Beyond the Standard Proton?

for Queens' Mathematical Society, 31st January 2024

James Moore, University of Cambridge

PBSP

erc
European Research Council
Established by the European Commission

Talk overview

1. Background: *Quantum chromodynamics, parton distributions, and all that...*

2. Fitting parton distributions: *A visit to the sausage factory*

3. Beyond the standard proton

4. Conclusions/questions

1. - Introduction: Quantum chromodynamics, parton distributions, and all that...

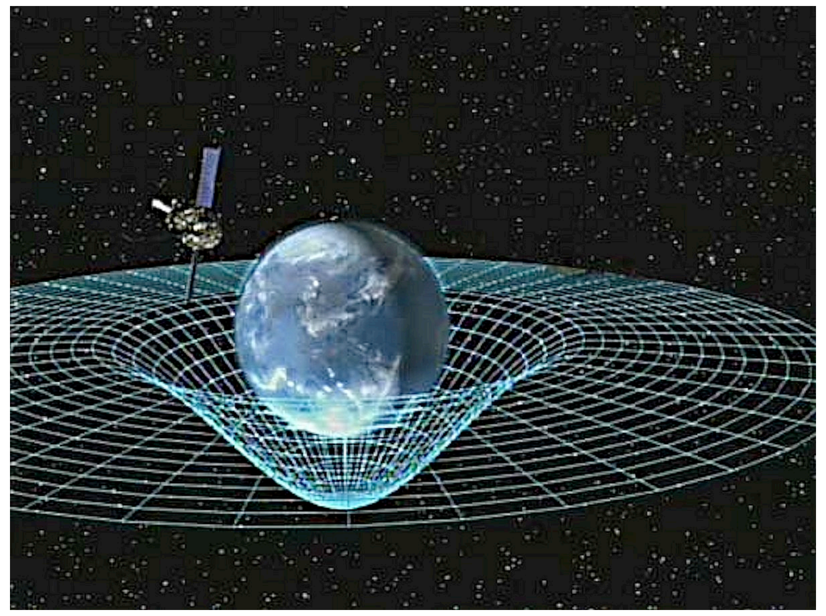
A lightning introduction to particle physics

- Modern particle physics is based on the **Standard Model**, which is a specific example of a **quantum field theory (QFT)**.
- QFTs are described in terms of the following data:

A lightning introduction to particle physics

- Modern particle physics is based on the **Standard Model**, which is a specific example of a **quantum field theory (QFT)**.
- QFTs are described in terms of the following data:

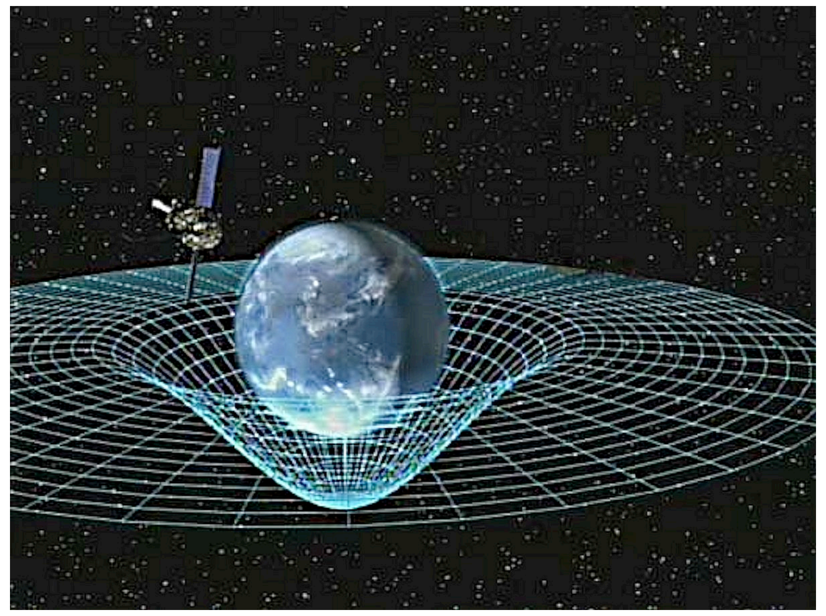
A **spacetime** - normally 1+3 dimensional Minkowski spacetime



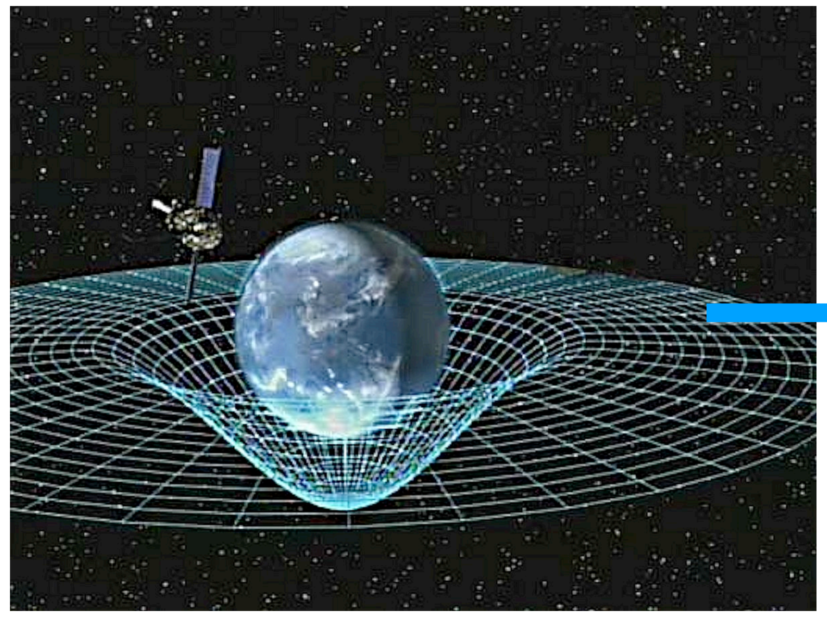
A lightning introduction to particle physics

