High precision Monte Carlo Generators at Past, Present and Future Colliders

Andrzej Siódmok

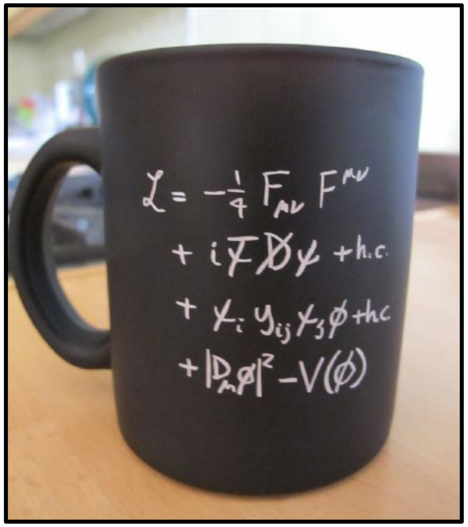
NCN, Poland Grant No. 2019/34/E/ST2/00457

Standard Model

There is a **huge gap** between a one-line formula of a fundamental theory, like the Lagrangian of the SM, and the experimental reality that it implies

Theory

Standard Model Lagrangian



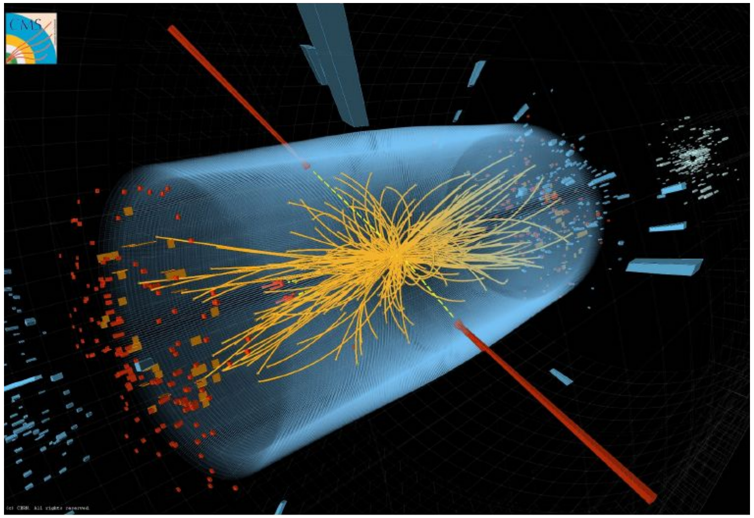
Data makes you smarter

It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong.

Richard P. Feynman

Experiment

LHC event

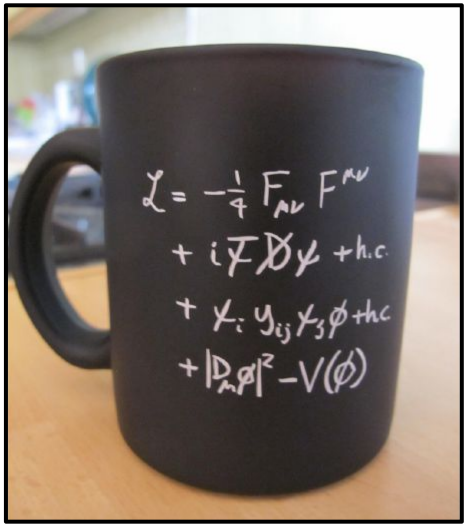


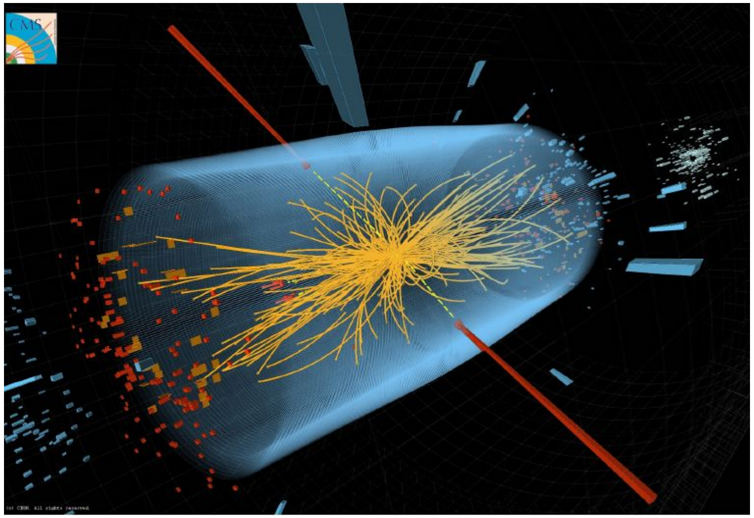
Standard Model

There is a **huge gap** between a one-line formula of a fundamental theory, like the Lagrangian of the SM, and the experimental reality that it implies

Theory
Standard Model Lagrangian

Experiment
LHC event





- MC event generators are designed to bridge the that **gap**
- “Virtual collider” ⇒ Direct comparison with data

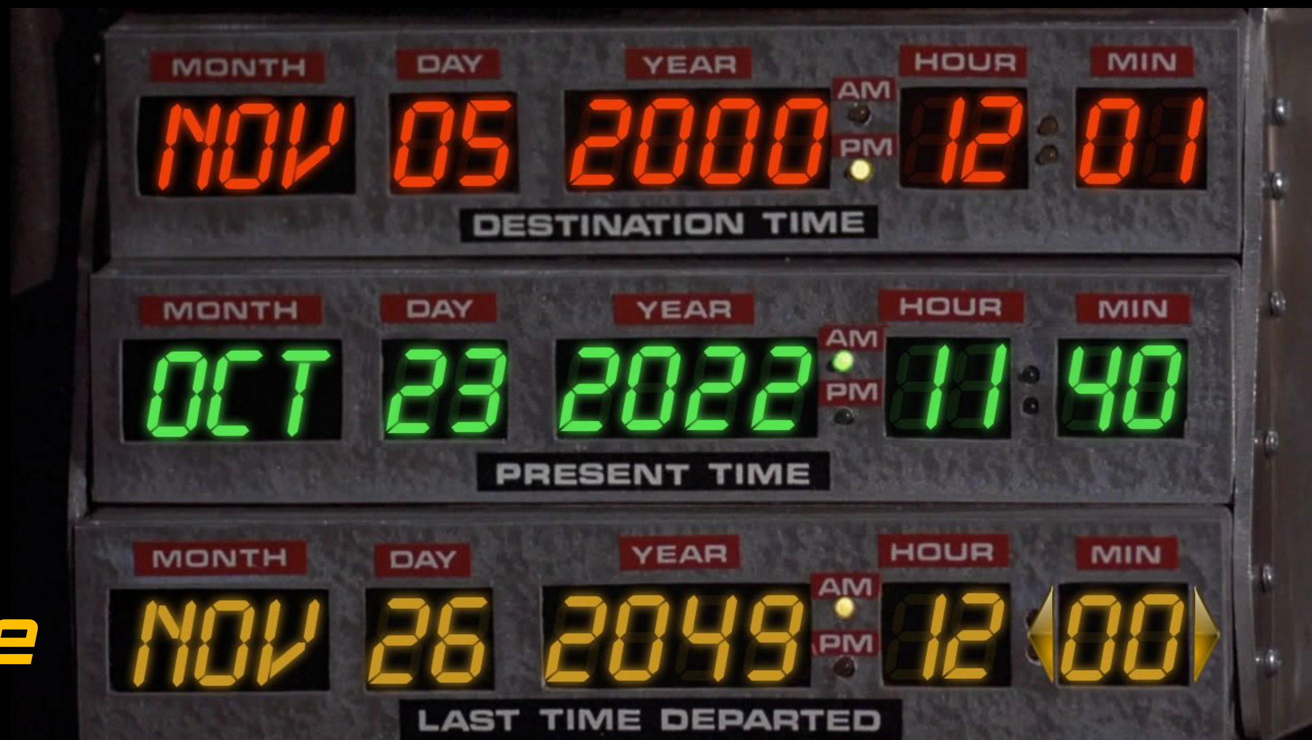
⇩

Almost all **HEP measurements and discoveries** in the modern era have **relied on MCEG**, most notably the discovery of the Higgs boson.

Published papers by ATLAS, CMS, LHCb: **2252**
Citing at least 1 of 3 existing MCEG: **1888** (**84%**)

Outline
Generators

1. LEP
2. LHC
3. Future-ee

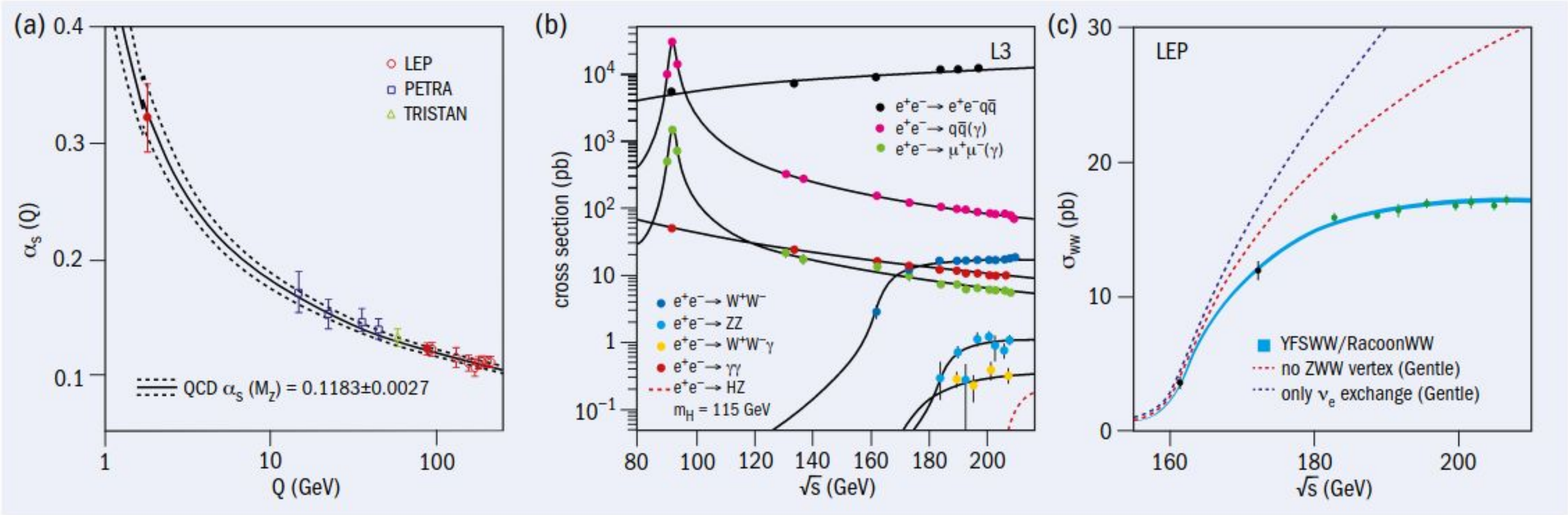


1. LEP

LEP's electroweak leap, CERNCOURIER

CERNCOURIER.COM

FEATURE LEP'S PHYSICS LEGACY



1. LEP legacy MC

Precise process-oriented MC
[main focus on QED/EW]

- BHLUMI (low angle Bhabha), BHWIDE:

$$e^+e^- \to e^+e^-(n\gamma)$$

- KKMC:

$$e^+e^- \to f\bar{f}(n\gamma),\ f = \mu, \tau, q, \tau \to X$$

- KORALW, YFSWW, RacoonWW

$$e^+e^- \to W^+W^-(n\gamma) \to 4f(n\gamma)$$

$\vdots$

General purpose MC
[main focus on QCD]

- Herwig 6
- PYTHIA 6/JETSET
  Ariadne

Specialized programs
- PHOTOS - *universal Monte Carlo for QED radiative corrections*
- TAUOLA - *tau decay library*

1. LEP legacy MC

Example:
Precise process-oriented MC
[S. Jadach, W. Placzek, M. Skrzypek, B.F.L. Ward, Z. Was]

$$e^+e^- \to W^+W^-(n\gamma) \to 4f(n\gamma)$$

KORALW, YFSWW

The YFS formalism provides a robust method for resumming the emission of real and virtual photons in the soft limit to all orders. This resummation can be further improved by including exact fixed-order expression in a systematic way.

YFSWW
Simplified Process
(Double-Resonant W)
$\bigoplus$
As Much Rad. Corr. As Possible (Needed)

KORALW
Full Process
(All 4f Channels)
$\bigoplus$
Simplified Rad. Corr. (ISR, Coulomb, ...)

δ^{NL}_{WW}
- $\mathcal{O}(\alpha)$ NL EW Corr.
- "Screened" Coul. Corr. (Approximation For Non-Factorizable Corr.)

