3.2 Distributions and new physics

Fitting distributions and effective-theory coefficients together, and how new physics can hide in a fit.

1. The question

Every parton distribution function (PDF) set was fitted under an assumption about the theory, which enters through the hard cross-sections, the coupling α_S and the evolution. A fit that takes $\alpha_S(M_Z) = 0.118$ returns distributions that are only consistent with that value. Using them to measure α_S from a new process biases the answer towards 0.118 (Moore, section 1.4 [1]).

The same argument applies to parameters that the Standard Model sets to zero. A global fit computes every cross-section in the Standard Model. If nature contains a new heavy particle, some of the data in the fit are shifted by it. The fit has no parameter for that shift, but it has a flexible function of x, and it will use it. Two questions follow (Moore, section 1.4 [1]; the second is the subject of chapter 5 of the thesis):

1. How do bounds on new physics change when the distributions are fitted together with the new-physics parameters, compared with holding Standard Model distributions fixed?
2. If new physics is present in the data, how far do distributions fitted under the Standard Model assumption absorb it, and what happens to a later search that uses them?

A first study with deep inelastic data alone found that such corrections can be partly absorbed by the distributions, and can be separated from them through their different dependence on energy [12]. This chapter answers both questions for the case studied in chapter 3 of the thesis and the paper it is based on [11]: heavy new physics in the high-mass tail of the Drell–Yan process, described by two parameters $\hat{W}$ and $\hat{Y}$.

Sections 2 to 5 derive the theory: effective field theories, the two operators, and a change in the Drell–Yan cross-section that grows as the square of the lepton-pair mass. Section 6 derives the factor by which a bound loosens when distributions and coefficients are fitted together. Sections 7 to 10 give the method and results of the thesis, quoted exactly, for current data and for the high-luminosity LHC (HL-LHC). Section 11 covers a light new particle, the SIMUnet method, and our own fits. The lowest-order Drell–Yan cross-section and the luminosity $\mathcal{L}_{q\bar{q}}$ come from chapter 1.6, the Fisher matrix and flat directions from chapter 2.3.

2. Integrating out a heavy particle

A heavy particle exchanged at low energy looks like a point interaction, with a strength suppressed by the particle's mass squared. This section derives that in the smallest model that shows it.

The model

Take one light fermion ψ of mass m and one heavy real scalar ϕ of mass $M \gg m$, coupled with strength g (Moore, Eq. 3.1 [1]):

$$\mathcal{L} = \tfrac{1}{2}(\partial\phi)^2 - \tfrac{1}{2}M^2\phi^2 + \bar{\psi}\,(i\partial\!\!\!/ - m)\,\psi - g\,\phi\,\bar{\psi}\psi. \tag{3.2.1}$$

At energies $E \ll M$ no ϕ particle can be produced. It still acts as a virtual particle exchanged between fermions.

Solving for the heavy field

The Lagrangian is quadratic in ϕ, so ϕ can be eliminated exactly. Its equation of motion is $(\partial^2 + M^2)\phi = -g\,\bar\psi\psi$. When every momentum flowing through ϕ is small compared with M, the derivative is a small correction:

$$\phi = -\frac{g}{\partial^2 + M^2}\,\bar\psi\psi = -\frac{g}{M^2}\,\bar\psi\psi + O\Big(\frac{\partial^2}{M^4}\Big). \tag{3.2.2}$$

At low energy the heavy field is locked to the fermion density. Substituting Eq. (3.2.2) back into Eq. (3.2.1) gives a Lagrangian for ψ alone (derivation D1):

$$\mathcal{L}_{\text{eff}} = \bar\psi\,(i\partial\!\!\!/ - m)\,\psi + \frac{g^2}{2M^2}\,(\bar\psi\psi)^2 + O\Big(\frac{1}{M^4}\Big). \tag{3.2.3}$$

The same result follows from the Feynman diagram. Exchange of ϕ with momentum q contributes the propagator

$$\frac{i}{q^2 - M^2} = -\frac{i}{M^2}\Big(1 + \frac{q^2}{M^2} + \frac{q^4}{M^4} + \dots\Big). \tag{3.2.4}$$

The first term does not depend on q. In position space it is a delta function: the two vertices sit at the same point. The later terms are the $O(1/M^4)$ corrections of Eq. (3.2.3).

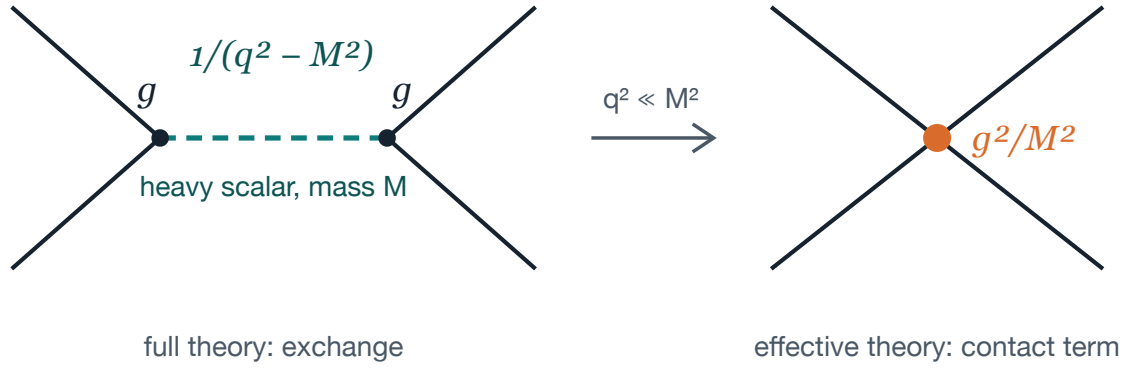

Figure 1. Matching. On the left the two fermion pairs are joined by the propagator of a heavy particle. When the momentum q it carries is far below its mass M, the propagator is the constant $-1/M^2$ and the diagram cannot be told from a point interaction of four fermions, on the right. The strength of the point interaction is fixed by the coupling and mass of the particle that was removed.

What the result shows

Three features of Eq. (3.2.3) are general.

- **The new term is not renormalisable.** A fermion field has mass dimension $\frac{3}{2}$, so $(\bar\psi\psi)^2$ has dimension 6 and its coefficient has dimension -2. An effective theory may contain such a term, because it is only claimed to hold for $E \ll M$.
- **The corrections are ordered by dimension.** Each further term in Eq. (3.2.4) adds two derivatives and two powers of $1/M$. An operator of dimension d comes with $1/M^{d-4}$ and has relative size $(E/M)^{d-4}$ at energy E.
- **Mass and coupling appear only in a ratio.** Low-energy measurements determine g^2/M^2 and nothing else.

Building the low-energy theory from a known high-energy one is the *top-down* use of an effective theory. Fermi's theory of the weak interaction, the Standard Model with the W and Z bosons integrated out, is the standard example (Moore, section 3.1.1 [1]).

---

3. The Standard Model as an effective theory

Here the construction is turned round. The high-energy theory is unknown, and the effective theory parametrises whatever it might do at accessible energies.

Bottom-up

Suppose only the low-energy theory $\bar{\psi}(i\partial\!\!\!/ - m)\psi$ is known. Whatever lies above it, its low-energy effects are described by adding every operator built from ψ and derivatives that respects the symmetries assumed to hold at high energy, ordered by dimension and divided by powers of a scale Λ. Up to dimension 6 the list reduces to one operator (Moore, Eqs. 3.10 to 3.13 [1]):

$$\mathcal{L}^{\mathrm{EFT}} = \bar{\psi}\,(i\partial\!\!\!/ - m)\,\psi - \frac{g_6}{\Lambda^2}\,(\bar{\psi}\psi)^2 - \dots, \qquad \frac{g_6}{\Lambda^2} = -\frac{g^2}{2M^2}. \tag{3.2.5}$$

The second relation is the matching to the model of section 2, from Eq. (3.2.3). The thesis prints it as $-g^2/M^2$ (derivation D1). The dimensionless number g_6 is a *Wilson coefficient*. A measurement of it constrains every model at once, each through its own matching relation.

Operators can be redundant. Terms with $\partial^2\psi$ were dropped from Eq. (3.2.5) because the free equation of motion turns $\partial^2\psi$ into $-m^2\psi$, with an error of higher order in $1/\Lambda$. Operators related by integration by parts are also equivalent. A minimal independent set is a *basis*, and only the coefficients of a basis can be determined from data. The coefficients run with scale, which the thesis and this chapter neglect.

The SMEFT

Apply this to the Standard Model. The Standard Model Effective Field Theory (SMEFT) is (Moore, Eq. 3.16 [1])

$$\mathcal{L}_{\mathrm{SMEFT}} = \mathcal{L}_{\mathrm{SM}} + \sum_{d\geq 5} \sum_{i=1}^{n_d} \frac{c_d^{(i)}}{\Lambda^{d-4}}\,\mathcal{O}_d^{(i)}. \tag{3.2.6}$$

The operators $\mathcal{O}_d^{(i)}$ are built from Standard Model fields and respect Lorentz and gauge symmetry. At dimension 5 there are only the lepton-number-violating Weinberg operators, which are strongly constrained and usually ignored. At dimension 6 there are $n_6 = 2499$ independent operators, organised in the Warsaw basis [2].

