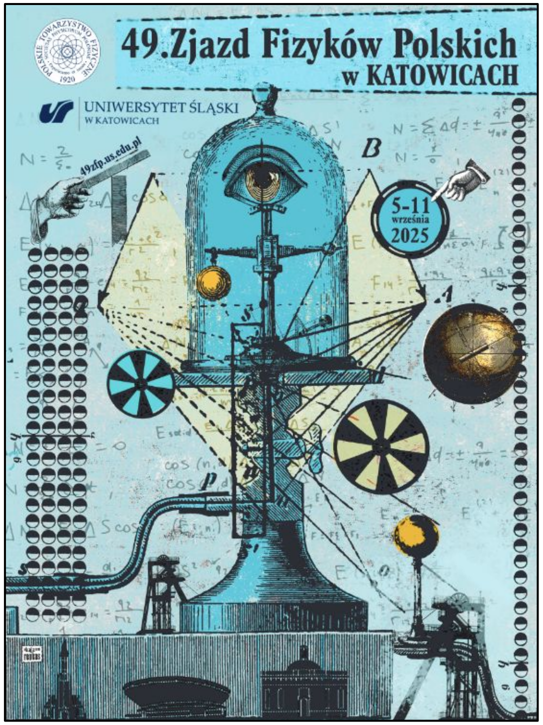
Symulacje dla odkryć w fizyce cząstek

Generatory Monte Carlo w erze LHC i przyszłego zderzacza e^+e^-



Andrzej Siódmok

NATIONAL SCIENCE CENTRE POLAND

JAGIELLONIAN UNIVERSITY IN KRAKÓW

"Grupa Twórcza Kwark"

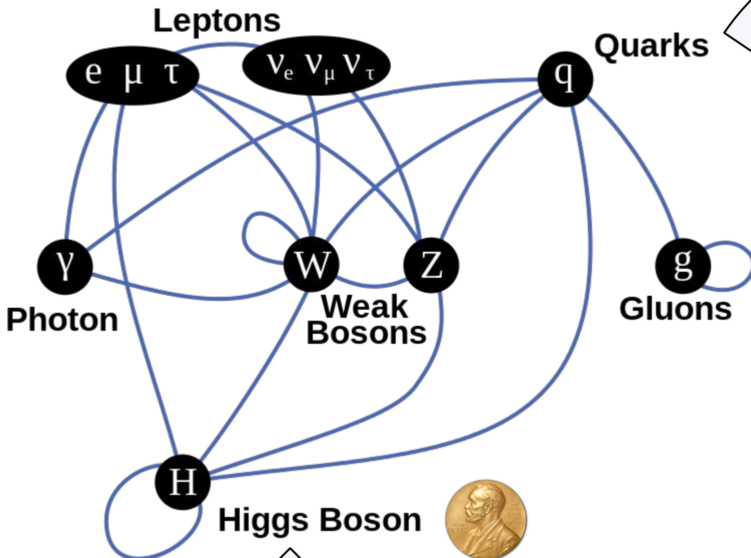
MCnet

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

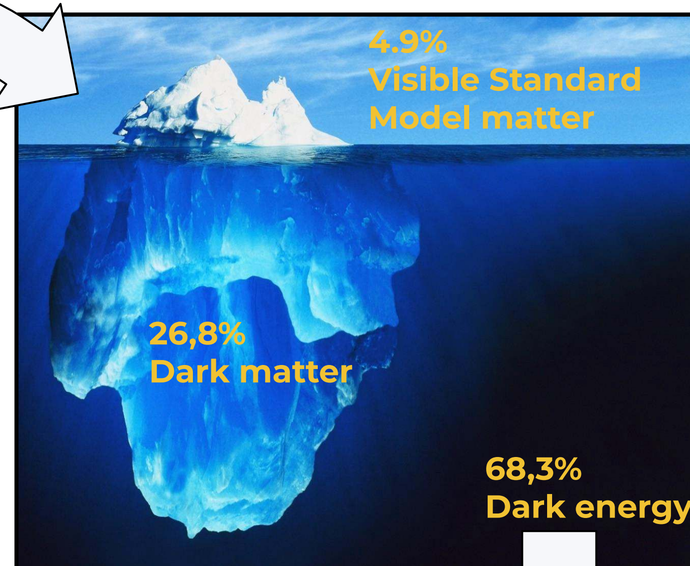
Motivation

Standard Model

very successful theory



Physics beyond SM must exist



LHC

European Strategy for Particle Physics

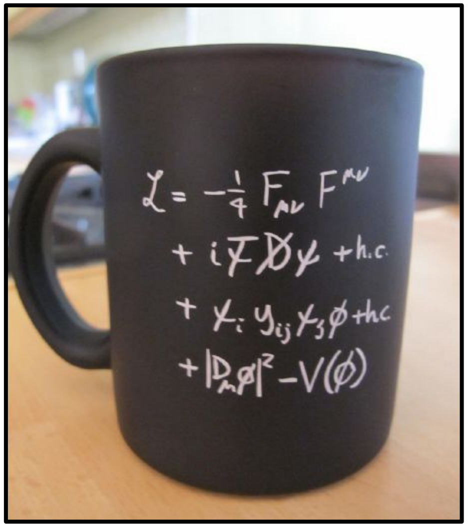
“Europe's top priority should be the exploitation of the full potential of the LHC”

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



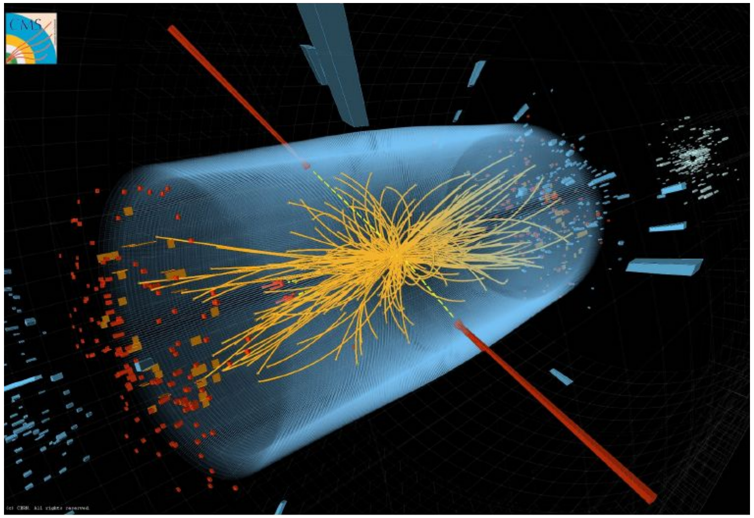
Data makes you smarter

> 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

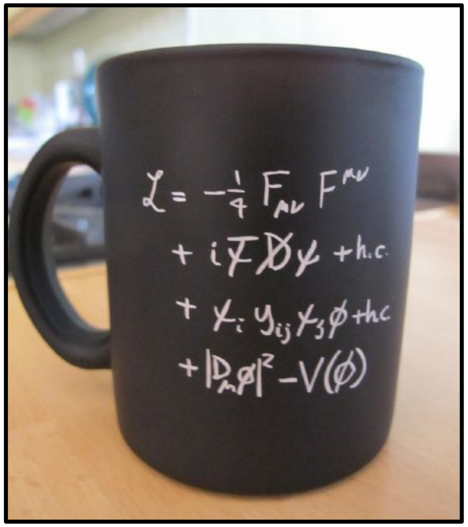


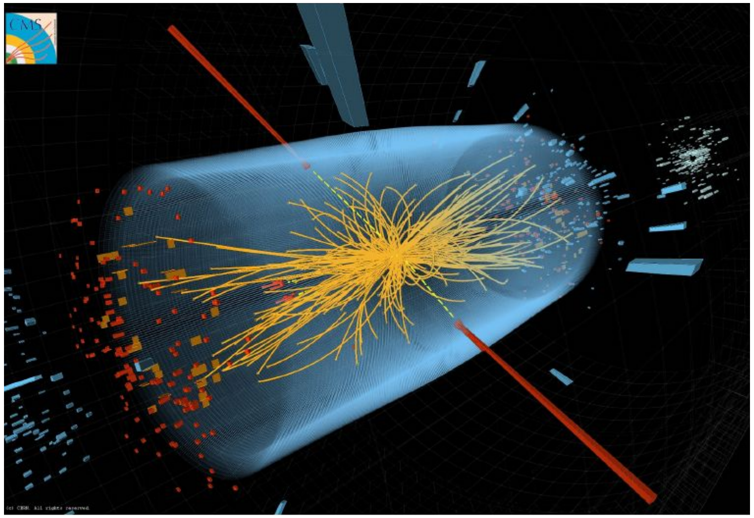
Standard Model

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.

(Herwig and Sherpa)@ Pythia

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

The landscape of Monte Carlo Event Generators in Poland

Polish research groups have a long and well-known tradition in preparing (or participating in preparation) of widely used MC simulation.

Landscape of Monte Carlo Event Generators in Poland

Polish research groups are working on (alphabetical order):

BHLUMI: YFS MC for Bhabha scattering at low angles **(IFJ,UJ)** [CPC 102 (1997) 229-251]

BHWIDE: YFS exponentiated Monte Carlo for Bhabha scattering at wide angles **(IFJ,UJ)** [Phys.Lett.B 390 (1997) 298-308]

CARLOMAT: MC for automatic computation of leading order cross sections of multiparticle reactions **(UŚ)** [CPC, 276 (2022) 108330]

EKHARA: MC generator for e+e- to e+e-pi0 and e+e- to e+e- pi+pi- processes and more **(UŚ)** [CPC 234 (2019) 245-255]

EpIC: MC event generator for exclusive processes based on generalised parton distributions **(NCBJ)** [Eur.Phys.J.C 82 (2022)]

GF-CAIN: MC generator for Gamma Factory **(UJ)** [Annalen Phys. 534(3), 2100250 (2022)]

Herwig: one of three existing general purpose MC **(UJ)** [Eur.Phys.J.C 84 (2024)]

HDECAY: a Program for Higgs Boson Decays in the Standard Model and its Supersymmetric Extension **(UW)** [CPC 238 (2019) 214-231]

KaTie: parton-level event generation with kT-dependent initial states **(IFJ)** [CPC 224 (2018) 371-380]

KKMCee: YFS Monte Carlo event generator KKMCee for lepton and quark pair production in lepton colliders [CPC 283 (2023) 108556]

KKMChh: YFS Monte Carlo event generator KKMCee for lepton and quark pair production in hadron colliders [PoS ICHEP2024 (2025)]

KoralW: YFS MC for all four-fermion final states in e+ e- collisions **(IFJ, UJ)** [CPC 140 (2001) 475-512]

KORALZ: YFS MC for the lepton or quark pair production at lepton colliders **(IFJ)** [CPC 124 (2000) 233-237]

KrkMC: studies on NLO parton shower Monte Carlo and KrkNLO matching **(IFJ, UJ)** [JHEP 01 (2025) 062]