WW-Process
- YFS $\mathcal{O}(\alpha^3)$LL ISR
- Coulomb Correction
- "Naive" QCD Corr.
- Full CKM Matrix
- W-BR's Incl. Rad Corr.
- Anomalous TGC's
- FSR by PHOTOS
- τ Decays by TAUOLA
- Hadronization by JETSET
- Semi-An. Code: KORWAN

δ_{4f}
- Non-WW 4f Contrib.
- YFS $\mathcal{O}(\alpha^3)$LL ISR

1. LEP legacy MC

Example:
Precise process-oriented MC

$$e^+e^- \to W^+W^-(n\gamma) \to 4f(n\gamma)$$

RaccoonWW

RacoonWW
a Monte Carlo generator for
4-fermion production at e+e- colliders

Authors

Ansgar Denner	*Universität Würzburg, Germany*
Stefan Dittmaier	*Universität Freiburg, Germany*
Markus Roth	
Doreen Wackeroth	*SUNY at Buffalo, USA*

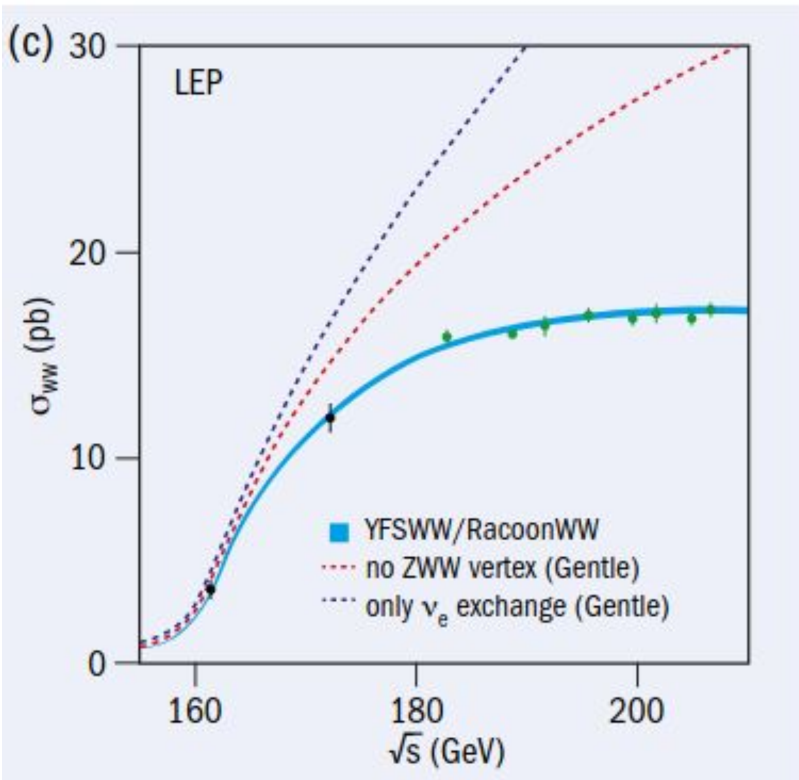
RacoonWW provides

- lowest-order predictions for
  - all processes e+e- -> 4 fermions
  - all processes e+e- -> 4 fermions + photon
- radiative corrections to e+e- -> WW -> 4f including
  - the full electroweak O(alpha) corrections in double-pole approximation
  - higher-order leading-logarithmic initial-state radiation
  - complete off-shell Coulomb singularity
- an Improved Born Approximation (IBA) with leading universal corrections

1. LEP legacy MC

Example

Process-oriented precision MC



$$e^+e^- \to W^+W^-(n\gamma) \to 4f(n\gamma)$$

YFSWW:
[Jadach, S., Płaczek, W., Skrzypek, M., Ward, B., & Wąs, Z., CPC 2001, 140(3)]

RacoonWW:
[Denner, A., Dittmaier, S., Roth, M., & Wackeroth, D. CPC 2003, 153(3)]

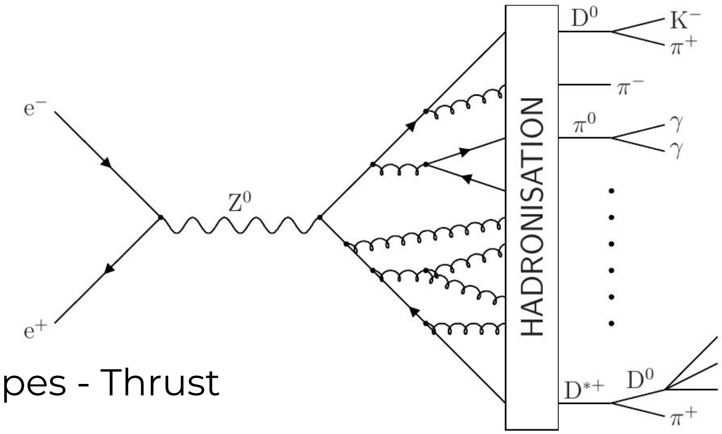
- 0.3% difference due to different treatment of QED: YFS vs Collinear Resummation
- important to have at least 2 independent MC!

The only tools capable to calculate QED+EW Standard Model predictions for the total cross section and distributions of the e+e− → W+W− process. They were also used to extract (fit) the mass of the W boson from data.

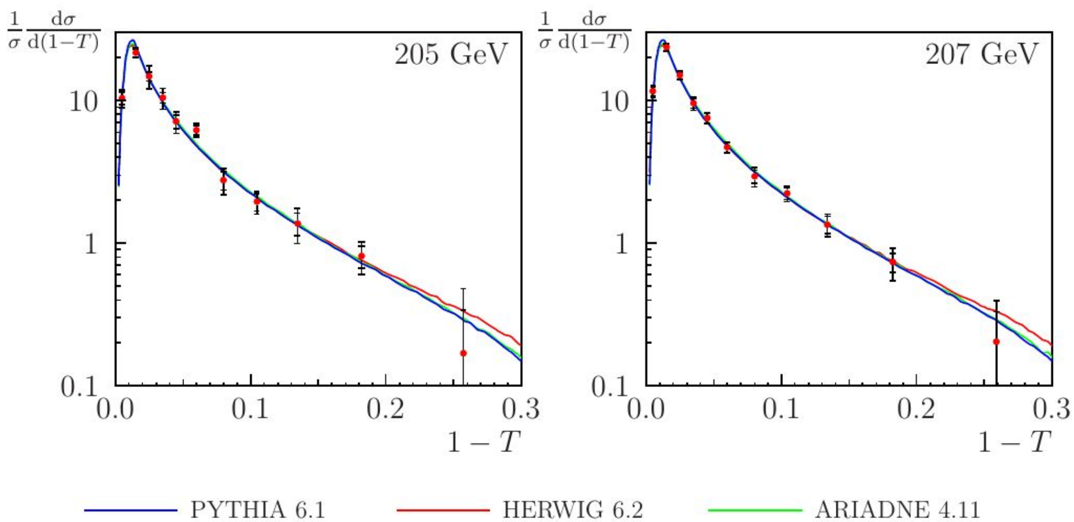
1. LEP legacy MC

Example:
General purpose MC

" The Monte Carlo programs used in our analysis to simulate multihadronic events are $\mathcal{KK}$2f 4.01/4.13 [46], PYTHIA 6.125 [47], HERWIG 6.2 [48] and ARIADNE 4.11 [49]. "

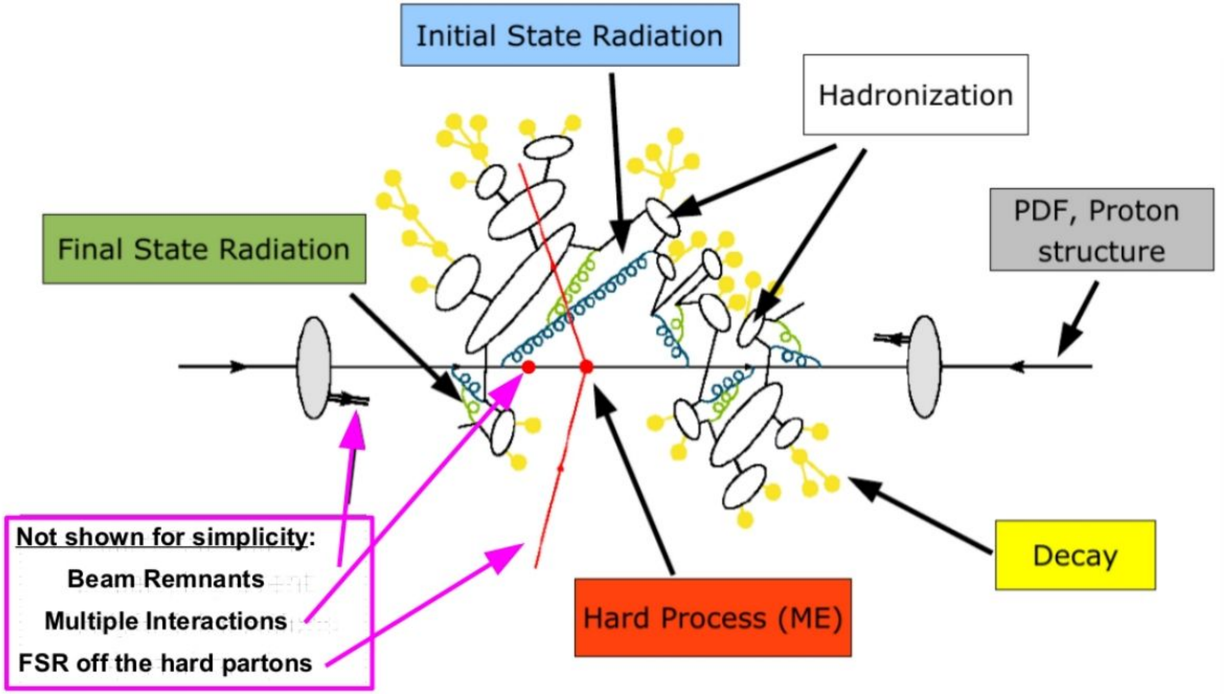


Hadronic Event Shapes - Thrust



LHC: QCD machine

Main progress in General purpose MC



taken from Stefan Gieseke©

The general approach is the same in different programs but the models and approximations used are different.

2. LHC MC

General purpose MC the Workhorses of the LHC:

Herwig 7: Successor to EARWIG (begun in 1984). Originated in coherence studies: angular ordering parton shower. Cluster model. Last version: **Herwig 7.2** [Bellm, Bewick, Ravasio, Gieseke, Grellscheid, Kirchgaesser, Masouminia. Nail, Papaefstathiou, Platzer, Rauch, Reuschle, Richardson, Seymour, Siodmok, Webster, *Eur.Phys.J.C* 80 (2020)]

PYTHIA 8: Successor to JETSET (begun in 1978). Originated in hadronization studies: Lund String. Last version: **PYTHIA 8.3** ["A comprehensive guide to the physics and usage of PYTHIA 8.3" 315 pages! Bierlich, C., Chakraborty, S., Desai, N., Gellersen, L., Helenius, I., Ilten, P., Lönnblad, L., Mrenna, S., Prestel, S., Preuss, C. T., Sjöstrand, T., Skands, P., Utheim, M., & Verheyen, R.]

Sherpa 2: Begun in 2000. Originated in "matching" of matrix elements to showers: CKKW. Last version **Sherpa 2.2**: [Bothmann, E., Chahal, G. S., Höche, S., Krause, J., Krauss, F., Kuttimalai, S., Liebschner, S., Napoletano, D., Schönherr, M., Schulz, H., Schumann, S., & Siegert, F. (2019) SciPostPhys.7.3.034]

2. LHC MC

General purpose MC the Workhorses of the LHC:

How big was the progress?

Herwig 7: Successor to EARWIG (begun in 1984). Originated in coherence studies: angular ordering parton shower. Cluster model. Last version: **Herwig 7.2** [Bellm, Bewick, Ravasio, Gieseke, Grellscheid, Kirchgaesser, Masouminia. Nail, Papaefstathiou, Platzer, Rauch, Reuschle, Richardson, Seymour, Siodmok, Webster, *Eur.Phys.J.C* 80 (2020)]

PYTHIA 8: Successor to JETSET (begun in 1978). Originated in hadronization studies: Lund String. Last version: **PYTHIA 8.3** ["A comprehensive guide to the physics and usage of PYTHIA 8.3" 315 pages! Bierlich, C., Chakraborty, S., Desai, N., Gellersen, L., Helenius, I., Ilten, P., Lönnblad, L., Mrenna, S., Prestel, S., Preuss, C. T., Sjöstrand, T., Skands, P., Utheim, M., & Verheyen, R.]

Sherpa 2: Begun in 2000. Originated in "matching" of matrix elements to showers: CKKW. Last version **Sherpa 2.2**: [Bothmann, E., Chahal, G. S., Höche, S., Krause, J., Krauss, F., Kuttimalai, S., Liebschner, S., Napoletano, D., Schönherr, M., Schulz, H., Schumann, S., & Siegert, F. (2019) SciPostPhys.7.3.034]

2. LHC MC

General purpose MC the Workhorses of the LHC:

How big was the progress?

C++

Herwig 7: Successor to EARWIG (begun in 1984). Originated in coherence studies: angular ordering parton shower. Cluster model. Last version: **Herwig 7.2** [Bellm, Bewick, Ravasio, Gieseke, Grellscheid, Kirchgaesser, Masouminia. Nail, Papaefstathiou, Platzer, Rauch, Reuschle, Richardson, Seymour, Siodmok, Webster, *Eur.Phys.J.C* 80 (2020)]

PYTHIA 8: Successor to JETSET (begun in 1978). Originated in hadronization studies: Lund String. Last version: **PYTHIA 8.3** ["A comprehensive guide to the physics and usage of PYTHIA 8.3" 315 pages! Bierlich, C., Chakraborty, S., Desai, N., Gellersen, L., Helenius, I., Ilten, P., Lönnblad, L., Mrenna, S., Prestel, S., Preuss, C. T., Sjöstrand, T., Skands, P., Utheim, M., & Verheyen, R.]