A dimension-6 operator interferes with the Standard Model amplitude. For a four-fermion operator in a two-to-two process at energy E, the Standard Model amplitude is a dimensionless coupling and the new one is $c\,E^2/\Lambda^2$:

$$\frac{\mathcal{A}_6}{\mathcal{A}_{\mathrm{SM}}} \sim \frac{c}{g_{\mathrm{SM}}^2}\,\frac{E^2}{\Lambda^2}. \tag{3.2.7}$$

The relative effect grows with energy, so a modest measurement at high energy can compete with a precise one at low energy (Moore, section 3.2 [1]; [3]). Section 5 puts numbers to this. The growth stops at $E \sim \Lambda$, where every term in Eq. (3.2.6) is as large as the first and the expansion fails.

4. The oblique parameters W and Y

This section defines the two-parameter scenario of the thesis and derives its form as a contact interaction.

Definition

In universal theories, new physics affects light fermions only through the propagators of the electroweak gauge bosons. Expanding the gauge-boson self-energies up to order q^4 leaves four parameters that correspond to dimension-6 operators: $\hat{S}$, $\hat{T}$, $\hat{W}$ and $\hat{Y}$ [4, 5]. $\hat{T}$ and $\hat{S}$ enter at order q^0 and q^2 and are well measured at LEP. $\hat{W}$ and $\hat{Y}$ enter at order q^4, so theirs is the effect that grows at the LHC [3]. They are defined by (Moore, section 3.3, footnote 2 [1])

$$\mathcal{L}_{\text{SMEFT}} \supset -\frac{\hat{W}}{4m_W^2}\left(D_\rho W^a_{\mu\nu}\right)^2 - \frac{\hat{Y}}{4m_W^2}\left(\partial_\rho B_{\mu\nu}\right)^2. \tag{3.2.8}$$

Here $W^a_{\mu\nu}$ and $B_{\mu\nu}$ are the field strengths of the $SU(2)_L$ and hypercharge gauge fields and D_ρ is the covariant derivative. The W mass m_W makes $\hat{W}$ and $\hat{Y}$ dimensionless.

Four-fermion form

The gauge fields obey $D^\mu W^a_{\mu\nu} = g\, J^a_{L\nu} + \dots$ and $\partial^\mu B_{\mu\nu} = g_Y J_{Y\nu} + \dots$, where g and g_Y are the two gauge couplings and the J are the fermion currents they couple to. Using these in Eq. (3.2.8) gives (derivation D2; Moore, Eqs. 3.17 and 3.18 [1])

$$\begin{aligned} \mathcal{L}_{\text{SMEFT}} &\supset -\frac{g^2\,\hat{W}}{2m_W^2}\, J^a_{L\mu} J^{a\mu}_L - \frac{g_Y^2\,\hat{Y}}{2m_W^2}\, J_{Y\mu} J^\mu_Y, \\ J^{a\mu}_L &= \sum_{f=q,l} \bar{f}\,\frac{\sigma^a}{2}\gamma^\mu f, \qquad J^\mu_Y = \sum_{f=q,l,u,d,e} Y_f\, \bar{f}\gamma^\mu f. \end{aligned} \tag{3.2.9}$$

The sums run over the left-handed doublets q, l and the right-handed singlets u, d, e of all three generations. σ^a are the Pauli matrices, and the hypercharges are $Y_f = \frac{1}{6}, -\frac{1}{2}, \frac{2}{3}, -\frac{1}{3}, -1$ for q, l, u, d, e. Like Eq. (3.2.3), this is a product of two currents with a coefficient of dimension -2. In the Warsaw basis it is a fixed combination of seven operators, $\mathcal{O}^{(3)}_{lq}$ for $\hat{W}$ and $\mathcal{O}_{ld}, \mathcal{O}_{lu}, \mathcal{O}^{(1)}_{lq}, \mathcal{O}_{ed}, \mathcal{O}_{eu}, \mathcal{O}_{qe}$ for $\hat{Y}$ (Moore, Eqs. 3.19 and 3.20 [1]).

Two readings

As a change to a propagator. The terms in Eq. (3.2.8) are quadratic in the gauge fields with two extra derivatives, and add $\hat{W} q^4/m_W^2$ to the inverse propagator. To first order (derivation D2),

$$\frac{1}{q^2 - m_V^2} \;\longrightarrow\; \frac{1}{q^2 - m_V^2} - \frac{\hat{W}}{m_W^2}, \tag{3.2.10}$$

for the $SU(2)_L$ bosons, and the same with $\hat{Y}$ for the hypercharge boson. A constant added to a propagator is a contact term, as in Eq. (3.2.4).

As a heavy particle. A heavy vector of mass M coupled with strength g_* to J^a_L, or to J_Y, gives $-(g_*^2/2M^2)\, J\cdot J$ when integrated out as in section 2. Comparing with Eq. (3.2.9):

$$\hat{W} = \frac{g_*^2}{g^2}\,\frac{m_W^2}{M^2} \quad \text{(triplet)}, \qquad \hat{Y} = \frac{g_*^2}{g_Y^2}\,\frac{m_W^2}{M^2} \quad \text{(hypercharge vector)}. \tag{3.2.11}$$

The second relation is Moore's Eq. (5.10) for a flavour-universal Z' [1]. With $g_* = g$, a value $\hat{W} = 10^{-4}$ corresponds to $M = m_W/\sqrt{\hat{W}} = 8.0$ TeV, and $\hat{W} = 6 \times 10^{-4}$ to 3.3 TeV. Both parameters are positive in these models.

Earlier bounds from 8 TeV LHC data, obtained with Standard Model PDFs, were $\hat{W} \in [-3, 15] \times 10^{-4}$ (ATLAS) and $[-5, 22] \times 10^{-4}$ (CMS) for $\hat{Y} = 0$, and $\hat{Y} \in [-4, 24] \times 10^{-4}$ (ATLAS) and $[-7, 41] \times 10^{-4}$ (CMS) for $\hat{W} = 0$, all at 95% confidence (Moore, Eqs. 3.21 and 3.22, from [3]).

---

5. High-mass Drell–Yan: why the effect grows with energy

This section derives the factor by which $\hat{W}$ and $\hat{Y}$ change the lowest-order Drell–Yan cross-section.

Helicity amplitudes

With fermion masses neglected, each combination of quark helicity X and lepton helicity Y (each left or right) scatters independently. Chapter 1.6 derived $q\bar{q} \to l^+l^-$ through a photon. Adding the Z boson and the contact terms of Eq. (3.2.9), the cross-section at partonic energy squared $\hat{s}$ is (derivation D3)

$$\hat{\sigma}(q\bar{q} \to l^+l^-) = \frac{\hat{s}}{144\pi} \sum_{X,Y} \left| A^q_{XY}(\hat{s}) \right|^2, \tag{3.2.12}$$

$$A^q_{XY}(\hat{s}) = \frac{e^2 Q_q Q_l}{\hat{s}} + \frac{g_Z^{q_X}\, g_Z^{l_Y}}{\hat{s} - m_Z^2} - \frac{g^2\, T^3_{q_X} T^3_{l_Y}\, \hat{W} + g_Y^2\, Y_{q_X} Y_{l_Y}\, \hat{Y}}{m_W^2}. \tag{3.2.13}$$

Here Q is the electric charge in units of e, T^3 the third component of weak isospin ($\pm\frac{1}{2}$ for left-handed fermions, 0 for right-handed), $Y = Q - T^3$ the hypercharge, and $g_Z^f = (g/\cos\theta_W)(T_f^3 - Q_f \sin^2\theta_W)$ the coupling to the Z. The factor $1/144\pi$ contains the colour average $1/N_c$. With the photon term alone, Eq. (3.2.12) is the result $4\pi\alpha^2 e_q^2/3N_c\hat{s}$ of chapter 1.6. The Z width is neglected, which is valid away from the resonance.

Equations (3.2.12) to (3.2.16) are derived here. The thesis gives no closed formula and computes the factors numerically. Equation (3.2.13) adds to the Standard Model amplitude the contact terms of Eq. (3.2.9) and nothing else. In that form it is the tree-level reweighting factor of Torre, Ricci and Wulzer, who define $\hat{W}$ and $\hat{Y}$ as the coefficients of the contact terms (Eqs. 11 and 12 of [7]). Farina et al. start from Eq. (3.2.8) and give the modified propagators (Eq. 1 of [3]). Their constant terms are those of Eq. (3.2.13). Their pole terms are also changed, by a relative amount of order $\hat{W}$ and $\hat{Y}$ that does not grow with energy: the coupling and mass shifts of derivation D2. In the check script that difference is below 2% of the growing term at 1 TeV and falls as $1/\hat{s}$.

The high-energy limit

For $\hat{s} \gg m_Z^2$ the photon and Z terms combine into exchange of the unmixed gauge fields, because $e^2 Q_q Q_l + g_Z^q g_Z^l = g^2 T_q^3 T_l^3 + g_Y^2 Y_q Y_l$ identically. Then

$$A^q_{XY}(\hat{s}) \;\longrightarrow\; \frac{1}{\hat{s}} \left[g^2\, T^3_{q_X} T^3_{l_Y} \left(1 - \hat{W}\frac{\hat{s}}{m_W^2}\right) + g_Y^2\, Y_{q_X} Y_{l_Y} \left(1 - \hat{Y}\frac{\hat{s}}{m_W^2}\right) \right]. \tag{3.2.14}$$

Each gauge interaction is rescaled by a factor that departs from one in proportion to $\hat{s}$, which is Eq. (3.2.7) with the coefficient worked out. Keeping the interference term only, the ratio of the cross-section to its Standard Model value is

$$K_q - 1 \simeq -\frac{2\hat{s}}{m_W^2}\left(w_q\,\hat{W} + y_q\,\hat{Y}\right), \qquad w_q = \frac{\sum_{XY} G_{XY}\, g^2 T^3_{q_X} T^3_{l_Y}}{\sum_{XY} G_{XY}^2}, \quad y_q = 1 - w_q, \tag{3.2.15}$$

with $G_{XY} = g^2 T^3_{q_X} T^3_{l_Y} + g'^2 Y_{q_X} Y_{l_Y}$. The weights depend only on $\sin^2\theta_W$. For $\sin^2\theta_W = 0.2312$ they are $w_u = 0.54$, $y_u = 0.46$ for up quarks and $w_d = 0.86$, $y_d = 0.14$ for down quarks.

