Measuring QCD Splittings with Invertible Neural Networks

Theo Heimel

February 2021

Recently published:
S. Bieringer, A. Butter, TH, S. Höche, U. Köthe, T. Plehn, S. Radev
Measuring QCD Splittings with Invertible Networks
arXiv:2012.09873

Machine learning at the LHC

- Vast amount of LHC data
  & well-understood Monte Carlo generators
  → Machine learning is valuable tool in LHC physics
- **Classification**
  → Tagging with sub-jet data: jet-images, four-vectors
- **Anomaly detection**
  → From model-driven approach to data-driven approach
- **Simulations**
  → Accelerating and substituting Monte Carlo generators
  → Phase space sampling, detector effects, unweighting, ...
- **Precision measurements**
  → Estimating uncertainties
  → Use high-dimensional data

Flow networks at the LHC

- Many applications for normalizing flows in LHC physics
- Phase space generation

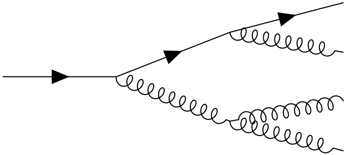
  [Bothmann, Janssen, Knobbe, Schmale, Schumann, 2001.05478]

  [Gao, Isaacson, Krause, 2001.05486]

  [Gao, Höche, Isaacson, Krause, Schulz, 2001.10028]

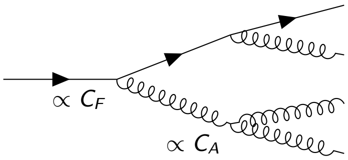
  [Chen, Klimek, Perelstein, 2009.07819]
- Event generation [Verheyen, Stienen, 2011.13445]
- Anomaly detection [Nachman, Shih, 2001.04990]
- Density estimation [Brehmer, Cranmer, 2003.13913]
- Parton shower unfolding [Bellagente et al., 2006.06685]
- **Our project:**
  INNs for precision measurements of QCD splittings

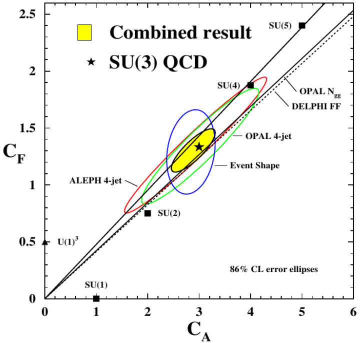
Parton showers



- Parton showers are part of every LHC analysis
- Described by *qqg* and *ggg* IA in soft-collinear limit
  $\rightarrow$ Leading order: simplify to a set of splitting kernels
  $\rightarrow$ Corrections are active field of research
  [Hartgring, Laenen, Skands, 1303.4974] [Li, Skands, 1611.00013]
  [Höche, Krauss, Prestel, 1705.00982] [Dulat, Höche, Prestel, 1805.03757]
  [Dasgupta, Dreyer, Hamilton, Monni, Salam, 1805.09327]
  [Dasgupta, Dreyer, Hamilton, Monni, Salam, Soyez, 2002.11114]
- No need to understand parton densities
  $\rightarrow$ **Great way to study fundamental QCD properties**

LEP results for QCD casimirs





- Splitting probabilities depend on QCD casimirs
- Combined LEP measurement [Kluth, hep-ex/0309070]

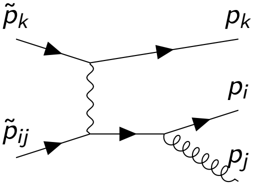
$$C_A = 2.89 \pm 0.21,\ \ C_F = 1.30 \pm 0.09$$

- **Can we measure beyond log-enhanced terms?**
- **Can we use low-level sub-jet observables?**

Parameterizing the splitting functions

- Parton showers in leading collinear approximation

$$\overline{|\mathcal{M}_{n+1}|^2} \simeq \frac{2g_s^2}{\tilde{p}_{ij}^2}\, \hat{P}(z,y)\, \overline{|\mathcal{M}_n|^2}$$



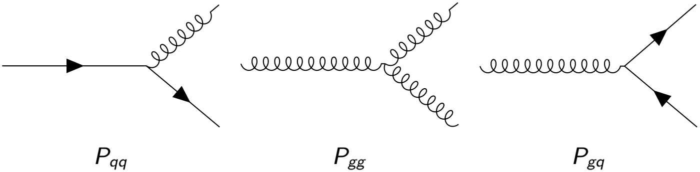
z: energy fraction
y: momentum transfer
$yz(1-z) \propto p_T$