Sherpa 2: Begun in 2000. Originated in "matching" of matrix elements to showers: CKKW. Last version **Sherpa 2.2**: [Bothmann, E., Chahal, G. S., Höche, S., Krause, J., Krauss, F., Kuttimalai, S., Liebschner, S., Napoletano, D., Schönherr, M., Schulz, H., Schumann, S., & Siegert, F. (2019) SciPostPhys.7.3.034]

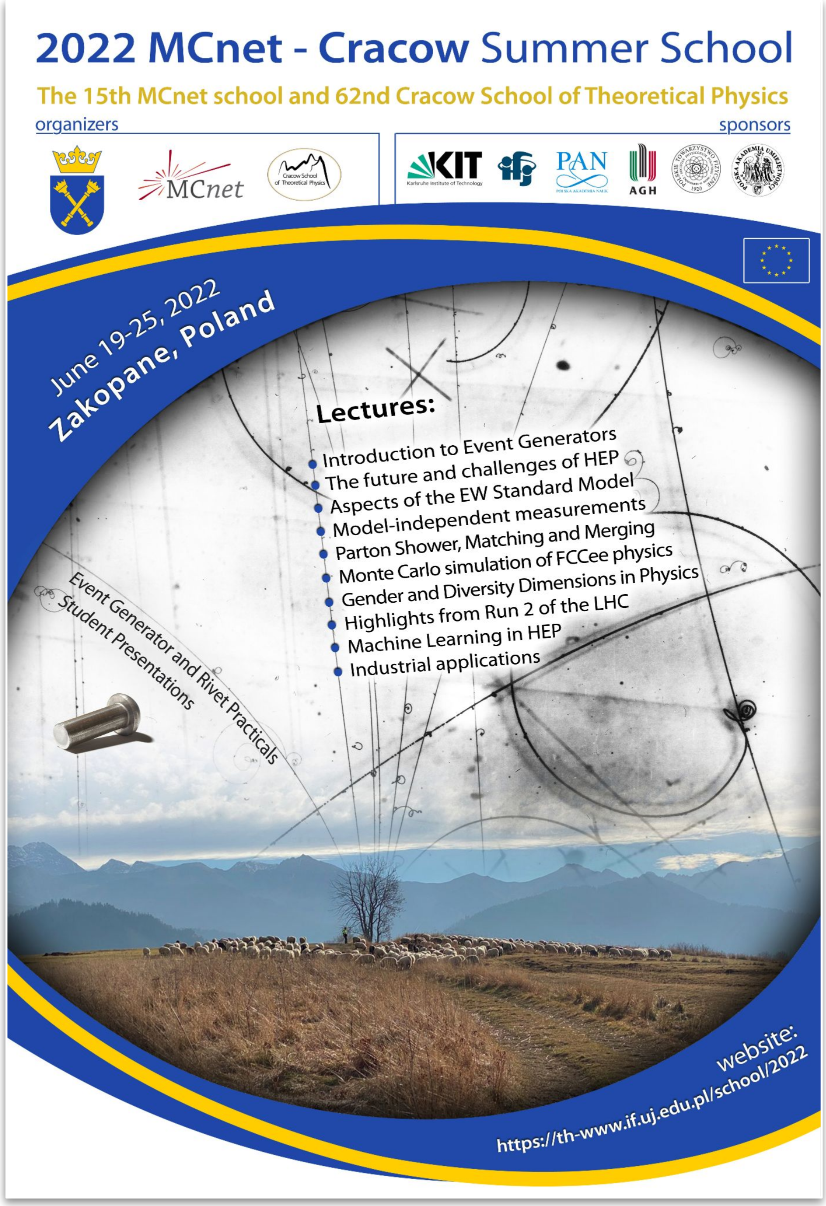
2. LHC MC

Current release series	Hard matrix elements	Shower algorithms	NLO Matching	Multijet merging	MPI	Hadronization	Shower variations
Herwig 7	Internal, libraries, event files	QTilde, Dipoles	Internally automated	Internally automated	Eikonal	Clusters, (Strings)	Yes
Pythia 8	Internal, event files	Pt ordered, DIRE, VINCIA	External	Internal, ME via event files	Interleaved	Strings	Yes
Sherpa 2	Internal, libraries	CSShower, DIRE	Internally automated	Internally automated	Eikonal	Clusters, Strings	Yes

[Table from S. Platzer]

2. LHC MC


MONTH
DAY
YEAR
AM
PM
HOUR
MIN
OCT 23 2022 11 40
PRESENT TIME


2022 MCnet - Cracow Summer School
The 15th MCnet school and 62nd Cracow School of Theoretical Physics
organizers
sponsors
MCnet
KIT
PAN
AGH
June 19-25, 2022
Zakopane, Poland
Lectures:
Introduction to Event Generators
The future and challenges of HEP
Aspects of the EW Standard Model
Model-independent measurements
Parton Shower, Matching and Merging
Monte Carlo simulation of FCCee physics
Gender and Diversity Dimensions in Physics
Highlights from Run 2 of the LHC
Machine Learning in HEP
Industrial applications
Event Generator and Rivet Practicals
Student Presentations
website:
https://th-www.if.uj.edu.pl/school/2022

SIMPLIFICATIONS in perturbative QCD techniques

NEW IDEAS of treatment non-perturbative QCD

NARODOWE CENTRUM NAUKI
NCN: 2019/34/E/ST2/00457

QCD ex-Machina

QCD ex-Machina - building blocks

1. Development a **novel technique KrkNLO**, designed specifically to reduce the complexity **of NLO matching to Parton Showers**;

   KrkNLO method: proof of concept for Z and H boson production
   [Jadach, Nail, Placzek, Sapeta, Siodmok, Skrzypek **Eur.Phys.J**. C77 (2017) no.3, 164, **Eur.Phys.J.** C76 (2016) no.12, 649, **JHEP** 1510 (2015) 052]

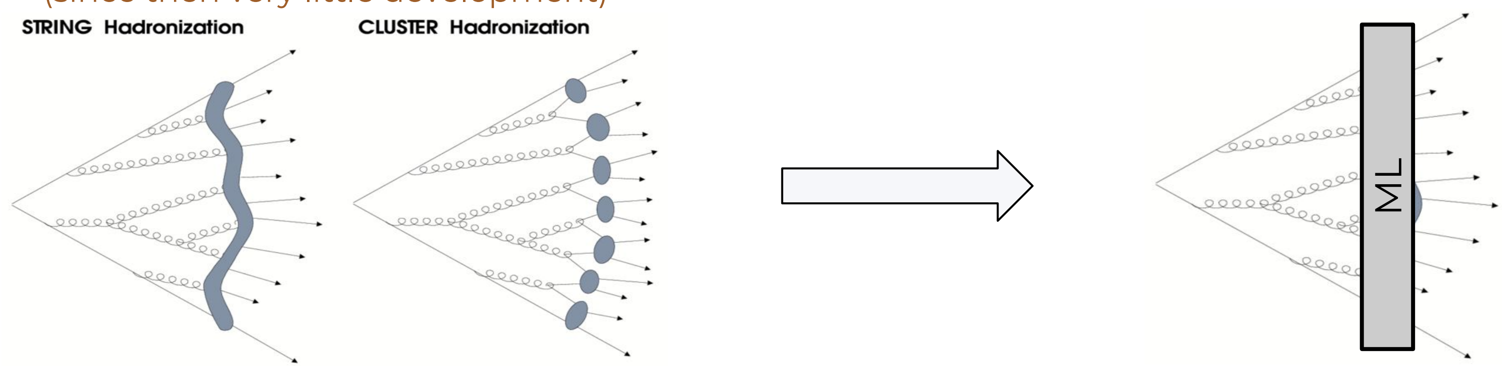
2. **A new method** of merging **Electroweak** corrections with **QCD**;

   Simplicity of KrkNLO method for inclusion of precise EW calculation from KKMC.

3. **Pioneering ideas of using Machine Learning** (ML) to improve non-perturbative hadronization.

   Early 1980's
   (since then very little development)

   Early 2020's
   (lot of progress in ML)



4. QCD ex-machina will be implemented in a GPMC Herwig, and will **significantly increase the precision** of the **generator**.

QCD ex-Machina - KrkNLO

1. Development a **novel technique KrkNLO**, designed specifically to reduce the complexity **of NLO matching to Parton Showers**;

   KrkNLO method: proof of concept for Z and H boson production
   [Jadach, Nail, Placzek, Sapeta, Siodmok, Skrzypek **Eur.Phys.J**. C77 (2017) no.3, 164, **Eur.Phys.J.** C76 (2016) no.12, 649, **JHEP** 1510 (2015) 052]

Based on a new MC factorization scheme which is tailored for GPMC

↓ More processes (WW, gamma gamma, ...) automatization

↓ Universality

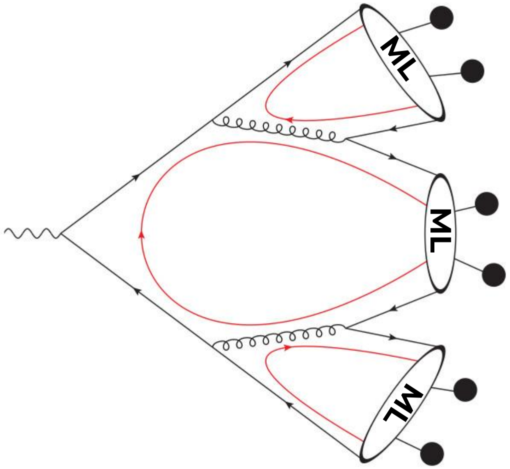
↓ Higher precision

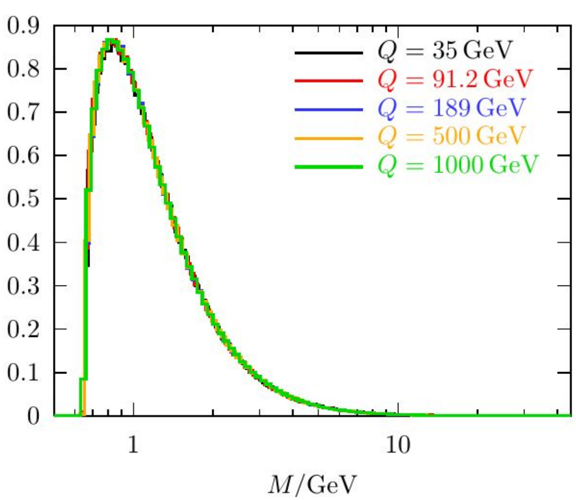
↓ MC@NLO+KrkNLO [Nason, Salam *JHEP* 01 (2022)]

Towards a Deep Learning Model for Hadronization

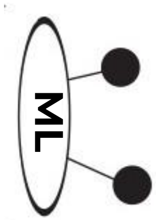
The philosophy of the model: use information from perturbative QCD as an input for hadronization.

QCD **pre-confinement** discovered by Amati & Veneziano:





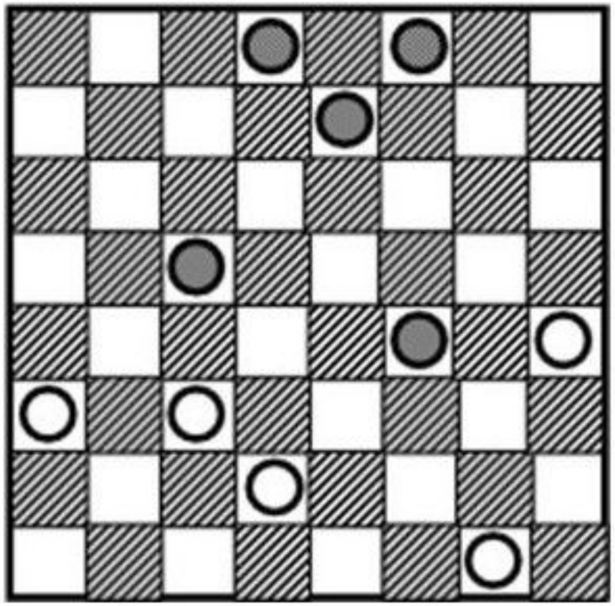
- QCD provide pre-confinement of colour
- Colour-singlet pair end up close in phase space and form highly excited hadronic states, the clusters
- Pre-confinement states that the spectra of clusters are independent of the hard process and energy of the collision
- Peaked at low mass (1-10 GeV) typically decay into 2 hadrons
- **Towards a Deep Learning Model for Hadronization**
  [Ghosh, Nachman, Siodmok, Yu, arXiv: 2203.12660]
  1st step: generate kinematics of a cluster decay:



- **How?**
  Use Generative Adversarial Networks (**GAN**)

Adversarial Networks

Arthur Lee Samuel (1959) wrote a program that learnt to play checkers well enough to beat him.

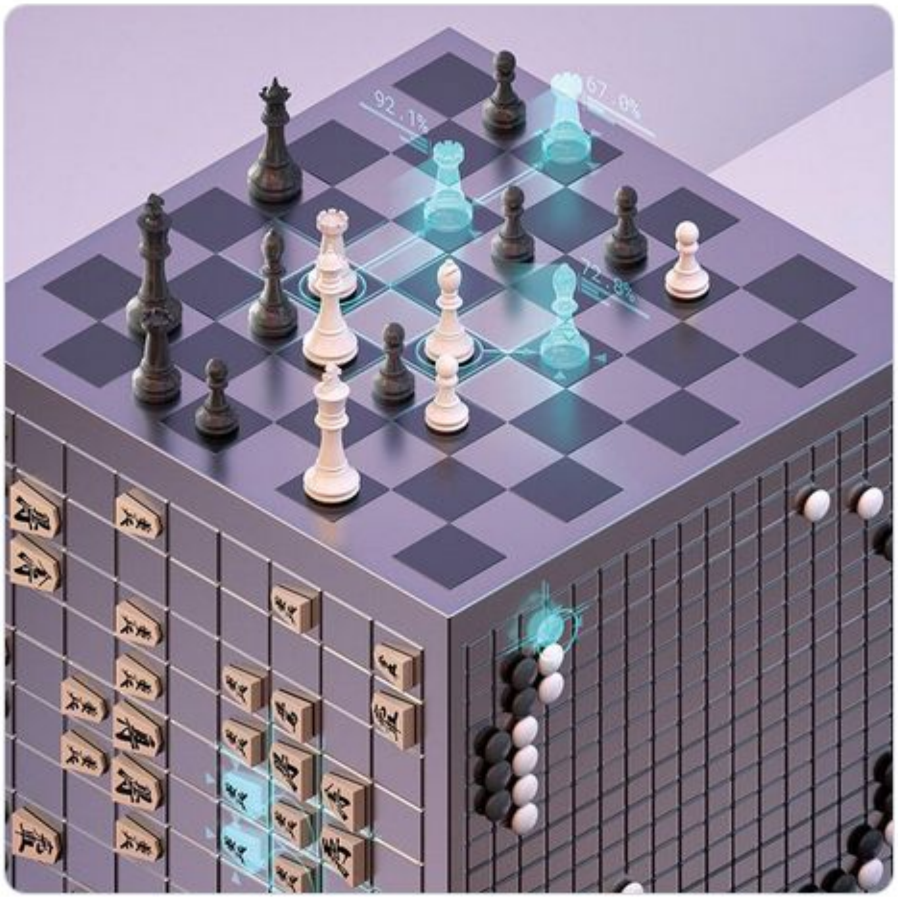


- He popularized the term **"machine learning"** in 1959.
- The program chose its move based on a **minimax** strategy, meaning it made the move assuming that the opponent was trying to optimize the value of the same function from its point of view.
- He also had it play thousands of **games against itself** as another way of learning.