MadGraph: automated tree-level and NLO differential cross sections **(IFJ)** polarized matrix elements [JHEP 04 (2020) 082]

MINCAS: Markov chain MC of QCD cascades in quark-gluon plasma in heavy-ion collisions at the LHC **(IFJ, UJ)** [EPJC 79 (2019) 317]

NuWro MC: MC simulation tools used for neutrino oscillations **(UWr)**. [see J. Sobczyk talk]

PHOKARA: radiative return at flavour factories, e+e- to hadrons with energetic photon **(UŚ)** [Phys. Rev. D100 (2019)]

PHOTOS: universal Monte Carlo for QED radiative corrections **(IFJ)** [CPC 283 (2023) 108592]

SHERPA: one of three existing general purpose MC **(UJ)** [JHEP 12 (2024) 156]

TAUOLA: Monte Carlo event generator TAUOLA for tau lepton decays **(IFJ)** [CPC 232 (2018) 220-236]

WHIZARD: MC for multi-particle scattering cross sections and simulated event samples **(UW)** [EPJ Web Conf. 315 (2024)]

WINHAC: MC WINHAC for the charged-current Drell-Yan process based on the YFS method **(IFJ, UJ)** [Eur.Phys.J.C 29 (2003)]

TERAZ POLSKA

Outline
Generators

1. LEP

2. LHC

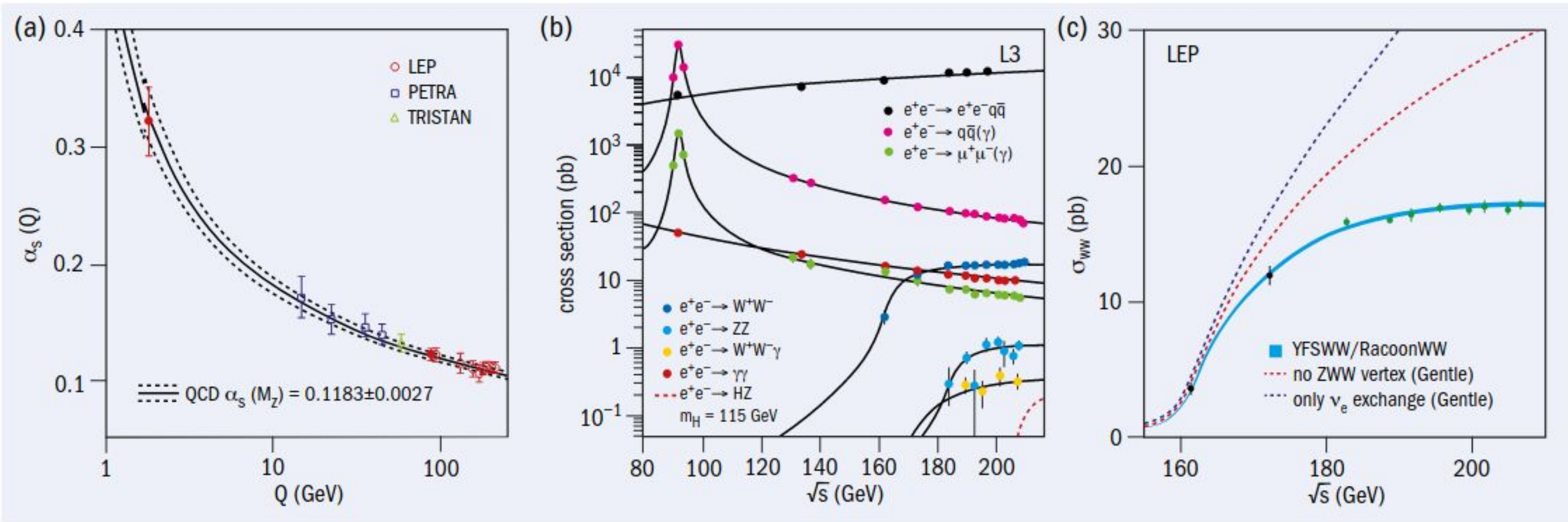
3. FCC-ee

1. LEP

LEP's electroweak leap, CERNCOURIER

CERNCOURIER.COM

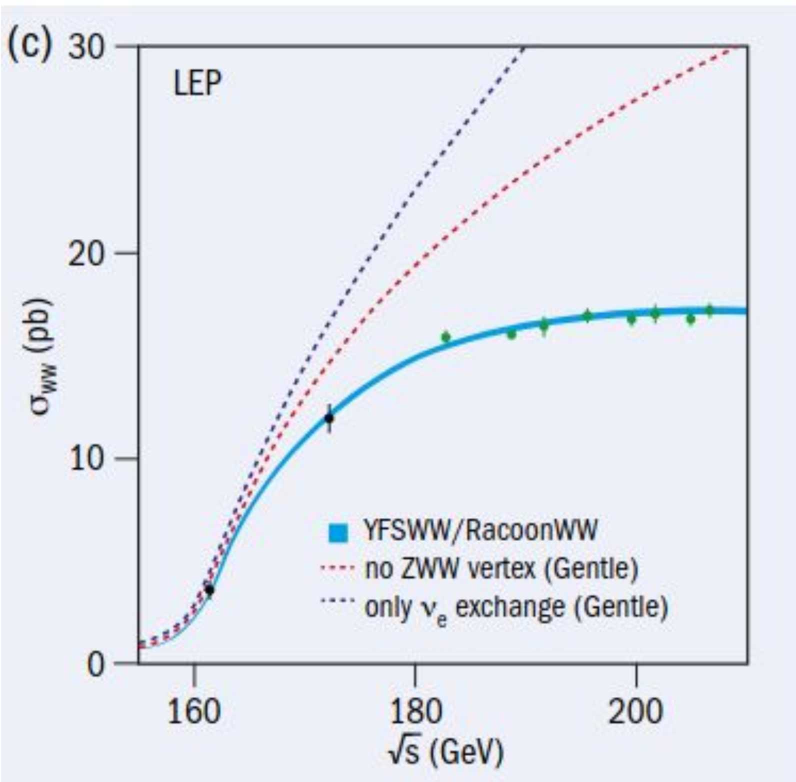
FEATURE LEP'S PHYSICS LEGACY



1. LEP legacy MC

FEATURE LEP'S PHYSICS LEGACY

Example



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

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

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]

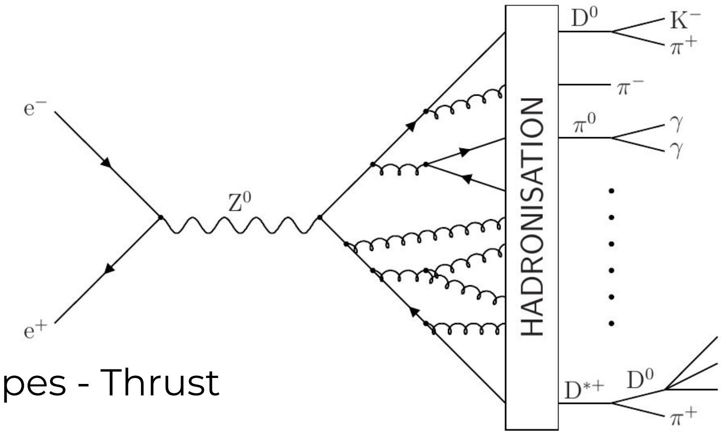
- Herwig 6
- PYTHIA 6/JETSET
  Ariadne

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

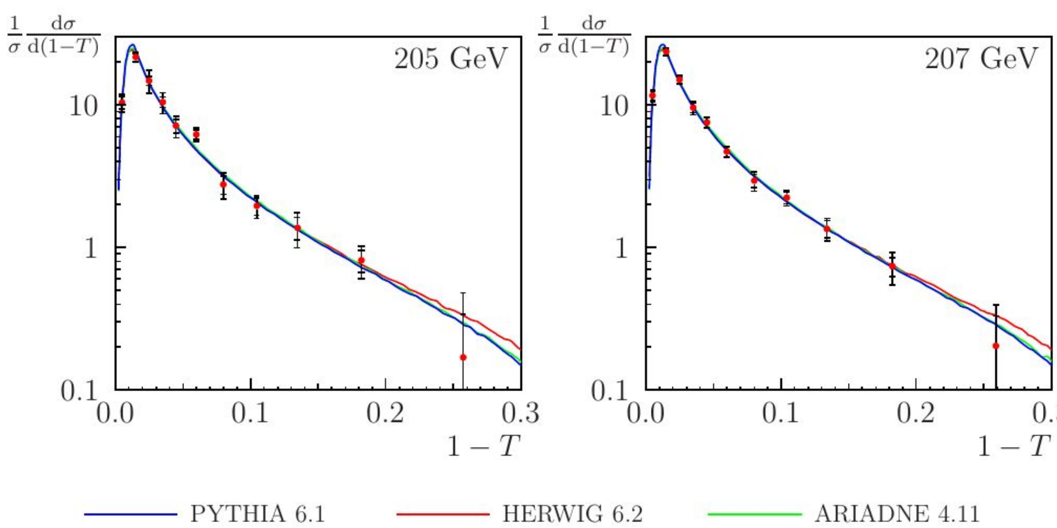
1. LEP legacy MC

Example:
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]."