For the charged-current process $u\bar{d} \to l^+\nu$ only the W contributes. With the contact term alone the factor needs no expansion in $\hat{s}$, and is the charged-current factor of [7]:

$$K_{\mathrm{CC}} = \left[1 - \hat{W}\,\frac{\hat{s} - m_W^2}{m_W^2}\right]^2 . \tag{3.2.16}$$

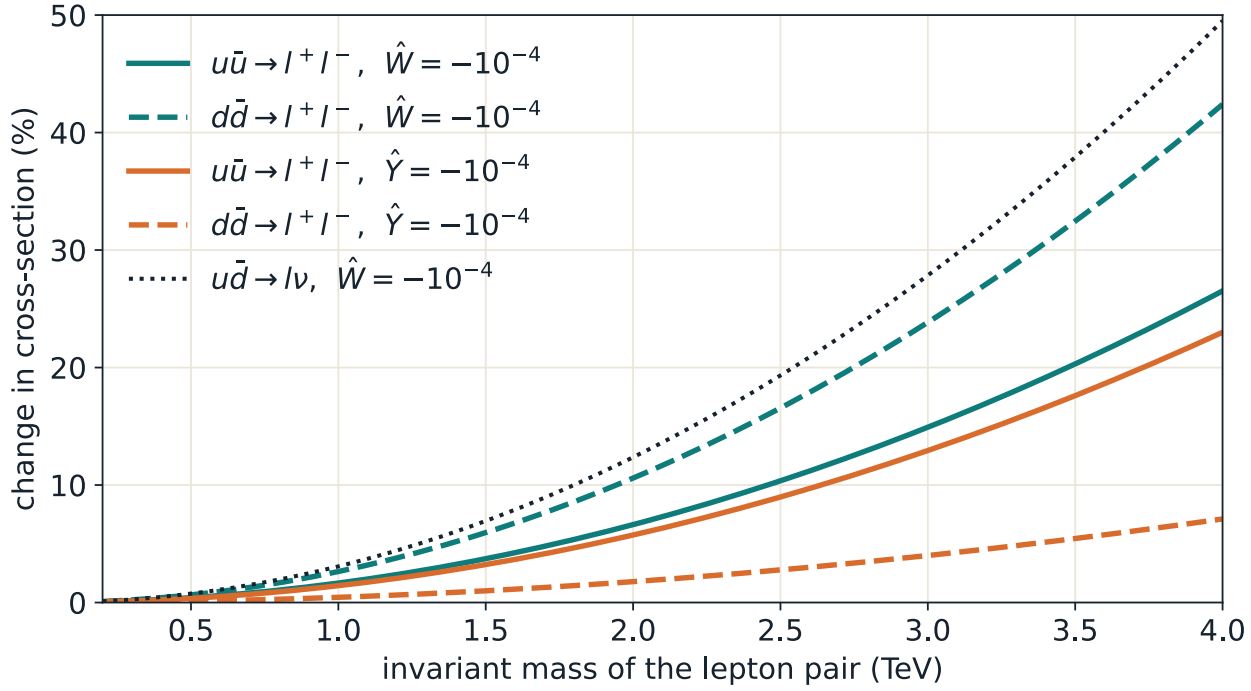

Figure 2. The energy growth of Eq. (3.2.15), computed from Eqs. (3.2.12) and (3.2.13) at parton level with the interference term only. Each curve is the change in the cross-section for one parameter set to -10^{-4} and the other to zero. At 1 TeV the changes are between 0.4% and 3%, at 4 TeV between 7% and 50%. The down-quark channel is almost blind to $\hat{Y}$, and the charged-current channel depends on $\hat{W}$ alone.

Numbers

With $\hat{s} = m_{ll}^2$, the square of the lepton-pair mass, $\hat{s}/m_W^2$ is 155 at 1 TeV and 1390 at 3 TeV. Three consequences:

- **Size.** For $\hat{W} = -10^{-3}$ the up-quark channel changes by $+17\%$ at 1 TeV and the down-quark channel by $+27\%$. Corrections of this size in the highest bins of the ATLAS and CMS data are comparable with the experimental errors (Moore, Fig. 3.3 [1]; the legends of that figure give $\hat{W}, \hat{Y} = -10^{-3}$ with a positive correction, the accompanying text $+10^{-3}$).
- **Energy against precision.** A measurement of relative precision δ at mass m_{ll} is sensitive to $\hat{W} \approx \delta\, m_W^2/(2w\, m_{ll}^2)$. With $w = 0.6$, 10% at 1 TeV reaches 5×10^{-4}. That is below the few per mille at which the precise measurements of LEP constrained these parameters, through the terms that do not grow with energy [3]. At 3 TeV, 10% reaches 6×10^{-5}.
- **Deep inelastic scattering.** There the exchanged momentum is space-like, $q^2 = -Q^2$, and in Eq. (3.2.13) with $\hat{s} \to -Q^2$ the sign of the interference reverses. Its size is of order $2w|\hat{W}|Q^2/m_W^2$, or 7×10^{-4} at $Q^2 = 4 \times 10^4$ GeV^2 for $|\hat{W}| = 10^{-4}$. The full calculation in the thesis finds it well below one per cent across the HERA

range (Moore, Fig. 3.2 [1]), so the constraints come from Drell–Yan. That figure, for $\hat{W}, \hat{Y} = -10^{-4}$, is drawn with a positive colour scale and does not show the change of sign.

In a fit

The thesis keeps the best Standard Model prediction and multiplies it bin by bin by a factor that carries the dependence on the Wilson coefficients c_n (Moore, Eqs. 3.31 to 3.34 [1]):

$$\begin{aligned} d\sigma_{\mathrm{SMEFT}} &= d\sigma_{\mathrm{SM}} \times K_{\mathrm{EFT}}, \qquad K_{\mathrm{EFT}} = 1 + \sum_n c_n\, R^{(n)} + \sum_{n,m} c_n c_m\, R^{(n,m)}, \\ R^{(n)} &= \frac{\mathcal{L}_{ij} \otimes d\hat{\sigma}_{ij}^{(n)}}{\mathcal{L}_{ij} \otimes d\hat{\sigma}_{ij,\mathrm{SM}}}. \end{aligned} \tag{3.2.17}$$

$\mathcal{L}_{ij}$ is the luminosity of partons i and j, and $d\hat{\sigma}^{(n)}$ is the interference of operator n with the Standard Model. The ratios R are computed once at lowest order with a reference PDF set and held fixed. Their dependence on the PDF set is at the per-mille level. In the $\hat{W}, \hat{Y}$ analysis only the linear terms are kept.

---

6. How a fit absorbs new physics

This section shows why distributions and Wilson coefficients are entangled in a high-mass tail, and by how much a bound loosens when they are fitted together.

One energy, one function

From chapter 1.6, the lowest-order mass spectrum at a collider of energy $\sqrt{s}$ is a known factor times the quark–antiquark luminosity at $\tau = M^2/s$. With Eq. (3.2.17), the prediction for bin i at mass M_i has the form

$$\begin{aligned} T_i(\theta, c) &= T_i^{\mathrm{SM}}(\theta)\,\big[1 + c\, s_i\big], \\ T_i^{\mathrm{SM}}(\theta) &\propto \mathcal{L}_{q\bar{q}}\big(\tau_i, M_i^2; \theta\big), \qquad s_i \propto \frac{M_i^2}{m_W^2} = \frac{s}{m_W^2}\,\tau_i. \end{aligned} \tag{3.2.18}$$

θ stands for the PDF parameters and c for $\hat{W}$ or $\hat{Y}$. At fixed s, the mass and τ are the same variable, and the tail measures one function of τ: the product of the luminosity and the bracket. A reduction of the cross-section that grows as τ can come from $c > 0$, or equally from a luminosity that falls a little faster at large τ.

The luminosity at large τ is set by antiquarks at large x, where the distributions are least well known. In the baseline fit of the thesis its PDF uncertainty is about 5% at 2 TeV (Moore, section 3.6.2 [1]), so the fit has room.

Three things break the degeneracy.

- **Other data at the same x.** A change at large x made to fit the tail must also describe deep inelastic and fixed-target data at the same x and low scales, where $c\, s_i$ is negligible. The scales are tied by evolution, which is logarithmic (Eq. 1.4.18) and cannot generate a power of M^2.
- **A second collider energy.** The same τ at a different s is a different mass. The luminosity term is unchanged and the new-physics term is not.
- **Other channels.** The charged-current tail depends on $\hat{W}$ alone, Eq. (3.2.16), and on a different combination of flavours.

The Snowmass report makes the first point for jets: a PDF explanation of a high-p_T excess must also describe data at the same x and lower p_T (section II.B.3 of [6]).