Adversarial Networks

DeepMind @DeepMind · Dec 6, 2018

The full peer-reviewed @sciencemagazine evaluation of #AlphaZero is here - a single algorithm that creatively masters chess, shogi and Go through self-play deepmind.com/blog/alphazero…



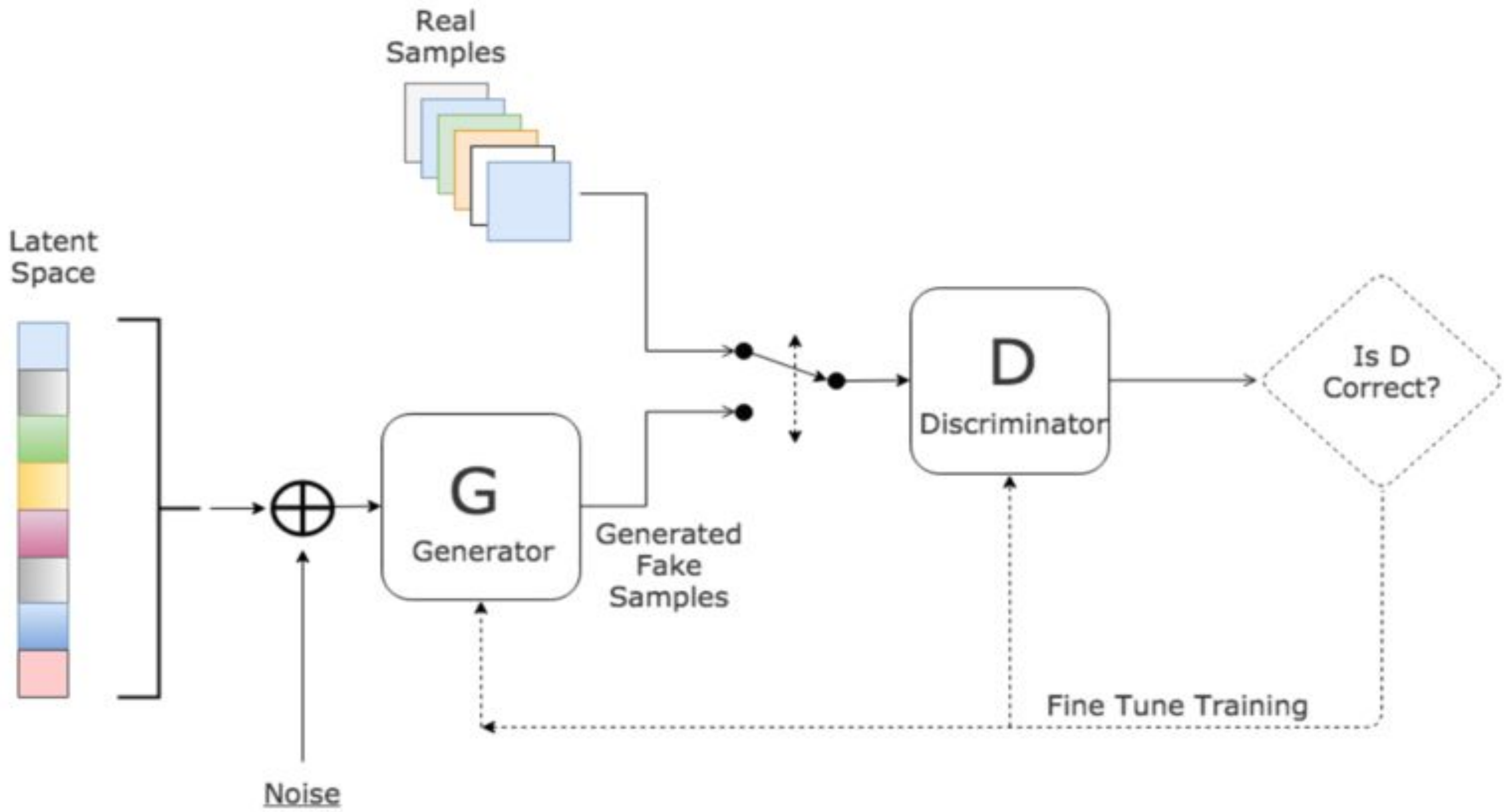
Demis Hassabis

CBE FRS FREng FRSA

By playing **games against itself**, AlphaGo Zero surpassed the strength of AlphaGo Lee in three days by winning 100 games to 0.

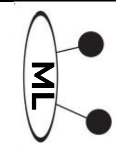
Generative Adversarial Network (GAN)

[Goodfellow et al. “Generative adversarial nets”. arxiv:1406.2661]



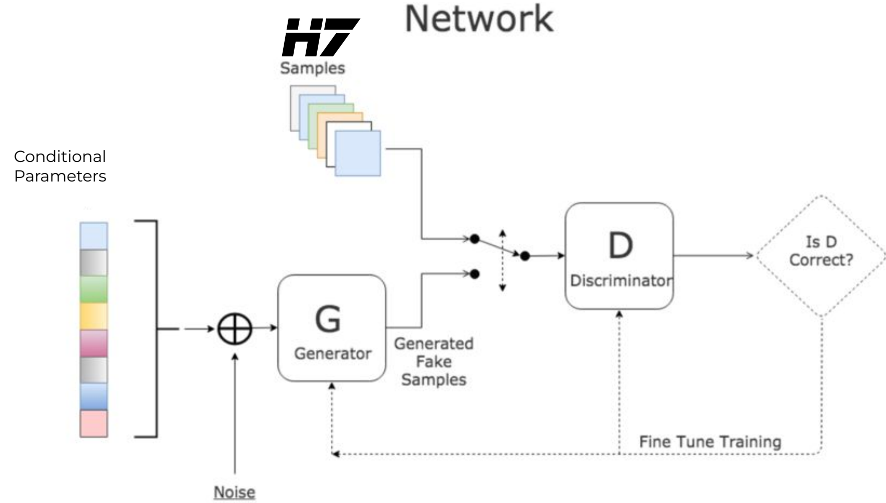
Towards a Deep Learning Model for Hadronization

1st step: generate kinematics of a cluster decay to 2 hadrons



How?

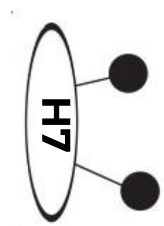
Generative Adversarial Network



Training data:

e^+e^- collisions at $\sqrt{s} = 91.2$ GeV

Cluster $(E,\ p_x,\ p_y,\ p_z)$



$\pi^0(E,\ p_x,\ p_y,\ p_z)$

$\pi^0(E,\ p_x,\ p_y,\ p_z)$

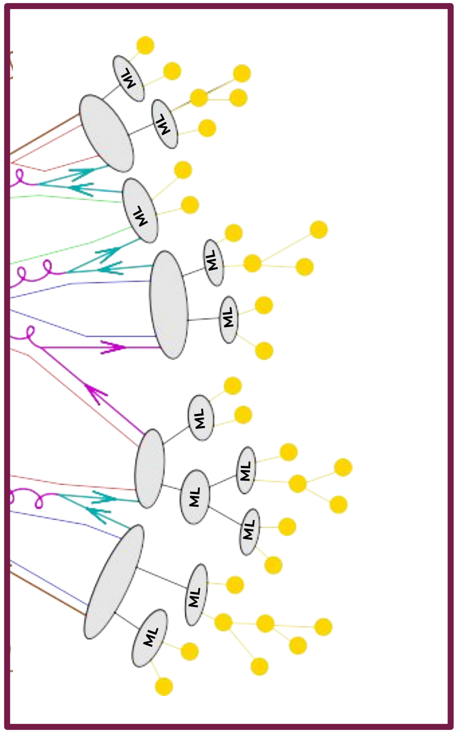
Pert = 0/1 memory of quarks direction

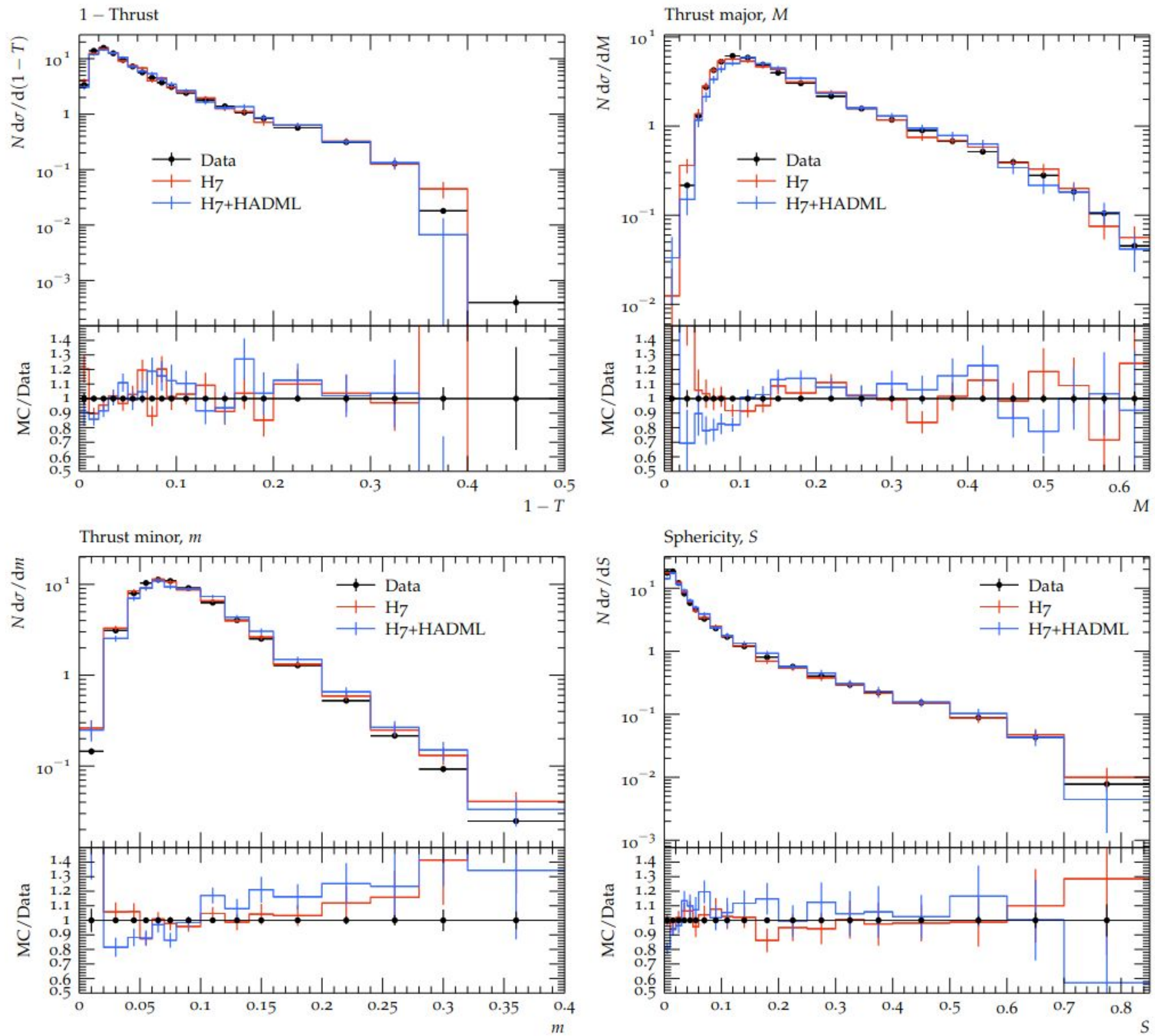
Results [Ghosh, Nachman, Siodmok, Yu, arXiv: 2203.12660]

Full-event Validation

(Full events using HADML integrated into Herwig 7)

LEP DELPHI **Data**





- The ultimate goal of is to train the ML model directly on data to improve hadronization models

2. LHC MC

Precise process-oriented MC
[main focus on QED/EW]

Precision Monte Carlos:
Theory predictions with <0.100% precision, so far only **KKMCee** and **BHLUMI** qualify, FCCee will require 0.001%.

KKMCee: $e^+e^- \to f\bar{f}(n\gamma),\ f = \mu, \tau, q, \tau \to X$

Rewritten to C++ [Jadach, Ward, Was, Yost, Siodmok, CPC 2022]

- Resumed (exponentiated) multi photon effects at the AMPLITUDE level (CEEX scheme) keeping (exponentiated) initial-final state interferences.
- Non-soft complete QED complete up to 3-rd order LO, NLO 2-nd order, in the initial and final states,
- Complete (longitudinal and transverse) spin polarisation for the incoming beams and outgoing fermions (mandatory for tau pairs) including spin correlations.
- It is **intended to be a starting point for the future improvements**, which will be mandatory for the future high precision lepton collider projects.
- Validated against of Fortran version
- A number of improvements in the Monte Carlo algorithm

BHLUMI: did not change from LEP but it was used

[Jadach and Janot, Phys. Letters B803 (2020) 135319] LEP data reanalyzed:

$$N_\nu = 2.9840 \pm 0.0082 \ \to 2.9963 \pm 0.0074$$

Specialized programs [see Z. Wąs talk]

Photos

- re-written to C++,
- emission of lepton pair was introduced,
- several processes, like Z,W,B meson decays emission kernels based on complete first order matrix element were introduced into fixed order and multiple photon mode of Photos operation.

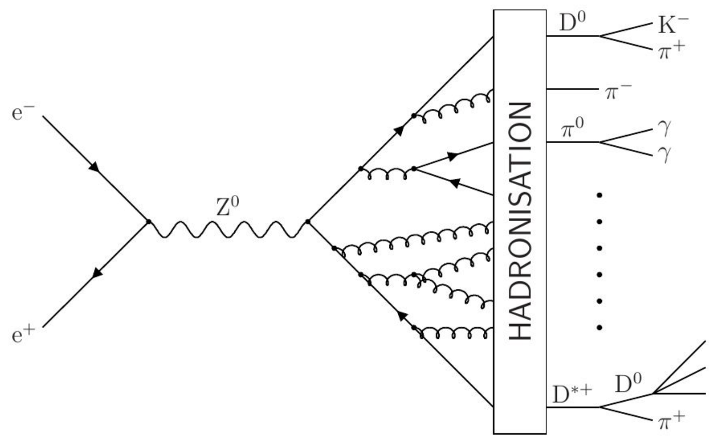
Tauola

- Multiple new tau decay modes, of new physics and of Standard Model were introduced into tauola.
- New version is now installed in Belle 2 software.
- Hopefully, new parametrization of hadronic currents for tau decay channels will become available to broader community in the forthcoming years.