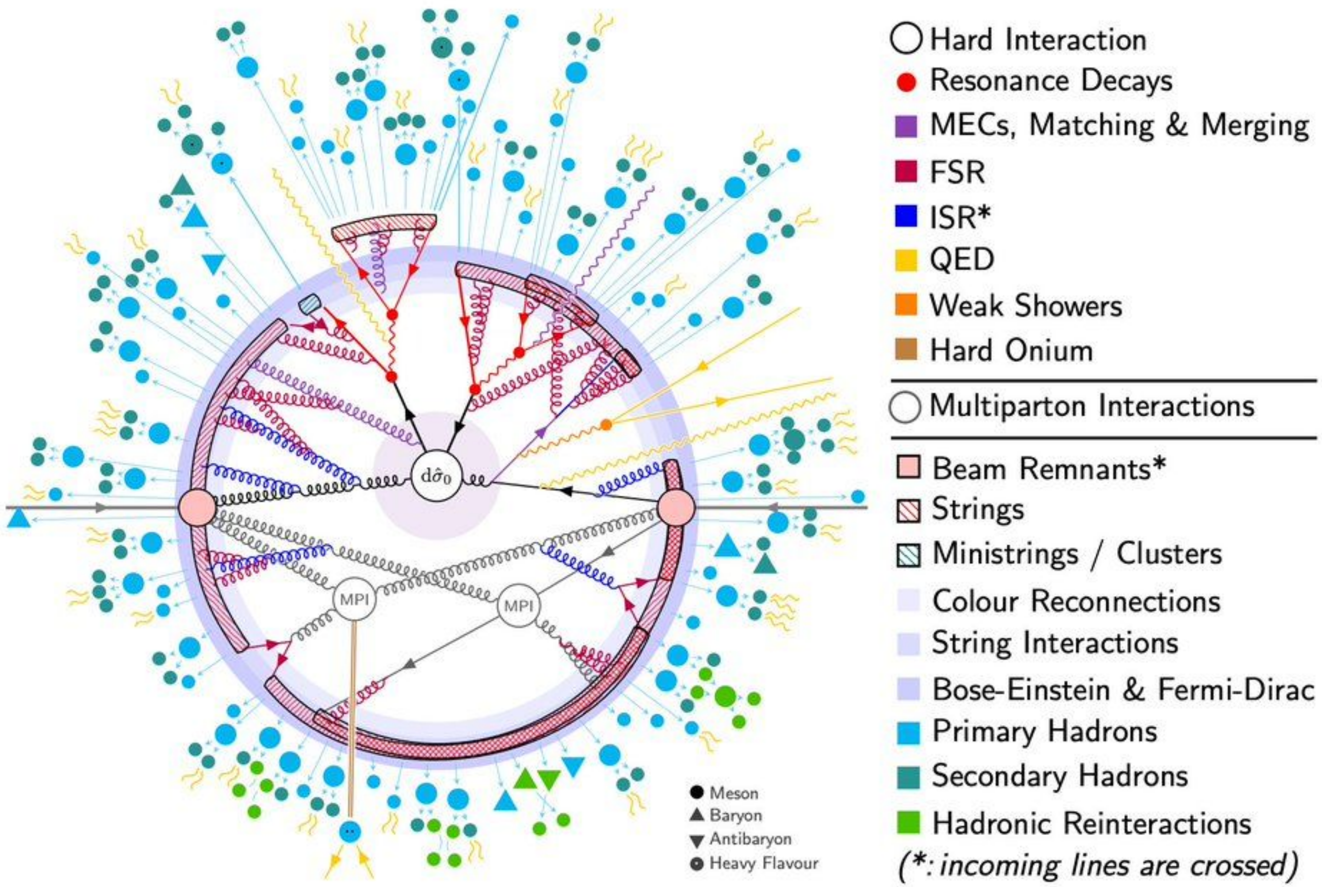
Hadronic Event Shapes - Thrust



2. LHC MC

LHC: QCD machine

Main progress in General purpose MC



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 7.3** [Bewick, Ravasio, Gieseke, Kiebacher,, Masouminia. Nail, Papaefstathiou,, Richardson, Samitz, Seymour, **AS,** Whitehead, *Eur.Phys.J.C* 84 (2024)]

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., SciPost Phys.Codeb. (2022)]

Sherpa 2: Begun in 2000. Originated in "matching" of matrix elements to showers: CKKW. Last version **Sherpa 3**: [Bothmann, Flower, Gutschow, Höche, Hoppe, Isaacson, Knobbe, Krauss, Meinzinger, Napoletano, Price, Reichelt, Schönherr, Schumann, Siegert, JHEP 12(2024)]

2. LHC MC

General purpose MC the Workhorses of the LHC:

How much progress has been made?

C++

Herwig 7: Successor to EARWIG (begun in 1984). Originated in coherence studies: angular ordering parton shower. Cluster model. Last version: **Herwig 7.3** [Bewick, Ravasio, Gieseke, Kiebacher,, Masouminia. Nail, Papaefstathiou,, Richardson, Samitz, Seymour, **AS**, Whitehead, *Eur.Phys.J.C* 84 (2024)]

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., SciPost Phys.Codeb. (2022)]

Sherpa 2: Begun in 2000. Originated in ”matching” of matrix elements to showers: CKKW. Last version **Sherpa 3**: [Bothmann, Flower, Gutschow, Höche, Hoppe, Isaacson, Knobbe, Krauss, Meinzinger, Napoletano, **Price**, Reichelt, Schönherr, Schumann, Siegert, JHEP 12(2024)]

2. LHC MC

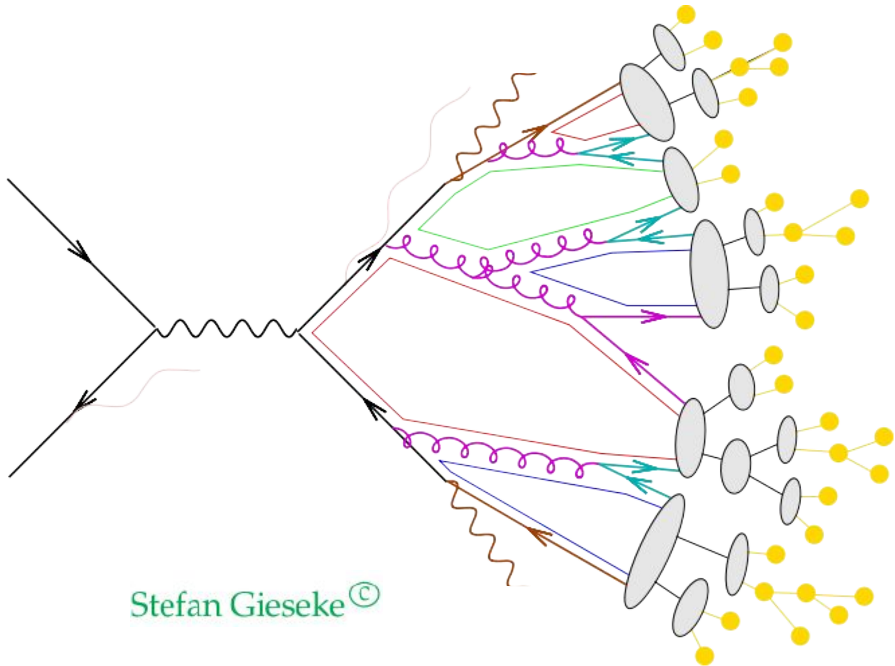
QCD correctly describes strong interactions in each energy range but its complex mathematical structure makes it very difficult to obtain precise predictions (Millennium Prize Problem $1,000,000)

High energy

- **perturbative QCD**
- **in theory we know what to do**
- **in practice very difficult**

Low energy

- **non-perturbative QCD**
- **we don't know what to do**
- **phenomenological models (with many free parameters)**



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

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]

NLO Parton Shower:

- NLO corrections to hard process in QCD shower -- proof of concept
  [S. Jadach , M. Jezabek, A. Kusina, W. Placzek, M. Skrzypek, APPB 43 (2012)]
  ....
- ”Shower Thoughts About Precision LHC Event simulations” – STAPLE
  ERC Grant Starting Grant 2025, Rene Poncelet, IFJ PAN

2. LHC MC

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, **AS**, 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
[“Factorisation schemes for proton PDFs” Delorme, Kusina, **AS**, Whitehead, EPJC 85 (2025)]

More processes (WW, gamma gamma, …) automatization
[P. Sarmah, **AS**, J. Whitehead JHEP 2025 062]

Universality

Higher precision

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

KrKNLO - di-photon (1st process not used to define MC PDF)

[P. Sarmah, **AS**, J. Whitehead JHEP 2025 062]

KrkNLO matching

$$
\begin{aligned}
\mathrm{d}\hat{\sigma}_{q\bar{q}}^{\mathrm{KrkNLO}(1)}[\mathcal{O}] = \mathrm{d}\Phi_m \frac{1}{2\hat{s}_{12}} \Bigg[& \boxed{\mathrm{B}_{q\bar{q}}(\Phi_m)\,\Delta^{(1)}\Big|_{p_\mathrm{T}^\mathrm{cut}}^{Q(\Phi_m)}} \\
& + \Big\{ \mathrm{V}_{q\bar{q}}(\Phi_m;\mu_\mathrm{R}) + \mathrm{I}_{q\bar{q}}(\Phi_m;\mu_\mathrm{R}) + \Delta_0^\mathrm{Krk}\,\mathrm{B}_{q\bar{q}}(\Phi_m) \Big\} \Delta^{(0)}\Big|_{p_\mathrm{T}^\mathrm{cut}}^{Q(\Phi_m)} \Bigg] \Theta_\mathrm{cut}\left[\Phi_m\right] \mathcal{O}(\Phi_m) \\
& + \mathrm{d}\Phi_{m+1} \frac{1}{2\hat{s}_{12}} \left[\frac{\mathrm{R}_{q\bar{q}}(\Phi_{m+1})}{\mathrm{B}_{q\bar{q}}(\tilde{\Phi}_m^{\mathrm{II}_1})\,\mathrm{S}^{q_1 g}(x)\,\Theta_\mathrm{cut}\left[\tilde{\Phi}_m^{\mathrm{II}_1}\right] + \mathrm{B}_{q\bar{q}}(\tilde{\Phi}_m^{\mathrm{II}_2})\,\mathrm{S}^{q_2 g}(x)\,\Theta_\mathrm{cut}\left[\tilde{\Phi}_m^{\mathrm{II}_2}\right]} \right. \\
& \left. \times \left(\sum_{i=1}^{2} \Theta_\mathrm{cut}\left[\tilde{\Phi}_m^{\mathrm{II}_i}\right] \Theta_{p_\mathrm{T}^\mathrm{cut}}^{Q(\tilde{\Phi}_m^{\mathrm{II}_i})}(\tilde{\Phi}_m^{\mathrm{II}_i})\,\Delta^{(0)}\Big|_{p_{\mathrm{T},1}}^{Q(\tilde{\Phi}_m^{\mathrm{II}_i})} \Delta^{(0)}\Big|_{p_\mathrm{T}^\mathrm{cut}}^{p_{\mathrm{T},1}}(\Phi_{m+1})\,\mathrm{B}_{q\bar{q}}(\tilde{\Phi}_m^{\mathrm{II}_i})\,\mathrm{S}^{q_i g}(x) \right) \right] \mathcal{O}(\Phi_{m+1})
\end{aligned}
$$

for all Born events **do** shower

  if first emission generated, from kernel (α) **then**

    $w \leftarrow w \times \dfrac{R(\Phi_{m+1})}{P_m^{(\alpha)}(\Phi_{m+1})}$

  end if

  $w \leftarrow w \times \left[1 + \dfrac{\alpha_\mathrm{s}(\mu_\mathrm{R})}{2\pi}\left(\dfrac{\mathrm{V}(\Phi_m;\mu_\mathrm{R})}{\mathrm{B}(\Phi_m)} + \dfrac{\mathrm{I}(\Phi_m;\tilde{\mu}_\mathrm{R})}{\mathrm{B}(\Phi_m)} + \Delta_0^\mathrm{FS}\right)\right]$

end for all

KrKNLO - di-photon (1st process not used to define MC PDF)

[P. Sarmah, **AS**, J. Whitehead JHEP 2025 062]