- Splitting function for gluon radiation:

$$P_{qq}(z,y) = C_F\left[D_{qq}\frac{2z(1-y)}{1-z(1-y)} + F_{qq}(1-z) + C_{qq}yz(1-z)\right]$$

leading term   finite term   rest term

Parameterizing the splitting functions



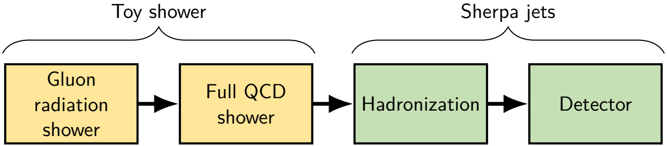
$$P_{qq}(z,y) = C_F\left[D_{qq}\frac{2z(1-y)}{1-z(1-y)} + F_{qq}(1-z) + C_{qq}yz(1-z)\right]$$

$$P_{gg}(z,y) = 2C_A\Bigg[D_{gg}\left(\frac{z(1-y)}{1-z(1-y)} + \frac{(1-z)(1-y)}{1-(1-z)(1-y)}\right) + F_{gg}z(1-z) + C_{gg}yz(1-z)\Bigg]$$

$$P_{gq}(z,y) = T_R\left[F_{qq}\left(z^2+(1-z)^2\right) + C_{gq}yz(1-z)\right]$$

Simulation setup

- Benchmark scenario: $e^+e^- \to Z \to q\bar{q}$
  $\to$ comparison with LEP measurements
- Drawback: small maximum p_T of $m_Z/2$
  $\to$ less splitting information than at the LHC
- Start on parton level, then more realistic simulations

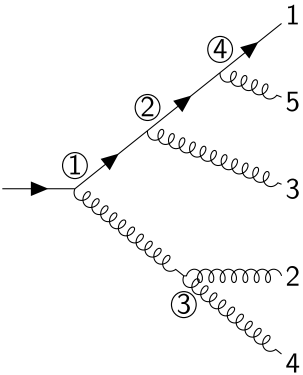


Toy shower generator

- Shower only off one hard quark
  → no additional jet clustering step needed
- 1 GeV lower p_T cutoff
- Network sensitive to order of jet constituents
  → need meaningful sorting scheme
- Implementation-specific "truth sorting"
  → information backdoor
- Approximate it with k_T sorting
  → **reconstruct splitting order with k_T algorithm**

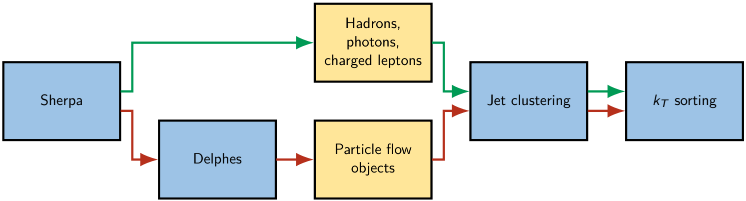
k_T sorting algorithm

- k_T algorithm:
  Combine sub-jets with minimal
  $d_{ij} = \Delta R_{ij}^2 \min(p_{T,i}^2, p_{T,j}^2)$
  [Ellis, Soper, 9305266]
- Start with first splitting
- If not visited yet:
  Visit hardest sub-jet
  Then visit second sub-jet
- Follow hardest sub-jets
- Repeat for second splitting, third splitting, ...



Sherpa simulation setup

- Modified Sherpa with our splitting parameterization
- Always generate full shower (no gluon-radiation shower)
- Ignore pile-up and ISR
- Simulate hadronization and optionally detector effects

