Hide and seek: how PDFs can conceal new physics

Maeve Madigan
Heidelberg University
Data driven era of particle physics

vast quantity of data  →  precision measurements
Beautiful agreement between data and the Standard Model.
Beautiful agreement between data and the Standard Model

New channels probed in Run II
Beautiful agreement between data and the Standard Model

New channels probed in Run II

Where is the new physics beyond the Standard Model?

Evidence comes from neutrino oscillations, dark matter, ….
ATLAS Search for a new heavy gauge boson decaying into a lepton + missing transverse momentum

1706.04786
ATLAS Search for a new heavy gauge boson decaying into a lepton + missing transverse momentum

1706.04786

W or W'
Indirect searches for new physics

Could BSM particles be heavy and out of reach?

\[ \Lambda_{NP} \gg E \]

\[ \frac{d\sigma}{dm_{\ell\ell}} \]

new physics

\[ m_{\ell\ell}^{\text{max}} \]

SM

e.g. high-mass Drell-Yan tails
Indirect searches for new physics

Could BSM particles be heavy and out of reach?

\[ \Lambda_{NP} \gg E \]

Indirect searches benefit from **precision** measurements.

*e.g. high-mass Drell-Yan tails*
The Standard Model Effective Field Theory

Assume new physics is heavy: $\Lambda \gg E$

Integrate out the new physics particle to obtain interactions of the SM fields.

Assume SM symmetries continue to hold and write down all possible interactions of SM fields:

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \frac{C^{(5)}}{\Lambda} \mathcal{O}^{(5)} + \sum_i \frac{C_i^{(6)}}{\Lambda^2} \mathcal{O}_i^{(6)} + \ldots$$

Compute observables as a systematically improvable expansion in $E/\Lambda$
At dimension 6: 2499 operators

<table>
<thead>
<tr>
<th>$X^3$</th>
<th>$H^6$ and $H^4D^2$</th>
<th>$\psi^2H^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{O}_{GG}^G$</td>
<td>$f^{ABC}G^{A\mu}<em>\nu G^{B\nu}</em>\rho G^{C\mu}_\sigma$</td>
<td>$\mathcal{O}_{H}^G$</td>
</tr>
<tr>
<td>$\mathcal{O}_{GG}^H$</td>
<td>$f^{ABC}G^{A\mu}<em>\nu G^{B\nu}</em>\rho G^{C\mu}_\sigma$</td>
<td>$\mathcal{O}_{H H}$</td>
</tr>
<tr>
<td>$\mathcal{O}_{GG}^D$</td>
<td>$f^{ABC}G^{A\mu}<em>\nu G^{B\nu}</em>\rho G^{C\mu}_\sigma$</td>
<td>$\mathcal{O}_{H D}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$X^2 H^2$</th>
<th>$\psi^2 X H$</th>
<th>$\psi^2 H^2 D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{O}_{GG}^G$</td>
<td>$H^1_H G^A_{\mu\nu}G^A_{\mu\nu}$</td>
<td>$\mathcal{O}_{G G}$</td>
</tr>
<tr>
<td>$\mathcal{O}_{GG}^H$</td>
<td>$H^1_H G^A_{\mu\nu}G^A_{\mu\nu}$</td>
<td>$\mathcal{O}_{G H}$</td>
</tr>
<tr>
<td>$\mathcal{O}_{GG}^D$</td>
<td>$H^1_H W^I_{\mu\nu}W^{I\mu\nu}$</td>
<td>$\mathcal{O}_{G D}$</td>
</tr>
<tr>
<td>$\mathcal{O}_{GG}^G$</td>
<td>$H^1_H W^I_{\mu\nu}W^{I\mu\nu}$</td>
<td>$\mathcal{O}_{G G}$</td>
</tr>
<tr>
<td>$\mathcal{O}_{GG}^H$</td>
<td>$H^1_H B_{\mu\nu}B^{\mu\nu}$</td>
<td>$\mathcal{O}_{G B}$</td>
</tr>
<tr>
<td>$\mathcal{O}_{GG}^D$</td>
<td>$H^1_H B_{\mu\nu}B^{\mu\nu}$</td>
<td>$\mathcal{O}_{G D}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$(LL)(LL)$ and $(RR)(RR)$ and $(LR)(RL)$</th>
<th>$B$-violating</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{O}_{LL}$</td>
<td>$(\bar{q}_p c_H) (\bar{q}_p c_H)$</td>
</tr>
<tr>
<td>$\mathcal{O}_{LR}$</td>
<td>$(\bar{q}_p c_H) (\bar{q}_p c_H)$</td>
</tr>
<tr>
<td>$\mathcal{O}_{LR}$</td>
<td>$(\bar{q}_p c_H) (\bar{q}_p c_H)$</td>
</tr>
<tr>
<td>$\mathcal{O}_{LR}$</td>
<td>$(\bar{q}_p c_H) (\bar{q}_p c_H)$</td>
</tr>
</tbody>
</table>
The global approach

The SMEFT framework connects different sectors of observables measured at the LHC.

We need to take a **global approach**, including as many relevant datasets as possible.

2012.02779, J. Ellis, MM, K. Mimasu, V. Sanz, T. You
Combination of Higgs, top, diboson and electroweak observables constraining 34 coefficients of the dimension-6 SMEFT
The top sector after Run II

Z. Kassabov et. al., 2303.06159

Top quark data available in 2023

Top quark data available in 2021
Looking forward

Run II data already provides precise constraints on the top quark sector of the SMEFT.

As constraints improve, subleading effects may no longer be negligible.
Looking forward

Run II data already provides precise constraints on the top quark sector of the SMEFT.