KrkNLO matching

$$\mathrm{d}\hat{\sigma}_{q\bar{q}}^{\mathrm{KrkNLO}(1)}[\mathcal{O}] = \mathrm{d}\Phi_m \frac{1}{2\hat{s}_{12}} \Bigg[\mathrm{B}_{q\bar{q}}(\Phi_m)\, \Delta^{(1)}\Big|_{p_{\mathrm{T}}^{\mathrm{cut}}}^{Q(\Phi_m)} + \Big\{ \mathrm{V}_{q\bar{q}}(\Phi_m;\mu_{\mathrm{R}}) + \mathrm{I}_{q\bar{q}}(\Phi_m;\mu_{\mathrm{R}}) + \Delta_0^{\mathrm{Krk}}\, \mathrm{B}_{q\bar{q}}(\Phi_m) \Big\} \Delta^{(0)}\Big|_{p_{\mathrm{T}}^{\mathrm{cut}}}^{Q(\Phi_m)} \Bigg] \Theta_{\mathrm{cut}}\left[\Phi_m\right] \mathcal{O}(\Phi_m)$$

$$+ \mathrm{d}\Phi_{m+1} \frac{1}{2\hat{s}_{12}} \Bigg[\frac{\mathrm{R}_{q\bar{q}}(\Phi_{m+1})}{\mathrm{B}_{q\bar{q}}(\tilde{\Phi}_m^{\mathrm{II}_1})\, \mathrm{S}^{q_1 g}(x)\, \Theta_{\mathrm{cut}}\left[\tilde{\Phi}_m^{\mathrm{II}_1}\right] + \mathrm{B}_{q\bar{q}}(\tilde{\Phi}_m^{\mathrm{II}_2})\, \mathrm{S}^{q_2 g}(x)\, \Theta_{\mathrm{cut}}\left[\tilde{\Phi}_m^{\mathrm{II}_2}\right]} \times \left(\sum_{i=1}^{2} \Theta_{\mathrm{cut}}\left[\tilde{\Phi}_m^{\mathrm{II}_i}\right] \Theta_{p_{\mathrm{T}}^{\mathrm{cut}}}^{Q(\tilde{\Phi}_m^{\mathrm{II}_i})}(\tilde{\Phi}_m^{\mathrm{II}_i})\, \Delta^{(0)}\Big|_{p_{\mathrm{T},1}}^{Q(\tilde{\Phi}_m^{\mathrm{II}_i})} \Delta^{(0)}\Big|_{p_{\mathrm{T}}^{\mathrm{cut}}}^{p_{\mathrm{T},1}}(\Phi_{m+1})\, \mathrm{B}_{q\bar{q}}(\tilde{\Phi}_m^{\mathrm{II}_i})\, \mathrm{S}^{q_i g}(x) \right) \Bigg] \mathcal{O}(\Phi_{m+1})$$

for all Born events **do** shower
  if first emission generated, from kernel (α) **then**
    $w \leftarrow w \times \dfrac{R(\Phi_{m+1})}{P_m^{(\alpha)}(\Phi_{m+1})}$
  end if
  $w \leftarrow w \times \left[1 + \dfrac{\alpha_s(\mu_{\mathrm{R}})}{2\pi}\left(\dfrac{\mathrm{V}(\Phi_m;\mu_{\mathrm{R}})}{\mathrm{B}(\Phi_m)} + \dfrac{\mathrm{I}(\Phi_m;\tilde{\mu}_{\mathrm{R}})}{\mathrm{B}(\Phi_m)} + \Delta_0^{\mathrm{FS}}\right)\right]$
end for all

KrKNLO - di-photon (1st process not used to define MC PDF)

[P. Sarmah, **AS**, J. Whitehead JHEP 2025 062]

KrkNLO matching

$$\mathrm{d}\hat{\sigma}_{q\bar{q}}^{\mathrm{KrkNLO}(1)}[\mathcal{O}] = \mathrm{d}\Phi_m \frac{1}{2\hat{s}_{12}} \Bigg[\mathrm{B}_{q\bar{q}}(\Phi_m)\, \Delta^{(1)}\Big|_{p_{\mathrm{T}}^{\mathrm{cut}}}^{Q(\Phi_m)}$$

$$+ \Big\{ \mathrm{V}_{q\bar{q}}(\Phi_m;\mu_{\mathrm{R}}) + \mathrm{I}_{q\bar{q}}(\Phi_m;\mu_{\mathrm{R}}) + \Delta_0^{\mathrm{Krk}}\, \mathrm{B}_{q\bar{q}}(\Phi_m) \Big\} \Delta^{(0)}\Big|_{p_{\mathrm{T}}^{\mathrm{cut}}}^{Q(\Phi_m)} \Bigg] \Theta_{\mathrm{cut}}\left[\Phi_m\right] \mathcal{O}(\Phi_m)$$

$$+ \mathrm{d}\Phi_{m+1} \frac{1}{2\hat{s}_{12}} \Bigg[\frac{\mathrm{R}_{q\bar{q}}(\Phi_{m+1})}{\mathrm{B}_{q\bar{q}}(\tilde{\Phi}_m^{\mathrm{II}_1})\, \mathrm{S}^{q_1 g}(x)\, \Theta_{\mathrm{cut}}\left[\tilde{\Phi}_m^{\mathrm{II}_1}\right] + \mathrm{B}_{q\bar{q}}(\tilde{\Phi}_m^{\mathrm{II}_2})\, \mathrm{S}^{q_2 g}(x)\, \Theta_{\mathrm{cut}}\left[\tilde{\Phi}_m^{\mathrm{II}_2}\right]}$$

$$\times \left(\sum_{i=1}^{2} \Theta_{\mathrm{cut}}\left[\tilde{\Phi}_m^{\mathrm{II}_i}\right] \Theta_{p_{\mathrm{T}}^{\mathrm{cut}}}^{Q(\tilde{\Phi}_m^{\mathrm{II}_i})}(\tilde{\Phi}_m^{\mathrm{II}_i})\, \Delta^{(0)}\Big|_{p_{\mathrm{T},1}}^{Q(\tilde{\Phi}_m^{\mathrm{II}_i})} \Delta^{(0)}\Big|_{p_{\mathrm{T}}^{\mathrm{cut}}}^{p_{\mathrm{T},1}} (\Phi_{m+1})\, \mathrm{B}_{q\bar{q}}(\tilde{\Phi}_m^{\mathrm{II}_i})\, \mathrm{S}^{q_i g}(x) \right) \Bigg] \mathcal{O}(\Phi_{m+1})$$

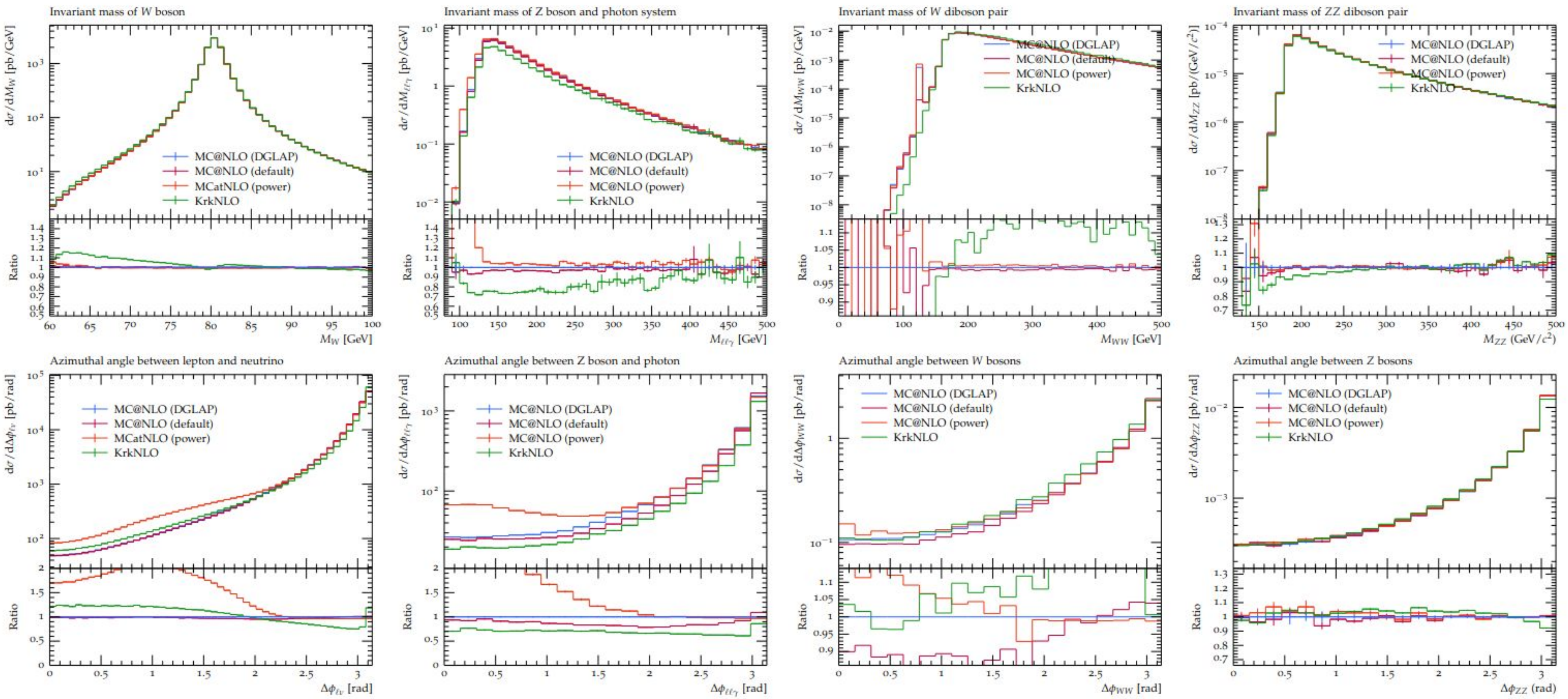
```
for all Born events do shower
  if first emission generated, from kernel (α) then
    w ← w × R(Φ_{m+1}) / P_m^{(α)}(Φ_{m+1})
  end if
  w ← w × [1 + α_s(μ_R)/(2π) ( V(Φ_m; μ_R)/B(Φ_m) + I(Φ_m; μ̃_R)/B(Φ_m) + Δ_0^{FS} )]
end for all
```

Phenomenology (coming soon)

Massive vector boson phenomenology with KrkNLO

Pratixan Sarmah,[a] Andrzej Siódmok,[a] James Whitehead[a]