- Modern particle physics is based on the **Standard Model**, which is a specific example of a **quantum field theory (QFT)**.
- QFTs are described in terms of the following data:

A **spacetime** - normally 1+3 dimensional Minkowski spacetime



A set of **fields** - operator-valued distributions on the spacetime



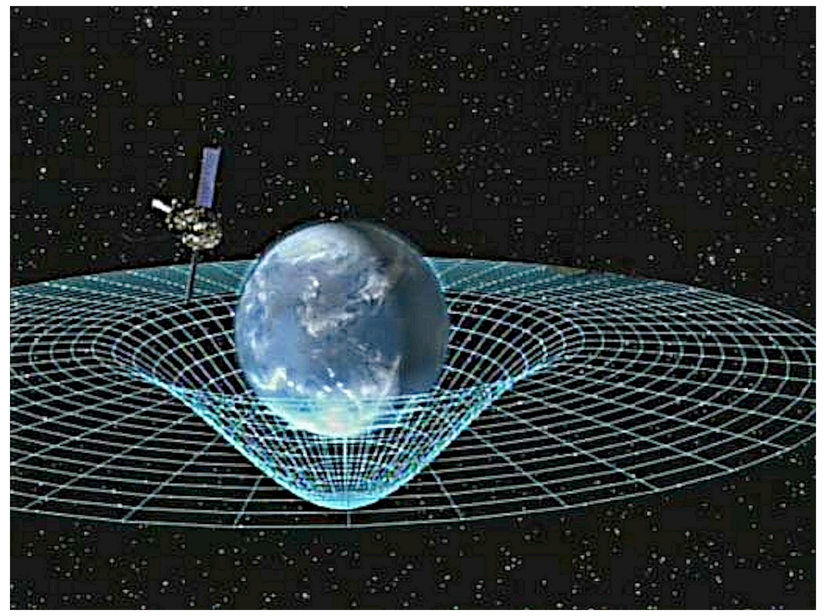
ϕ

operators on a certain Hilbert space

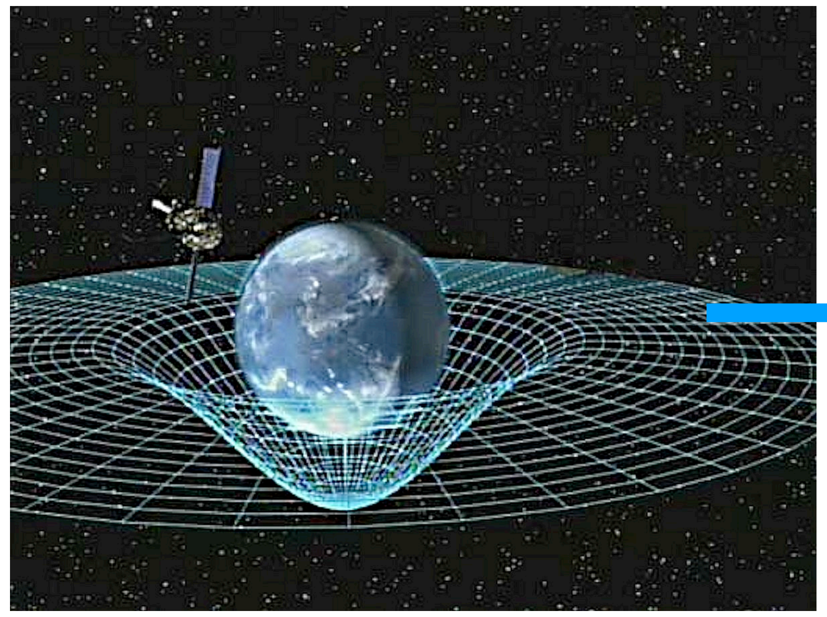
A lightning introduction to particle physics

- Modern particle physics is based on the **Standard Model**, which is a specific example of a **quantum field theory (QFT)**.
- QFTs are described in terms of the following data:

A **spacetime** - normally 1+3 dimensional Minkowski spacetime



A set of **fields** - operator-valued distributions on the spacetime



ϕ → operators on a certain Hilbert space

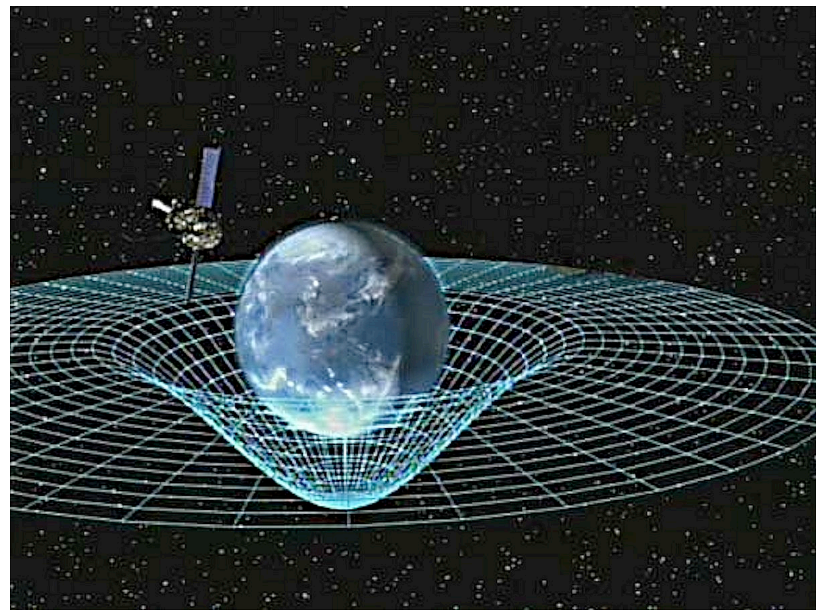
A **Lagrangian density** - describes the **dynamics** of the fields, and depends on **parameters**, e.g. masses and interaction strength

$$\mathscr{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{1}{4!}g\phi^4$$

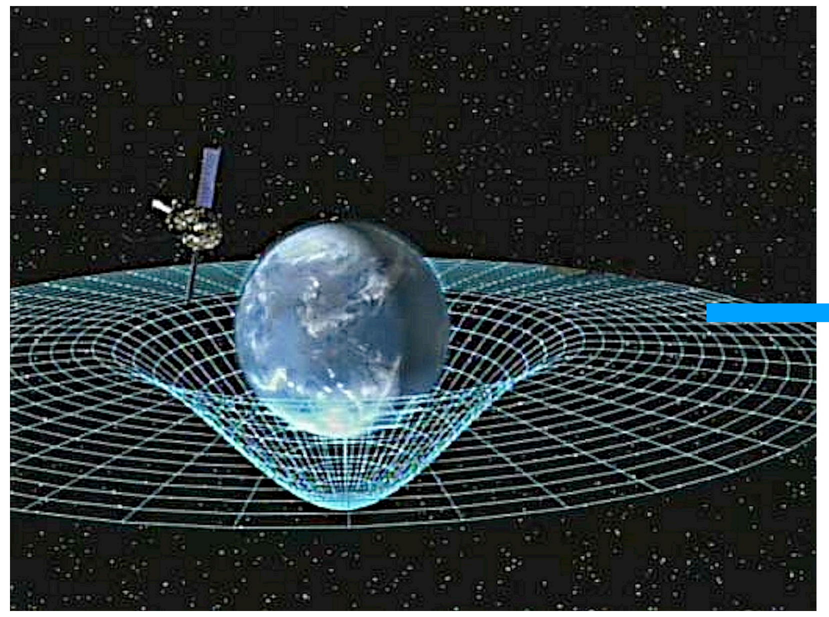
A lightning introduction to particle physics

- Modern particle physics is based on the **Standard Model**, which is a specific example of a **quantum field theory (QFT)**.
- QFTs are described in terms of the following data:

A **spacetime** - normally 1+3 dimensional Minkowski spacetime



A set of **fields** - operator-valued distributions on the spacetime



ϕ → operators on a certain Hilbert space

A **Lagrangian density** - describes the **dynamics** of the fields, and depends on **parameters**, e.g. masses and interaction strength

$$\mathscr{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{1}{4!}g\phi^4$$

A **renormalisation scheme** - relates parameters in the Lagrangian density to physically observable quantities

