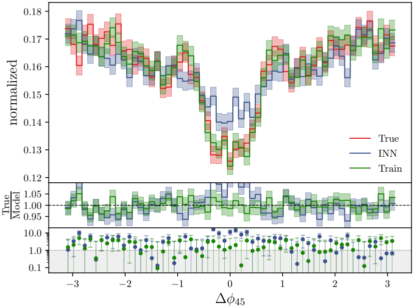
INN-generator [Butter, Heimel, TP,... (prelim)]

– normalizing flow to phase space
  bijective, stable mapping
  Jacobian known

– training on 2M n-jet events

– goal 1% precision

$\Rightarrow$ challenges remaining...

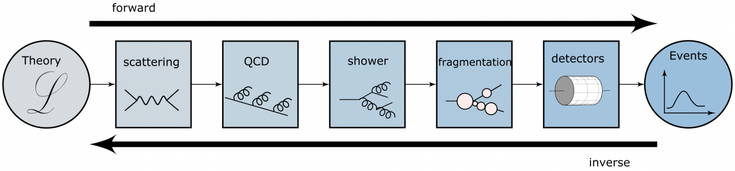


INNverting event simulations

Inverting LHC simulations

– unfolding QCD-shower to hard parton standard [jet algorithm]
  unfolding detector common
  unfolding top-quark decays useful
  matrix element method for hypothesis test

$\Rightarrow$ systematic approach through generative network



INNverting event simulations

Inverting LHC simulations

– unfolding QCD-shower to hard parton standard [jet algorithm]
  unfolding detector common
  unfolding top-quark decays useful
  matrix element method for hypothesis test

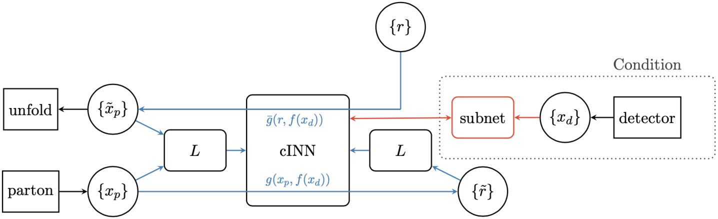
$\Rightarrow$ systematic approach through generative network

Conditional INN

– standard INN/cINN setup: parton-level events from $\{r\}$

– maximum likelihood loss

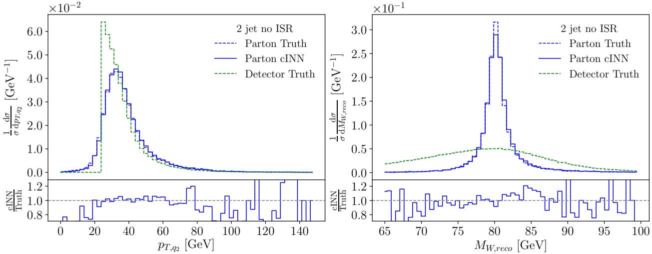
$$L = -\left\langle \log p(\theta | x_p, x_d) \right\rangle_{x_p, x_d}$$
$$= -\left\langle \log p(g(x_p, x_d)) + \log \left| \frac{\partial g(x_p, x_d)}{\partial x_p} \right| \right\rangle_{x_p, x_d} - \log p(\theta) + \text{const.}$$



INNverting event simulations

Hard process $q\bar{q} \to ZW \to (\ell\ell)\ (jj)$

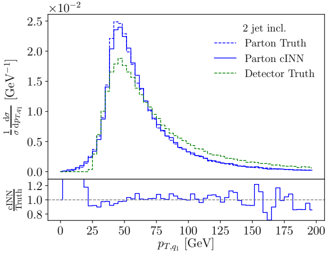
- stochastic inverse problem
  model assumption like in forward direction
- invert detector effects

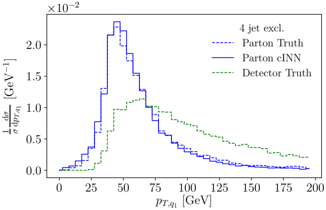


INNverting event simulations

Hard process $q\bar{q} \to ZW \to (\ell\ell)\ (jj)$

- stochastic inverse problem
  model assumption like in forward direction
- invert detector effects
- invert QCD jet radiation [matrix element vs parton shower]