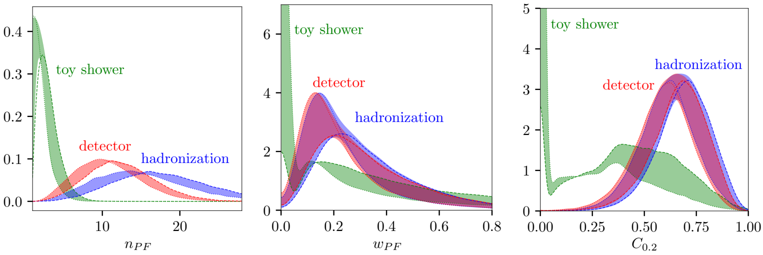


High-level jet observables

Number of constituents [Frye et al., 1704.06266]	n_{PF}
Girth of radiation distribution [Gallicchio et al., 1010.3698]	$w_{\mathrm{PF}} = \frac{\sum_i p_{T,i} \Delta R_{i,\mathrm{jet}}}{\sum_i p_{T,i}}$
Effect of soft constituents [CMS]	$p_T D = \frac{\sqrt{\sum_i p_{T,i}^2}}{\sum_i p_{T,i}}$
Two-point energy correlator [Larkoski et al., 1305.0007]	$C_{0.2} = \frac{\sum_{ij} E_{T,i} E_{T,j} (\Delta R_{ij})^{0.2}}{\sum_i E_{T,i}^2}$
Largest fraction of p_T of a single constituent [Pumplin]	x_{max}
Lowest number of constituents with 95% of the total jet p_T [Pumplin]	N_{95}

Toy shower vs. hadronization and detector

- Vary D_{qq} from 0.5 (dotted line) to 2 (dashed line)



- **Large change from partons to hadronization level**
- Small change from hadronization to detector level

How to measure the splitting parameters?

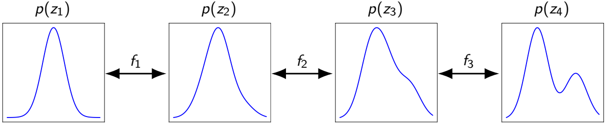
- Classical approach: Fit HLOs
- Advantage of machine learning-based method:
  $\rightarrow$ works for high-dimensional data
  $\rightarrow$ **use low-level observables (four-momenta)**
- Learning the uncertainties
  $\rightarrow$ normalizing flows to sample from posterior distribution

$$P(\text{ parameters } | \text{ measurements })$$

- Need to **condition flow network on measurement data**

Normalizing flows

- Invertible mapping between probability distributions
  [Kobyzev, Prince, Brubaker, 1908.09257]



- Transformation with change of variables formula:

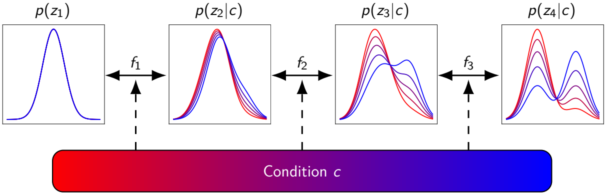
$$p(z_n) = p(z_1) \prod_{i=1}^{N} \left| \det \frac{\partial f_i}{\partial z_i} \right|^{-1}$$

- Add parameter to condition transformation:

$$z_{i+1} = f_i(z_i; c) \qquad \text{and} \qquad z_i = \bar{f}_i(z_{i+1}; c)$$

Normalizing flows

- Invertible mapping between probability distributions
  [Kobyzev, Prince, Brubaker, 1908.09257]



- Transformation with change of variables formula:

$$p(z_n) = p(z_1) \prod_{i=1}^{N} \left| \det \frac{\partial f_i}{\partial z_i} \right|^{-1}$$

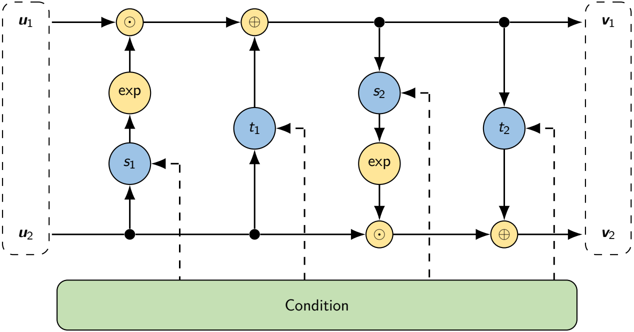
- Add parameter to condition transformation:

$$z_{i+1} = f_i(z_i; c) \qquad \text{and} \qquad z_i = \bar{f}_i(z_{i+1}; c)$$

Invertible neural networks

- Invertible coupling block

  [Ardizzone, Lüth, Kruse, Rother, Köthe, 1907.02392]