$$m_{\text{phys}} = f_1(m^2(\epsilon), g(\epsilon))$$

$$g_{\text{phys}} = f_2(m^2(\epsilon), g(\epsilon))$$

A lightning introduction to particle physics

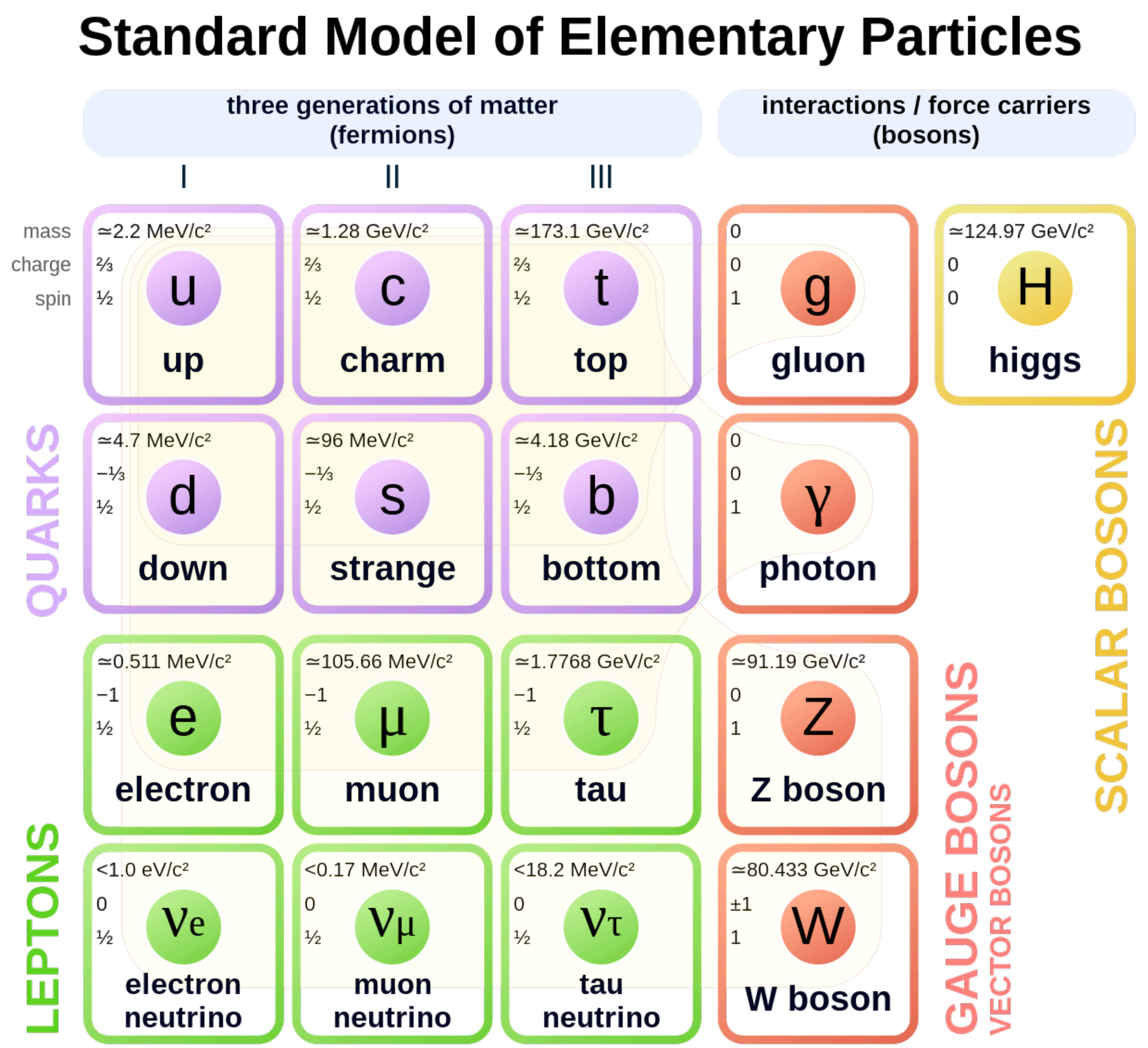
- For the **Standard Model (SM)** of particle physics, the ingredients are:

A lightning introduction to particle physics

- For the **Standard Model (SM)** of particle physics, the ingredients are:
  - **Minkowski spacetime**

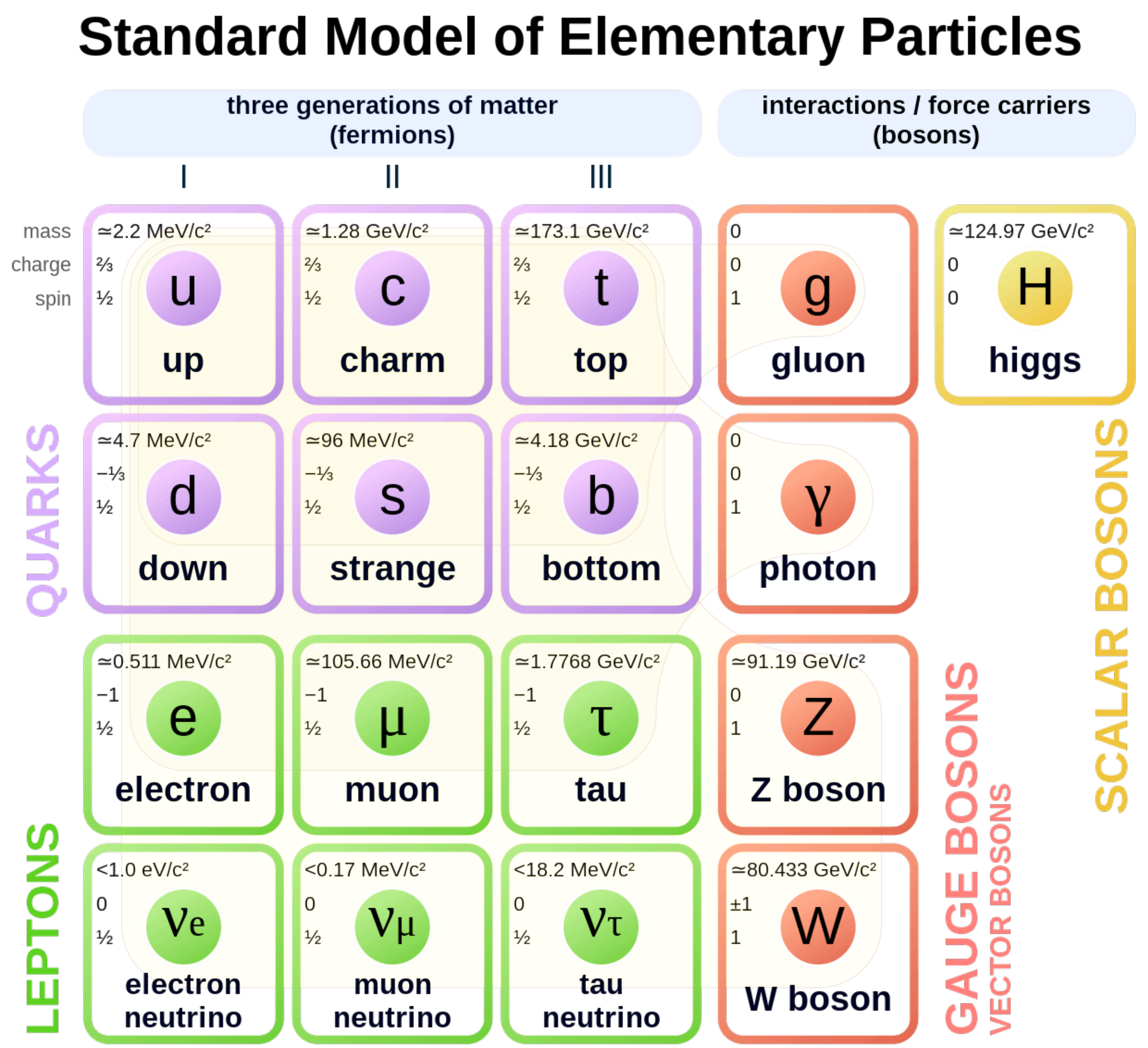
A lightning introduction to particle physics