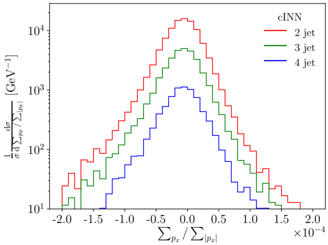




INNverting event simulations

Hard process $q\bar{q} \to ZW \to (\ell\ell)\ (jj)$

- stochastic inverse problem
  model assumption like in forward direction
- invert detector effects
- invert QCD jet radiation [matrix element vs parton shower]
- precision from momentum conservation



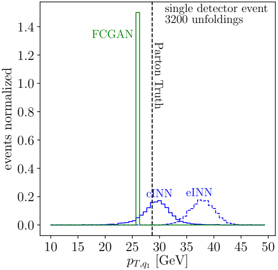
INNverting event simulations

Hard process $q\bar{q} \to ZW \to (\ell\ell)\ (jj)$

- stochastic inverse problem
  model assumption like in forward direction
- invert detector effects
- invert QCD jet radiation [matrix element vs parton shower]
- precision from momentum conservation

Proper statistical inversion

- distribution
  single detector event
  3200 unfoldings to partonic phase space



INNverting event simulations

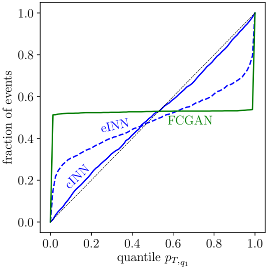
Hard process $q\bar{q} \to ZW \to (\ell\ell)\,(jj)$

- stochastic inverse problem
  model assumption like in forward direction
- invert detector effects
- invert QCD jet radiation [matrix element vs parton shower]
- precision from momentum conservation

Proper statistical inversion

- distribution
  single detector event
  3200 unfoldings to partonic phase space
- calibration
  1500 detector-parton event pairs
  60 unfoldings per pair, in quantiles
  truth within given quantile for fraction of pairs

$\Rightarrow$ Probability distribution in parton phase space!

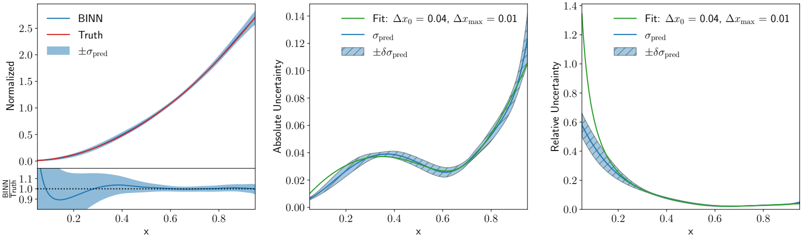


One number not a prediction

Bayesian INN

- generate events with error bars
  learn density and uncertainty maps over phase space
- 2D toy models: wedge ramp, kicker ramp, Gaussian ring

⇒ Uncertainty estimate works...



...and we see how the network learns!

One number not a prediction

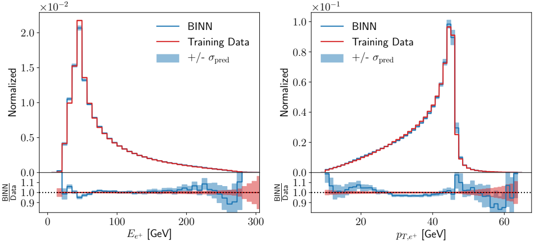
Bayesian INN

- generate events with error bars
  learn density and uncertainty maps over phase space
- 2D toy models: wedge ramp, kicker ramp, Gaussian ring

⇒ Uncertainty estimate works...

LHC toy process: $pp \to e^+ e^-$

- kinematic distributions with errors



⇒ Crucial step for LHC-simulations

Machine learning for LHC in Heidelberg

Thank you to all collaborators here and elsewhere!

Lynton Ardizzone, Mathias Backes, Pierre Baldi, Marco Bellagente, Lukas Blecher, Sebastian Bieringer, Johann Brehmer, Anja Butter, Sascha Diefenbacher, Barry Dillon, Manuel Haußmann, Theo Heimel, Jessica Howard, Sander Hummerich, Gregor Kasieczka, Fabian Keilbach, Ulli Köthe, Tobias Krebs, Michel Luchmann, Ben Nachman, Stefan Radev, Armand Rousselot, Michael Russell, Christof Sauer, Torben Schell, Peter Sorrensen, Natalie Soybelman, Jennifer Thompson, Sophia Vent, Lorenz Vogel, Daniel Whiteson, Ramon Winterhalder