W | Z+gamma | WW | ZZ

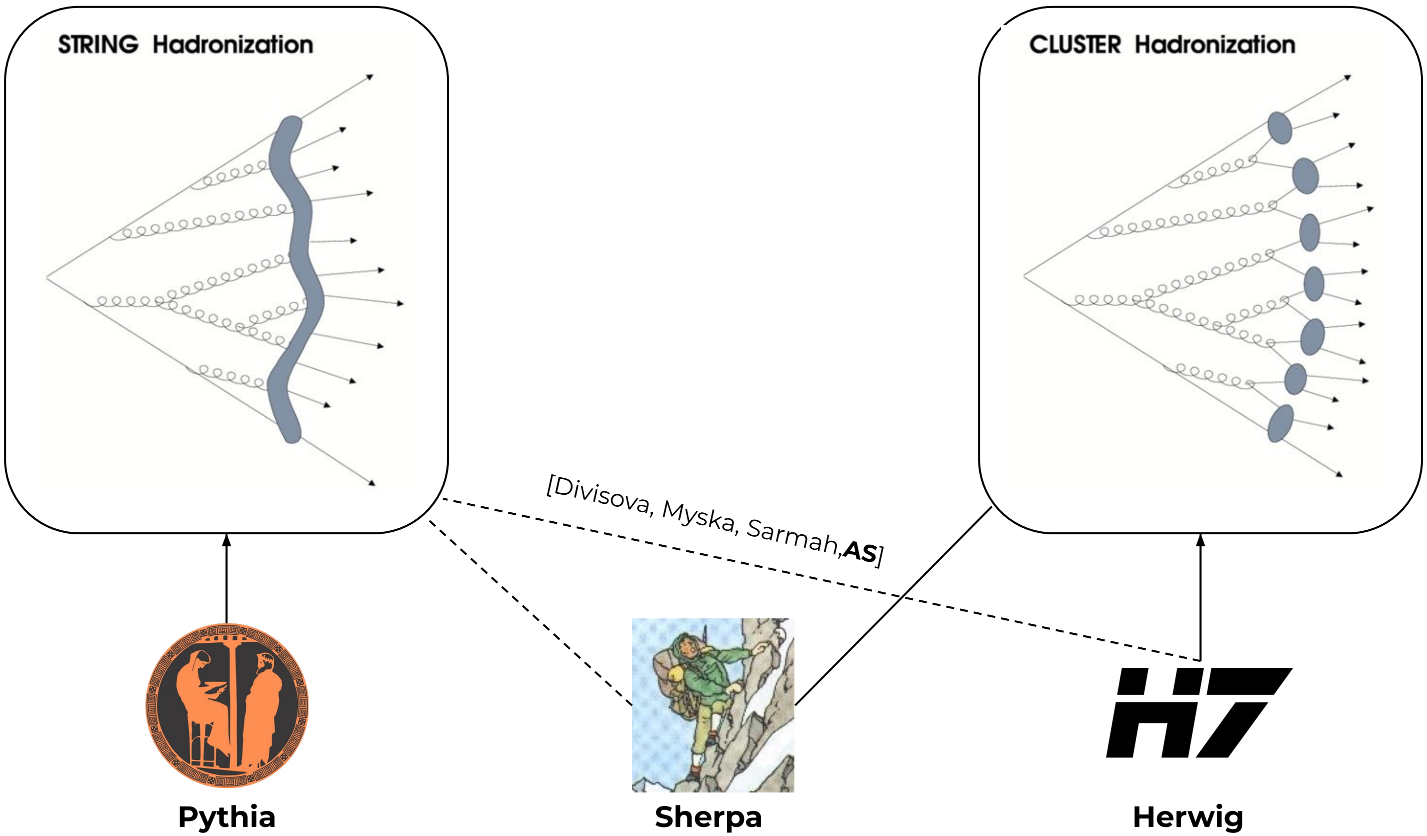


2. LHC MC

Hadronization:

➔ Increased control of perturbative corrections ⇒ more often LHC measurements are limited by non-perturbative components, such as hadronization.

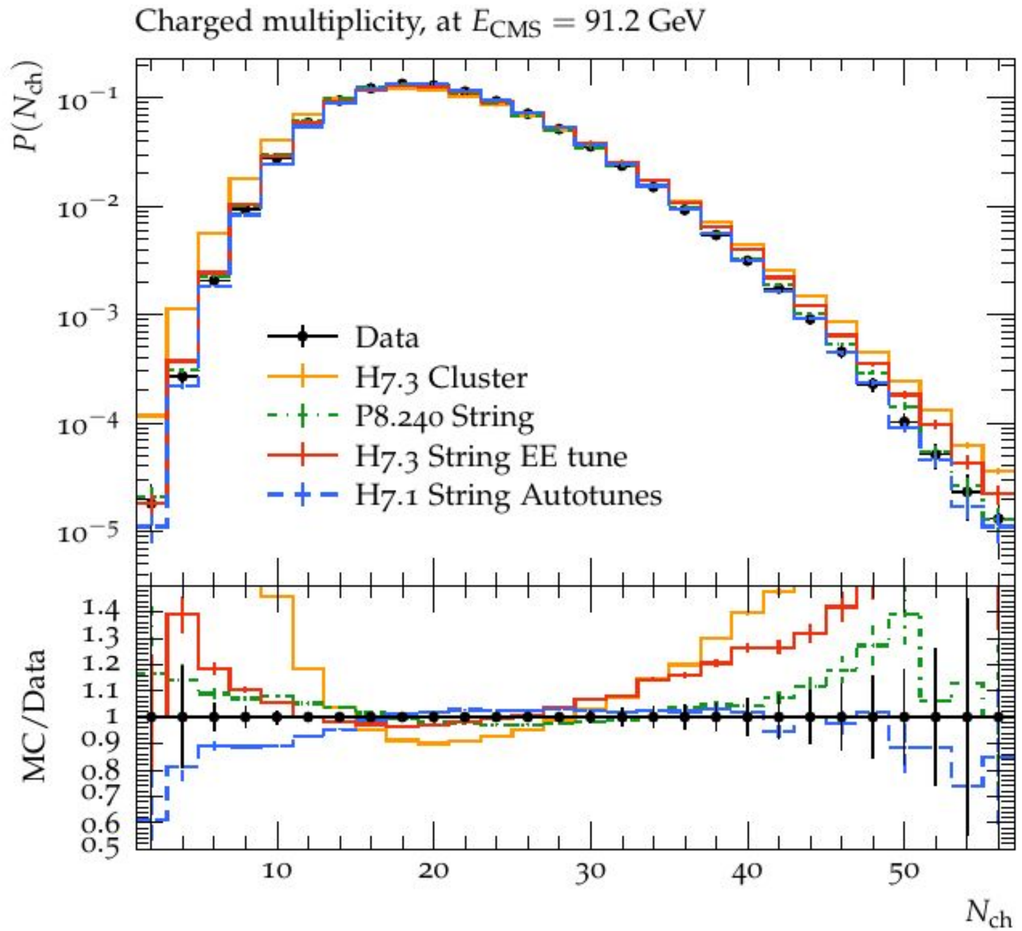
- W mass measurement using a new method [Freytsis at al. JHEP 1902 (2019) 003]
- Extraction of the strong coupling in [M. Johnson, D. Maître, Phys.Rev. D97 (2018) no.5]
- Top mass [S. Argyropoulos, T. Sjöstrand, JHEP 1411 (2014) 043]
- ...



2. LHC MC

Herwig 7 with the Lund String Model: Tuning and Comparative Hadronization Studies

Michaela Divisova[a,1], Miroslav Myska[b,1], Pratixan Sarmah[c,2], Andrzej Siódmok[d,2,3]

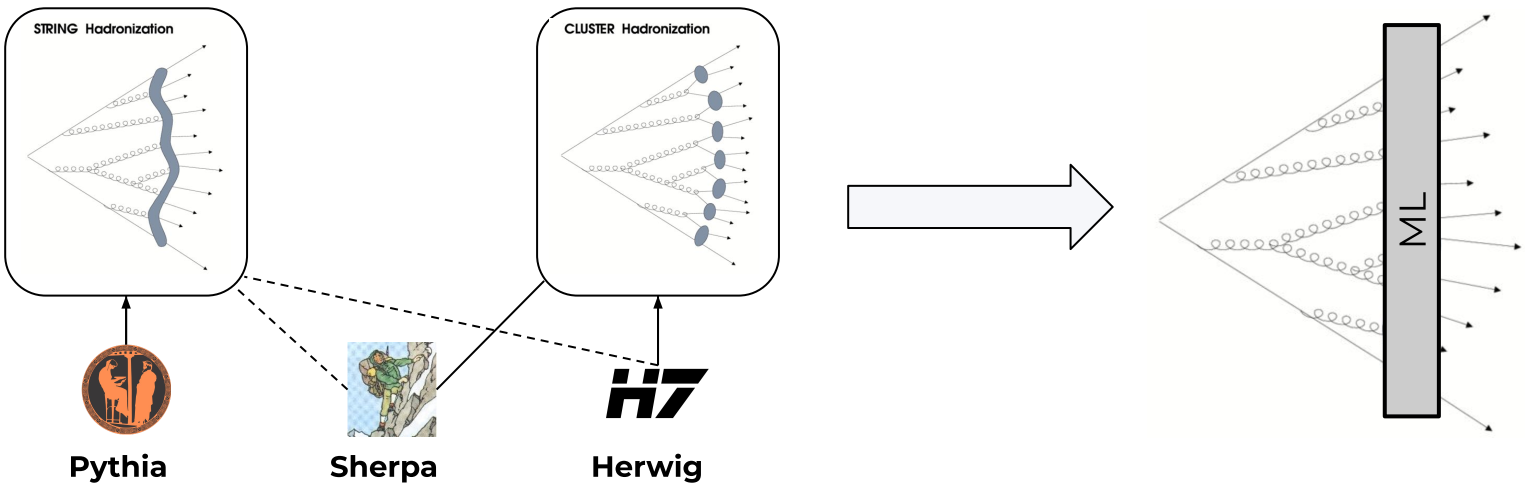


2. LHC MC

Hadronization:

Early 1980's
(limited progress)

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) for hadronization.