- For the **Standard Model (SM)** of particle physics, the ingredients are:
  - **Minkowski spacetime**
  - Fields of **special types** for each of the particles we **observe in Nature**: *photons, W and Z bosons, gluons, quarks, leptons,* and *the Higgs boson*



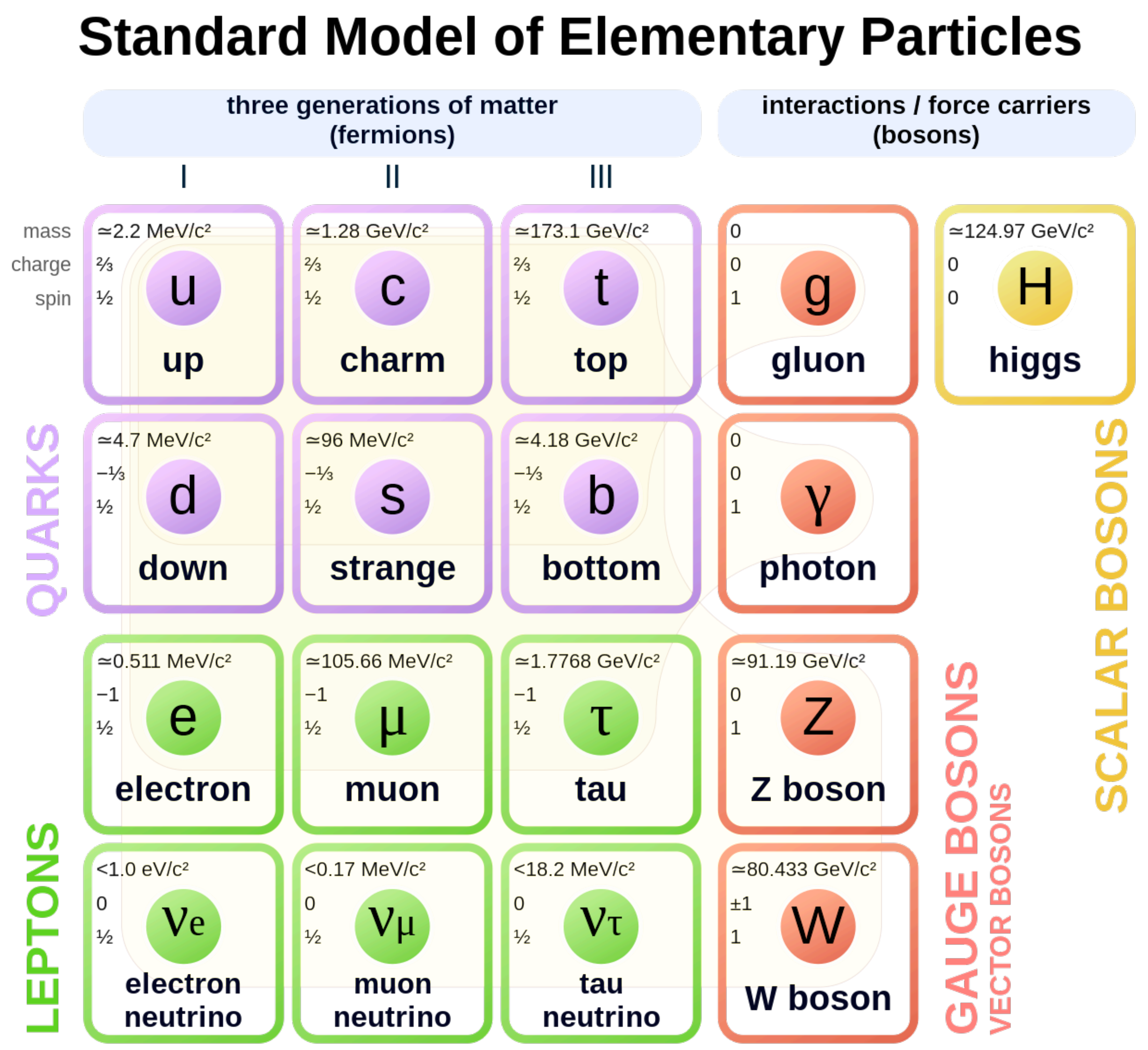
A lightning introduction to particle physics