3. Future ee factory

General purpose MC QCD developments

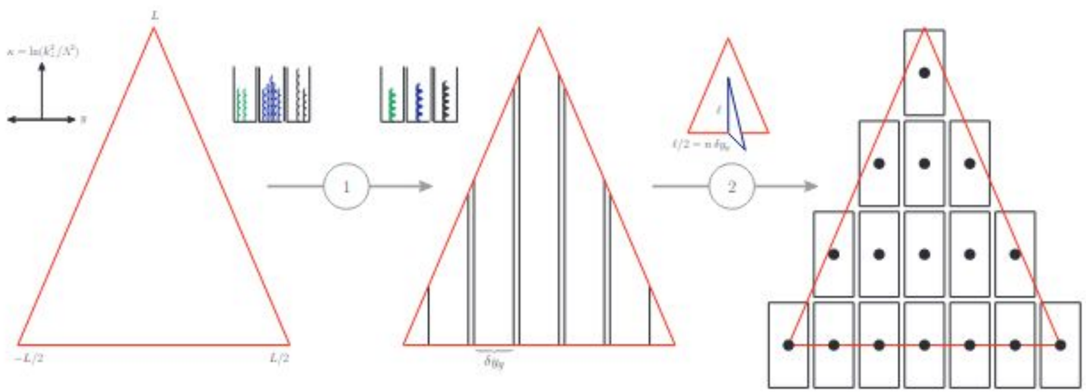
- Better Parton Shower
  - NNLO + NLO Parton Shower?
  - Amplitude evolution?
  - Quantum Computers



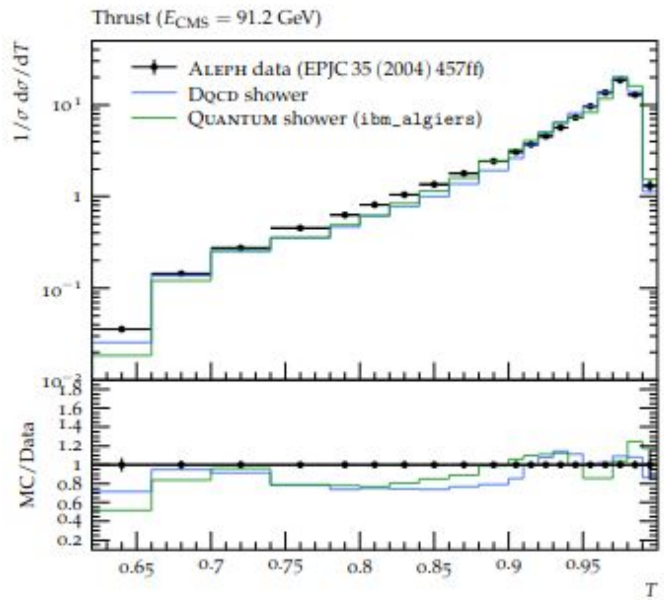
- Better Hadronization
  - lattice QCD?
  - ML?
  - improved string and cluster
  - new measurements

Collider Events on a Quantum Computer

Gösta Gustafson,[a] Stefan Prestel,[a] Michael Spannowsky,[b] Simon Williams[c]



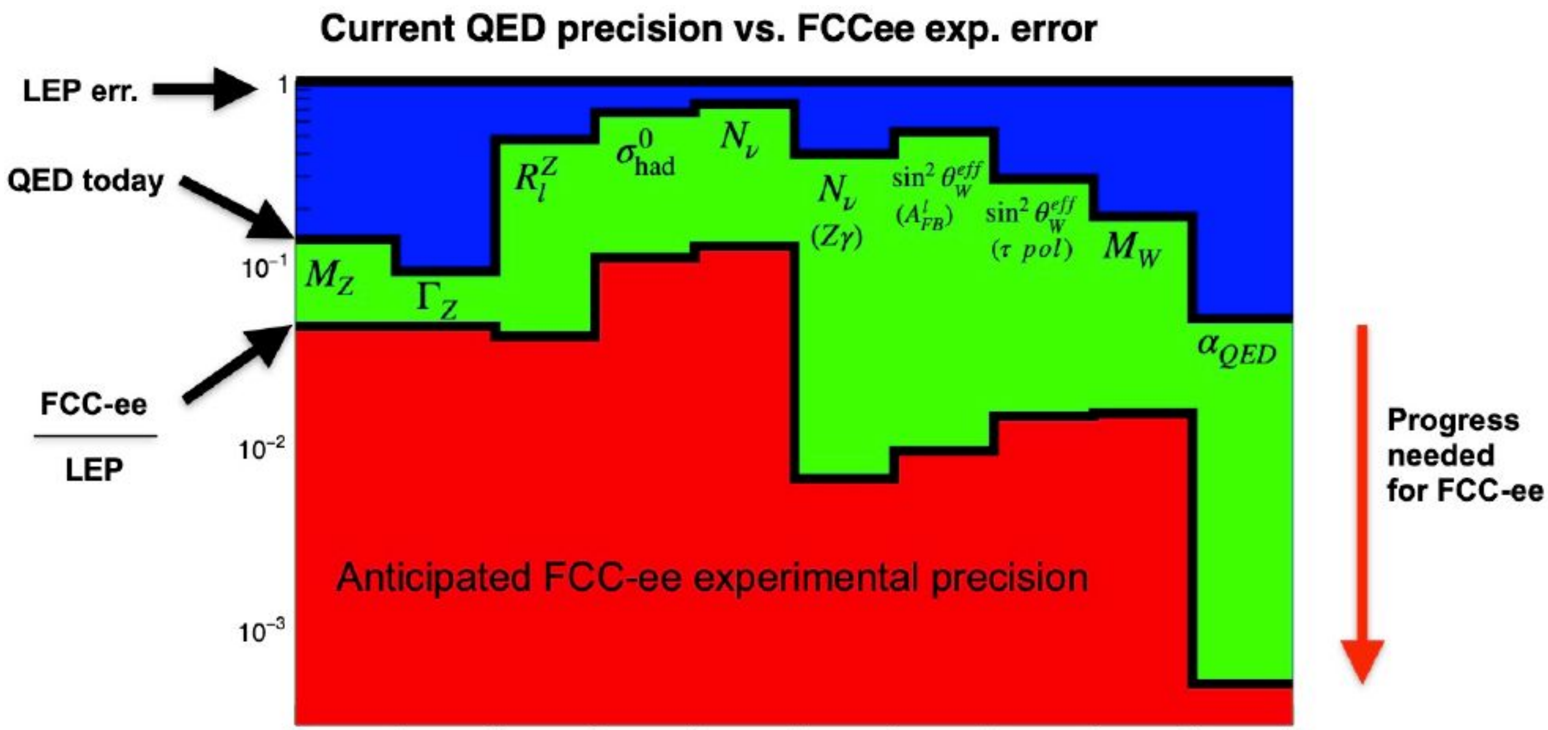
Simplified parton showers using the Discrete QCD method



“This is the first time a Noisy Intermediate-Scale Quantum (NISQ) device has been used to simulate realistic high-energy particle collision events” [2022]

3. Future ee factory

Future e+e- machine will be precision factory (luminosity up to 10^5 higher then LEP)



Most sensitive to QED radiation observables
[Jadach, Skrzypek arXiv:1903.09895]

- The present precision of QED theoretical predictions would severely limit the analysis of precise measurements at FCC-ee.
- To properly confront the data with theoretical predictions of similar accuracy demands a huge progress in precision MC calculations!
- Needed factor 6-200 improvement with respect to LEP.

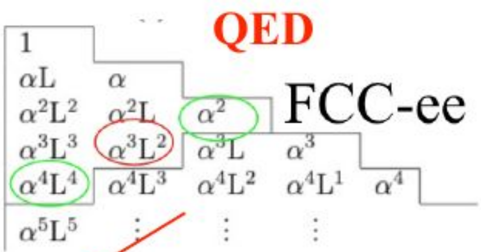
3. Future-ee

Example: KKMCee

Using the notation of [1303], the CEEX total cross section for the fermion pair production process at an electron collider, $\mathrm{e}^-(p_a) + \mathrm{e}^+(p_b) \to \mathrm{f}(p_c) + \bar{\mathrm{f}}(p_d) + \gamma(k_1), \dots, \gamma(k_n)$ reads as follows

$$\sigma^{(r)} = \sum_{n=0}^{\infty} \frac{1}{n!} \int d\tau_n(p_1 + p_2;\ p_3, p_4,\ k_1, \dots, k_n)$$
$$\times\, e^{2\alpha \Re B_4(p_a,\dots,p_d)} \frac{1}{4} \sum_{\text{spin}} \left| \mathfrak{M}_n^{(r)}(p, k_1, k_2, \dots k_n) \right|^2,$$

CEEX QED+EW matrix element in CEEX



where $\mathfrak{M}_n^{(r)}$ are the CEEX spin amplitudes, $d\tau_n$ is the standard LIPS, the virtual form factor B_4 is factorized (exponentiated) and the real emission spin independent soft factors $\mathfrak{s}$ are also factorized out. The momenta $p_1, \dots$ of the fermions are denoted collectively as p. The spin amplitudes read

Now in KKMC/CEEX ⟵

$$\mathfrak{M}_n^{(r)}(p, k_1, k_2, k_3, \dots, k_n) = \prod_{s=1}^{n} \mathfrak{s}(k_s) \Big\{ \hat{\beta}_0^{(r)}(p) + \sum_{j=1}^{n} \frac{\hat{\beta}_1^{(r)}(p, k_j)}{\mathfrak{s}(k_j)} + \sum_{j_1 < j_2} \frac{\hat{\beta}_2^{(r)}(p, k_{j_1}, k_{j_2})}{\mathfrak{s}(k_{j_1})\mathfrak{s}(k_{j_2})}$$
$$+ \sum_{j_1<j_2<j_3} \frac{\hat{\beta}_3^{(r)}(k_{j_1}, k_{j_2}, k_{j_3})}{\mathfrak{s}(k_{j_1})\mathfrak{s}(k_{j_2})\mathfrak{s}(k_{j_3})} + \sum_{j_1<j_2<\dots<j_r} \frac{\hat{\beta}_r^{(r)}(k_{j_1}, k_{j_2}, \dots, k_{j_r})}{\mathfrak{s}(k_{j_1})\mathfrak{s}(k_{j_2})\dots\mathfrak{s}(k_{j_r})} + \cdots \Big\}, \qquad (26)$$

$\mathcal{O}(\alpha^3)$ ⟵ $\mathcal{O}(\alpha^r)$

To be added for FCCee

$\mathcal{O}(\alpha^2)$ 2-loop **EW**

$\mathcal{O}(\alpha^2)$ 1-loop **EW**

[see Janusz Gluza Talk]

such that the subtracted amplitudes $\hat{\beta}_j^{(r)}$ are IR-finite. In the $\mathcal{O}(\alpha^2)$ $(r = 2)$ implementation of KKMC we define

$$\hat{\beta}_0^{(2)}(p) = \mathfrak{M}_0^{(2)}(p) = \left[e^{-\alpha B_4(p)} \mathcal{M}_0^{(2)}(p) \right] \Big|_{\mathcal{O}(\alpha^2)}, \qquad (27)$$

which includes QED and EW virtual corrections. In the future implementation of the $\mathcal{O}(\alpha^2)$ EW corrections, they would also enter into to the $2 \to 3$ non-soft components:

$$\hat{\beta}_1^{(2)}(p, k_1) = \mathfrak{M}_1^{(2)}(p, k_1) - \hat{\beta}_0^{(1)}(p)\mathfrak{s}(p, k_1), \quad \mathfrak{M}_1^{(2)}(p, k_1) = e^{-\alpha B_4(p)} \mathcal{M}_1^{(2)}(p, k_1)\Big|_{\mathcal{O}(\alpha^2)}. \qquad (28)$$

Watch out! The existing EW $\mathcal{O}(\alpha^2)$ calculations for $e^+e^- \to f\bar{f}$ are only for *inclusive* quantities.

Conclusions

1. Event generators crucial at the LEP times, since the start of LHC and most likely also in the Future.
2. General Purpose MC : Qualitatively predictive already 25 years ago and quantitatively steady progress, continuing today.
3. Precise process-oriented MC, LEP <0.100% precision, however FCCee will require 0.001%.
4. FCC-ee: Upgrade of LEP legacy MCs is good but limited strategy.
   For factor 50-150 improvement in precision one needs new innovative projects.
5. Future:
   - continuous dialogue with experimental community
   - more powerful computational techniques (ML) and computers (Quantum Computers)
   - new ideas (amplitude level Parton Showers, ML hadronization...)
6. One should avoid "monopoly" of a single MC for a given process/observable. The best would be (at least) two MCs of similar high quality developed independently by two or more groups of authors. [examples: YFSWW3 + RACOONWW, Herwig, Pythia, Sherpa]

Thank you for the attention!

Backup slides

Thrust, $\boldsymbol{T}$ is defined by

$$T = \max_{\hat{\boldsymbol{n}}} \left(\frac{\sum_i |\boldsymbol{p}_i \cdot \hat{\boldsymbol{n}}|}{\sum_i |\boldsymbol{p}_i|} \right) \quad , \tag{1.24}$$

where the thrust axis $\hat{\boldsymbol{n}}_T$ is defined as the unit 3-vector $\hat{\boldsymbol{n}}$ which maximises the expression. For a perfectly 'pencil-like' two-jet event, the thrust axis lies parallel to the jets, so $|\boldsymbol{p}_i \cdot \hat{\boldsymbol{n}}_T| = |\boldsymbol{p}_i|$, yielding $T = 1$. In the case of a 'spherical' event, with an infinite number of particles distributed isotropically in the final state, the thrust becomes a ratio of solid angle integrals:

$$T = \frac{\int |\cos\theta|\, \mathrm{d}\Omega}{\int \mathrm{d}\Omega} = \frac{2\pi}{4\pi} = \frac{1}{2} \quad . \tag{1.25}$$

It can be shown that all events satisfy $\frac{1}{2} < T < 1$. All the other event shapes considered here approach zero in the two-jet limit; for consistency, we will therefore define the observable $y = 1 - T$ which shares this property.

The concept of thrust was already in use before the advent of QCD. In 1964 [18], a "principal axis" equivalent to $\hat{\boldsymbol{n}}_T$ was proposed for the analysis of jets observed in hadron collisions, though the origin of the jets was hitherto unexplained. Later, in 1977 [19], it was recognised that this "maximum directed momentum" represented a calculable quantity in perturbative QCD.

