Lecture 2: Phenomenological aspects of Drell-Yan and Higgs production

Valerio Bertone IRFU, CEA, Université Paris-Saclay



September 09, 2021, MCnet-CTEQ School

Outline

- Historical importance of Drell-Yan production:
 - the discovery of charm, bottom, and Z/W weak bosons.
- Modern phenomenological implications of Drell-Yan production:
 - constraints on PDFs and the mass of the W boson,
 - importance of higher-order corrections,
 - resummation at work.
- Higgs production:
 - $gg \rightarrow H + X$ at fixed order and resummed,
 - some phenomenological implications of $gg \rightarrow H + X \rightarrow \gamma\gamma + X$

How much Drell-Yan and Higgs?

Drell-Yan production (Z and W):

- large cross section,
- clean final states (leptons and missing $E_{\rm T}$),
- an additional jet (*e.g.* for measuring the q_T) leaves the cross section large,
- $\sim 30(10)\%$ of W(Z)'s decay leptonically,
- this all allows for precise measurements.
- theoretically well-understood.

Higgs production:

- cross section much smaller ($gg \rightarrow H$ largest),
- harder to measure,
- harder to extract information.



How much Drell-Yan and Higgs?

Drell-Yan production (Z and W):

- large cross section,
- clean final states (leptons and missing $E_{\rm T}$),
- an additional jet (*e.g.* for measuring the q_T) leaves the cross section large,
- $\sim 30(10)\%$ of W(Z)'s decay leptonically,
- this all allows for precise measurements.
- theoretically well-understood.

Higgs production:

- cross section much smaller ($gg \rightarrow H$ largest),
- harder to measure,
- harder to extract information.



How much Drell-Yan and Higgs?

Drell-Yan production (Z and W):

- large cross section,
- clean final states (leptons and missing $E_{\rm T}$),
- an additional jet (*e.g.* for measuring the q_T) leaves the cross section large,
- $\sim 30(10)\%$ of W(Z)'s decay leptonically,
- this all allows for precise measurements.
- theoretically well-understood.

Higgs production:

- cross section much smaller ($gg \rightarrow H$ largest),
- harder to measure,
- harder to extract information.



The Drell-Yan mass spectrum



The Drell-Yan mass spectrum



The discovery of the charm: the J/ψ

In 1970, Glashow, Iliopoulos and Maiani (GIM) postulated the existence of a new quark favour: the **charm quark** [*Phys.Rev.D* 2 (1970) 1285-1292]:

- originally conceived to explain the suppression of FCNC processes (they only occur at one loop),
- on top of *u*, *d*, and *s*, GIM introduced a **4th quark** flavour to complete the second generation of quarks.
- Like other quarks, the charm can form bound states (resonances).
- In 1974, the BNL observed a very narrow resonance at m_{e+e⁻} = 3.1 GeV in the invariant mass of the electron pair in p+Be → e⁺ + e⁻ + X (Drell-Yan): the J(J/ψ) hadron was observed for the fist time and thus the charm discovered.
- The valence structure of the J/ψ is $c\overline{c}$. Therefore, the mass of the J/ψ suggests $m_c \sim 1.5$ GeV. The current PDG value is $m_c(m_c) = 1.27 \pm 0.02$ GeV.



The discovery of the bottom: the Υ

In 1977, the E288 experiment at Fermilab observed a resonance at $m_{\mu^+\mu^-} = 9.5$ GeV in the invariant mass of the muon pair in $p+(\text{Cu, Pt}) \rightarrow \mu^+ + \mu^- + X$ (**Drell-Yan**): the Υ hadron was observed for the first time and the **bottom** (or **beauty**) discovered.

- The valence structure of the Υ is $b\overline{b}$. Therefore, the mass of the Υ suggests $m_b \sim 4.7$ GeV. The current PDG value is $m_b(m_b) = 4.18^{+0.03}_{-0.02}$ GeV.
- The existence of a fifth quark flavour immediately triggered the hypothesis of a **sixth quark**, the **top**, to
- The top quark was discovered later in 1995 at Fermilab by the the **Tevatron** experiments CDF and D0.
- The presence of a third family and the consequent 3×3 mixing matrix (CKM) between up- and down-type quarks introduced the possibility of **CP violation** in the Standard Model.