The linear model

To quantify this, linearise. Let r_i be the relative deviation of the prediction for data point i from a reference, and let the experimental covariance of the relative deviations be C:

$$r_i = \sum_k A_{ik}\,\theta_k + b_i\,c, \qquad F_{\theta\theta} = A^T C^{-1} A, \quad F_{\theta c} = A^T C^{-1} b, \quad F_{cc} = b^T C^{-1} b. \tag{3.2.19}$$

A_{ik} is the response of point i to PDF parameter k, and b_i its response to the Wilson coefficient ($b_i = s_i$ in the tail, zero elsewhere). F is the Fisher matrix of chapter 2.3 in block form. The one-sigma uncertainty on c is then (derivation D4)

$$\sigma_c^{\text{fixed}} = \frac{1}{\sqrt{F_{cc}}}, \qquad \sigma_c^{\text{sim}} = \frac{1}{\sqrt{F_{cc}\,(1-\rho^2)}}, \qquad \rho^2 = \frac{F_{c\theta}\,F_{\theta\theta}^{-1}\,F_{\theta c}}{F_{cc}}. \tag{3.2.20}$$

The first is for distributions held fixed with no uncertainty, the second for a simultaneous fit. ρ^2 lies between 0 and 1 and is the fraction of the data's sensitivity to c that a change of the distributions can imitate.

> **The broadening factor.** Fitting the distributions together with a Wilson coefficient loosens its bound by $1/\sqrt{1-\rho^2}$. When the data that constrain c are also the only data that constrain the relevant part of the distributions, $\rho^2 \to 1$ and the bound disappears. This is a flat direction of the joint fit, of the kind studied in chapter 2.3.

A toy with one tail

The check script builds the smallest version. It has twelve bins from 0.5 to 4 TeV at $\sqrt{s} = 14$ TeV with errors from 0.5% to 7%, a three-parameter change of the luminosity $\theta_1 + \theta_2\,x + \theta_3\,x^2$ in $x = \sqrt{\tau}$, a coefficient $c = \hat{W} \times 10^5$ with $s_i = -1.2\,M_i^2/m_W^2$, and twelve low-energy points at the same x with errors from 2% to 30% and no sensitivity to c. The inputs are illustrative. Errors on c are in units of 10^{-5}.

DATA IN THE FIT	FIXED-PDF ERROR	SIMULTANEOUS ERROR	BROADENING	DEGENERACY
tail only	0.97	unbounded	infinite	$\rho^2 = 1$
tail and low-energy data	0.97	5.05	5.2	$\rho^2 = 0.963$
low-energy errors three times larger	0.97	14.9	15.3	$\rho^2 = 0.996$
a second collider energy added	0.97	3.28	3.4	$\rho^2 = 0.912$

With the tail alone the degeneracy is exact, because $s_i \propto x_i^2$ is one of the three shapes available to the luminosity. The low-energy data make the bound finite. The second energy in the last row is a tail at the same x and half the mass, with errors three times larger.

7. Simultaneous fits by scanning

The thesis performs the joint fit by a scan: one complete PDF fit for each value of the Wilson coefficients.

Data and theory

The fit uses deep inelastic and Drell–Yan data only. Jets and top production are left out because other SMEFT operators could affect them, and the study varies only $\hat{W}$ and $\hat{Y}$ (Moore, section 3.4.1 [1]). The Drell–Yan data are 659 low-mass and on-shell points with no sensitivity to the two parameters (the printed total of Moore's Table 3.1; its rows add to 660), and 270 points from five high-mass neutral-current measurements by ATLAS and CMS at 7, 8 and 13 TeV, whose highest bins end between 1.5 and 3 TeV (Moore, Table 3.2).

Standard Model predictions are at NNLO in QCD with NLO electroweak corrections. The latter are essential in the tail: for the CMS 13 TeV combined channel, χ^2 per point falls from 19.7 to 1.91 when they are included (Moore, Table 3.3). The SMEFT enters through Eq. (3.2.17), for the high-mass sets and for the deep inelastic data sets whose reach in Q exceeds 120 GeV.

Two profiles

For each point c of a grid in the parameter space, two numbers are computed (Moore, section 3.4.4 [1]):

$$\chi^2_{\text{eftp}}(c) = \min_{\theta} \chi^2(\theta, c), \qquad \chi^2_{\text{smp}}(c) = \chi^2(\theta_{\text{SM}}, c). \tag{3.2.21}$$

The first comes from a full NNPDF fit of the distributions with the Wilson coefficients held at c. These are the *SMEFT PDFs*. The second uses the baseline distributions θ_{SM}, fitted once with $c = 0$. χ^2_{eftp} is the profile of the joint χ^2 over the PDF parameters. In the linear model its curvature in c is $F_{cc}(1-\rho^2)$ and that of χ^2_{smp} is F_{cc}, as in Eq. (3.2.20).

Near its minimum each profile is fitted with a quadratic form:

$$\begin{aligned} \chi^2(c) &= \chi^2_0 + \sum_{n,m} \left(c_n - c_n^{(0)}\right) H_{nm} \left(c_m - c_m^{(0)}\right), \\ \Delta\chi^2 &= \frac{\left(c - c^{(0)}\right)^2}{\delta c^2} \qquad \text{for one parameter.} \end{aligned} \tag{3.2.22}$$

In this form, which is that of the thesis, H_{nm} is half the second-derivative matrix that chapter 2.3 calls the Hessian. The interval is read off at $\Delta\chi^2 = 1$ for 68% and $\Delta\chi^2 = 4$ for 95% (strictly 95.45%; the thesis states the second value only). With linear new-physics terms the quadratic form is exact for χ^2_{smp} and approximate for χ^2_{eftp}, because the distributions respond to c.

Two uncertainties

PDF uncertainty in the fixed-PDF bound. The bound from χ^2_{smp} with the central distributions ignores PDF uncertainty. To include it, the thesis repeats the parabola for each replica k of the baseline set and quotes the outermost limits of the 68% envelope of the intervals $[c^{(k)}_{\min}, c^{(k)}_{\max}]$ (Moore, Eq. 3.37). The simultaneous bound contains the PDF uncertainty by construction.

Finite number of replicas. Each $\chi^2_{\text{eftp}}(c)$ is computed from the mean prediction of N_{rep} replicas. Its fluctuation is found by bootstrap: resample the replicas with replacement $N_{\text{res}} \simeq 10^4$ times, recompute the mean prediction T_l and its χ^2 for each resample l, and take

$$\sigma_{\chi^2}(c) = \mathrm{std}_l\Big[\chi^2\big(T_l(c)\big)\Big]. \tag{3.2.23}$$

These are the error bars on the scan points (Moore, Eq. 3.40 and Fig. 3.5).

The scan is expensive. The analysis of current data uses 21 values of $\hat{Y}$ and 17 of $\hat{W}$, each a full fit, and the cost rules out more than a few parameters. Section 11 describes the method that replaced it.

---

8. Results with current data

With the data available at the time, the interplay is visible and small. The fits are one-dimensional, varying $\hat{W}$ or $\hat{Y}$ with the other at zero. Both parameters scale as q^4, so a neutral-current mass spectrum constrains only the combination in Eq. (3.2.15). The flat direction this leaves in the $(\hat{W}, \hat{Y})$ plane can only be lifted by charged-current data (Moore, section 3.5 [1]).

The comparison is expressed through two numbers, for $\hat{W}$ and likewise for $\hat{Y}$ (Moore, Eqs. 3.43 and 3.44):

$$\begin{aligned} \text{shift} &= \hat{W}^{(0)}\Big|_{\text{SMEFT PDF}} - \hat{W}^{(0)}\Big|_{\text{SM PDF}}, \\ \text{broadening} &= \frac{\delta\hat{W}|_{\text{SMEFT PDF}} - \delta\hat{W}|_{\text{SM PDF}}}{\delta\hat{W}|_{\text{SM PDF}}}. \end{aligned} \tag{3.2.24}$$

The results are Moore's Table 3.4, in units of 10^{-4}. The last column is the broadening relative to the fixed-PDF bound without and with PDF uncertainty.

PARAMETER, LEVEL	FIXED, NO PDF UNCERTAINTY	FIXED, WITH PDF UNCERTAINTY	SIMULTANEOUS	BROADENING
$\hat{W}$, 68%	$[-3.0, 2.2]$	$[-4.3, 3.8]$	$[-3.5, 2.4]$	$+13\%$, -27%
$\hat{W}$, 95%	$[-5.5, 4.7]$	$[-6.8, 6.3]$	$[-6.4, 5.3]$	$+15\%$, -11%
$\hat{Y}$, 68%	$[-4.4, 4.7]$	$[-6.7, 7.5]$	$[-3.4, 6.9]$	$+13\%$, -27%
$\hat{Y}$, 95%	$[-8.8, 9.2]$	$[-11.1, 12.0]$	$[-8.3, 11.8]$	$+12\%$, -13%

Four things can be read from the table.

- **Both parameters are consistent with zero.**
- **Against the central-PDF bound, the simultaneous fit is looser** by 12% to 15%. In the linear model a width ratio of 1.15 corresponds to $\rho^2 = 0.24$.
- **Against the bound with PDF uncertainty, it is tighter** by 11% to 27%. The two procedures combine the PDF uncertainty with the tail's own constraint differently, and the thesis reports the outcome without breaking it down further.
- **The best fit of $\hat{Y}$ moves** by $+1.6 \times 10^{-4}$, about a third of its 68% uncertainty. That of $\hat{W}$ moves by -0.2×10^{-4}.

The distributions barely respond. For $\hat{W}$ and $\hat{Y}$ near the ends of the 95% intervals, the shift in every parton luminosity is smaller than its one-sigma uncertainty, and is largest in the quark–antiquark luminosity above 500 GeV (Moore, Fig. 3.7). The high-mass data are few: at 13 TeV there is one measurement, from 5.1 fb^{-1}.