1. LEP legacy MC

Example:
Precise process-oriented MC

$$e^+e^- \to W^+W^-(n\gamma) \to 4f(n\gamma)$$

KORALW, YFSWW

[S. Jadach, W. Placzek, M. Skrzypek, B.F.L. Ward, Z. Was]

YFSWW

Simplified Process

(Double-Resonant W)

KORALW

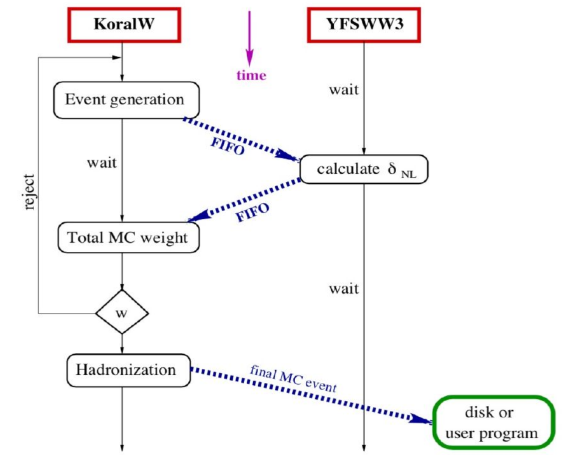
Full Process

(All 4f Channels)

Merge of KoralW and YFSWW3 = KandY

Possible because the underlying photonic distribution is the same YFS-ISR in both codes. All other photonic effects are included as weights. So are the $\mathcal{O}(\alpha)$ EW corrs.

Concurent realization of $\sigma_{K/Y}$ with ”named pipes”



Works effectively as a single MC event generator

KrkNLO - motivation

- **Why would you like another method of NLO+PS matching?**
  - The method is extremely simple (can be applied on event record).
  - No negative weight events.
  - In angular ordered PS - no need for a truncated shower.
  - Simple at NLO $\Rightarrow$ you may hope that pushing the method to NNLO+NLO PS should be possible.

KrkNLO - basic idea

Lets consider

DY cross section at NLO in collinear $\overline{\mathrm{MS}}$ factorization for the $q\bar{q}$ channel:

$$\sigma^1_{\mathrm{DY}} - \sigma^B_{\mathrm{DY}} \quad = \quad \sigma^B_{\mathrm{DY}}\, D_1^{\overline{\mathrm{MS}}}(x_1,\mu^2) \otimes \frac{\alpha_s}{2\pi} C_q^{\overline{\mathrm{MS}}}(z) \otimes D_2^{\overline{\mathrm{MS}}}(x_2,\mu^2)\,,$$

where

$$C_q^{\overline{\mathrm{MS}}}(z) = C_F \left[4\,(1+z^2) \left(\frac{\ln(1-z)}{1-z} \right)_+ - 2\frac{1+z^2}{1-z} \ln z + \delta(1-z) \left(\frac{2}{3}\pi^2 - 8 \right) \right].$$

All solutions for NLO + PS matching which use $\overline{\mathrm{MS}}$ PDFs, need to implement collinear remnant term of the type $4\,(1+z^2)\left(\frac{\ln(1-z)}{1-z}\right)_+$ that are technical artefacts of $\overline{\mathrm{MS}}$ scheme.

The implementation is not easy since those terms correspond to the collinear limit but Monte Carlo lives in 4 dimensions and not in the phase space restricted by $\delta(k_T^2)$.
The idea behind the MC scheme is to absorb those terms to PDF.

KrkNLO - the method

1. Take a parton shower that covers the (α, β) phase space completely (no gaps, no overlaps) and produces emissions according to approx. real matrix element K.
2. Upgrade the real emissions to exact ME R by reweighting the PS events by $W_R = R/K$.
3. We define the coefficion function $C^R(z) = \int (R - K)$. To avoid unphysical artifacts of $\overline{\mathrm{MS}}$.
4. Transform PDF for $\overline{\mathrm{MS}}$ scheme to this new physical MC factorization scheme.
5. As a result the virtual+soft correction, Δ_{S+V}, is just a constant, without x-depended collinear remnant terms now. Multiply the whole result by $1 + \Delta_{S+V}$ to achieve complete NLO accuracy.

KrkNLO - example DY

- Our approach to NLO+PS matching (example: Drell-Yan)

Real part:

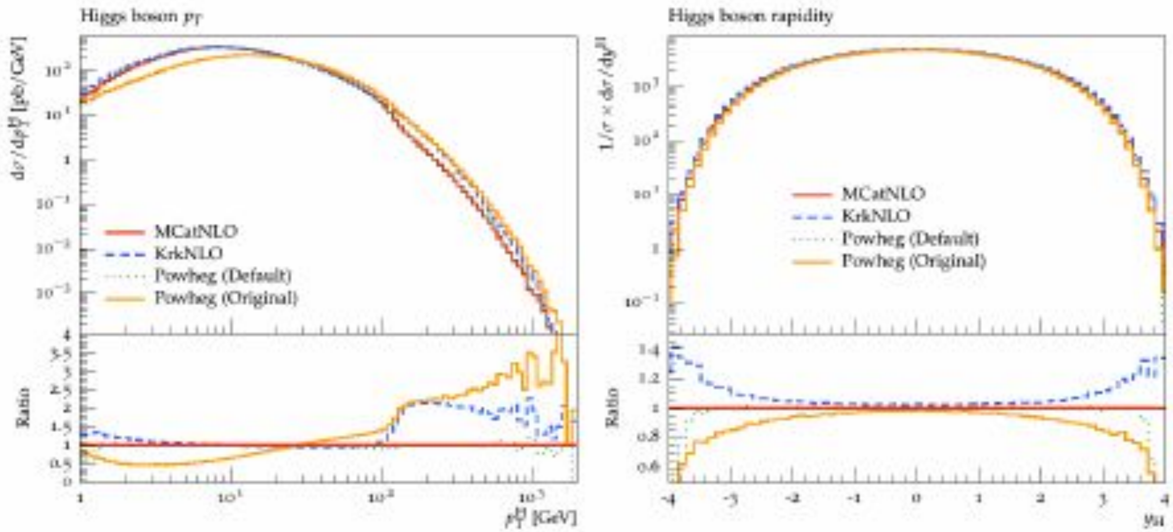
$$W_R^{q\bar{q}}(\alpha,\beta) = 1 - \frac{2\alpha\beta}{1+(1-\alpha-\beta)^2}$$
$$W_R^{qg}(\alpha,\beta) = 1 + \frac{\alpha(2-\alpha-2\beta)}{1+2(1-\alpha-\beta)(\alpha+\beta)}$$

Virtual + soft:

$$W_{V+S}^{q\bar{q}} = \frac{\alpha_s}{2\pi} C_F \left[\frac{4}{3}\pi^2 - \frac{5}{2}\right]$$
$$W_{V+S}^{qg} = 0$$

- PDF in MC factorization scheme - full definition

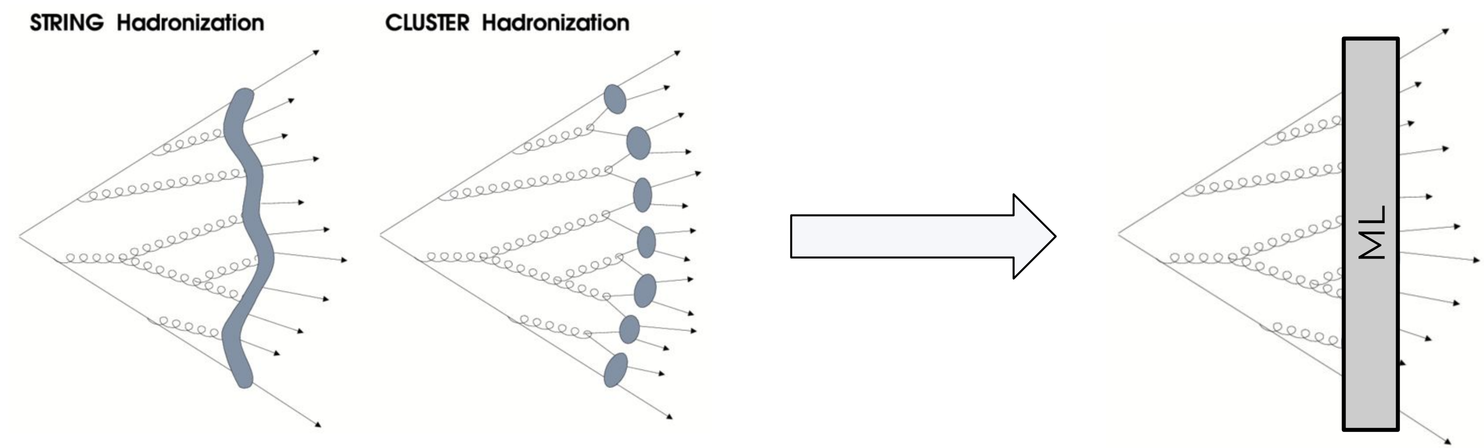
- KrkNLO for the Higgs boson production



Non-perturbative QCD

Hadronization: Early 1980's
(since then very little development)

Early 2020's
(lot of progress in ML)



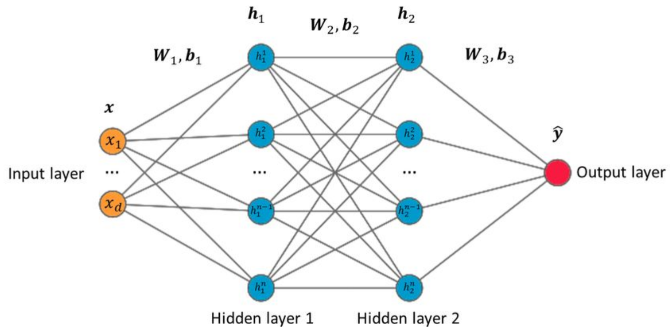
- ➔ Increased control of perturbative corrections ⇒ more often the precision of LHC measurements is limited by MCEG's non-perturbative components, such as hadronization.
- ➔ Hadronization (phenomenological models with many free parameters ~ 30 parameters)
- ➔ Hadronization is a fitting problem ML is proved to be well suited for such a problems.

Idea of using Machine Learning (ML) to improve hadronization.

NNPDF used successfully ML to nonperturbative Parton Density Functions (PDF)
Hadronization is closely related to so-called fragmentation functions (FF). Early on, FFs were considered the counterpart of PDFs. While PDFs are understood as probability densities for finding partons, with a given momentum, inside colour-neutral particles, FFs (or hadronization) were understood as probability densities for finding colour-neutral particles from partons.

Architecture: conditional GAN

(each a fully connected, hidden size 256, a batch normalization layer, LeakyReLU activation function)



Generator

Input

Cluster $(E,\ p_x,\ p_y,\ p_z)$ and 10 noise features sampled from a Gaussian distribution

Output (in the cluster frame)

ϕ - polar angle
θ - azimuthal angle

we reconstruct the four vectors of the two outgoing hadrons

Discriminator

Input

ϕ and θ labeled as signal (generated by Herwig) or background (generated by Generator)

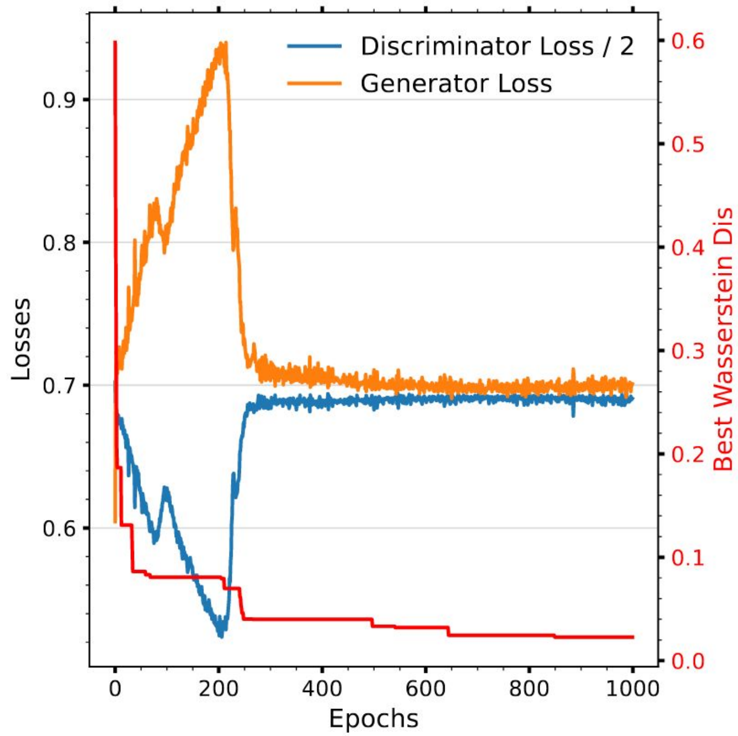
Output

Classification.

Training

cluster's four vector and angular variables are scaled to be between -1 and 1 (tanh activation function as the last layer of the Generator)

- **Discriminator** and the **Generator** are trained separately and alternately by two independent Adam optimizers with a learning rate of 10^{-4}, for 1000 epochs

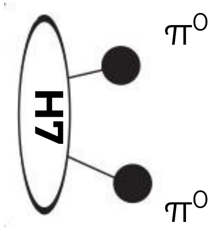


- **The best model** for events with partons of Pert = 0, is found at the epoch 849 with a total Wasserstein distance of 0.0228.