The discovery of the Z and W bosons

The massive Z and W are perhaps the most direct manifestation of the **spontaneous breaking** of the $SU(2)_L$ gauge symmetry of the standard model:

• In 1983 both the UA1 [*Phys. Lett.* B 122, 103 (1983)] and the UA2 [*Phys. Lett.* B 122 476 (1983)] experiments at the $Sp\overline{p}S$ collider at CERN announced the discovery of the W^{\pm} bosons in $p\overline{p} \rightarrow e^{\pm}\nu_{e}(\overline{\nu}_{e}) + X$ collisions (**charged-current Drell-Yan**) with mass around 80 GeV.



The discovery of the Z and W bosons

The massive Z and W are perhaps the most direct manifestation of the **spontaneous breaking** of the $SU(2)_L$ gauge symmetry of the standard model:

• In 1983 both the UA1 [*Phys. Lett.* B 122, 103 (1983)] and the UA2 [*Phys. Lett.* B 122 476 (1983)] experiments at the $Sp\overline{p}S$ collider at CERN announced the discovery of the W^{\pm} bosons in $p\overline{p} \rightarrow e^{\pm}\nu_{e}(\overline{\nu}_{e}) + X$ collisions (**charged-current Drell-Yan**) with mass around 80 GeV.



The discovery of the Z and W bosons

The massive Z and W are perhaps the most direct manifestation of the **spontaneous breaking** of the $SU(2)_L$ gauge symmetry of the standard model:

• In 1983 both the UA1 [*Phys. Lett.* B 122, 103 (1983)] and the UA2 [*Phys. Lett.* B 122 476 (1983)] experiments at the $Sp\overline{p}S$ collider at CERN announced the discovery of the W^{\pm} bosons in $p\overline{p} \rightarrow e^{\pm}\nu_{e}(\overline{\nu}_{e}) + X$ collisions (charged-current Drell-Yan) with mass around 80 GeV.



The Drell-Yan mass spectrum for discovery



Drell-Yan in PDF determinations



Fixed-target Drell-Yan and sea PDFs

Very recently the SeaQuest (E906) experiment at Fermilab has released data for the ratio of cross sections σ_{pd}/σ_{pp} [*Nature* 590 (2021) 7847, 561-565].

This ratio is sensitive to the ratio of sea quark PDFs:

Can you derive this relation?

Being a fixed-target experiment, **large values of** *x* are probed giving us access to the sea quark PDFs in a region that is presently **poorly known**.

 $\frac{\sigma_{pd}}{\sigma_{pp}} \simeq 1 + \frac{a(x)}{\overline{u}(x)}$



Significant impact on the \overline{u} and \overline{d} PDFs at large x.

Currently unresolved tension with the older NuSea (E866) data.

Drell-Yan q_T distribution and the gluon PDF The q_T of the Z boson allows us to constraint the collinear gluon PDF:

- as we have seen in the previous lecture, collinear factorisation is reliable for $q_T \simeq Q$.
- In order for the Z to have a large q_T , it needs an object to recoil against. This is typically a jet. As a consequence, the relevant process is $pp \rightarrow Z + j + X$.
- One of the leading-order partonic cross sections contributing to this process is:



10²

M_x [GeV]

The impact on the gluon PDF is **significant**.

х

The photon PDF of the proton

If we promote the **photon** to be a **parton**, *i.e.* we allow the photon to contribute to the proton structure, then we need to allow for a **photon PDF**.



Therefore, Drell-Yan production regarded as the production of a lepton pair receives contribution also from the **photon** already at the **leading order**.

At small and middle leptonpair invariant masses $m_{\ell\ell}$ the photon-initiated contribution is suppressed by the photon PDF.

At high $m_{\ell\ell}$ the photon PDF becomes relatively larger:

• probing large x.