9. Projections for the high-luminosity LHC

With HL-LHC statistics the tail becomes the dominant constraint on the large-x antiquarks and on the Wilson coefficients at once.

Pseudo-data

The projected data are neutral-current spectra in m_{ll} up to 4 TeV and charged-current spectra in the transverse mass m_T up to 3.5 TeV, at $\sqrt{s} = 14$ TeV and above 500 GeV. They assume 6 ab^{-1} in total from ATLAS and CMS and a five-fold reduction of systematic errors (Moore, sections 2.3.3 and 3.6.1 [1]; section 3.6.3 prints the same luminosity as 6 fb^{-1}). The central values are the baseline Standard Model prediction σ_i^{th}, fluctuated once:

$$\sigma_i^{\text{HL}} = \sigma_i^{\text{th}}\big(1 + \lambda\,\delta_{\mathcal{L}} + r_i\,\delta_i\big), \qquad \text{cov}_{ij} = \sigma_i^{\text{th}}\sigma_j^{\text{th}}\big(\delta_{\mathcal{L}}^2 + \delta_{ij}\,\delta_i^2\big). \tag{3.2.25}$$

λ and r_i are unit Gaussian random numbers. δ_i is the total relative error of bin i: statistical, $1/\sqrt{N_i}$ for N_i expected events, and reduced systematic, in quadrature. $\delta_{\mathcal{L}} = 1.5\%$ is the luminosity error, common to all bins of one experiment. The covariance follows from the first expression. The true value of both parameters is zero.

Added to a Standard Model fit, the pseudo-data cut the PDF uncertainty on the quark–antiquark luminosity at 2 TeV from about 5% to about 2.5% with the neutral-current data, and to about 1.5% with the charged-current data as well. The quark–quark and gluon–gluon luminosities are unaffected (Moore, Fig. 3.10).

Bounds

With charged-current data the flat direction is lifted and the fit is two-dimensional, on a grid of 35 points in the $(\hat{W}, \hat{Y})$ plane. The marginalised bounds, in units of 10^{-5}, are Moore's Table 3.5:

PARAMETER, LEVEL	FIXED, NO PDF UNCERTAINTY	FIXED, WITH PDF UNCERTAINTY	SIMULTANEOUS	BROADENING
$\hat{W}$, 68%	$[-0.7, 0.5]$	$[-1.0, 0.9]$	$[-4.5, 6.9]$	850%, 500%
$\hat{W}$, 95%	$[-1.0, 0.8]$	$[-1.4, 1.2]$	$[-8.1, 10.6]$	940%, 620%
$\hat{Y}$, 68%	$[-1.8, 3.2]$	$[-3.7, 4.7]$	$[-6.4, 8.0]$	190%, 70%
$\hat{Y}$, 95%	$[-3.4, 4.7]$	$[-5.3, 6.3]$	$[-11.1, 12.6]$	190%, 110%

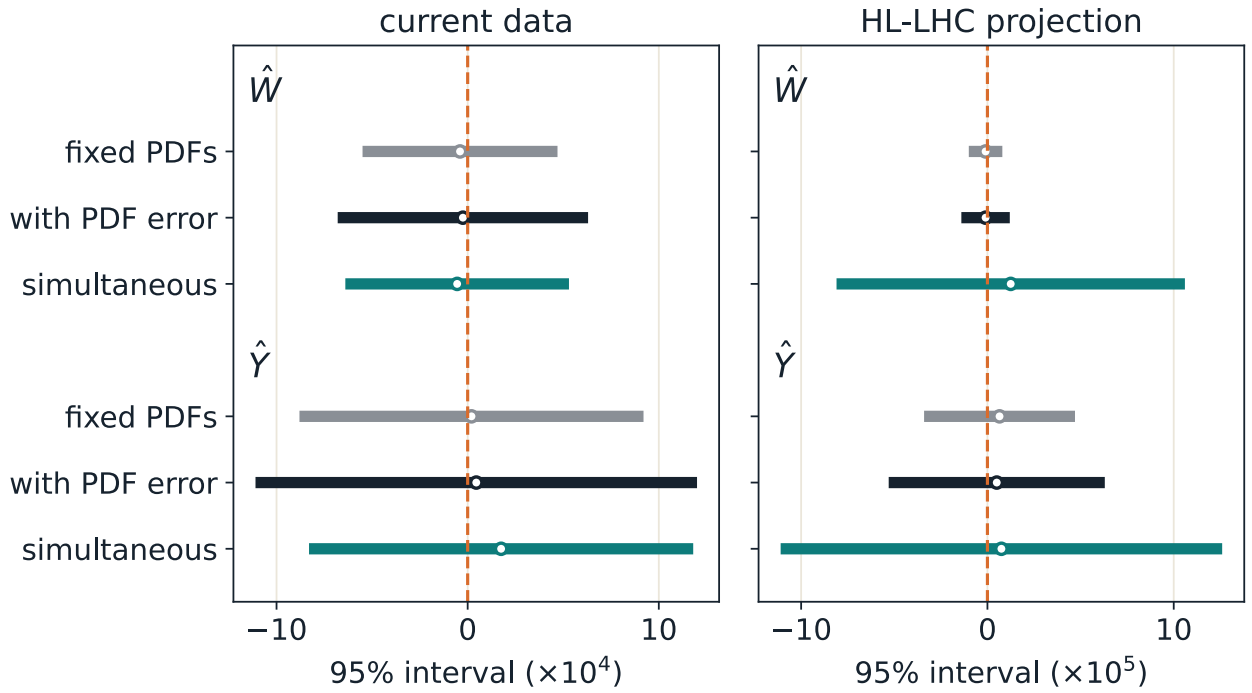

Figure 3. The 95% intervals of Moore's Tables 3.4 and 3.5 [1], drawn to scale. Left, current data, in units of 10^{-4}. Right, HL-LHC pseudo-data, in units of 10^{-5}. Grey: fixed baseline distributions, central value only. Dark: the same with PDF uncertainty. Teal: simultaneous fit. Circles mark the centres of the intervals and the dashed line is the Standard Model.

The thesis text quotes 500% for the broadening of the 95% bound on $\hat{W}$ with PDF uncertainty. Its Table 3.5 has 620% there and 500% in the 68% row, and the intervals agree with the table.

The simultaneous bound on $\hat{W}$ at 95% is wider than the central-PDF bound by a factor of 10.4, and that on $\hat{Y}$ by a factor of 2.9. In the linear model these correspond to $\rho^2 = 0.991$ and 0.88. Including PDF uncertainty in the fixed-PDF bound does not close the gap: the factors are still 7.2 and 2.0.

The reason is that of section 6. The baseline distributions were fitted to the same pseudo-data with $c = 0$, so their small uncertainty at large x is conditional on that assumption. When c is varied, the distributions follow it: the quark–antiquark luminosity of the SMEFT PDFs moves well outside the one-sigma band of the Standard Model PDFs, the width of the band is unchanged, and the χ^2 is better than with fixed PDFs (Moore, Fig. 3.12).

Conservative distributions

An alternative is to keep the distributions fixed but take them from a fit that excludes every high-mass set, current and projected. Moore's Table 3.6 gives the result at 95%, in units of 10^{-5} and with PDF uncertainty: $\hat{W} \in [-4.3, 3.1]$ and $\hat{Y} \in [-5.0, 13.7]$. These are closer to the simultaneous bounds, with remaining broadenings of 150% and 30% and best-fit shifts of 1.9 and -3.6. The thesis describes them as still slightly underestimated, and the Snowmass report repeats the point (section II.B.3 of [6]).

The fixed-PDF bounds are also tighter than an earlier projection [7], by about a factor of 4 for $\hat{W}$ and 2 for $\hat{Y}$, which the thesis attributes to its use of 13 TeV measurements as the reference, twice the luminosity, and a more optimistic reduction of systematic errors.

10. Contaminated distributions

Section 9 asked what happens to a bound when there is no new physics. This section asks what happens to a signal when there is.

The experiment

Regenerate the HL-LHC pseudo-data with the Standard Model prediction multiplied by K_{EFT} at $(\hat{W}, \hat{Y}) = (4, 8) \times 10^{-5}$, and repeat the analysis. The result is Moore's Table 3.7 (95%, units of 10^{-5}, PDF uncertainty included):

PARAMETER	FIXED PDFS FITTED TO ALL DATA	FIXED CONSERVATIVE PDFS	SIMULTANEOUS	INJECTED
$\hat{W}$	$[-1.5,\ 1.2]$	$[3.1,\ 5.0]$	$[-5.3,\ 9.0]$	4
$\hat{Y}$	$[-3.1,\ 8.7]$	$[5.8,\ 13.6]$	$[-0.2,\ 26.7]$	8

- **Fixed distributions that have seen the data miss the signal.** The interval for $\hat{W}$ is centred near zero and excludes the true value. The distributions were fitted to the contaminated tail with $c = 0$ and absorbed the effect.
- **The simultaneous fit moves towards the signal.** Both intervals contain the injected values and have shifted upward from Table 3.5. They still contain zero.
- **Conservative distributions find it.** Both intervals contain the injected values and exclude zero. For $\hat{W}$ the interval is narrower than the simultaneous one by a factor of 7.5.

These statements are read from the table. The thesis text says that the injected point lies outside the bounds of the first column. In the table that holds for $\hat{W}$ only: $\hat{Y} = 8$ lies inside $[-3.1, 8.7]$. The caption of the table calls the injected signal spurious.