As constraints improve, subleading effects may no longer be negligible.
Parton distribution functions

A proton-proton collision
Parton distribution functions

PDFs parameterise the quark and gluon constituents of the proton

Peskin and Schroeder, pg 565
Parton distribution functions

PDFs parameterise the quark and gluon constituents of the proton

\[ x f(x) \]

\( u, d, g, \bar{u}, \bar{d}, s, s, c, c \)

\( 0 \leq x \leq 1 \)

\[ J. \text{ Botts et. al, Phys. Lett. B304 159 (1993)} \]
Parton distribution functions

PDFs parameterise the quark and gluon constituents of the proton

NNPDF4.0 NNLO Q = 3.2 GeV

NNPDF4.0, [2109.02653]
Parton distribution functions

\[ \sigma = \int_0^1 dx_1 \int_0^1 dx_2 \sum_{q_1, q_2} f_{q_1}(x_1, Q^2) f_{q_2}(x_2, Q^2) \hat{\sigma}(x_1, x_2) \]
Parton distribution functions

\[ \sigma = \int_0^1 dx_1 \int_0^1 dx_2 \sum_{q_1, q_2} f_{q_1}(x_1, Q^2) f_{q_2}(x_2, Q^2) \hat{\sigma}(x_1, x_2) \]

The SMEFT enters here:  
\[ \hat{\sigma} = \hat{\sigma}_{SM} + \frac{C}{\Lambda^2} \hat{\sigma}_{SMEFT} + ... \]
Parton distribution functions

\[ \sigma = \int_0^1 dx_1 \int_0^1 dx_2 \sum_{q_1, q_2} f_{q_1}(x_1, Q^2) f_{q_2}(x_2, Q^2) \hat{\sigma}(x_1, x_2) \]

Both PDFs and SMEFT are determined by fitting from data
PDF-EFT Interplay

<table>
<thead>
<tr>
<th>PDF fits</th>
<th>SMEFT Fits</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMEFT parameters are kept fixed:</td>
<td>PDF parameters are fixed:</td>
</tr>
</tbody>
</table>

\[
\sigma(\bar{c}, \theta) = f_1(\theta) \otimes f_2(\theta) \otimes \hat{\sigma}(\bar{c})
\]

\[
\sigma(c, \bar{\theta}) = f_1(\bar{\theta}) \otimes f_2(\bar{\theta}) \otimes \hat{\sigma}(c)
\]
## PDF-EFT Interplay

<table>
<thead>
<tr>
<th>PDF fits</th>
<th>SMEFT Fits</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SMEFT parameters are kept fixed:</strong></td>
<td><strong>PDF parameters are fixed:</strong></td>
</tr>
<tr>
<td>[ \sigma(\bar{c}, \theta) = f_1(\theta) \otimes f_2(\theta) \otimes \hat{\sigma}(\bar{c}) ]</td>
<td>[ \sigma(c, \bar{\theta}) = f_1(\bar{\theta}) \otimes f_2(\bar{\theta}) \otimes \hat{\sigma}(c) ]</td>
</tr>
<tr>
<td>Typically PDF fits assume the SM: [ \bar{c} = 0 ]</td>
<td></td>
</tr>
</tbody>
</table>

Wilson coefficients: c

PDF parameters: \( \theta \)
## PDF-EFT Interplay

<table>
<thead>
<tr>
<th>PDF fits</th>
<th>SMEFT Fits</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SMEFT parameters are kept fixed:</strong></td>
<td><strong>PDF parameters are fixed:</strong></td>
</tr>
<tr>
<td>[ \sigma(\bar{c}, \theta) = f_1(\theta) \otimes f_2(\theta) \otimes \hat{\sigma}(\bar{c}) ]</td>
<td>[ \sigma(c, \bar{\theta}) = f_1(\bar{\theta}) \otimes f_2(\bar{\theta}) \otimes \hat{\sigma}(c) ]</td>
</tr>
</tbody>
</table>

- Typically PDF fits assume the SM: \( \bar{c} = 0 \)
- PDFs used in SMEFT fits rely on SM assumptions

Wilson coefficients: \( c \)

PDF parameters: \( \theta \)
Data overlap

Often the data used in PDF fits are also used in EFT fits.

This overlap will grow as we take the global approach to constraining the SMEFT.
Data overlap

Often the data used in PDF fits are also used in EFT fits.

This overlap will grow as we take the global approach to constraining the SMEFT.

Data included in NNPDF4.0, [2109.02653]:

- Fixed-target DIS
- Collider DIS
- Fixed-target DY
- Collider gauge boson production
- Collider gauge boson production+jet
- Z transverse momentum
- Top-quark pair production
- Single-inclusive jet production
- Di-jet production
- Direct photon production
- Single top-quark production
- Black edge: new in NNPDF4.0
**Data overlap**

Often the data used in PDF fits are also used in EFT fits.

This overlap will grow as we take the global approach to constraining the SMEFT.

- e.g. Top quark data used to fit the SMEFT in the global fit of 2012.02779, J. Ellis, MM, K. Mimasu, V. Sanz, T. You
Overview

1. PDF-EFT interplay in high-mass Drell-Yan

2. Can PDFs absorb new physics?

3. Simultaneous PDF and SMEFT determination in the top sector
PDF-EFT interplay in high-mass Drell-Yan

Greljo et. al 2104.02723
PDF-EFT interplay in high-mass Drell-Yan

Energy-growing 4-fermion operators manifest as a smooth distortion of the high-mass tail:
PDF-EFT interplay in high-mass Drell-Yan

Energy-growing 4-fermion operators manifest as a smooth distortion of the high-mass tail:

\[ \frac{d\sigma}{dm_{\ell\ell}} \]

Constraints on 4-fermion operators of the SMEFT:

Farina et. al 1609.08157
Constraints on the large-\(x\) region of the \(u\) and \(d\) PDFs:

\[ u \text{ at } 100.0 \text{ GeV} \]

![Graph showing ratio to DIS+DY for u at 100.0 GeV](image)

Greljo et. al 2104.02723

Constraints on 4-fermion operators of the SMEFT:

![Graph showing constraints on 4-fermion operators of the SMEFT](image)

Farina et. al 1609.08157
Simultaneous PDF and SMEFT fit
Simultaneous PDF and SMEFT fit

Data

Deep inelastic scattering + Drell-Yan

- including high-mass DY:

<table>
<thead>
<tr>
<th>Exp.</th>
<th>√s (TeV)</th>
<th>Ref.</th>
<th>L (fb⁻¹)</th>
<th>Channel</th>
<th>1D/2D</th>
<th>n_{dat}</th>
<th>m_{W̃}^{max} (TeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATLAS</td>
<td>7</td>
<td>[120]</td>
<td>4.9</td>
<td>e⁻⁻e⁺</td>
<td>1D</td>
<td>13</td>
<td>[1.0, 1.5]</td>
</tr>
<tr>
<td>ATLAS (*)</td>
<td>8</td>
<td>[86]</td>
<td>20.3</td>
<td>ℓ⁻⁻ℓ⁺</td>
<td>2D</td>
<td>46</td>
<td>[0.5, 1.5]</td>
</tr>
<tr>
<td>CMS</td>
<td>7</td>
<td>[121]</td>
<td>9.3</td>
<td>μ⁻⁻μ⁺</td>
<td>2D</td>
<td>127</td>
<td>[0.2, 1.5]</td>
</tr>
<tr>
<td>CMS (*)</td>
<td>8</td>
<td>[87]</td>
<td>19.7</td>
<td>ℓ⁻⁻ℓ⁺</td>
<td>2D</td>
<td>41</td>
<td>[1.5, 2.0]</td>
</tr>
<tr>
<td>CMS (*)</td>
<td>13</td>
<td>[122]</td>
<td>5.1</td>
<td>e⁻⁻e⁺, μ⁻⁻μ⁺</td>
<td>1D</td>
<td>43, 43</td>
<td>[1.5, 3.0]</td>
</tr>
<tr>
<td>CMS</td>
<td>13</td>
<td>[122]</td>
<td>5.1</td>
<td>ℓ⁻⁻ℓ⁺</td>
<td>1D</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>270</td>
<td>(313)</td>
</tr>
</tbody>
</table>

+ High Luminosity projections
Simultaneous PDF and SMEFT fit

Data

Deep inelastic scattering + Drell-Yan

- including high-mass DY:

<table>
<thead>
<tr>
<th>Exp.</th>
<th>$\sqrt{s}$ (TeV)</th>
<th>Ref.</th>
<th>$\mathcal{L}$ (fb$^{-1}$)</th>
<th>Channel</th>
<th>1D/2D</th>
<th>$n_{\text{dat}}$</th>
<th>$m_W^{\text{exp}}$ (TeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATLAS</td>
<td>7</td>
<td>[120]</td>
<td>4.9</td>
<td>$e^-e^+$</td>
<td>1D</td>
<td>13</td>
<td>[1.0, 1.5]</td>
</tr>
<tr>
<td>ATLAS (*)</td>
<td>8</td>
<td>[86]</td>
<td>20.3</td>
<td>$\ell^-\ell^+$</td>
<td>2D</td>
<td>46</td>
<td>[0.5, 1.5]</td>
</tr>
<tr>
<td>CMS</td>
<td>7</td>
<td>[121]</td>
<td>9.3</td>
<td>$\mu^-\mu^+$</td>
<td>2D</td>
<td>127</td>
<td>[0.2, 1.5]</td>
</tr>
<tr>
<td>CMS (*)</td>
<td>8</td>
<td>[87]</td>
<td>19.7</td>
<td>$\ell^-\ell^+$</td>
<td>1D</td>
<td>41</td>
<td>[1.5, 2.0]</td>
</tr>
<tr>
<td>CMS (*)</td>
<td>13</td>
<td>[122]</td>
<td>5.1</td>
<td>$e^-e^+, \mu^-\mu^+$</td>
<td>1D</td>
<td>43, 43</td>
<td>[1.5, 3.0]</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>270</td>
<td>313</td>
<td></td>
</tr>
</tbody>
</table>

+ High Luminosity projections

Theory benchmarks

Electroweak oblique parameters $\hat{W}, \hat{Y}$

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} - \frac{g^2\hat{W}}{4m_W^2} \mathcal{O}_{tq}^{(3)} - \frac{g^2\hat{Y}}{m_W^2} \left( Y_t Y_d \mathcal{O}_{td} + Y_t Y_u \mathcal{O}_{tu} + Y_e Y_q \mathcal{O}_{tq}^{(1)} + Y_e Y_d \mathcal{O}_{ed} + Y_e Y_u \mathcal{O}_{eu} + Y_e Y_q \mathcal{O}_{qe} \right)$$
Simultaneous PDF and SMEFT fit

Data

Deep inelastic scattering + Drell-Yan

- including high-mass DY:

<table>
<thead>
<tr>
<th>Exp.</th>
<th>$\sqrt{s}$ (TeV)</th>
<th>Ref.</th>
<th>$\mathcal{L}$ (fb$^{-1}$)</th>
<th>Channel</th>
<th>1D/2D</th>
<th>$n_{\text{dat}}$</th>
<th>$m_{W^\prime}$ (TeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATLAS</td>
<td>7</td>
<td>[120]</td>
<td>4.9</td>
<td>$e^-e^+$</td>
<td>1D</td>
<td>13</td>
<td>[1.0, 1.5]</td>
</tr>
<tr>
<td>ATLAS (*)</td>
<td>8</td>
<td>[86]</td>
<td>20.3</td>
<td>$\ell^-\ell^+$</td>
<td>2D</td>
<td>46</td>
<td>[0.5, 1.5]</td>
</tr>
<tr>
<td>CMS</td>
<td>7</td>
<td>[121]</td>
<td>9.3</td>
<td>$\mu^-\mu^+$</td>
<td>2D</td>
<td>127</td>
<td>[0.2, 1.5]</td>
</tr>
<tr>
<td>CMS (*)</td>
<td>8</td>
<td>[87]</td>
<td>19.7</td>
<td>$\ell^-\ell^+$</td>
<td>1D</td>
<td>41</td>
<td>[1.5, 2.0]</td>
</tr>
<tr>
<td>CMS (*)</td>
<td>13</td>
<td>[122]</td>
<td>5.1</td>
<td>$e^-\ell^+, \mu^-\mu^+$</td>
<td>1D</td>
<td>43, 43</td>
<td>[1.5, 3.0]</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>270 (313)</td>
<td></td>
</tr>
</tbody>
</table>