High-mass Drell-Yan data^{\mathbb{E}} enables us to determine the photon PDF.



The photon PDF of the proton

In [*Eur.Phys.J.C* 77 (2017) 6, 400] the high-mass ATLAS 8 TeV data was used to extract the photon PDF:

• QED/EW effects included up to NLO.



The mass of the W

A precise measurement of the Wmass would not only provide a strong test of the Standard Model but would also allow us to constrain possible extensions to it.

The most precise measurements of the *W* mass is achieved by fitting the p_T^{ℓ} and m_T in Drell-Yan production $pp \to W^{\pm} + X \to \ell^{\pm} \nu_{\ell}(\overline{\nu}_{\ell}) + X.$



and thus measure M_W .

 \mathbf{X} ALEPH 80.440 ±0.051 + M^+W^- DELPHI 80.336±0.067 L3 80.270 ±0.055 **OPAL** 80.415 ± 0.052 e^+ LEP2 80.376 ±0.033 χ^2 /dof = 49/41 CDF 80.389 ± 0.019 \varkappa 80.383 ± 0.023 D0 + $\overline{\ell^{\pm}}
u_{\ell} (\overline{
u}_e ll)$ 80.387 ±0.016 Tevatron χ^2 /dof = 4.2/6 World av. (old) 80.385 ±0.015 80.370 ± 0.019 7 TeV ATLAS \uparrow dd World av. (new) 80.379 ±0.012 80.2 80.6 80.4 M_w [GeV]

Higher-order corrections: the mass spectrum

Our current understanding of the invariant mass spectrum in Drell-Yan is very good.





NNLO corrections significantly improve the agreement with data.

Higher-order corrections: the rapidity

[JHEP 12 (2017) 059]



- NNLO + NLO EW predictions.
- Very good agreement between data and theory also in lower and higher invariant mass bins.

Higher-order corrections: q_T **distribution**

The fully differential NNLO (*i.e.* $O(\alpha_s^3)$) corrections to the cross section for $pp \rightarrow Z+jet$ was presented in [*Phys. Rev. Lett.* 117 (2016) 2, 022001]. This calculation allows us to compute the q_T of the Z to NNLO accuracy.



NNLO corrections are moderate but significant in the fixed-order domain $(q_T \sim Q, y \sim 1)$: • improve the agreement with data (see next slide). Can you guess what happens if y becomes very large?

Sizeable reduction of scale uncertainties except (as expected) at low q_T and large rapidity.

Higher-order corrections: q_T **distribution**

NNLO corrections improve the agreement with data all across the board:

- for $q_{\rm T} \sim Q$ the agreement with data is now excellent,
- for $q_T \ll Q$, NNLO partly captures the double-log behaviour and provides qualitative improvements in the description of the shape of the data: **resummation still needed**.



In order to exploit q_T resummation at $q_T \ll Q$ and still benefit of the fixed-order calculation at $q_T \sim Q$, one needs to adopt a matched procedure like CSS:

$$\frac{d\sigma}{dQdydq_T} = W(Q, y, q_T) + Y(Q, y, q_T)$$
Resummation Power corrections of $\alpha_s^n \ln^{2n-1}(Q/q_T)$ of $(q_T/Q)^m$

We know how to compute W (see previous lecture), how do we compute Y? We know that Y contains powers of (q_T / Q) only:

• W contains all the logarithmically enhanced terms resummed up to some order:

$$W^{N^{l}LL}(Q, y, q_{T}) = \frac{1}{q_{T}} \sum_{m=0}^{l} \sum_{n=\lfloor m/2 \rfloor}^{\infty} W^{[m,n]} \alpha_{s}^{n} \ln^{2n-m-1} \left(\frac{Q}{q_{T}}\right)$$

• The N^{*p*}LO fixed-order calculation contains log terms and power corrections up to α_s^{p+1} :

$$FO^{N^{p}LO}(Q, y, q_{T}) = \frac{1}{q_{T}} \sum_{n=1}^{p+1} \alpha_{s}^{n} \sum_{k=1}^{2n-1} F^{[k,n]} \ln^{k} \left(\frac{Q}{q_{T}}\right) + \text{power corrections}$$