The mechanism

The linear model of section 6 reproduces the first two results in closed form. Let the true coefficient be c^* and the true PDF parameters be the reference point. A fit that assumes $c = 0$ returns, on average (derivation D5),

$$\langle \hat{\theta} \rangle = F_{\theta\theta}^{-1}\, F_{\theta c}\, c^{*}. \tag{3.2.26}$$

The distributions are pulled away from the truth in proportion to the signal. Fit c next with those distributions held fixed, to the same data:

$$\langle \hat{c} \rangle = (1 - \rho^{2})\, c^{*}, \qquad \text{quoted error} = \frac{1}{\sqrt{F_{cc}}}. \tag{3.2.27}$$

Only the fraction $1 - \rho^2$ of the signal survives, and the interval is narrow and in the wrong place. The fit quality gives little warning. With N_{dat} points and p PDF parameters, the minimum χ^2 of the Standard Model fit has the expectation

$$\left\langle \chi^{2}_{\mathrm{min}} \right\rangle = N_{\mathrm{dat}} - p + \left(\frac{c^{*}}{\sigma_{c}^{\mathrm{sim}}} \right)^{2}. \tag{3.2.28}$$

What can be hidden. The χ^2 of a Standard Model fit rises by exactly the squared significance that a simultaneous fit would assign to the signal. What the simultaneous fit cannot resolve, the distributions absorb with no loss of fit quality.

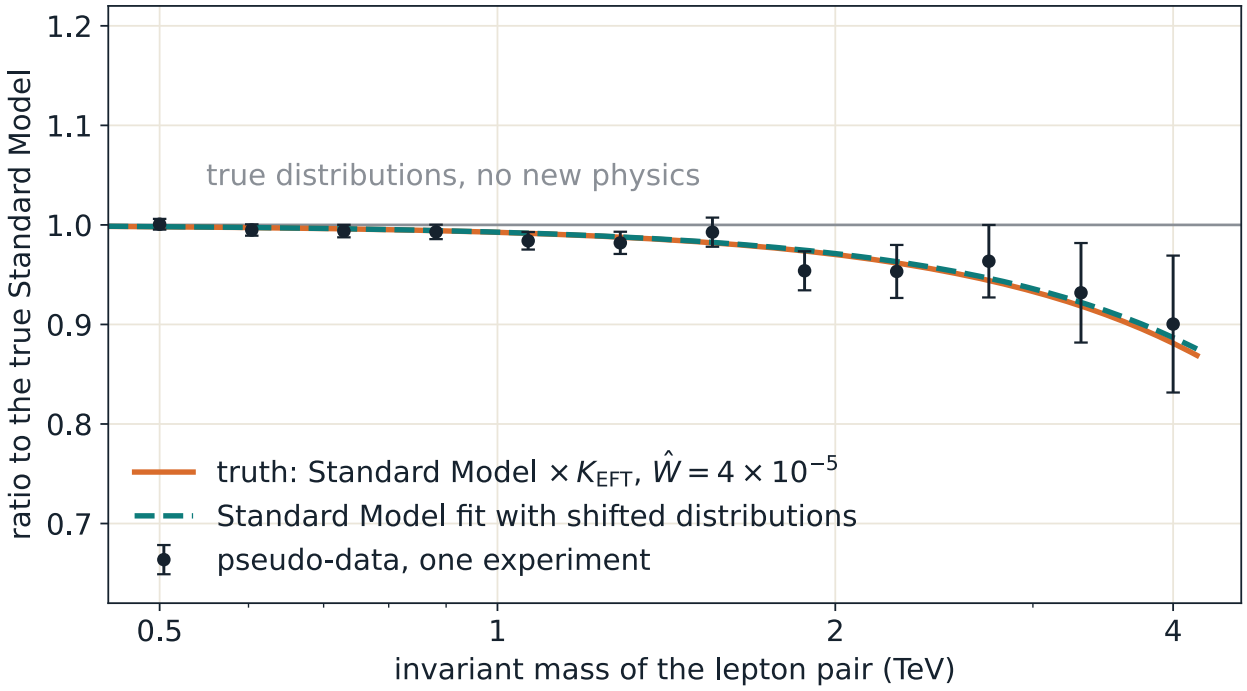

Figure 4. The toy of section 6 with a signal $\hat{W} = 4 \times 10^{-5}$ injected. Orange: the true prediction, falling below the Standard Model as the mass rises. Points: one set of pseudo-data. Dashed: the mean result of a fit that assumes the Standard Model and adjusts the distributions. It follows the signal across the whole range, with a luminosity that is wrong by up to 12% at 4 TeV. The inputs are illustrative.

In the toy, $\rho^2 = 0.963$ and the fixed-PDF error is 0.97. With exactly known distributions the injected $c^* = 4$ would be a four-sigma effect. Over 40,000 pseudo-experiments the two-step analysis returns $\hat{c} = 0.15 \pm 0.97$, and its 95% interval contains the true value in none of them. The mean χ^2 per degree of freedom of the Standard Model fit is 1.03. The simultaneous fit returns 4.0 ± 5.1 and covers the truth 95% of the time.

Conservative sets, and a caution

Conservative distributions are unbiased in the linear model, since they never saw the tail. With their PDF uncertainty propagated in full, the error on c is at least as large as that of the simultaneous fit (derivation D5). In the toy it is 5.2 against 5.1. The much narrower conservative intervals of Table 3.7 are therefore not what a linear, Gaussian treatment predicts. The thesis defers a precise definition of a conservative set to its chapter 5, where pseudo-data are generated from explicit heavy Z' and W' models, which reduce to $\hat{Y}$ and $\hat{W}$ through Eq. (3.2.11).

With real data, the thesis adds, improper treatment of the distributions could produce a spurious signal as well as hide a true one (Moore, section 3.6.3 [1]).

11. Other routes: a dark photon, SIMUnet, and this project

Light new physics: a dark photon in the proton

An effective theory assumes the new particle is heavy. Chapter 2 of the thesis treats the opposite case: a light vector boson B, a dark photon, with mass m_B between 2 and 80 GeV, coupled to all quarks equally with strength $g_B/3$ and a dark fine-structure constant α_B of order 10^{-3} [1, 8]:

$$\mathcal{L}_{\text{int}} = \tfrac{1}{3}\, g_B \sum_q \bar{q}\, \not{B}\, q. \tag{3.2.29}$$

Such a particle cannot be integrated out at LHC energies. It is radiated from quarks as the photon is, and becomes a parton. The evolution equations of chapter 1.4 gain a distribution $B(x,\mu^2)$ above the threshold $\mu = m_B$ (Moore, Eq. 2.2):

$$\begin{aligned} \mu^2 \frac{\partial q_i}{\partial \mu^2} &= \big[\text{QCD and QED terms}\big] + P_{q_i B} \otimes B, \\ \mu^2 \frac{\partial B}{\partial \mu^2} &= \sum_j \left(P_{B q_j} \otimes q_j + P_{B \bar{q}_j} \otimes \bar{q}_j \right) + P_{BB} \otimes B. \end{aligned} \tag{3.2.30}$$

The new physics now enters through the distributions and leaves the hard cross-section alone. The quark distributions are reduced, mainly at large x, and the observable is again the high-mass Drell–Yan tail. For consistency at order α_B the evolution is at NNLO in QCD and NLO in QED.

The thesis builds 21 PDF sets on a grid of (m_B, α_B), compares each with HL-LHC neutral-current pseudo-data, and sets a 95% bound where $\Delta\chi^2 = 3.8$. The projected sensitivity is competitive with existing bounds, in particular for m_B between 2 and 6 GeV and between 25 and 50 GeV (Moore, Fig. 2.7). The dark photon is added to a fixed baseline, so this is not yet a simultaneous fit.

SIMUnet

SIMUnet replaces the scan of section 7 by a single fit [9]. The NNPDF neural network, which outputs the distributions at the initial scale, is given one further layer. That layer multiplies the Standard Model prediction for each data point by the factor of Eq. (3.2.17), and the Wilson coefficients are its trainable weights (Moore, section 4.3.1 and Eq. 4.8):

$$T_i\big(\hat{\theta}\big) = T_i^{\text{SM}}(\theta) \times K_i(c), \qquad \hat{\theta} = (\theta, c). \tag{3.2.31}$$

One minimisation of the loss function of chapter 2.2 determines θ and c together, and the Monte Carlo replica method gives the uncertainty on both.

Chapter 4 of the thesis applies SIMUnet to the top-quark data of LHC Run II, with 20 Wilson coefficients fitted at linear order together with the distributions [1, 10]. The top data suppress the gluon at large x by up to 10% in a Standard Model fit, and the pull is about halved in the joint fit. The bounds on the coefficients change little: the broadening is at most about 10%, for the chromomagnetic coefficient c_{tG} with top-pair data alone. As in section 8, the interplay matters now for the distributions and not yet for the bounds.

The thesis also records a limitation: with quadratic terms $c_n c_m$ kept, the Monte Carlo replica method does not reproduce the correct distribution of the coefficients (Moore, section 4.7 and chapter 6). The equivalence of the replica ensemble and the Bayesian posterior holds only for a linear model with Gaussian errors (chapter 2.3).

In this project

Our fits put the MSHT20 parametrisation, with 52 parameters at $Q_0 = 1.65$ GeV, into the Colibri code, and predict every data point in the Standard Model with NNPDF theory 40000000 (NNLO, $\alpha_S(M_Z) = 0.118$).