- For the **Standard Model (SM)** of particle physics, the ingredients are:
  - **Minkowski spacetime**
  - Fields of **special types** for each of the particles we **observe in Nature**: *photons, W and Z bosons, gluons, quarks, leptons,* and *the Higgs boson*
  - A Lagrangian density of a special type, called a **gauge theory** (with gauge group $SU(3) \times SU(2) \times U(1)$)



A lightning introduction to particle physics

- For the **Standard Model (SM)** of particle physics, the ingredients are:
  - **Minkowski spacetime**
  - Fields of **special types** for each of the particles we **observe in Nature**: *photons, W and Z bosons, gluons, quarks, leptons,* and *the Higgs boson*
  - A Lagrangian density of a special type, called a **gauge theory** (with gauge group $SU(3) \times SU(2) \times U(1)$)
  - A suitable renormalisation scheme (usually **dimensional regularisation** with **on-shell mass renormalisation** of heavy particles, and **$\overline{\text{MS}}$ subtraction** for everything else)



Quantum chromodynamics for the general reader

- The SM Lagrangian can be broken into three main sectors: **quantum electrodynamics**, the **weak sector** and **quantum chromodynamics** (QCD).
- QCD involves the **quark** and **gluon** fields, and describes the **strong force** that **binds composite particles** together.
- The **Lagrangian density** for QCD is:

covariant derivative

$$\mathscr{L} = -\frac{1}{4} G^a_{\mu\nu} G^{a,\mu\nu} + \sum_q \bar{q}(i\gamma_\mu D^\mu - m_q)q$$

field strength tensors for **eight gluon fields**

sum over **six quark fields**

quark masses

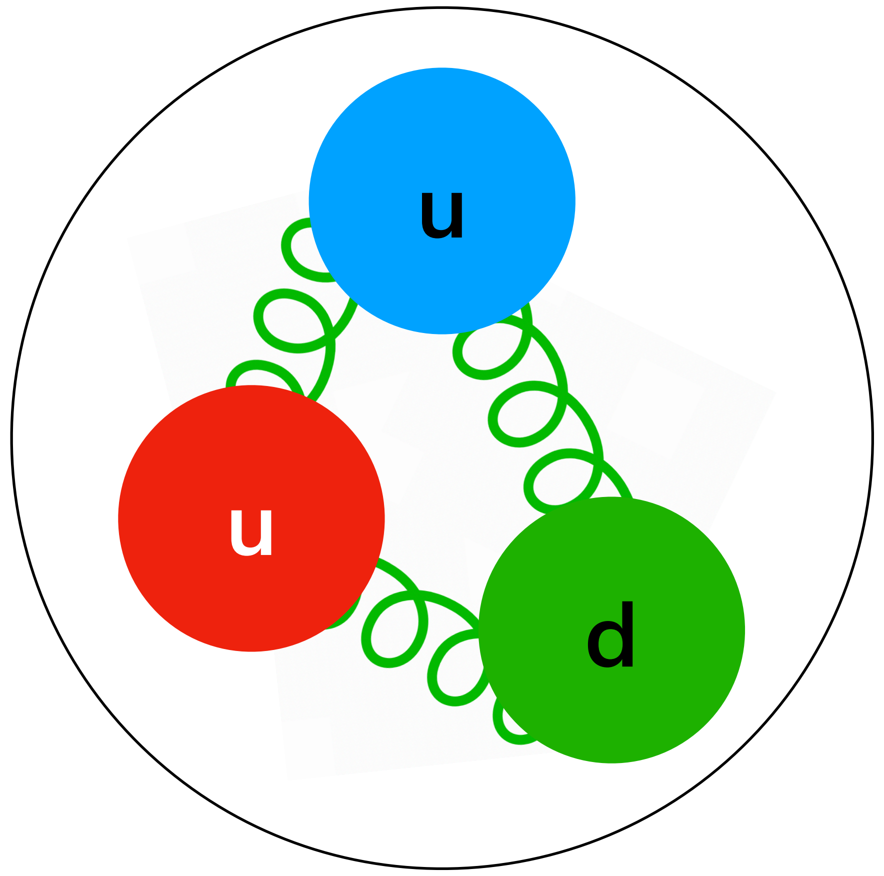
Quantum chromodynamics for the general reader

$$\mathcal{L} = -\frac{1}{4}G^a_{\mu\nu}G^{a,\mu\nu} + \sum_q \bar{q}(i\gamma_\mu D^\mu - m_q)q$$

- From the **QCD Lagrangian**, we *should* be able to prove some things we see experimentally:

  1. **Strongly bound quark states exist**, for example the **proton**, **neutron**, **pion**...