+ High Luminosity projections

Theory benchmarks

Electroweak oblique parameters $\hat{W}, \hat{Y}$

$L_{\text{SMEFT}} = L_{\text{SM}} - \frac{g_2^2 \hat{W}}{4 m_W^2} O_{tq}^{(3)} - \frac{g_2^2 \hat{Y}}{m_W^2} \left( Y_{lY_d} O_{ld} + Y_{lY_u} O_{lu} + Y_{eY_d} O_{ed} + Y_{eY_u} O_{eu} + Y_{eY_q} O_{qe} \right)$

parametrises the impact of a flavour universal $W^\prime$

parametrises the impact of a flavour universal $Z^\prime$
PDF-EFT interplay in high-mass Drell-Yan

Energy-growing 4-fermion operators manifest as a smooth distortion of the high-mass tail:

\[
\mathcal{A} \sim \mathcal{A}_\text{SM} + C \frac{\mathcal{E}^2}{\Lambda^2}
\]

E.g. \(Z'\)

\[
\mathcal{L}_{\text{SMEFT}}^{Z'} = \mathcal{L}_{\text{SM}} - \frac{g'^2 \hat{Y}}{2m_W^2} J_{Y,\mu}^\mu J_{Y,\mu}
\]

\[
J_L^\mu = \sum_f Y_f \bar{f} \gamma^\mu f
\]

Impacts only neutral-current DY:

\[
\frac{d\sigma}{dm_{\ell\ell}}
\]
PDF-EFT interplay in high-mass Drell-Yan

Energy-growing 4-fermion operators manifest as a smooth distortion of the high-mass tail:

\[ \mathcal{A} \sim \mathcal{A}_{\text{SM}} + C \frac{E^2}{\Lambda^2} \]

E.g. \[ \mathcal{L}_{\text{SMEFT}}^{W'} = \mathcal{L}_{\text{SM}} - \frac{g^2 W}{2m_W^2} J_L^\mu J_{L,\mu} \]

\[ J_L^\mu = \sum_{f_L} \bar{f}_L T^0 \gamma^\mu f_L \]

Impacts both neutral and charged-current DY:

\[ \frac{d\sigma}{d m_{\ell\ell}}, \frac{d\sigma}{d m_T}, pp \rightarrow \ell^+ \ell^-, pp \rightarrow \ell \nu \]
Simultaneous PDF and SMEFT fit results
Simultaneous PDF and SMEFT fit results

Excluding HL-LHC projections for NC and CC Drell-Yan:

PDF fits under the assumption of nonzero SMEFT coefficients:

We see a moderate shift of the PDF central values, and no change to the PDF uncertainties.
Simultaneous PDF and SMEFT fit results

Excluding HL-LHC projections for NC and CC Drell-Yan:

SMEFT constraints are stable:

moderate shifts when using SMEFT vs SM PDFs

Farina et al (1609.08157)
this work, SM PDFs
this work, SMEFT PDFs
Simultaneous PDF and SMEFT fit results

Including HL-LHC projections for NC and CC Drell-Yan:

PDF fits under the assumption of nonzero SMEFT coefficients:

We see a **large shift** of the PDF central values, in some cases beyond PDF uncertainties.
Neglecting PDF-EFT interplay leads to a significant overestimate of the EFT constraints.
Conservative PDFs

Could we improve the SM PDF fits by removing the high-mass data from PDF fits?

• not in the spirit of global fits

• still have a theoretical inconsistency due to SM assumptions

• **but** much easier than doing a simultaneous PDF-SMEFT fit
Simultaneous PDF and SMEFT fit results

Including HL-LHC projections for NC and CC Drell-Yan:

Neglecting PDF-EFT interplay leads to a significant overestimate of the EFT constraints.
Simultaneous PDF and SMEFT fit results

Neglecting PDF-EFT interplay leads to a significant overestimate of the EFT constraints.

what does this mean for searches for new physics?
Can PDFs Absorb New Physics?

CDF collaboration measured a deviation at high transverse momentum

However, this was not new physics

- deviation went away with improvements to large-x gluon PDFs
What if no new physics is observed... …because it has been absorbed by the PDFs?
Contaminated PDFs

Assume that we know the true underlying law of nature: SM + UV model

$$T = T(\theta_{\text{SM}}, \theta_{\text{NP}})$$

closely follows the closure test methodology developed by NNPDF, 1410.8849
Contaminated PDFs
closely follows the closure test methodology developed by NNPDF, 1410.8849

Assume that we know the true underlying law of nature: SM + UV model

\[ T = T(\theta_{\text{SM}}, \theta_{\text{NP}}) \]

Generate Monte Carlo pseudodata according to this underlying law:

\[ D \sim \mathcal{N}(T(\theta_{\text{SM}}, \theta_{\text{NP}}), \Sigma) \]
Contaminated PDFs

closely follows the closure test methodology developed by NNPDF, 1410.8849

Assume that we know the true underlying law of nature: \( \text{SM + UV model} \)

\[ T = T(\theta_{\text{SM}}, \theta_{\text{NP}}) \]

Generate Monte Carlo pseudodata according to this underlying law:

\[ D \sim \mathcal{N}(T(\theta_{\text{SM}}, \theta_{\text{NP}}), \Sigma) \]

Perform a PDF fit: fit only the SM parameters \( \theta_{\text{SM}} \) using the NNPDF4.0 methodology 2109.02653
Contaminated PDFs closely follows the closure test methodology developed by NNPDF, 1410.8849

Assume that we know the true underlying law of nature: SM + UV model

\[ T = T(\theta_{\text{SM}}, \theta_{\text{NP}}) \]

Generate Monte Carlo pseudodata according to this underlying law:

\[ D \sim \mathcal{N}(T(\theta_{\text{SM}}, \theta_{\text{NP}}), \Sigma) \]