- **The distributions are conditional.** Their uncertainties assume that no data point in the fit carries new physics.
- **A Wilson coefficient is one more parameter.** In a parametric fit, Eq. (3.2.31) adds c to the parameter vector, and the Hessian, Monte Carlo and Bayesian methods of chapter 2.3 apply to the enlarged vector unchanged.
- **Flat directions are the same phenomenon.** A value of ρ^2 near one is a flat direction of the joint fit. The diagnostics for flat directions inside PDF space, the eigenvalues of the Fisher matrix and the disagreement between the three methods (chapter 3.4), would also reveal a PDF–SMEFT degeneracy.

Derivations

D1. Integrating out the heavy scalar

Derives (3.2.3)

The equation of motion. Varying Eq. (3.2.1) with respect to ϕ gives $-\partial^2\phi - M^2\phi - g\,\bar\psi\psi = 0$, so

$$\phi = -\frac{g}{\partial^2 + M^2}\,\bar\psi\psi = -\frac{g}{M^2}\Big(1 - \frac{\partial^2}{M^2} + \dots\Big)\bar\psi\psi.$$

The expansion is valid when the operator acts on fields whose momenta satisfy $q^2 \ll M^2$. Its first term is Eq. (3.2.2).

Substituting. Write $S = \bar\psi\psi$ and keep the leading term, $\phi = -gS/M^2$. The kinetic term $\frac{1}{2}(\partial\phi)^2$ is of order $(\partial S)^2/M^4$ and is dropped. The other two terms give

$$-\tfrac{1}{2}M^2\phi^2 - g\,\phi\,S = -\frac{g^2S^2}{2M^2} + \frac{g^2S^2}{M^2} = \frac{g^2}{2M^2}\,S^2,$$

which is the contact term of Eq. (3.2.3).

The path integral. The thesis instead does the Gaussian integral over ϕ, by a shift of the field (Moore, Eqs. 3.2 to 3.8 [1]). Completing the square in $-\frac{1}{2}\phi\,(\partial^2 + M^2)\,\phi - g\phi S$ gives a factor $\exp\left[\frac{i}{2}g^2\int S\,(\partial^2 + M^2)^{-1}S\right]$. The check script verifies the one-variable analogue, $\int d\phi\, e^{-M^2\phi^2/2 - g\phi S} \big/ \int d\phi\, e^{-M^2\phi^2/2} = e^{g^2S^2/2M^2}$.

Checking the coefficient with a diagram. In the full theory, two vertices $-ig$ and the propagator of Eq. (3.2.4) at $q^2 \to 0$ give an amplitude factor $(-ig)^2 \cdot (-i/M^2) = ig^2/M^2$ for each way of routing the fermion lines. In the effective theory, a term $c\,(\bar\psi\psi)^2$ gives $2ic$ for the same routing, because either of its two bilinears can attach to a given fermion pair. Equating them, $c = g^2/2M^2$.

A remark on the source. Moore's Eqs. (3.5) to (3.8) and (3.13) carry the contact coefficient as g^2/M^2. His Eq. (3.1) is Eq. (3.2.1), and integrating ϕ out of it gives $g^2/2M^2$ by all three routes above. Nothing later depends on the factor.

D2. From gauge-boson operators to contact terms

Derives (3.2.9), (3.2.10), (3.2.11)

Reducing the operator. Consider the $\hat{W}$ term of Eq. (3.2.8). Integrate by parts, $(D_\rho W^a_{\mu\nu})^2 \to -W^{a\mu\nu} D^2 W^a_{\mu\nu}$. The Bianchi identity $D_\rho W_{\mu\nu} + D_\mu W_{\nu\rho} + D_\nu W_{\rho\mu} = 0$ lets D^ρ be moved onto the field-strength indices. Commuting covariant derivatives produces terms with three field strengths, which do not contribute to four-fermion processes at this order. After a second integration by parts,

$$\left(D_\rho W^a_{\mu\nu}\right)^2 = 2\left(D^\mu W^a_{\mu\nu}\right)^2 + O(W^3).$$

For an Abelian field in momentum space, $F_{\mu\nu} = q_\mu A_\nu - q_\nu A_\mu$, this is the identity $q^2 F_{\mu\nu}F^{\mu\nu} = 2\,(q^\mu F_{\mu\nu})^2$, which the check script verifies for random vectors.

The equation of motion. In the Standard Model, $D^\mu W^a_{\mu\nu} = g\,(J^a_{L\nu} + J^a_{H\nu})$, where J_L is the fermion current of Eq. (3.2.9) and J_H the current of the Higgs doublet. As in section 3, using the lowest-order equation of motion inside a dimension-6 operator changes the result only at higher order. Hence

$$-\frac{\hat{W}}{4m_W^2}\cdot 2\,g^2\left(J^a_L + J^a_H\right)^2 = -\frac{g^2\,\hat{W}}{2m_W^2}\,J^a_{L\mu}J^{a\mu}_L + \dots$$

The fermion term is Eq. (3.2.9). The terms with J_H change the masses and couplings of the W and Z by a relative amount of order $\hat{W}$ and do not grow with energy. The $\hat{Y}$ term is the same calculation for an Abelian field, with $\partial^\mu B_{\mu\nu} = g_Y J_{Y\nu} + \dots$.

The propagator. The same result follows without the equation of motion. For transverse modes the quadratic Lagrangian $-\frac{1}{4}W^2_{\mu\nu} + \frac{1}{2}m_V^2 W_\mu^2$ is $\frac{1}{2}\tilde{W}^\mu(-q^2 + m_V^2)\tilde{W}_\mu$ in momentum space. Two more derivatives multiply the kinetic part by $\hat{W}q^2/m_W^2$, so the inverse propagator $q^2 - m_V^2$ becomes $q^2 - m_V^2 + \hat{W}q^4/m_W^2$. To first order in $\hat{W}$,

$$\frac{1}{q^2 - m_V^2 + \hat{W}q^4/m_W^2} \simeq \frac{1}{q^2 - m_V^2} - \frac{\hat{W}}{m_W^2}\,\frac{q^4}{(q^2 - m_V^2)^2},$$
$$\frac{q^4}{(q^2 - m_V^2)^2} = 1 + \frac{2m_V^2}{q^2 - m_V^2} + \frac{m_V^4}{(q^2 - m_V^2)^2}.$$

The term 1 gives Eq. (3.2.10). The other two have the pole structure of the original propagator and of its square. They are the shifts of coupling and mass found above from J_H.

The sign. Exchange of a vector of mass m_V coupled as $g\,W_\mu J^\mu$ gives the effective interaction $\frac{1}{2}g^2\,J_\mu(q^2 - m_V^2)^{-1}J^\mu$, which is $-(g^2/2m_V^2)J\cdot J$ at low energy by the argument of D1. The replacement (3.2.10) adds $-(g^2\hat{W}/2m_W^2)J\cdot J$ to it, which is Eq. (3.2.9) again. A heavy vector of coupling g_* and mass M gives $-(g_*^2/2M^2)J\cdot J$, and hence Eq. (3.2.11).

D3. Drell–Yan with a contact term

Derives (3.2.12), (3.2.14), (3.2.15), (3.2.16)

One helicity combination. In the interaction $\frac{1}{2}g^2 J\cdot J/(q^2 - m_V^2)$ of D2, the cross term between a quark and a lepton current is $g_q g_l/(q^2 - m_V^2)$ times $(\bar{q}\gamma^\mu q)(\bar{l}\gamma_\mu l)$. Summing the photon, the Z and the contact terms of Eq. (3.2.9) in the same way, the interaction of a quark of helicity X with a lepton of helicity Y is $A^q_{XY}(\hat{s})\,(\bar{q}_X\gamma^\mu q_X)(\bar{l}_Y\gamma_\mu l_Y)$ with A^q_{XY} as in Eq. (3.2.13).

Squaring. For $q_X(p_1)\,\bar q(p_2) \to l_Y(k_1)\,\bar l(k_2)$ with massless fermions the traces contain a chirality projector. The check script evaluates them with explicit Dirac matrices. With θ the angle between $\vec k_1$ and $\vec p_1$,

$$\sum_{\text{spins}} |\mathcal{M}_{XY}|^2 = |A_{XY}|^2 \times \begin{cases} 16\,(p_1\cdot k_2)(p_2\cdot k_1) = \hat s^2(1+\cos\theta)^2, & X = Y, \\ 16\,(p_1\cdot k_1)(p_2\cdot k_2) = \hat s^2(1-\cos\theta)^2, & X \neq Y. \end{cases}$$

Different helicity combinations do not interfere.

Integrating. Both angular factors integrate to $\int d\Omega\,(1\pm\cos\theta)^2 = 16\pi/3$. Averaging over the four initial spin states and using $d\hat\sigma/d\Omega = \overline{\sum}|\mathcal{M}|^2/64\pi^2\hat s$ as in chapter 1.6,

$$\hat\sigma = \frac{1}{64\pi^2\hat s}\cdot\frac{\hat s^2}{4}\cdot\frac{16\pi}{3}\sum_{XY}|A_{XY}|^2 = \frac{\hat s}{48\pi}\sum_{XY}|A_{XY}|^2.$$

The colour average multiplies this by $1/N_c = 1/3$, which gives Eq. (3.2.12). For the photon alone, all four amplitudes are $e^2Q_qQ_l/\hat s$ and $\hat\sigma = e^4Q_q^2/12\pi\hat s N_c = 4\pi\alpha^2Q_q^2/3\hat s N_c$.