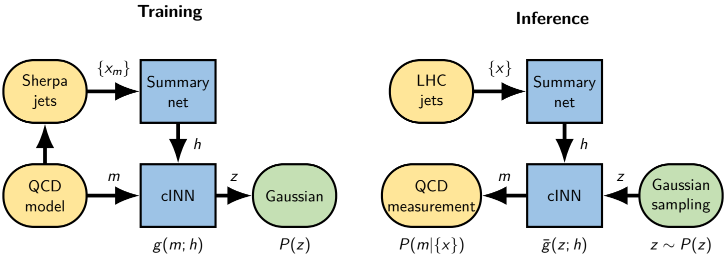


  s_1, t_1, s_2, t_2: fully-connected sub-networks
  (u_1, u_2): input vector split in two
  (v_1, v_2): output vector split in two

- cINN: chain of coupling blocks with permutations and bijective clamping

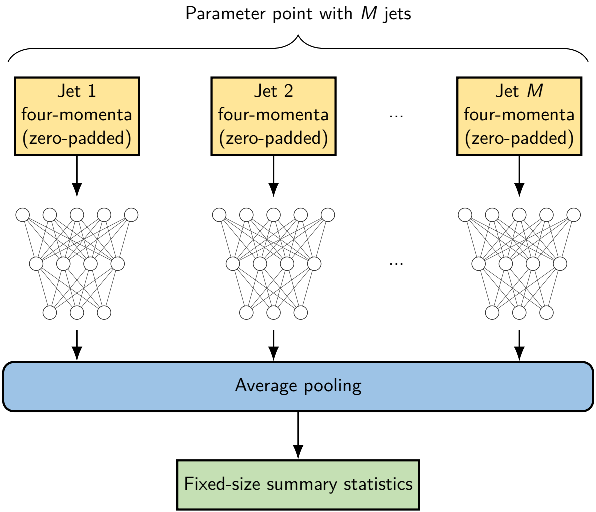
Full measurement setup

- Combine cINNs and summary network to extract posterior distributions
  [Radev, Mertens, Voss, Ardizzone, Köthe, 1907.02392]



- Varying numbers of jets M per parameter point
  $\rightarrow \sqrt{M/M_{\text{max}}}$ as additional condition for cINN
- Caveat: maximal training M also limits inference M
  $\rightarrow$ can't combine probability distributions
  $\rightarrow$ not efficient to analyze millions of jets

Summary network



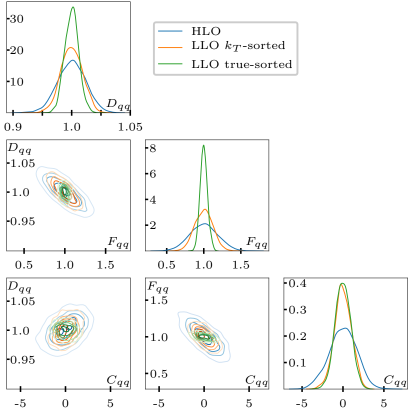
Gluon radiation shower measurements

Testing the performance:

1. Leading-order SM QCD:

$$D_{qq} = F_{qq} = 1, \; C_{qq} = 0$$

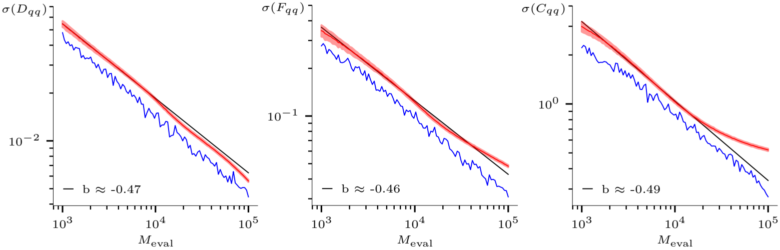
2. Generate 10000 jets
3. Estimate posterior



- Hierarchical structure of $\{D_{qq}, F_{qq}, C_{qq}\}$
- Best results for truth sorting: Information backdoor
- k_T-sorted **LLO better than HLO**

Scaling of measurement errors

- Vary the number of jets M during training
  $\rightarrow$ posterior width is function of M_{eval}
- Each M_{eval}: Test with 200 sets of M_{eval} SM-like jets