  2. Quarks **must** always be **confined** in **bound states**.



- But... *no-one knows how to do it! (∃ a $1 million prize!)*

Quantum chromodynamics for the general reader

- Some progress has been made…

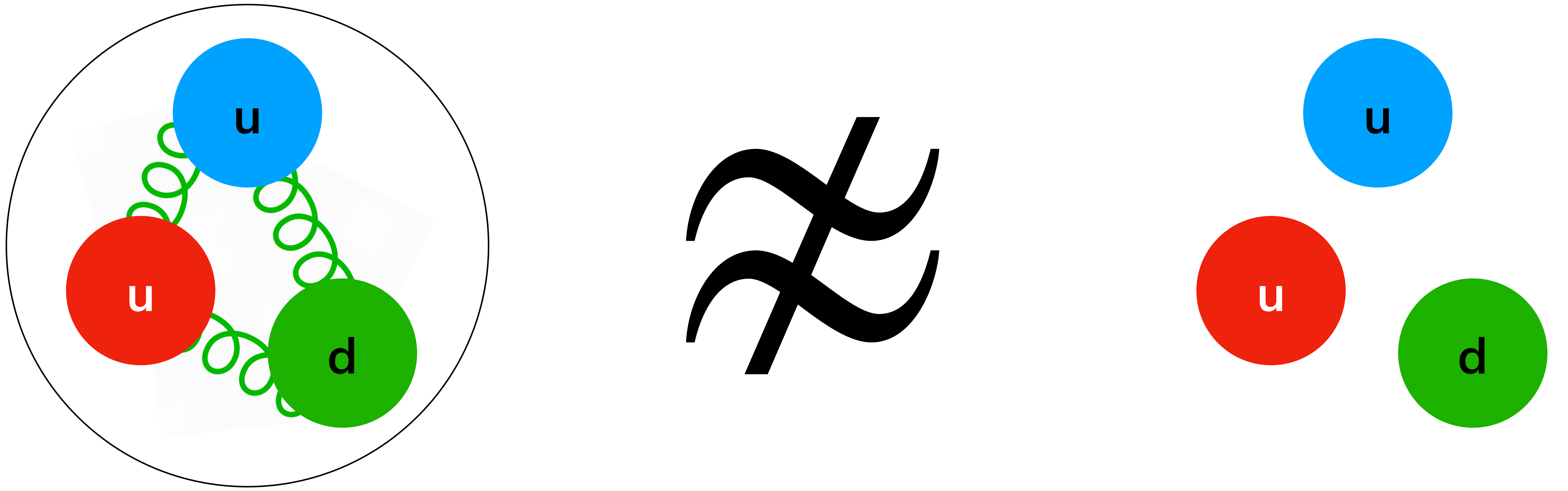
  1. At low energies, simulations using **lattice** versions of QCD (where spacetime is discretised in order to regulate the QFT) predict the existence of e.g. the **proton**.

  2. In **model theories**, e.g. certain theories in 1+1 dimensions, or **supersymmetric** theories, it is possible to prove **confinement**, and derive the existence of bound states.

- These are **limited in scope** though. How do we make SM predictions for **particle accelerators** in 1+3 dimensions, where e.g. protons **collide** at **extremely high energies**? Do we just give up?

Perturbative QCD for the general reader

- **The solution:** *perturbative QCD.*
- Initially **sounds crazy**: normally in physics, **perturbation theory** is used for **weakly-interacting phenomena** which only **deviate in small ways** from **free theories** (where particles don't interact at all).
- Perturbation theory is good for **quantum electrodynamics** and the **weak sector**. But for QCD, the basic fields (quarks and gluons) are **strongly interacting** - it is a **terrible approximation** to treat them as free!



Perturbative QCD for the general reader

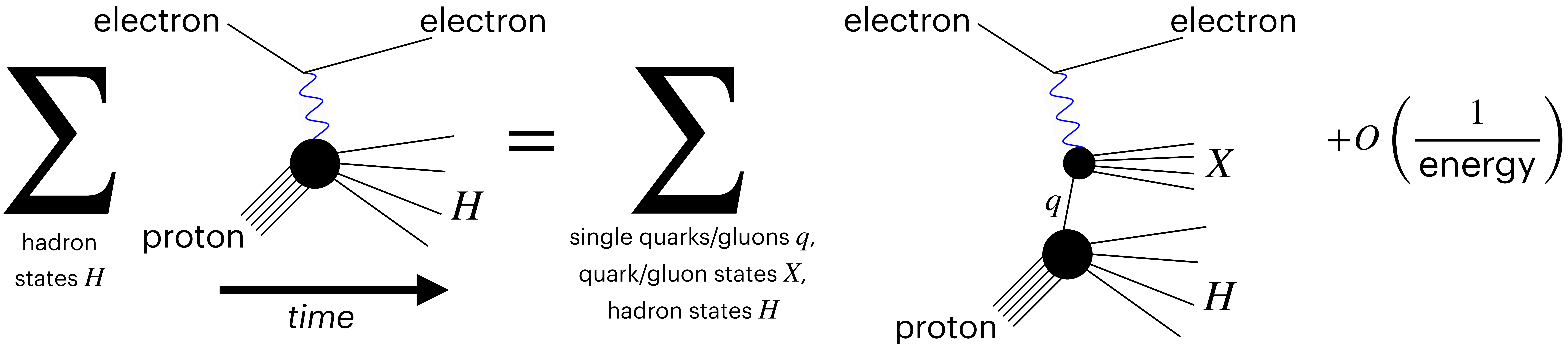
- This can be **partially overcome**, however:
  - If we study processes where we **sum over all final states** (*inclusive* processes), then **completeness relations** tell us it doesn't matter whether we use free quarks and gluons, or the proper bound states.