NNPDF used successfully ML to nonperturbative Parton Density Functions (PDF)

Recent progress: Machine learning hadronization

First steps for ML hadronization:

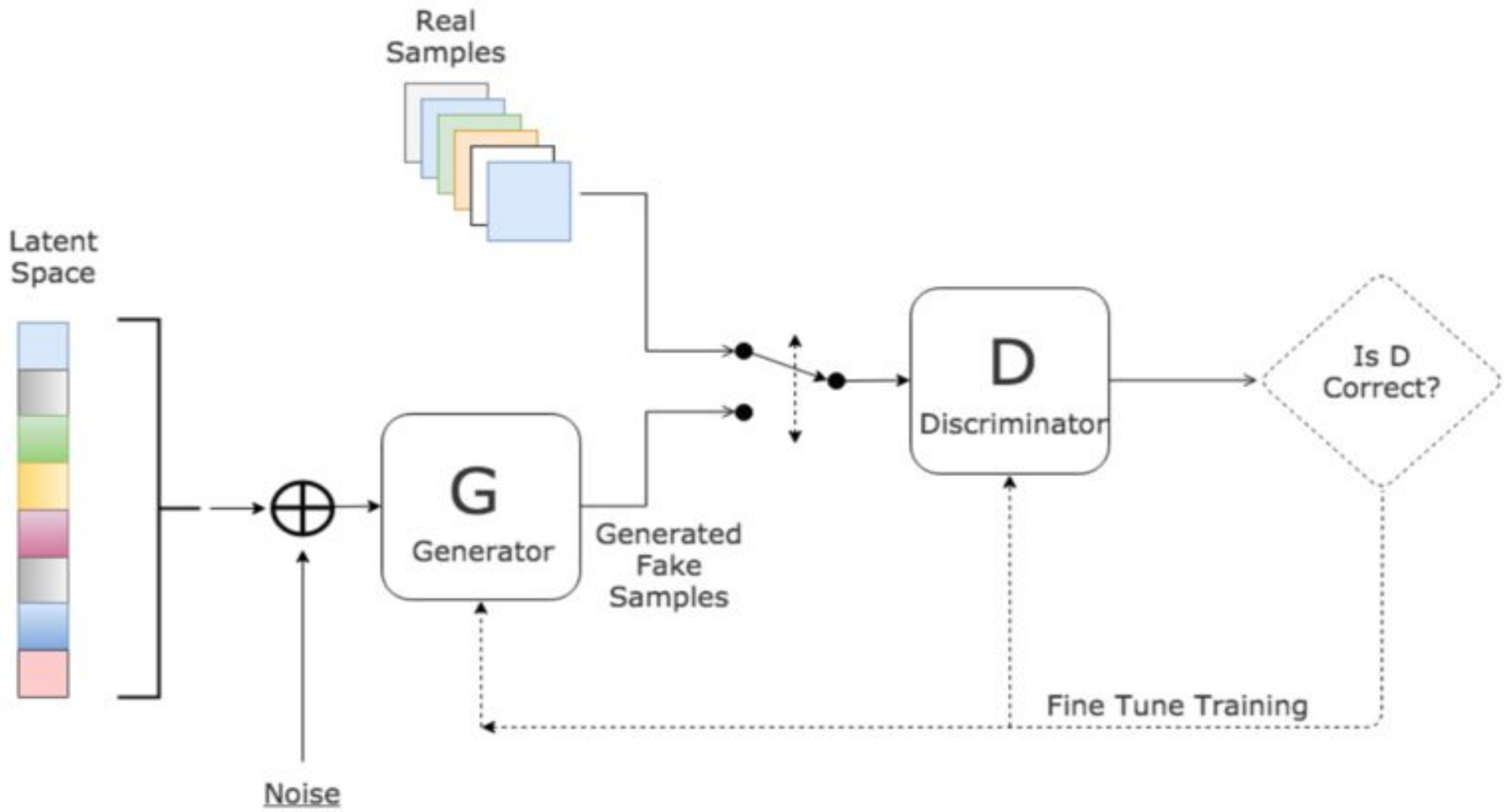
- HADML - [A. Ghosh, Xi. Ju, B. Nachman **AS**, *Phys.Rev.D* 106 (2022) 9]
- MLhad - [P. Ilten, T. Menzo, A. Youssef and J. Zupan, SciPost Phys. 14, 027 (2023)]

	MLhad	HADML
Deep generative model:	Variational Autoencoder	Generative Adversarial Networks
Trained on:	String model	Cluster model
Recent progress:	*"Reweighting Monte Carlo Predictions and Automated Fragmentation Variations in Pythia 8"* [Bierlich, Ilten, Menzo, Mrenna, Szewc, Wilkinson, Youssef, Zupan, 2308.13459] ...	*"Fitting a Deep Generative Hadronization Model"* [J. Chan, X. Ju, A. Kania, B. Nachman, V. Sangli and **AS**, JHEP 09 (2023) 084] ...

MLhadML (2411.02194)

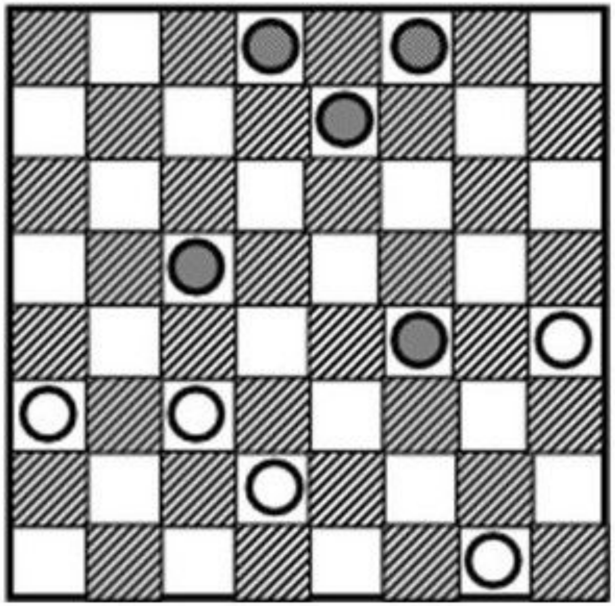
Generative Adversarial Network (GAN)

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



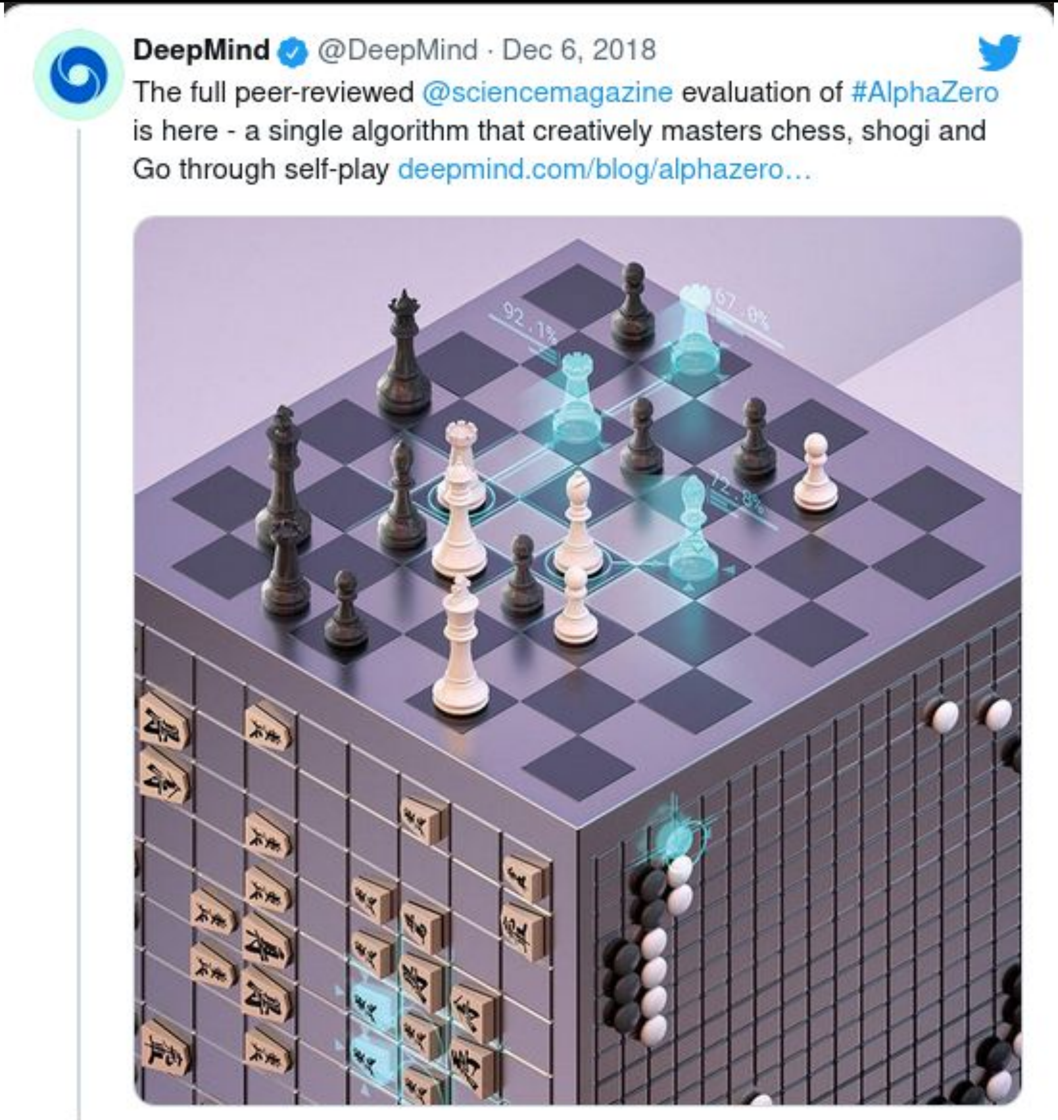
Adversarial Networks

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.

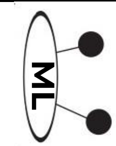
Adversarial Networks



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

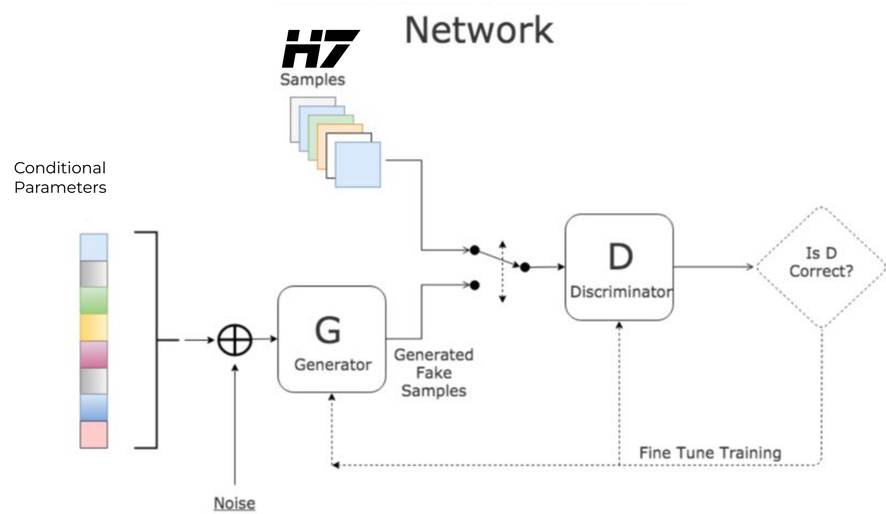
Towards a Deep Learning Model for Hadronization

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



How?