Perform a PDF fit: fit only the SM parameters \( \theta_{\text{SM}} \) using the NNPDF4.0 methodology

\[ 2109.02653 \]

PDF has **absorbed new physics** if the fit quality is good

\[ n_\sigma = \frac{\chi^2 - 1}{\sigma \chi^2} < 2 \]
Data

• We generate MC pseudodata for all datasets included in NNPDF 4.0

2109.02653

• Additionally, we include **HL-LHC** projections for neutral current and charged current DY

  as in Greljo et. al 2104.02723
BSM scenario: Z’

- Flavour universal Z’

\[ \mathcal{L}^\text{Z'}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} - \frac{g'^2 \hat{Y}}{2m_W^2} J^\mu Y_\mu \]

- Impacts NC DY

\[ J^\mu_L = \sum_f Y_f \bar{f} \gamma^\mu f \]
BSM scenario: $W'$

- Flavour universal $W'$
  \[ \mathcal{L}_{\text{SMEFT}}^{W'} = \mathcal{L}_{\text{SM}} - \frac{g^2 \hat{W}}{2m_W^2} J_L^\mu J_{L,\mu} \]

\[ J_L^\mu = \sum_{f_L} \bar{f}_L T^a \gamma^\mu f_L \]

- Impacts NC and CC DY

\[ pp \rightarrow l\nu \]
\[ M_{W'} = 13.8 \text{ TeV} \]
\[ \hat{W} = 8 \cdot 10^{-5} \]
Do our contaminated fits pass the selection criteria?

\[ n_\sigma = \frac{\chi^2 - 1}{\sigma \chi^2} \]

➡️ Z’ scenario

- HL-LHC HM DY 14 TeV - charged current - muon channel
- HL-LHC HM DY 14 TeV - charged current - electron channel
- HL-LHC HM DY 14 TeV - neutral current - muon channel
- HL-LHC HM DY 14 TeV - neutral current - electron channel

<table>
<thead>
<tr>
<th>Condition</th>
<th>Y</th>
<th>2\sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>baseline</td>
<td>0.0182</td>
<td>2.66</td>
</tr>
<tr>
<td>Y=5e-5</td>
<td>0.0777</td>
<td>2.64</td>
</tr>
<tr>
<td>Y=15e-5</td>
<td>0.0195</td>
<td>2.66</td>
</tr>
<tr>
<td>Y=25e-5</td>
<td>0.0265</td>
<td>2.64</td>
</tr>
<tr>
<td>7.75</td>
<td>7.88</td>
<td></td>
</tr>
</tbody>
</table>

Maeve Madigan | Hide and Seek: how PDFs can conceal new physics
University of Vienna 01.12.23
Do our contaminated fits pass the selection criteria?

![](image)

\[ n_\sigma = \frac{\chi^2 - 1}{\sigma \chi^2} \]

➡️ **Z' scenario**

---

No: PDFs do not absorb new physics

---

**Results:**

- **Baseline:**
  - HL-LHC HM DY 14 TeV - charged current - muon channel: 2.66
  - HL-LHC HM DY 14 TeV - charged current - electron channel: 0.371
  - HL-LHC HM DY 14 TeV - neutral current - muon channel: 2.64
  - HL-LHC HM DY 14 TeV - neutral current - electron channel: 0.361

**Significance:**

- 7.75
- 7.88

**Significance Levels:**

- **2σ**
Do our contaminated fits pass the selection criteria?

\[ n_\sigma = \frac{\chi^2 - 1}{\sigma \chi^2} \]

\( \rightarrow \) Z’ scenario

No: PDFs do not absorb new physics

unless the NP effects are negligible
Do our contaminated fits pass the selection criteria?

\[ n_\sigma = \frac{\chi^2 - 1}{\sigma \chi^2} \]

\( \Rightarrow \) \( W' \) scenario

- HL-LHC HM DY 14 TeV - charged current - muon channel
- HL-LHC HM DY 14 TeV - charged current - electron channel
- HL-LHC HM DY 14 TeV - neutral current - muon channel
- HL-LHC HM DY 14 TeV - neutral current - electron channel
Do our contaminated fits pass the selection criteria?

\[ n_\sigma = \frac{\chi^2 - 1}{\sigma \chi^2} \]

➡️ W’ scenario

\[ \hat{W} = 8 \times 10^{-5}, \quad M_{W'} \approx 14 \text{ TeV} \]

Yes: PDFs absorb new physics
**W’-contaminated PDFs**

### NC DY

\[ u\bar{u} + d\bar{d} \text{ luminosity} \]
\[ \sqrt{s} = 14 \text{ TeV} \quad ||y|| < 2.5 \]

![Diagram showing the ratio to baseline for NC DY process](image1)

### CC DY

\[ u\bar{d} + d\bar{u} \text{ luminosity} \]
\[ \sqrt{s} = 14 \text{ TeV} \quad ||y|| < 2.5 \]

![Diagram showing the ratio to baseline for CC DY process](image2)

Fewer constraints on the large-x antiquark PDFs allow freedom to shift away from the baseline.
W’-contaminated PDFs

Fewer constraints on the large-x antiquark PDFs allow freedom to shift away from the baseline.
Impact on Drell-Yan

- The data appears to agree well with the SM
- **The shift in the PDFs compensates the NP effects**
- The effects of NP are completely missed

Data: ‘true’ PDF $\otimes$ SM + W
Theory: contaminated PDF $\otimes$ SM

Excellent data-theory agreement
Z’-contaminated PDFs

Charged current DY is not impacted by the Z’ model

- CC DY data constrains the large-x quark and antiquark PDFs to be SM-like
- PDFs cannot shift enough to absorb NP effects in neutral current DY
Impact on Drell-Yan

Data: ‘true’ PDF $\otimes$ SM + Z’
Theory: contaminated PDF $\otimes$ SM

PDFs remain SM-like: discrepancy with Z’ in NC DY data
Impact on EW processes

The PDF then causes spurious NP effects in other observables e.g.