The gauge basis. With $e = g\sin\theta_W$, $g_Y = g\tan\theta_W$ and $Q = T^3 + Y$, the Z coupling is $g_Z^f = (g/\cos\theta_W)(T^3\cos^2\theta_W - Y\sin^2\theta_W)$. Multiplying out,

$$e^2QQ' + g_Z^f g_Z^{f'} = g^2\,T^3T^{3\prime} + g^2\tan^2\theta_W\,YY'\big(\sin^2\theta_W + \cos^2\theta_W\big) = g^2\,T^3T^{3\prime} + g_Y^2\,YY'.$$

The cross terms in T^3Y' cancel between the photon and the Z. For $\hat s \gg m_Z^2$ both propagators are $1/\hat s$ and Eq. (3.2.13) becomes Eq. (3.2.14).

The linear term. Write $A_{XY} = (G_{XY}/\hat s)(1-\epsilon_{XY})$ with $\epsilon_{XY} = (\hat s/m_W^2)\big(g^2T^3T^{3\prime}\hat W + g_Y^2YY'\hat Y\big)/G_{XY}$. To first order, $\sum|A|^2 \propto \sum G^2 - 2\sum G^2\epsilon$, and dividing by $\sum G^2$ gives Eq. (3.2.15). For the up quark, in units of g^2 and with $\tan^2\theta_W = 0.3007$, $G_{XY} = -0.275, -0.050, -0.100, -0.200$ for $(X,Y) = (L,L),(L,R),(R,L),(R,R)$. Only the first contains $T^3T^{3\prime} = -\tfrac14$, so $w_u = (0.275\times0.25)/0.128 = 0.54$. For the down quark $G_{XY} = 0.225, -0.050, 0.050, 0.100$ and $w_d = 0.86$.

Charged current. The amplitude for $u_L\bar d_L \to l^+\nu$ is proportional to $\tfrac12 g^2\big[(\hat s - m_W^2)^{-1} - \hat W/m_W^2\big]$ by Eq. (3.2.10), with a single helicity combination. The ratio to the Standard Model is Eq. (3.2.16).

D4. Fixed and profiled bounds in the linear model

Derives (3.2.20)

The joint χ^2. With the model of Eq. (3.2.19) and data whose relative deviations from the reference are d_i,

$$\chi^2(\theta, c) = (d - A\theta - bc)^T C^{-1}(d - A\theta - bc).$$

It is a quadratic form in (θ, c) whose second-derivative matrix is twice the Fisher matrix, $F = \begin{pmatrix} F_{\theta\theta} & F_{\theta c} \\ F_{c\theta} & F_{cc}\end{pmatrix}$. By chapter 2.3 the covariance of the fitted parameters is F^{-1}.

Fixed distributions. If θ is held at its true value, χ^2 as a function of c has second derivative $2F_{cc}$. The condition $\Delta\chi^2 = 1$ gives $\sigma_c = 1/\sqrt{F_{cc}}$.

Profiling. For given c the minimum over θ is at $\hat\theta(c) = F_{\theta\theta}^{-1}A^TC^{-1}(d - bc)$. As c changes by δc, the best-fit distributions move by $-F_{\theta\theta}^{-1}F_{\theta c}\,\delta c$. Substituting back, the profile $\chi^2_{\text{eftp}}(c) = \chi^2(\hat\theta(c), c)$ is a parabola with

$$\frac{1}{2}\frac{d^2\chi^2_{\text{eftp}}}{dc^2} = F_{cc} - F_{c\theta}F_{\theta\theta}^{-1}F_{\theta c} = F_{cc}\,(1-\rho^2).$$

This is the Schur complement of $F_{\theta\theta}$ in F, and its inverse is the cc element of F^{-1}. Profiling and marginalising agree for a Gaussian. Hence Eq. (3.2.20).

The bound on ρ^2. Write $u = C^{-1/2}b$ and let Π be the orthogonal projector onto the columns of $C^{-1/2}A$. Then $F_{cc} = |u|^2$ and $F_{c\theta}F_{\theta\theta}^{-1}F_{\theta c} = |\Pi u|^2$, so $\rho^2 = |\Pi u|^2/|u|^2$ lies in $[0,1]$. It is the squared cosine of the angle between the signal shape and the space of shapes the distributions can produce, and equals one when b is a combination of the columns of A.

D5. Fitting away a signal

Derives (3.2.26), (3.2.27), (3.2.28)

The Standard Model fit. Let the data be $d = b\,c^* + n$, with noise n of covariance C and the true PDF parameters at $\theta = 0$. A fit with $c = 0$ gives $\hat\theta = F_{\theta\theta}^{-1}A^TC^{-1}d$. Its mean is Eq. (3.2.26).

The residual. In the notation of D4 the weighted residual is $C^{-1/2}(d - A\hat\theta) = (1-\Pi)\,C^{-1/2}d$. The fit removes from the data everything that lies in the space of PDF shapes, including the part $\Pi u\,c^*$ of the signal.

The second step. Fitting c to the same data with $\hat\theta$ fixed gives $\hat c = u^T(1-\Pi)C^{-1/2}d/|u|^2$. Its mean is $c^*\,|(1-\Pi)u|^2/|u|^2 = (1-\rho^2)\,c^*$, the first part of Eq. (3.2.27). Its actual scatter is $\sqrt{1-\rho^2}/\sqrt{F_{cc}}$, smaller than the quoted $1/\sqrt{F_{cc}}$, because the PDF fit has already removed part of the noise.

The fit quality. $\chi^2_{\min} = |(1-\Pi)C^{-1/2}d|^2$. The noise contributes $N_{\text{dat}} - p$ on average, the trace of $1-\Pi$. The signal contributes $c^{*2}|(1-\Pi)u|^2 = c^{*2}F_{cc}(1-\rho^2) = (c^*/\sigma_c^{\text{sim}})^2$. This is Eq. (3.2.28).

Conservative distributions. Split the data into a tail, which alone has $b \neq 0$, and the rest. Let L and A_t be the Fisher matrices for θ of the rest and of the tail, so $F_{\theta\theta} = L + A_t$. Fitting θ to the rest and then c to the tail, with the PDF covariance L^{-1} propagated, gives an unbiased $\hat c$ with variance $1/F_{cc} + F_{c\theta}L^{-1}F_{\theta c}/F_{cc}^2$. The simultaneous fit is the minimum-variance linear unbiased estimator for this model, so its variance $1/\big(F_{cc} - F_{c\theta}(L+A_t)^{-1}F_{\theta c}\big)$ cannot be the larger. For a single PDF parameter the difference is proportional to $F_{cc}A_t - F_{c\theta}^2$, which is non-negative by the Cauchy–Schwarz inequality. The check script confirms the ordering in the toy and verifies Eqs. (3.2.26) to (3.2.28) by simulation.

References

1. J. M. Moore, *Parton Distributions in Beyond the Standard Model Theories*, PhD thesis, University of Cambridge. Chapter 3 is the primary source for sections 4 to 10; sections 1.4, 2, 4.3, 4.6 and 5.2 are used in sections 1 and 11. All bounds and data counts are quoted from its Tables 3.1 to 3.7. The derivations as set out here, the toy model and the figures are our own.
2. B. Grzadkowski, M. Iskrzyński, M. Misiak and J. Rosiek, “Dimension-six terms in the Standard Model Lagrangian”, JHEP 10 (2010) 085, arXiv:1008.4884.
3. M. Farina, G. Panico, D. Pappadopulo, J. T. Ruderman, R. Torre and A. Wulzer, “Energy helps accuracy: electroweak precision tests at hadron colliders”, Phys. Lett. B772 (2017) 210, arXiv:1609.08157.
4. M. E. Peskin and T. Takeuchi, “Estimation of oblique electroweak corrections”, Phys. Rev. D46 (1992) 381; G. Altarelli, R. Barbieri and S. Jadach, “Toward a model-independent analysis of electroweak data”, Nucl. Phys. B369 (1992) 3.

5. R. Barbieri, A. Pomarol, R. Rattazzi and A. Strumia, “Electroweak symmetry breaking after LEP-1 and LEP-2”, Nucl. Phys. B703 (2004) 127, arXiv:hep-ph/0405040.
6. S. Amoroso et al., “Snowmass 2021 whitepaper: Proton structure at the precision frontier”, arXiv:2203.13923, section II.B.3.
7. R. Torre, L. Ricci and A. Wulzer, “On the W&Y interpretation of high-energy Drell-Yan measurements”, JHEP 02 (2021) 144, arXiv:2008.12978.
8. M. McCullough, J. Moore and M. Ubiali, “The dark side of the proton”, JHEP 08 (2022) 019, arXiv:2203.12628.
9. S. Iranipour and M. Ubiali, “A new generation of simultaneous fits to LHC data using deep learning”, JHEP 05 (2022) 032, arXiv:2201.07240.
10. Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales Alvarado, J. Rojo and M. Ubiali, “The top quark legacy of the LHC Run II for PDF and SMEFT analyses”, JHEP 05 (2023) 205, arXiv:2303.06159.
11. A. Greljo, S. Iranipour, Z. Kassabov, M. Madigan, J. Moore, J. Rojo, M. Ubiali and C. Voisey, “Parton distributions in the SMEFT from high-energy Drell-Yan tails”, JHEP 07 (2021) 122, arXiv:2104.02723; M. Madigan and J. Moore, PoS EPS-HEP2021 (2022) 424, arXiv:2110.13204. The papers on which chapter 3 of the thesis is based.
12. S. Carrazza, C. Degrande, S. Iranipour, J. Rojo and M. Ubiali, “Can New Physics hide inside the proton?”, Phys. Rev. Lett. 123 (2019) 132001, arXiv:1905.05215. The first simultaneous fit, with deep inelastic data only.