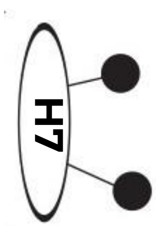
Generative Adversarial



Training data:

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

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



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

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

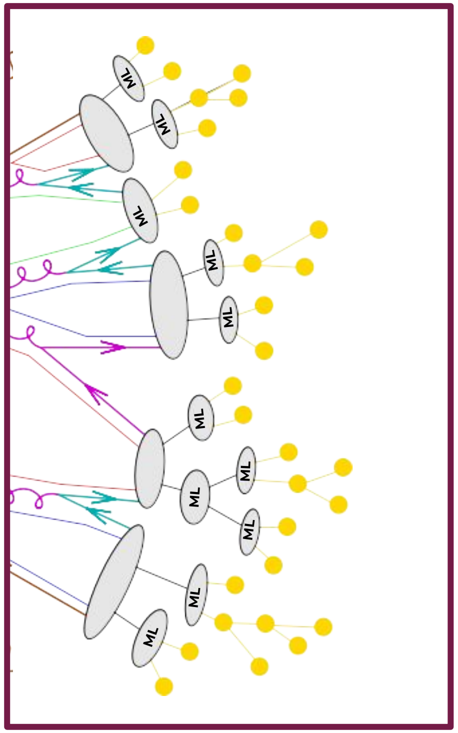
Pert = 0/1 memory of quarks direction

Results [Ghosh, Nachman, **AS**, Yu, Phys.Rev.D 106 (2022)]

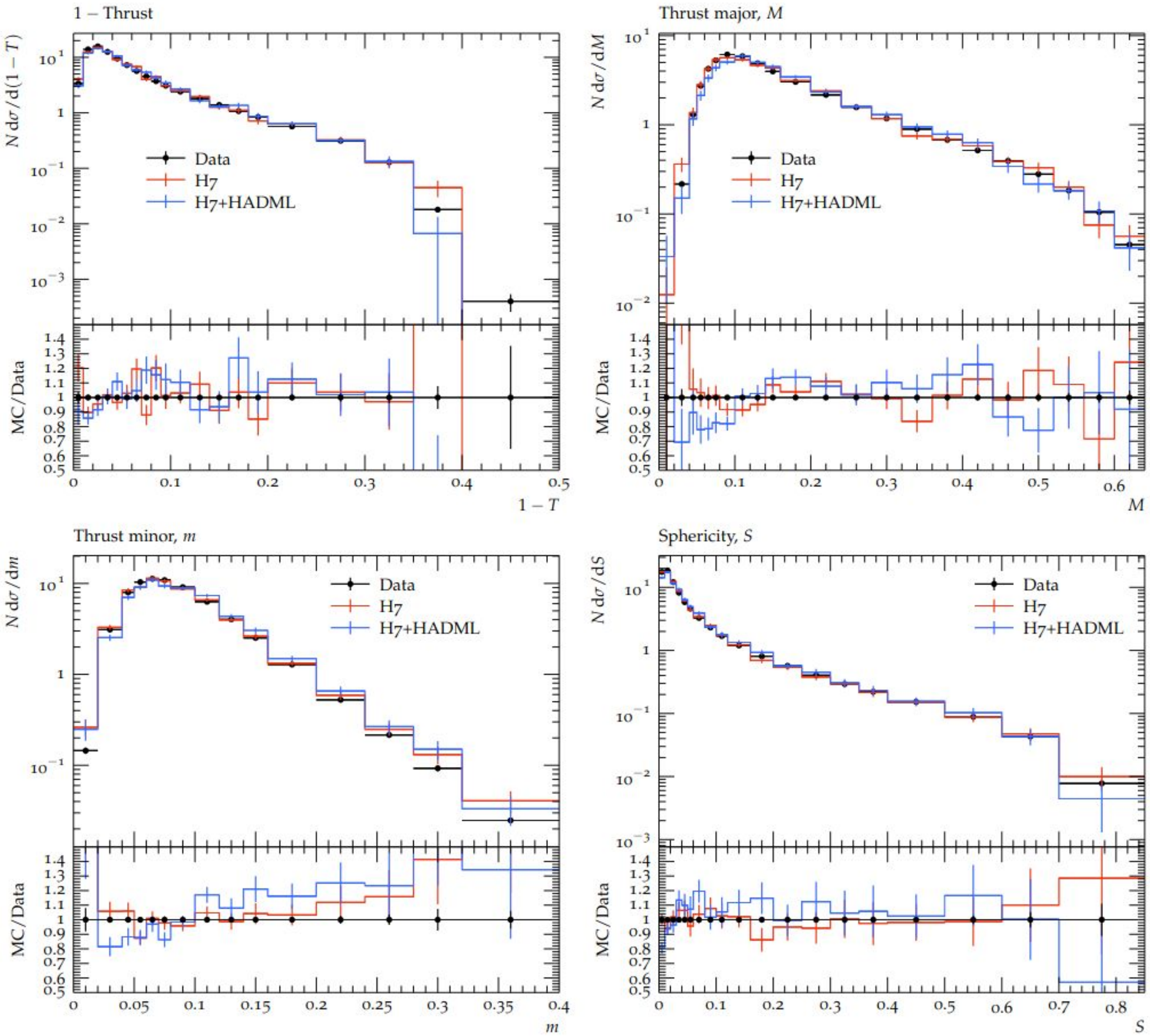
Full-event Validation

(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



HadML

A Novel Deep Generative Approach to Hadronisation

HADML

Towards a Deep Learning Model for Hadronization

Aishik Ghosh,[a,b] **Xiangyang Ju,**[b] **Benjamin Nachman,**[b,c] **and Andrzej Siodmok**[d]

[a] *Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA*
[b] *Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[c] *Berkeley Institute for Data Science, University of California, Berkeley, CA 94720, USA*
[d] *Jagellonian University, Krakow, Poland*

Phys.Rev.D 106 (2022)

Fitting a Deep Generative Hadronization Model

Jay Chan,[a,b] **Xiangyang Ju,**[b] **Adam Kania,**[e] **Benjamin Nachman,**[b,c] **Vishnu Sangli,**[d,b] **and Andrzej Siodmok**[d]

[a] *Department of Physics, University of Wisconsin-Madison, Madison, WI 53706, USA*
[b] *Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[c] *Berkeley Institute for Data Science, University of California, Berkeley, CA 94720, USA*
[d] *Department of Physics, University of California, Berkeley, CA 94720, USA*
[e] *Jagiellonian University, Krakow, Poland*

JHEP 09 (2023) 084

Integrating Particle Flavor into Deep Learning Models for Hadronization

Jay Chan,[a] **Xiangyang Ju,**[a] **Adam Kania,**[e] **Benjamin Nachman,**[b,c] **Vishnu Sangli,**[d,b] **and Andrzej Siodmok**[e]

[a] *Scientific Data Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[b] *Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[c] *Berkeley Institute for Data Science, University of California, Berkeley, CA 94720, USA*
[d] *Department of Physics, University of California, Berkeley, CA 94720, USA*
[e] *Jagiellonian University, Krakow, Poland*

Phys.Rev.D (2025)

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, **AS**, 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:

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

WINHAC [Jadach, Placzek]: W mass measurement

KKMChh [Jadach, Ward, Yost]: $Z \to \ell\ell + \gamma\gamma$

2. LHC MC

Specialized programs

Photos [N. Davidson, T. Przedzinski, Z. Was, CPC 199 (2016) 86-101]
[S. Antropov, Sw. Banerjee, Z. Was, J. Zaremba, 283 (2023) 108592]

- 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.

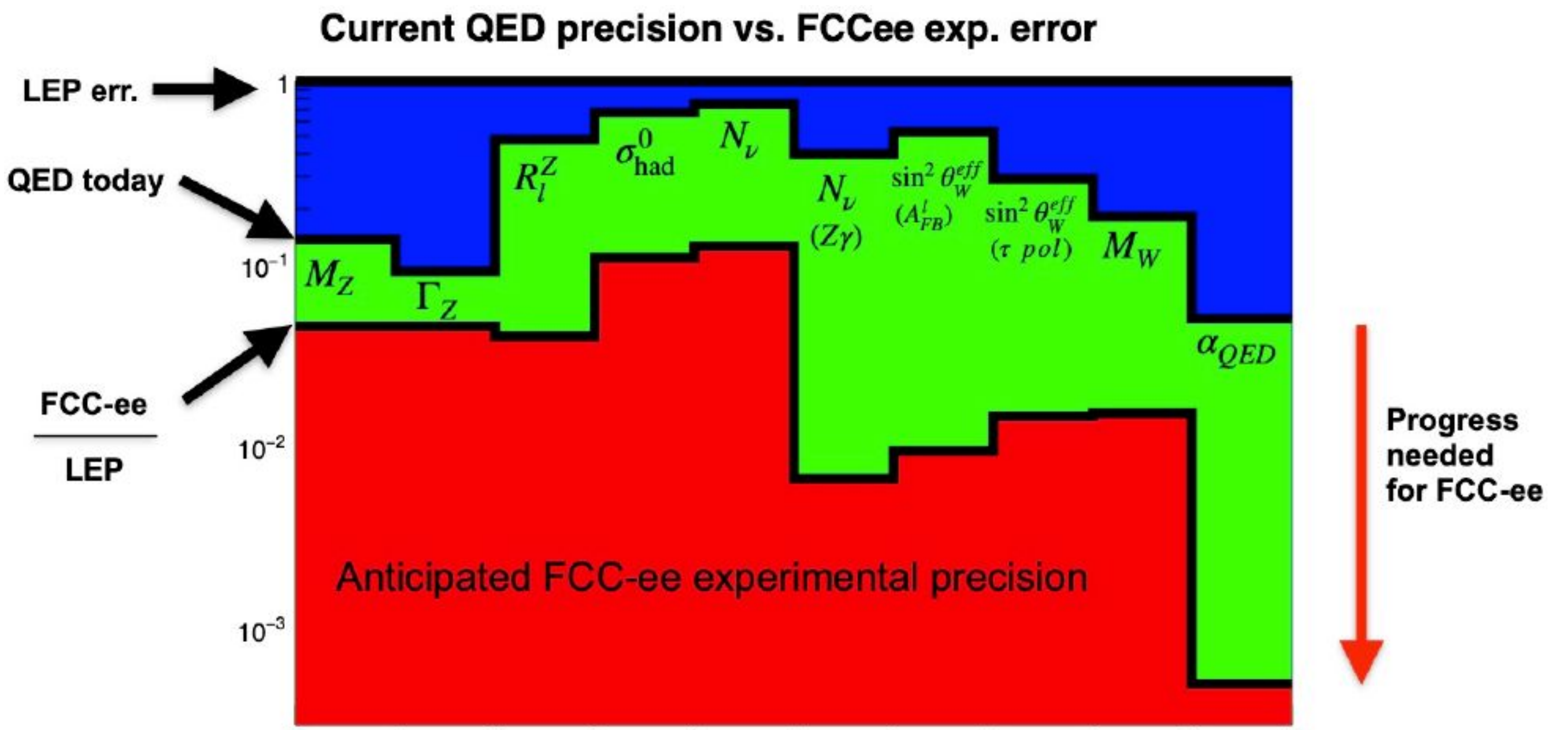
Tauola [S. Jadach, Z. Was, R. Decker, J. H. Kuhn, CPC 76 (1993) 361-380], ...,
[M. Chrzaszcz, T. Przedzinski, Z. Was, J. Zaremba, CPC 232 (2018) 220-236]

- 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.