\[ q\bar{q} \rightarrow W^+W^- \]

• Data appears to disagree with SM at 3σ

• However, \( W^+W^- \) is unaffected by \( W' \) model:
  
  the deviation is in the PDF

---

HL-LHC projections

Data: ‘true’ PDF ⊗ SM
Theory: contaminated PDF ⊗ SM

\[ x^2/n_{dat} = 2.194 \]
\[ n_a = 3.044 \]
Impact on EW processes

The PDF then causes **spurious NP effects** in other observables e.g.

\[ q\bar{q} \rightarrow WH \]

- Data appears to disagree with SM at 2\(\sigma\)
- However, \(WH\) is unaffected by \(W'\) model:
  - the deviation is in the PDF

**HL-LHC projections**

Data: 'true' PDF \(\otimes\) SM  
Theory: contaminated PDF \(\otimes\) SM

\[ \chi^2/n_{\text{dat}} = 1.774 \]
\[ n_g = 1.972 \]
\[ k_{\text{stat}} = 10 \]

statistics improved by a factor of 10
Opportunities to disentangle PDF and SMEFT effects
Opportunities to disentangle PDF and SMEFT effects

Ratio observables:
Opportunities to disentangle PDF and SMEFT effects

Ratio observables:

Low-energy precision measurements sensitive to high-\(x\) PDFs
Opportunities to disentangle PDF and SMEFT effects

Ratio observables:

Low-energy precision measurements sensitive to high-x PDFs

➡ add precision here:
Opportunities to disentangle PDF and SMEFT effects

Ratio observables:

Low-energy precision measurements sensitive to high-x PDFs

what about simultaneous PDF and SMEFT determinations?
The SIMUnet Methodology
S. Iranipour, M. Ubiali, 2201.07240

The SIMUUnet methodology

An extension of the NNPDF framework

- PDFs parameterised by a neural network

Ball et. al, NNPDF4.0, 2109.02653
The SIMUnet methodology

An extension of the NNPDF framework

- PDFs parameterised by a neural network
- Propagates uncertainties from data to NN parameters using the Monte Carlo replica method

\[ d_k \sim \mathcal{N}(d, \sigma) \]

\[ c_k = \arg \min_c \chi^2(c, d_k) \]
The SIMUnet methodology

An extension of the NNPDF framework

- PDFs parameterised by a neural network
- Propagates uncertainties from data to NN parameters using the Monte Carlo replica method

Additional layer incorporates SMEFT Wilson coefficients
The SIMUnet methodology

Train only the final layer: reproduce SMEFT fits

Figure 3.1. Schematic depiction of the SIMUnet methodology. The input nodes (shown in green) are Bjorken-\(x\) and its logarithm. The forward pass through the deep hidden layers (blue) are performed as in the NNPDF4.0 methodology \[^{2}\] to yield the output PDFs at the initial scale (red). The initial scale luminosity is then convoluted with the pre-computed FK-tables (shown in blue) to obtain the theoretical prediction (shown in red), which enters the figure of merit (3.1), which is minimised in the fit. The dependence on the parameters \(\{c_n\}\) is fed into theoretical prediction via the trainable edges of the combination layer. All trainable edges are shown by solid edges and are thus learned parameters determined through gradient descent, while dashed edges are non-trainable.

The values of \(\{c_n\}\) are associated with the weights of the trainable edges which determine the FK table, \(\Sigma\), as in Eq. (2.6). Such dependence enters the theoretical prediction via the bilinear product between \(\Sigma\) (\(\{c_n\}\)) and the initial scale PDFs, which in Eq. (2.3) we refer to as \(L_0\), where \(L_0\) indicates either the parametrization of one independent PDF at the initial scale or the product of two of them.

Letting \(\tilde{\mathbf{x}}\) denote the set of trainable neural network parameters (the weights and biases) that parameterize the PDFs and \(\{c_n\}\) the parameters that we fit alongside the PDFs, SIMUnet fits the joint \(\hat{\mathbf{\tilde{x}}} = \mathbf{\tilde{x}}[\{c_n\}]\) parameter set, by letting gradient descent determine their optimum value in order to minimize the figure of merit used in the fit, which is defined as:

\[
\mathcal{F}(\hat{\mathbf{\tilde{x}}}) = \frac{1}{N_{\text{data}}} \left( \mathbf{D}^T(\hat{\mathbf{\tilde{x}}}) \mathbf{T}(\hat{\mathbf{\tilde{x}}}) \right)^T \mathbf{cov}^{-1} \left( \mathbf{D}^T(\hat{\mathbf{\tilde{x}}}) \mathbf{T}(\hat{\mathbf{\tilde{x}}}) \right),
\]

with \(\mathbf{D}\) being the vector of experimental central values, \(\mathbf{T}\) the vector of theoretical predictions and \(\mathbf{cov}\) the covariance matrix encapsulating the experimental uncertainties and the
The SIMUnet methodology

Train only the PDF NN weights on all data: reproduce NNPDF
The SIMUnet methodology

Train only the PDF NN weights on all data except the SMEFT-affected sector: conservative PDFs
The SIMUnet methodology

Train everything: **simultaneous fit**

\[
\begin{align*}
\sum & \quad c_1 \\
& \quad \vdots \\
& \quad c_{N-1} \\
& \quad c_N \\
& \quad T \\
\end{align*}
\]

\[
L^0
\]

\[
\begin{align*}
x \quad h_1^{(1)} \\
& \quad h_2^{(1)} \\
& \quad h_3^{(1)} \\
& \quad h_4^{(1)} \\
& \quad h_5^{(1)} \\
& \quad \vdots \\
& \quad h_{20}^{(2)} \\
& \quad h_{25}^{(1)} \\
\end{align*}
\]

\[
\begin{align*}
f_1 \quad f_2 \\
& \quad f_3 \\
& \quad f_4 \\
& \quad f_5 \\
& \quad f_6 \\
& \quad f_7 \\
& \quad f_8 \\
\end{align*}
\]