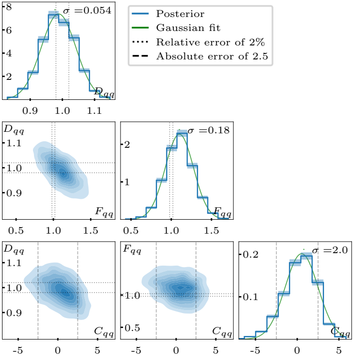


Red: Estimated measurement error
Blue: True measurement error
Black: fit to function $\sigma = a \cdot M^b$

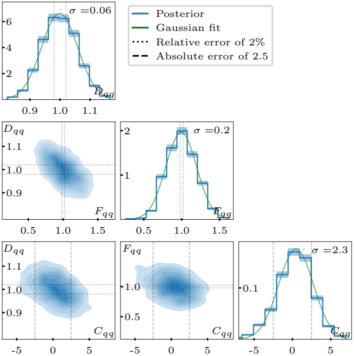
- Consistent error estimation
- **Extracts correct $1/\sqrt{M}$ scaling**

Measuring the gluon radiation parameters

- Measuring $\{D_{qq}, F_{qq}, C_{qq}\}$ (full shower)



Sherpa shower: hadronization

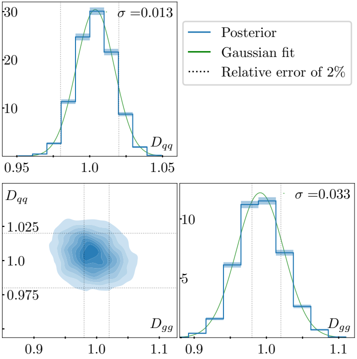


Hadronization + detector

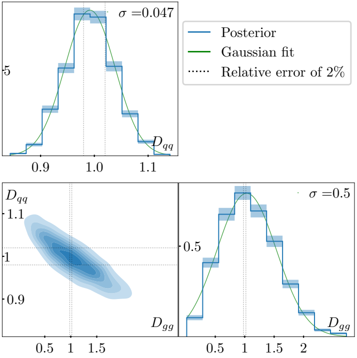
- **Hadronization has large impact on performance**
- Detector effects small
  $\rightarrow$ next slides: compare parton level $\leftrightarrow$ detector level

Measuring the leading soft-collinear terms

- Measuring $\{D_{qq}, D_{gg}\} \leftrightarrow$ Casimirs $\{C_F, C_A\}$



Toy shower

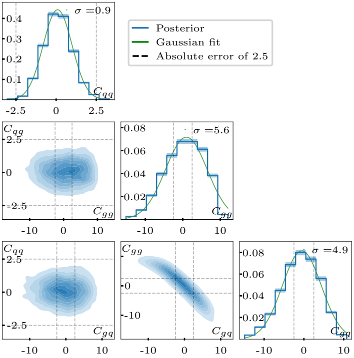


Hadronization + detector

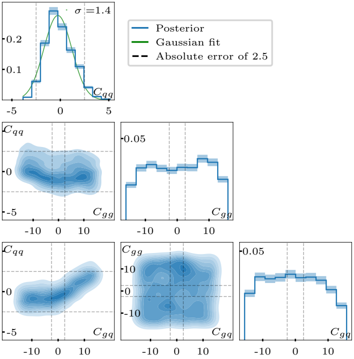
- $\sim 5\%$ error for D_{qq} (for only 10000 jets!)
  $\rightarrow$ **comparable with LEP result**
- Few gluon splittings in quark-initiated jets
  $\rightarrow$ D_{gg} performance breaks down after hadronization

Measuring the rest terms

- Measuring $\{C_{qq}, C_{gg}, C_{gq}\}$



Toy shower



Hadronization + detector

- **Measuring C_{qq} at the LHC within reach**
- Strong correlation of C_{gg} and C_{gq}
  $\rightarrow$ invisible after hadronization
  $\rightarrow$ need gluon-initiated jet

Outlook

- Parton showers everywhere at the LHC
  → need systematic way to understand fundamental QCD
  → parameterization of log-leading, finite and rest terms of splitting functions
  → ML-based method to measure them
- Our approach is promising first step
- Next steps:
  → repeat with gluon-initiated jets
  → use harder jets
  → use real experimental data
- Interpreting the learned summary statistics
- How can we work with higher numbers of jets?