• The log terms of W up to α_s^{p+1} have to match those in FO, it follows that:

$$\mathrm{FO}^{\mathrm{N}^{p}\mathrm{LO}}(Q, y, q_{T}) - \left[W^{\mathrm{N}^{l}\mathrm{LL}}(Q, y, q_{T})\right]_{\mathrm{exp. to } O(\alpha_{s}^{p+1})} = \mathrm{power \ corrections} = Y(Q, y, q_{T})$$

• The implementation of the **additive** matching. Others exist. matching procedure? 25

In order to exploit q_T resummation at $q_T \ll Q$ and still benefit of the fixed-order calculation at $q_T \sim \hat{Q}$, one needs to adopt a matched procedure like CSS:

$$\frac{d\sigma}{dQdydq_T} = W(Q, y, q_T) + Y(Q, y, q_T)$$
Resummation Power collections of $\alpha_s^n \ln^{2n-1}(Q/q_T)$ of $(q_T/Q)^m$

We know how to compute W (see previous lecture), how do we compute Y? We know that Y contains powers of (q_T / Q) only:

• W contains all the logarithmically enhanced terms resummed up to some order:

$$W^{N^{l}LL}(Q, y, q_{T}) = \frac{1}{q_{T}} \sum_{m=0}^{l} \sum_{n=\lfloor m/2 \rfloor}^{\infty} W^{[m,n]} \alpha_{s}^{n} \ln^{2n-m-1} \left(\frac{Q}{q_{T}}\right)$$

$$FO^{N^{p}LO}(Q, y, q_{T}) = \frac{1}{q_{T}} \sum_{n=1}^{p+1} \alpha_{s}^{n} \sum_{k=1}^{2n-1} F^{[k,n]} \ln^{k} \left(\frac{Q}{q_{T}}\right) + \text{power corrections}$$

• The log terms of W up to α_s^{p+1} have to match those in FO, it follows that:

$$FO^{N^{p}LO}(Q, y, q_{T}) - \left[W^{N^{l}LL}(Q, y, q_{T})\right]_{exp. \text{ to } O(\alpha_{s}^{p+1})} = \text{power corrections} = Y(Q, y, q_{T})$$

• The implementation of the **additive** matching. Others exist. matching procedure? 26

Resummation of the q_T **distribution**









RadISH + NNLOjet implements the matching procedure described below (except that the matching is **multiplicative** rather than additive [*JHEP* 12 (2018) 132]. Can you tell what is the difference?).



RadISH + NNLOjet implements the matching procedure described below (except that the matching is **multiplicative** rather than additive [*JHEP* 12 (2018) 132]. Can you tell what is the difference?).



RadISH + NNLOjet implements the matching procedure described below (except that the matching is **multiplicative** rather than additive [*JHEP* 12 (2018) 132]. Can you tell what is the difference?).

Non-perturbative effects in q_T distributions

As mentioned in the previous lecture, for $q_T \lesssim 1$ GeV non-perturbative effects, related to α_s becoming large and eventually hitting the Landau pole, become important.

The TMD view on q_T resummation allowed for a transparent way of parameterising nonperturbative effects into a function that can determined through fits to data (like PDFs). A Little reminder:

 $\frac{d\sigma}{dQ^2 dy dq_T^2} = \frac{2\pi \alpha_{\rm em}^2}{3sQ^2 N_c} \sum_q H_{ab}(\alpha_s(Q)) \int_0^\infty db \, bJ_0(bq_T) F_a(x_1, b; Q, Q^2) F_b(x_2, b; Q, Q^2)$ do this trick:

$$b_*(b) = \frac{b}{\sqrt{1 + b^2/b_{\max}^2}} \qquad F(x, b; \mu, \zeta) = \left[\frac{F(x, b; \mu, \zeta)}{F(x, b_*(b); \mu, \zeta)}\right] F(x, b_*(b); \mu, \zeta) = f_{\rm NP}(x, b, \zeta) F(x, b_*(b); \mu, \zeta)$$

so that:

$$\frac{d\sigma}{dQ^2 dy dq_T^2} = \frac{2\pi \alpha_{\rm em}^2}{3sQ^2 N_c} \sum_q H_{ab}(\alpha_s(Q)) \int_0^\infty db \, b J_0(bq_T) f_{\rm NP}(x_1, b, Q^2) f_{\rm NP}(x_2, b, Q^2)$$

×
$$F_a(x_1, b_*(b); Q, Q^2) F_b(x_2, b_*(b); Q, Q^2)$$

take small q_T data that, parameterise f_{NP} , e.g.:

$$f_{\rm NP}(x,b,\zeta) = \exp\left[g_1(b)\ln\left(\frac{\zeta}{Q_0^2}\right) + g_2(x,b)\right]$$

and determine it through a fit.

Non-perturbative effects in q_T distributions



Non-perturbative effects in q_T distributions



The main reason why inclusive Higgs production in gluon fusion fits this discussion is that, like Drell-Yan production, the final state is a **colour singlet**.

As a consequence, factorisation (collinear and TMD) works just as well as for Drell-Yan. The parallel is made particularly transparent in the $m_t \rightarrow \infty$ limit in which top-quark loops can be integrated out. As a consequence:



which amounts to introducing an additional term in the ($n_f = 5$) QCD Lagrangian:

$$\mathcal{L}_{gg\to H} = -\frac{\lambda}{4} G^{\mu\nu} G_{\mu\nu} H$$

thus making higher-order perturbative calculations more convenient.

The effective coupling $\lambda = \mathcal{O}(\alpha_s^2)$ receives perturbative corrections that are currently known to NNLO [Nucl. Phys. B510 (1998) 61–87]. The partonic cross sections for $pp \rightarrow H + j + X$ are also know to NNLO accuracy (*e.g.* [*JHEP* 10 (2016) 066]).

This enables a **full NNLO calculation** of the q_T of the Higgs for $q_T \sim M_{H^*}$.



A comparison at the level of normalised cross sections between theoretical predictions and the ATLAS and CMS data for $pp \rightarrow H + X \rightarrow \gamma\gamma + X$ is reassuring:

- higher-order corrections seem to get closer to the data,
- substantial theory-uncertainty reduction going from LO to NNLO,
- data accuracy though is still not very competitive.

Let us briefly explore what happens at low q_T .

Remarkably, the CSS and TMD-factorisation formulas apply (almost) out of the box:

• the only change w.r.t. Drell-Yan is replacing quarks with gluons.

$$\begin{aligned} \frac{d\sigma_{gg \to H+X}^{\mathrm{CSS}}}{dy dq_T^2} &\propto H_{qq \to H}(\alpha_s(M_H)) \int_0^\infty db \, bJ_0(bq_T) \\ &\times \sum_i \int_{x_1}^1 \frac{dy_1}{y_1} C_{gi}(y_1, \alpha_s(\mu_b)) f_i\left(\frac{x_1}{y_1}, \mu_b\right) \\ &\times \sum_j \int_{x_2}^1 \frac{dy_2}{y_2} C_{gj}(y_2, \alpha_s(\mu_b)) f_j\left(\frac{x_2}{y_2}, \mu_b\right) \\ &\times \exp\left\{-\int_{\mu_b^2}^{M_H^2} \frac{d\mu^2}{\mu^2} \left[A_g(\alpha_s(\mu)) \ln\left(\frac{M_H^2}{\mu^2}\right) + B_g(\alpha_s(\mu))\right]\right\} \\ &+ Y(y, q_T) \\ &\propto \underbrace{H_{qq \to H}(\alpha_s(M_H)) \int_0^\infty db \, bJ_0(bq_T) F_g(x_1, b, M_H, M_H^2) F_g(x_1, b, M_H, M_H^2) + Y(y, q_T)} \end{aligned}$$

In this form, the formula assumes that the gluons are **unpolarised**:

- f_i are the collinear unpolarised PDFs,
- F_g is the TMD unpolarised **gluon** distribution.