\[
\begin{align*}
\text{Input layer} & \quad h_1^{(1)} \\
& \quad h_2^{(1)} \\
& \quad h_3^{(1)} \\
& \quad h_4^{(1)} \\
& \quad h_5^{(1)} \\
& \quad \vdots \\
& \quad h_{20}^{(2)} \\
& \quad h_{25}^{(1)} \\
\end{align*}
\]

\[
\begin{align*}
\text{Hidden layer 1} & \quad h_1^{(2)} \\
& \quad h_2^{(2)} \\
& \quad h_3^{(2)} \\
& \quad \vdots \\
& \quad h_{20}^{(2)} \\
\end{align*}
\]

\[
\begin{align*}
\text{Hidden layer 2} & \quad f_1 \\
& \quad f_2 \\
& \quad f_3 \\
& \quad f_4 \\
& \quad f_5 \\
& \quad f_6 \\
& \quad f_7 \\
& \quad f_8 \\
\end{align*}
\]
Simultaneous PDF and SMEFT determination in the top sector

Kassabov et. al: 2303.06159
Top quark data provides important constraints on the large-\( x \) region of the gluon PDF.

This impact is largely driven by top quark pair production cross sections and differential distributions. e.g. Czakon et. al, 1303.7215, 1611.08609, 1912.08801

PDF-EFT interplay in the top sector

Potential for interplay between gluon PDF and coefficients modifying top quark pair production:
The top sector of the SMEFT

Impact of top quark pair production

Top quark data available in 2023
Top quark data available in 2021

Impact of top quark pair production

Top quark data available in 2023
Top quark data available in 2021

Impact of top quark pair production

Top quark data available in 2023
Top quark data available in 2021
Simultaneous fit

A simultaneous fit shows better agreement with the no-top fit:

- the impact of top data is diluted by the inclusion of the SMEFT

Uncertainties increase relative to the SM, all top data PDF fit

- reflecting the increase in number of fitted parameters
Simultaneous fit

A simultaneous fit shows better agreement with the no-top fit:

- the impact of top data is diluted by the inclusion of the SMEFT

Uncertainties increase relative to the SM, all top data PDF fit

- reflecting the increase in number of fitted parameters

\[
\sqrt{s} = 13 \text{ TeV}
\]
Simultaneous fit

Constraints on the Wilson coefficients are stable, despite differences in PDFs

\[ O(\Lambda^{-2}) \]

Conservative PDF fit - only SMEFT fit sees top data
Simultaneous fit: what about quadratic EFT effects?

\[ \sigma = \sigma_{\text{SM}} + \frac{C}{\Lambda^2} \sigma_{\text{lin}} + \frac{C^2}{\Lambda^4} \sigma_{\text{quad}} \]

\[ \mathcal{O}(\Lambda^{-4}) \]

Top quark data available in 2023

Top quark data available in 2021
The Monte Carlo Replica Method

Consider fitting 1 Wilson coefficient $c$ to 1 datapoint $\sigma_{\text{exp}}$: define

$$
\chi^2 = \frac{(\sigma(c) - \sigma_{\text{exp}})^2}{\delta\sigma^2}
$$

1. Resample: $\tilde{\sigma}_{\text{exp}} \sim \mathcal{N}(\sigma_{\text{exp}}, \delta\sigma)$

2. Minimise:

$$
\bar{c} = \arg \min_c \frac{(\sigma(c) - \tilde{\sigma}_{\text{exp}})^2}{\delta\sigma^2}
$$

3. Repeat, and treat the sample $\{\bar{c}\}$ as a sample from the Bayesian posterior $p(c|D)$

- Often used in the context of PDF fitting and SMEFT fitting, e.g. 2109.02653, 1901.05965
The Monte Carlo Replica Method

**Problem:** in the presence of a quadratic theory, often the minimum $\chi^2$ will be given by the same $\bar{c}$.

2. Minimise: $\bar{c} = \arg \min_c \frac{(\sigma(c) - \bar{\sigma}_{\text{exp}})^2}{\delta \sigma^2}$
The Monte Carlo Replica Method

**Problem:** in the presence of a quadratic theory, often the minimum $\chi^2$ will be given by the same $\bar{c}$.

\[ \sigma(c) = \sigma_{SM} + \sigma_{\text{lin } c} + \sigma_{\text{quad } c^2} \]

\[ C_{dt} \]

MC replica method

Bayesian
Conclusions

The discovery of new physics will rely on an unbiased and accurate understanding of the parton distribution functions.

Parton distribution functions have the potential to conceal new physics:

- Contaminated PDFs may translate signs of new physics into Higgs+EW processes
- Disentangling these effects post-fit is not guaranteed

PDF-EFT interplay is moderate in current LHC top and DY data, but may become significant at the HL-LHC

Simultaneous PDF and SMEFT determinations are crucial for the assessment of the extent of PDF-EFT interplay in current LHC data.

Public SIMUnet release coming soon!
Conclusions

The discovery of new physics will rely on an unbiased and accurate understanding of the parton distribution functions

Parton distribution functions have the potential to conceal new physics:

• Contaminated PDFs may translate signs of new physics into Higgs+EW processes

• Disentangling these effects post-fit is not guaranteed

PDF-EFT interplay is moderate in current LHC top and DY data, but may become significant at the HL-LHC

Simultaneous PDF and SMEFT determinations are crucial for the assessment of the extent of PDF-EFT interplay in current LHC data.

Public SIMUnet release coming soon!