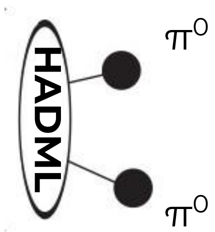
Results

Low-level Validation
(similar to training data)

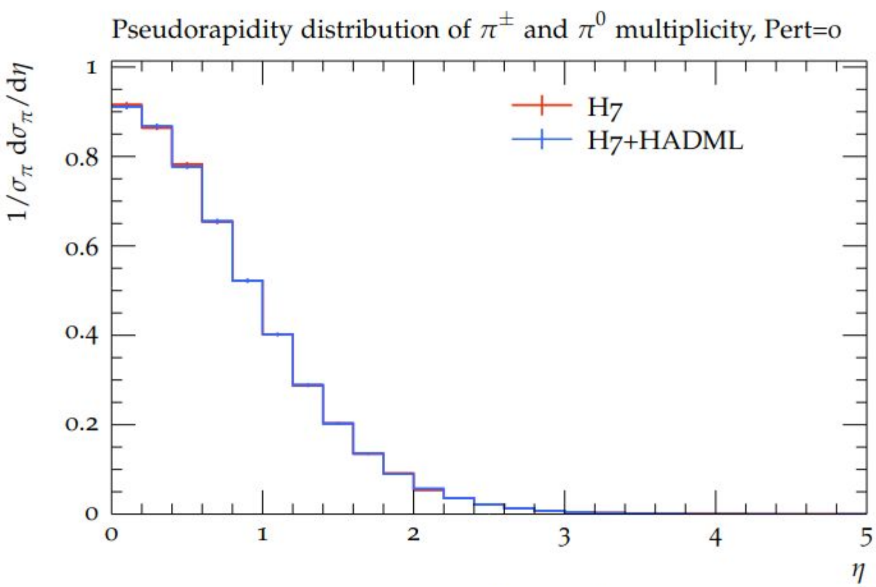
e^+e^- collisions at $\sqrt{s} = 91.2$ GeV

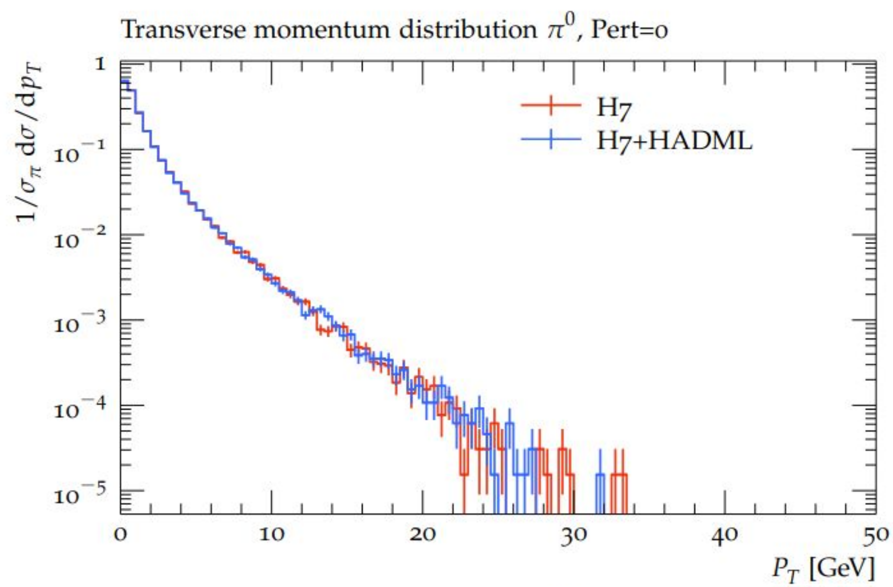


VS



π^0 kinematic variables



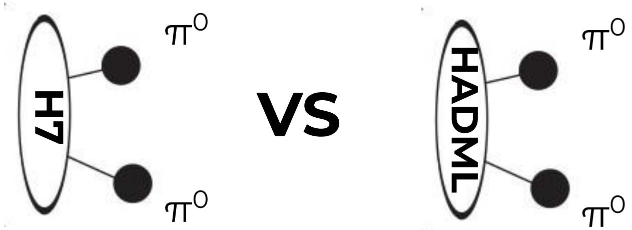


Pert = 0 (no memory of quark kinematics)

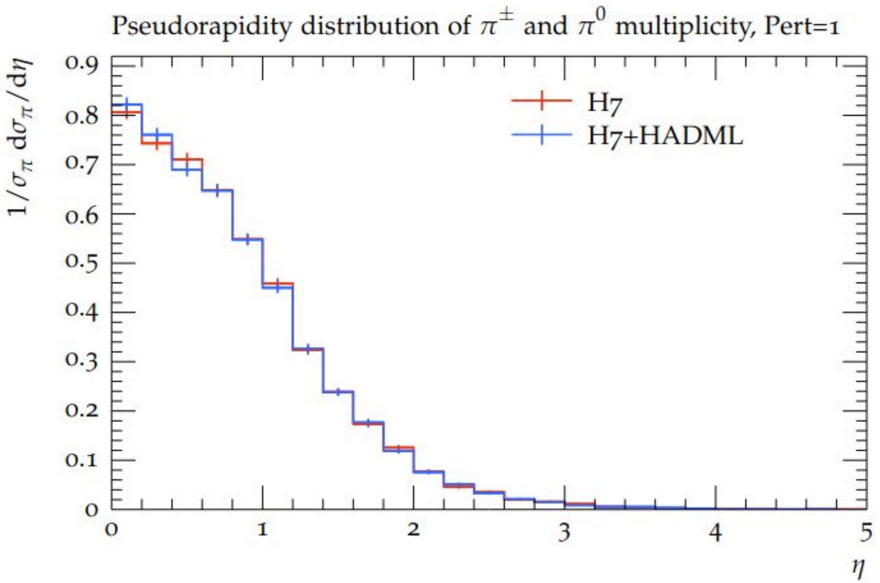
Results

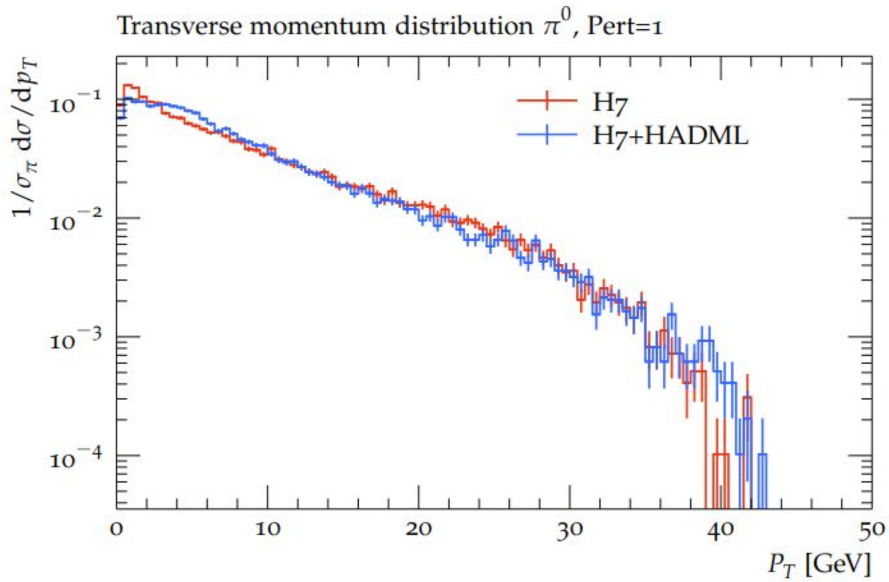
Low-level Validation
(similar to training data)

e^+e^- collisions at $\sqrt{s} = 91.2$ GeV



π^0 kinematic variables



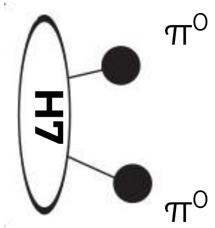


Pert = 1 (memory of quark kinematics)

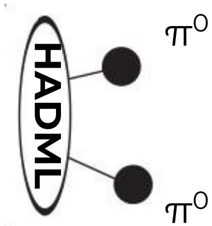
Results

Low-level Validation
(**beyond** training data **different energy**)

e^+e^- collisions at $\sqrt{s} = 192$ GeV

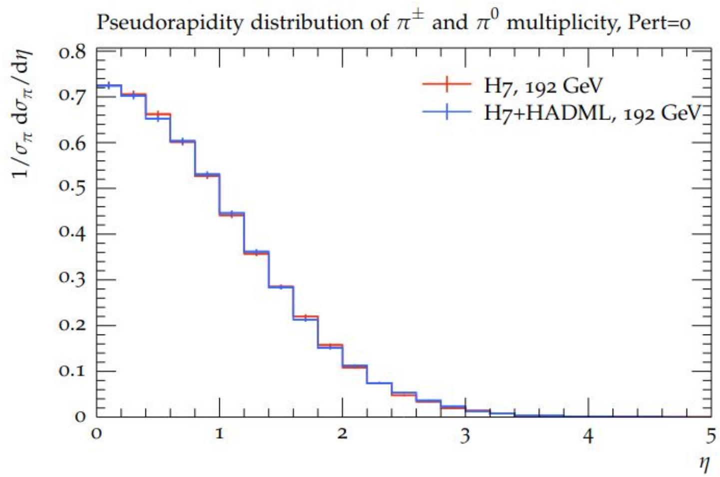


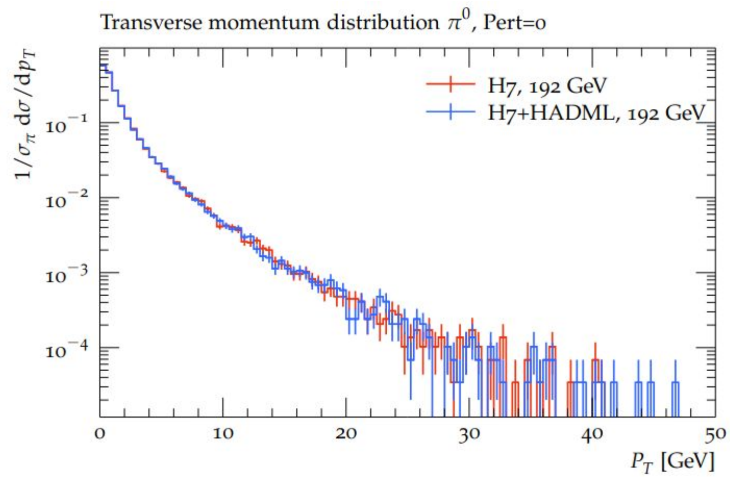
VS



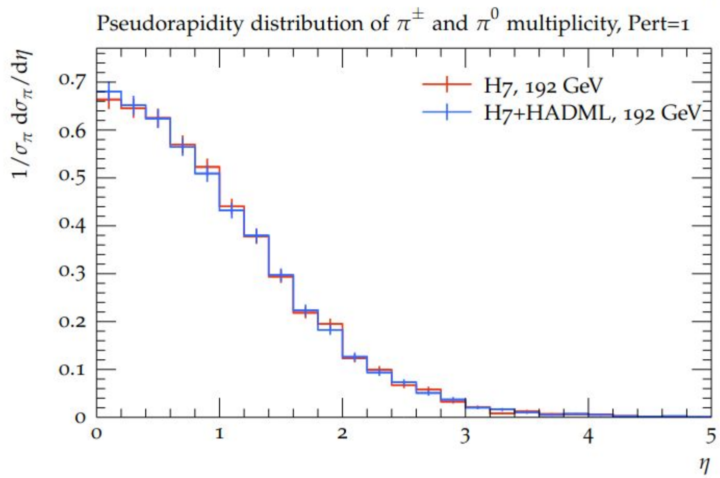
π^0 kinematic variables

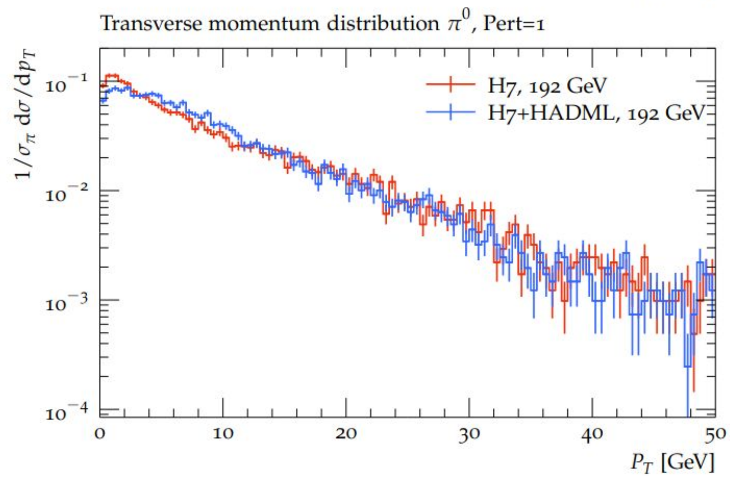
Pert = 0





Pert = 1

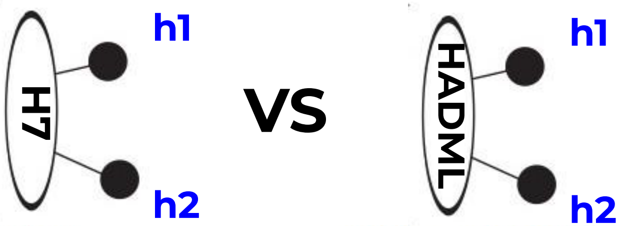




Results

Low-level Validation
(**beyond** training data **different hadrons**)