However, it turns out that also **linearly** polarised gluons contribute to the cross section:

• if b is small enough, the linearly polarised gluon distribution can be matched onto the unpolarised gluon PDF:

$$h_1^{\perp}(x,b,\mu_b,\mu_b^2) \underset{b \ll \Lambda_{QCD}^{-1}}{=} \sum_i \int_x^1 \frac{dy}{y} G_{gi}(y,\alpha(\mu_b)) f_i\left(\frac{x}{y},\mu_b\right)$$

- In fact, TMD factorisation introduces a new (non-perturbative) TMD distribution: the so-called **Boer-Mulders TMD** h_1^{\perp} that parameterises the distribution of a linearly polarised gluon inside an unpolarised hadron.
- The net result is that the factorisation formular is modified as follows:

1 CSS

$$\begin{split} \frac{d\sigma_{gg \to H+X}^{\text{CSS}}}{dy dq_T^2} &\to \frac{d\sigma_{gg \to H+X}^{\text{CSS}}}{dy dq_T^2} + (C \to G) \\ &\propto \quad H_{qq \to H}(\alpha_s(M_H)) \bigg[\int_0^\infty db \, bJ_0(bq_T) F_g(x_1, b, M_H, M_H^2) F_g(x_1, b, M_H, M_H^2) \\ &+ \quad \int_0^\infty db \, bJ_2(bq_T) h_1^\perp(x_1, b, M_H, M_H^2) h_1^\perp(x_1, b, M_H, M_H^2) \bigg] + Y(y, q_T) \\ \text{The TMD form allows one to introduce a non-perturbative component for } b \gtrsim \Lambda_{\text{QCD}}^{-1} \text{:} \\ h_1^\perp(x, b; \mu, \zeta) = \left[\frac{h_1^\perp(x, b; \mu, \zeta)}{h_1^\perp(x, b_*(b); \mu, \zeta)} \right] h_1^\perp(x, b_*(b); \mu, \zeta) = f_{\text{NP}}^\perp(x, b, \zeta) h_1^\perp(x, b_*(b); \mu, \zeta) \end{split}$$

However, it turns out that also **linearly** polarised gluons contribute to the cross section:

• if *b* is small enough, the linearly polarised gluon distribution can be matched onto the unpolarised gluon PDF:

$$h_1^{\perp}(x,b,\mu_b,\mu_b^2) \underset{b \ll \Lambda_{QCD}^{-1}}{=} \sum_i \int_x^1 \frac{dy}{y} G_{gi}(y,\alpha(\mu_b)) f_i\left(\frac{x}{y},\mu_b\right)$$

- In fact, TMD factorisation introduces a new (non-perturbative) TMD distribution: the so-called **Boer-Mulders TMD** h_1^{\perp} that parameterises the distribution of a linearly polarised gluon inside an unpolarised hadron.
- The net result is that the factorisation formular is modified as follows:





One can attempt to assess how well predictions (including linearly-polarised gluons) at $low-q_T$ compare to the LHC Higgs-production data:



Unfortunately, current data does not allow to say much on the accuracy of the formalism as well as on the TMD gluon distributions.

$$pp \to H(\to \gamma\gamma) + X$$

A carefully study of the Higgs q_T matching the resummed calculation with the fixedorder one for $pp \rightarrow H + X \rightarrow \gamma\gamma + X$ was done in [*JHEP* 12 (2018) 132] with realistic cuts on the photons:



Expected patterns:

- unreliability of the fixed-order calculation at low $p_T^{\gamma\gamma}$ (NLO (left) vs. NNLO (right)),
- reduction of the theoretical uncertainties going from NLO to NNLO,
- dominance of the resummation at low $p_{T}^{\gamma\gamma}$ in the matched calculation.

That's all folks! Thank you!