Thank you for listening!
Backup
Data

175 datapoints:

a superset of measurements in

fitmaker  SMEFiT  NNPDF

**tt**  **tt + V**  charge asymmetry $A_C$  **ttt̅, t̅bb**  single top, $tW$
SM
NLO QCD using MG5_aMC@NLO

Where available, NNLO QCD using k-factors from HighTea:

Czakon et. al, 2304.05993
https://www.precision.hep.phy.cam.ac.uk/hightea/
Theory

SM
NLO QCD using MG5_aMC@NLO

Where available, NNLO QCD using k-factors from HighTea: Czakon et. al, 2304.05993
https://www.precision.hep.phy.cam.ac.uk/hightea/

SMEFT
25 Wilson coefficients at NLO QCD using SMEFT@NLO Degrande et. al, 2008.11743
PDF-EFT interplay in high-mass Drell-Yan

**Data**

Deep inelastic scattering + Drell-Yan

- including high-mass DY:

<table>
<thead>
<tr>
<th>Exp.</th>
<th>$\sqrt{s}$ (TeV)</th>
<th>Ref.</th>
<th>$\mathcal{L}$ (fb$^{-1}$)</th>
<th>Channel</th>
<th>1D/2D</th>
<th>$n_{\text{dat}}$</th>
<th>$m_W^{\text{exp}}$ (TeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATLAS</td>
<td>7</td>
<td>[120]</td>
<td>4.9</td>
<td>$e^- e^+$</td>
<td>1D</td>
<td>13</td>
<td>[1.0, 1.5]</td>
</tr>
<tr>
<td>ATLAS (*)</td>
<td>8</td>
<td>[86]</td>
<td>20.3</td>
<td>$t^- t^+$</td>
<td>2D</td>
<td>46</td>
<td>[0.5, 1.5]</td>
</tr>
<tr>
<td>CMS</td>
<td>7</td>
<td>[121]</td>
<td>9.3</td>
<td>$\mu^- \mu^+$</td>
<td>2D</td>
<td>127</td>
<td>[0.2, 1.5]</td>
</tr>
<tr>
<td>CMS (*)</td>
<td>8</td>
<td>[87]</td>
<td>19.7</td>
<td>$\ell^- \ell^+$</td>
<td>1D</td>
<td>41</td>
<td>[1.5, 2.0]</td>
</tr>
<tr>
<td>CMS (*)</td>
<td>13</td>
<td>[122]</td>
<td>5.1</td>
<td>$e^- e^+, \mu^- \mu^+$</td>
<td>1D</td>
<td>43</td>
<td>[1.5, 3.0]</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>270 (313)</td>
<td></td>
</tr>
</tbody>
</table>

+ High Luminosity projections

**Theory benchmarks**

Electroweak oblique parameters $\hat{W}, \hat{Y}$

$$
\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} - \frac{g^2 \hat{W}}{4m_W^2} \mathcal{O}_{tq}^{(3)} - \frac{g^2 \hat{Y}}{m_W^2} \left( Y_t Y_d \mathcal{O}_{td} + Y_t Y_u \mathcal{O}_{tu} + Y_e Y_q \mathcal{O}_{eq} + Y_e Y_d \mathcal{O}_{ed} + Y_e Y_u \mathcal{O}_{eu} + Y_e Y_q \mathcal{O}_{qe} \right)
$$
Conservative PDFs

Could we improve the SM PDF fits by removing the high-mass data from PDF fits?

- not in the spirit of global fits
- still have a theoretical inconsistency due to SM assumptions
- **but** much easier than doing a simultaneous PDF-SMEFT fit
Conservative PDFs

Could we improve the SM PDF fits by removing the high-mass data from PDF fits?

- not in the spirit of global fits
- still have a theoretical inconsistency due to SM assumptions
- but much easier than doing a simultaneous PDF-SMEFT fit

See J. Ethier et. al, 2105.00006 for the use of conservative PDFs in a global SMEFT fit
Conservative PDFs:

• assume the SM
• are fit to data which does not receive large SMEFT corrections (i.e. no HL-LHC data, no high-mass DY data)
Conservative PDFs:

- assume the SM
- are fit to data which does not receive large SMEFT corrections (i.e. no HL-LHC data, no high-mass DY data)

Comparing green to orange:

- the constraints using SM conservative PDFs are closer to those using SMEFT PDFs
- still overestimating the constraints, especially in the $\hat{W}$ direction
SMEFT PDFs

Simultaneous PDF-EFT determination outputs a **SMEFT PDF**

Use this as an input to future SMEFT determinations as an approximation for a full PDF-SMEFT fit

- Captures the increase in PDF uncertainties due to the inclusion of SMEFT parameters in the fit
- Neglects PDF-SMEFT correlations captured by a full simultaneous determination
Simultaneous PDF-EFT determination outputs a **SMEFT PDF**

Use this as an input to future SMEFT determinations as an approximation for a full PDF-SMEFT fit

$$R_n = \frac{c_n^*}{\sigma_n}$$

SMEFT PDFs are a good approximation - small PDF-EFT correlation in the top sector
The Monte Carlo Replica Method

Consider fitting 1 Wilson coefficient $c$ to 1 datapoint $\sigma_{\text{exp}}$: define $\chi^2 = \frac{(\sigma(c) - \sigma_{\text{exp}})^2}{\delta \sigma^2}$

1. Resample: $\tilde{\sigma}_{\text{exp}} \sim \mathcal{N}(\sigma_{\text{exp}}, \Sigma)$

2. Minimise: $\bar{c} = \arg \min_c \frac{(\sigma(c) - \tilde{\sigma}_{\text{exp}})^2}{\delta \sigma^2}$

3. Repeat, and treat the sample $\{\bar{c}\}$ as a sample from the Bayesian posterior $p(c|D)$

- Often used in the context of PDF fitting and SMEFT fitting, e.g. 2109.02653, 1901.05965
The Monte Carlo Replica Method

**Problem:** in the presence of a quadratic theory, often the minimum $\chi^2$ will be given by the same $\bar{c}$.

2. Minimise:  
$$\bar{c} = \arg \min_c \frac{(\sigma(c) - \tilde{\sigma}_{\text{exp}})^2}{\delta \sigma^2}$$
The Monte Carlo Replica Method

Problem: in the presence of a quadratic theory, often the minimum $\chi^2$ will be given by the same $\bar{c}$. 

$$\sigma(c) = \sigma_{SM} + \sigma_{\text{lin}}c + \sigma_{\text{quad}}c^2$$

Problem: in the presence of a quadratic theory, often the minimum $\chi^2$ will be given by the same $\bar{c}$.