See: ECFA Higgs Factories: 1st Topical Meeting on Generators

3. Future-ee

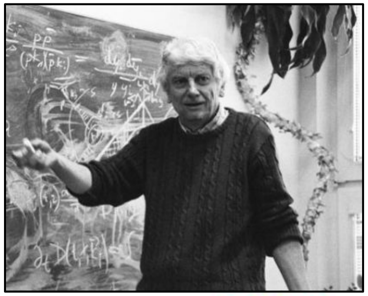
Future e+e- machine will be precision factory (luminosity up to 10^5 higher then LEP)



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.

Example: KKMCee **[slide from S. Jadach talk at Epiphany 2021]**



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

$$\mathfrak{M}_n^{(r)}(p,k_1,k_2,k_3,\dots,k_n) = \prod_{s=1}^{n} \mathfrak{s}(k_s) \Big\{ \hat\beta_0^{(r)}(p) + \sum_{j=1}^{n} \frac{\hat\beta_1^{(r)}(p,k_j)}{\mathfrak{s}(k_j)} + \sum_{j_1<j_2} \frac{\hat\beta_2^{(r)}(p,k_{j_1},k_{j_2})}{\mathfrak{s}(k_{j_1})\mathfrak{s}(k_{j_2})} + \sum_{j_1<j_2<j_3} \frac{\hat\beta_3^{(r)}(k_{j_1},k_{j_2},k_{j_3})}{\mathfrak{s}(k_{j_1})\mathfrak{s}(k_{j_2})\mathfrak{s}(k_{j_3})} + \sum_{j_1<j_2<\dots<j_r} \frac{\hat\beta_r^{(r)}(k_{j_1},k_{j_2},\dots,k_{j_r})}{\mathfrak{s}(k_{j_1})\mathfrak{s}(k_{j_2})\dots\mathfrak{s}(k_{j_r})} + \cdots \Big\}, \quad (26)$$

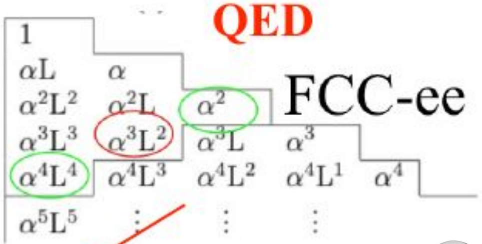
$\mathcal{O}(\alpha^3)$ ⟵ $\mathcal{O}(\alpha^r)$

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

$$\hat\beta_0^{(2)}(p) = \mathfrak{M}_0^{(2)}(p) = \left[e^{-\alpha B_4(p)} \mathcal{M}_0^{(2)}(p) \right]\Big|_{\mathcal{O}(\alpha^2)}, \quad (27)$$

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:

$$\hat\beta_1^{(2)}(p,k_1) = \mathfrak{M}_1^{(2)}(p,k_1) - \hat\beta_0^{(1)}(p)\mathfrak{s}(p,k_1), \quad \mathfrak{M}_1^{(2)}(p,k_1) = e^{-\alpha B_4(p)} \mathcal{M}_1^{(2)}(p,k_1)\Big|_{\mathcal{O}(\alpha^2)}. \quad (28)$$



To be added for FCCee

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

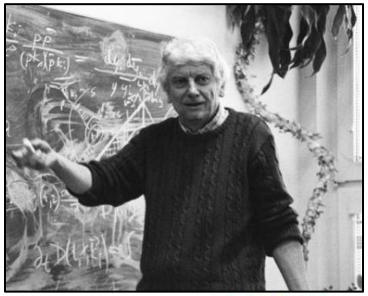
$\mathcal{O}(\alpha^2)$ 1-loop **EW**

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

JAGIELLONIAN UNIVERSITY IN KRAKÓW

3. Future-ee

Example: KKMCee **[slide from S. Jadach talk at Epiphany 2021]**



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

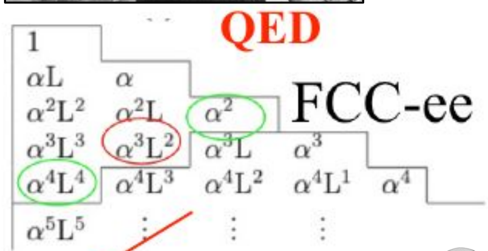
$$\sigma^{(r)} = \sum_{n=0}^{\infty} \frac{1}{n!} \int d\tau_n(p_1 + p_2;\ p_3, p_4,\ k_1, \ldots, k_n)$$
$$\times\ e^{2\alpha \Re B_4(p_a,\ldots,p_d)} \frac{1}{4} \sum_{\rm spin} \left| \mathfrak{M}_n^{(r)}(p, k_1, k_2, \ldots k_n) \right|^2,$$

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 $p_1, \ldots$ of the fermions are denoted collectively as p. The spin amplitudes read

Now in KKMC/CEEX ⟵

$$\mathfrak{M}_n^{(r)}(p, k_1, k_2, k_3, \ldots, k_n) = \prod_{s=1}^{n} \mathfrak{s}(k_s) \Big\{ \hat{\beta}_0^{(r)}(p) + \sum_{j=1}^{n} \frac{\hat{\beta}_1^{(r)}(p, k_j)}{\mathfrak{s}(k_j)} + \sum_{j_1<j_2} \frac{\hat{\beta}_2^{(r)}(p, k_{j_1}, k_{j_2})}{\mathfrak{s}(k_{j_1})\mathfrak{s}(k_{j_2})}$$
$$+ \sum_{j_1<j_2<j_3} \frac{\hat{\beta}_3^{(r)}(k_{j_1}, k_{j_2}, k_{j_3})}{\mathfrak{s}(k_{j_1})\mathfrak{s}(k_{j_2})\mathfrak{s}(k_{j_3})} + \sum_{j_1<j_2<\ldots<j_r} \frac{\hat{\beta}_r^{(r)}(k_{j_1}, k_{j_2}, \ldots, k_{j_r})}{\mathfrak{s}(k_{j_1})\mathfrak{s}(k_{j_2})\ldots\mathfrak{s}(k_{j_r})} + \cdots \Big\}, \quad \mathcal{O}(\alpha^3)\ \ \mathcal{O}(\alpha^r) \tag{26}$$



To be added for FCCee

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

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

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:

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

J. Gluza

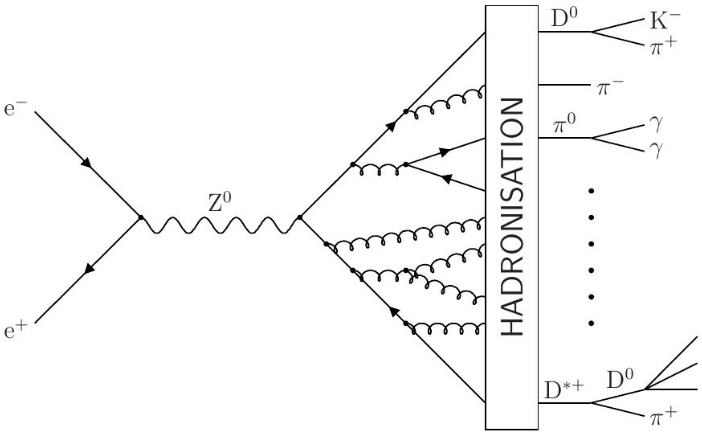
W. Płaczek, AS

M. Skrzypek, Z. Wąs

3. Future-ee

General purpose MC QCD developments

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

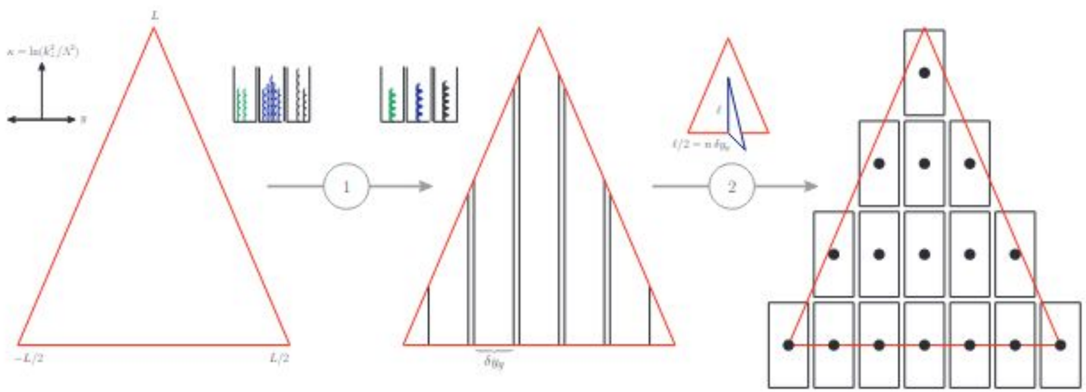


- Better Hadronization
  - lattice QCD?
  - ML?
  - improved string and cluster
  - new measurements

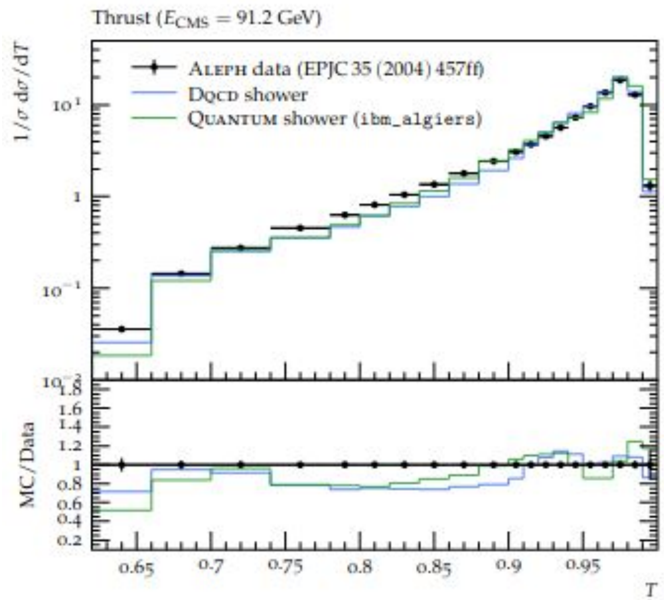
3. Future ee factory

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]

Conclusions

1. Monte Carlo generators are central to HEP.
2. Polish groups have strong position in development of key Monte Carlo generators used in HEP and our strategy should be to maintain the position
3. The progress in all kinds of generators will be needed for the current and future experiments
   - example FCCee: upgrade of LEP legacy MCs. However, for factor 50-150 improvement in precision one needs new innovative projects.
4. Will we be in better position in the future then at LEP?
   Not necessary MC generators are complex and in 30 years we might lost the know how...
   - inter-sectional character of the MCEG developers' (interface of experiment, theory, and computation) activity poses serious challenges in terms of career paths and funding opportunities
5. Future:
   - continuous dialogue with experimental community
   - more powerful computational techniques (ML) and 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 or Herwig, Pythia, Sherpa]