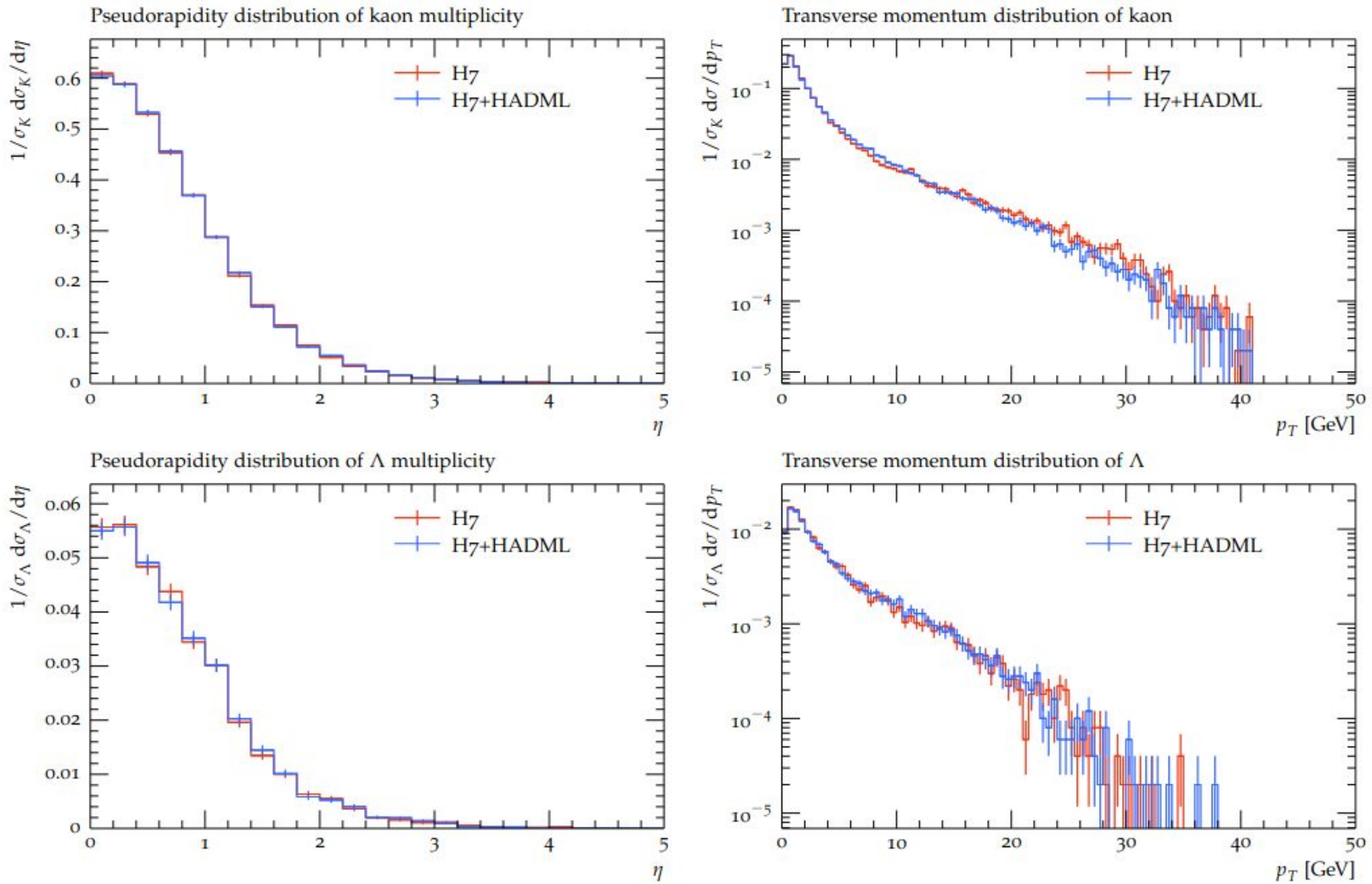
e^+e^- collisions at
$\sqrt{s} = 91.2$ GeV



h kinematic variables

Kaons

Lambda



Minimax Loss

In the paper that introduced GANs, the generator tries to minimize the following function while the discriminator tries to maximize it:

$$E_x[log(D(x))] + E_z[log(1 - D(G(z)))]$$

In this function:

- `D(x)` is the discriminator's estimate of the probability that real data instance x is real.
- E_x is the expected value over all real data instances.
- `G(z)` is the generator's output when given noise z.
- `D(G(z))` is the discriminator's estimate of the probability that a fake instance is real.
- E_z is the expected value over all random inputs to the generator (in effect, the expected value over all generated fake instances G(z)).
- The formula derives from the cross-entropy between the real and generated distributions.

The generator can't directly affect the `log(D(x))` term in the function, so, for the generator, minimizing the loss is equivalent to minimizing `log(1 - D(G(z)))`.

Wasserstein distance

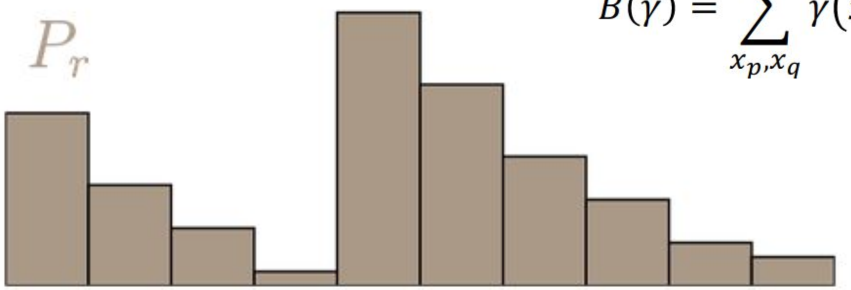
The Wasserstein distance

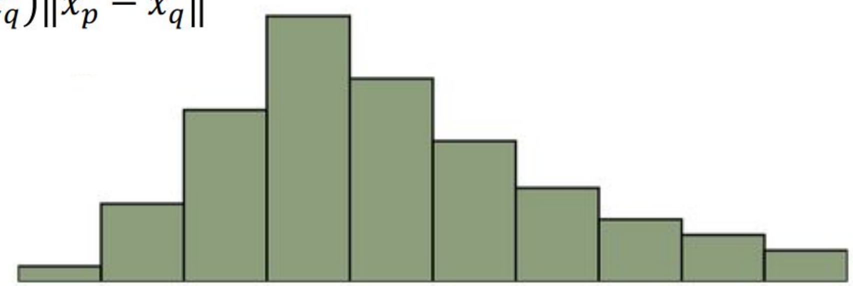
- For discrete probability distributions, the Wasserstein distance is called the earth mover's distance (EMD):
- EMD is the minimal total amount of work it takes to transform one heap into the other.

$$W(P,Q) = \min_{\gamma \in \Pi} B(\gamma)$$

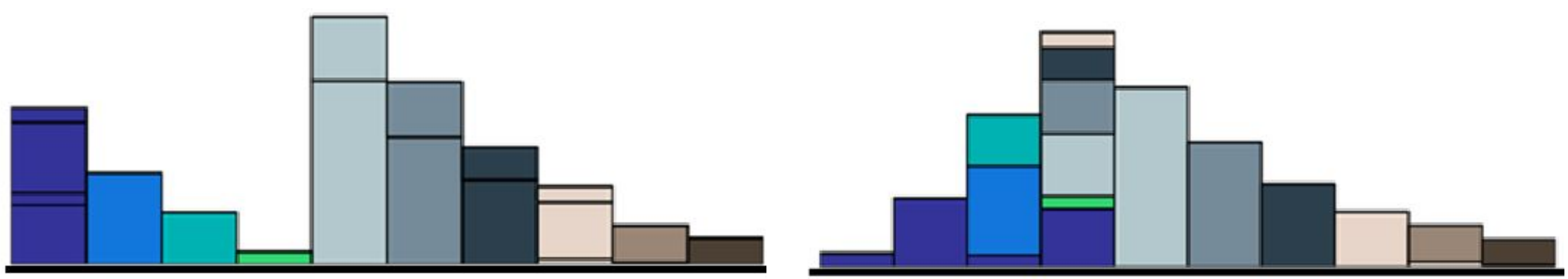
- Work is defined as the amount of earth in a chunk times the distance it was moved.

$$B(\gamma) = \sum_{x_p, x_q} \gamma(x_p, x_q) \|x_p - x_q\|$$

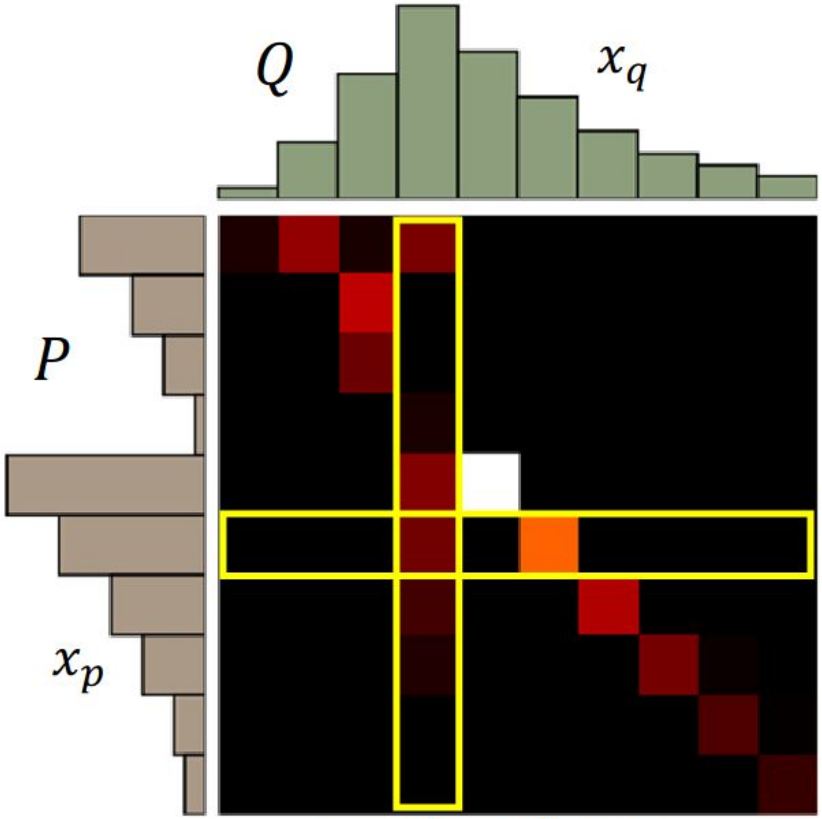




Best "moving plans" of this example



Wasserstein distance



moving plan γ
All possible plan Π

A "moving plan" is a matrix

The value of the element is the amount of earth from one position to another.

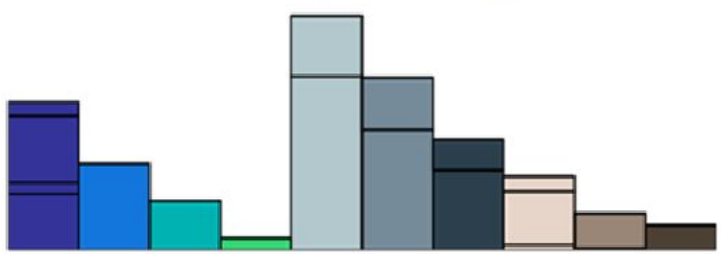
Average distance of a plan γ:

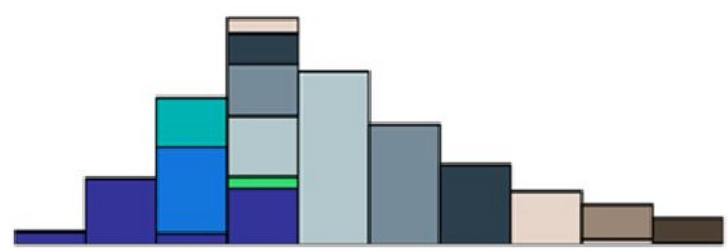
$$B(\gamma) = \sum_{x_p, x_q} \gamma(x_p, x_q) \| x_p - x_q \|$$

Earth Mover's Distance:

$$W(P, Q) = \min_{\gamma \in \Pi} B(\gamma)$$

The best plan





Effectively the strength of the QED

$(a)\quad 0.6\%$

1				
$\alpha \mathrm{L}$	α			
$\alpha^2\mathrm{L}^2$	$\alpha^2\mathrm{L}$	α^2		
$\alpha^3\mathrm{L}^3$	$\alpha^3\mathrm{L}^2$	$\alpha^3\mathrm{L}$	α^3	
$\alpha^4\mathrm{L}^4$	$\alpha^4\mathrm{L}^3$	$\alpha^4\mathrm{L}^2$	$\alpha^4\mathrm{L}^1$	α^4
$\vdots$	$\vdots$	$\vdots$		

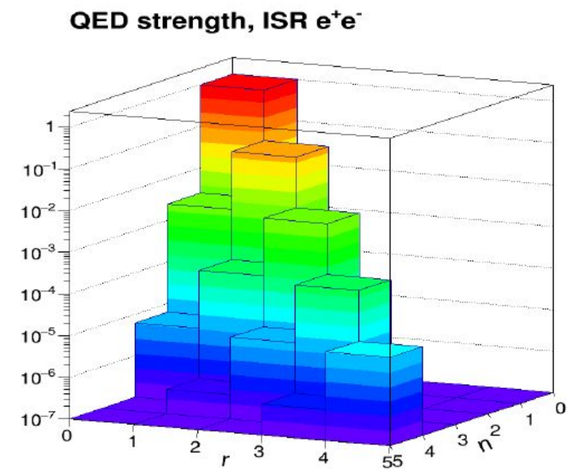
$(b)\quad 0.03\%$

1				
$\alpha \mathrm{L}$	α			
$\alpha^2\mathrm{L}^2$	$\alpha^2\mathrm{L}$	α^2		
$\alpha^3\mathrm{L}^3$	$\alpha^3\mathrm{L}^2$	$\alpha^3\mathrm{L}$	α^3	
$\alpha^4\mathrm{L}^4$	$\alpha^4\mathrm{L}^3$	$\alpha^4\mathrm{L}^2$	$\alpha^4\mathrm{L}^1$	α^4
$\vdots$	$\vdots$	$\vdots$		

$(c)\quad 0.001\%$

1				
$\alpha \mathrm{L}$	α			
$\alpha^2\mathrm{L}^2$	$\alpha^2\mathrm{L}$	α^2		
$\alpha^3\mathrm{L}^3$	$\alpha^3\mathrm{L}^2$	$\alpha^3\mathrm{L}$	α^3	
$\alpha^4\mathrm{L}^4$	$\alpha^4\mathrm{L}^3$	$\alpha^4\mathrm{L}^2$	$\alpha^4\mathrm{L}^1$	α^4
$\alpha^5\mathrm{L}^5$	$\vdots$	$\vdots$	$\vdots$	

$$\gamma_{nr} = \left(\frac{\alpha}{\pi}\right)^n \left(2\ln\frac{M_Z^2}{m_f^2}\right)^r,\ 0 \le r \le n,$$



- The (complete) order-by-order perturbative calculation in QED is definitely not the economic way to to obtain predictions for cross sections or asymmetries with the precision below 0.1%.
- Soft photon resummation is an absolute necessity, especially for resonant processes. It exists in at least three different variants (IEX, EEX and CEEX).
- Do not follow Bloch-Nordsieck to eliminate IR singularities!
- Resummation of collinear mass logarithms $\ln(s/m_f^2)$ is very useful, but in QED it is usually convenient to truncate it at some finite order.
- Approximation of small lepton mass $m_f^2/s \ll 1$ should be exploited for electron and muons as much as possible, but for τ lepton $\propto m_\tau^2/s, m_\tau^4/s^2$ terms may not be negligible at tree-level, while higher powers of this type in higher orders of α are probably irrelevant.