$$\sum_{\substack{\text{bound} \\ \text{states } H}} |H\rangle\langle H| = \sum_{\substack{\text{quark/gluon} \\ \text{states } X}} |X\rangle\langle X|$$

  - **Classic example:** electron-positron annihilation, $e^+e^- \to$ any hadrons

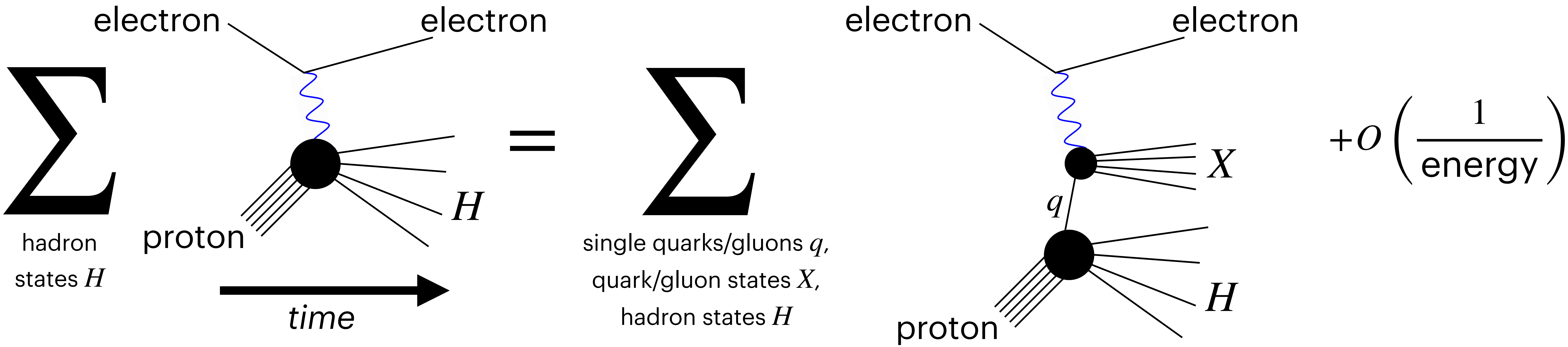
Perturbative QCD for the general reader

- This can be **partially overcome**, however:
    - If we have **specified hadrons** in the **initial state** though (or indeed final state), need more help. At **sufficiently high energies**, the **factorisation theorems** save us.
    - E.g. **deep inelastic scattering**, $e^- +$ proton $\rightarrow e^- +$ any hadron



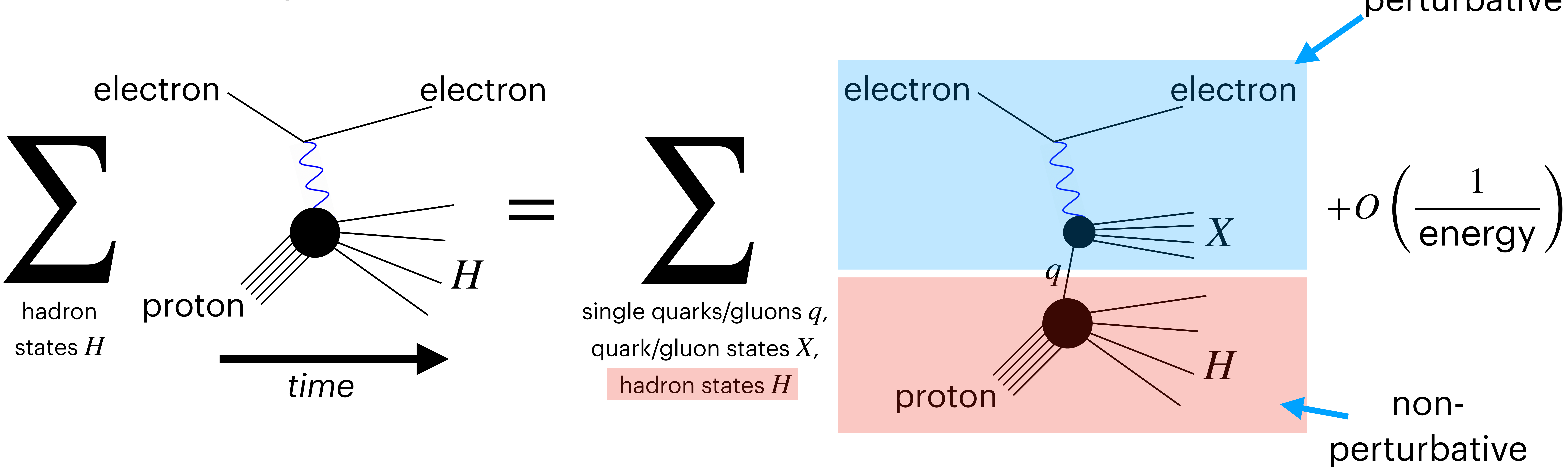
Perturbative QCD for the general reader

- The factorisation theorems separate the physics into a **calculable perturbative part**, and a **non-calculable**, **non-perturbative**, BUT ***universal*** part.



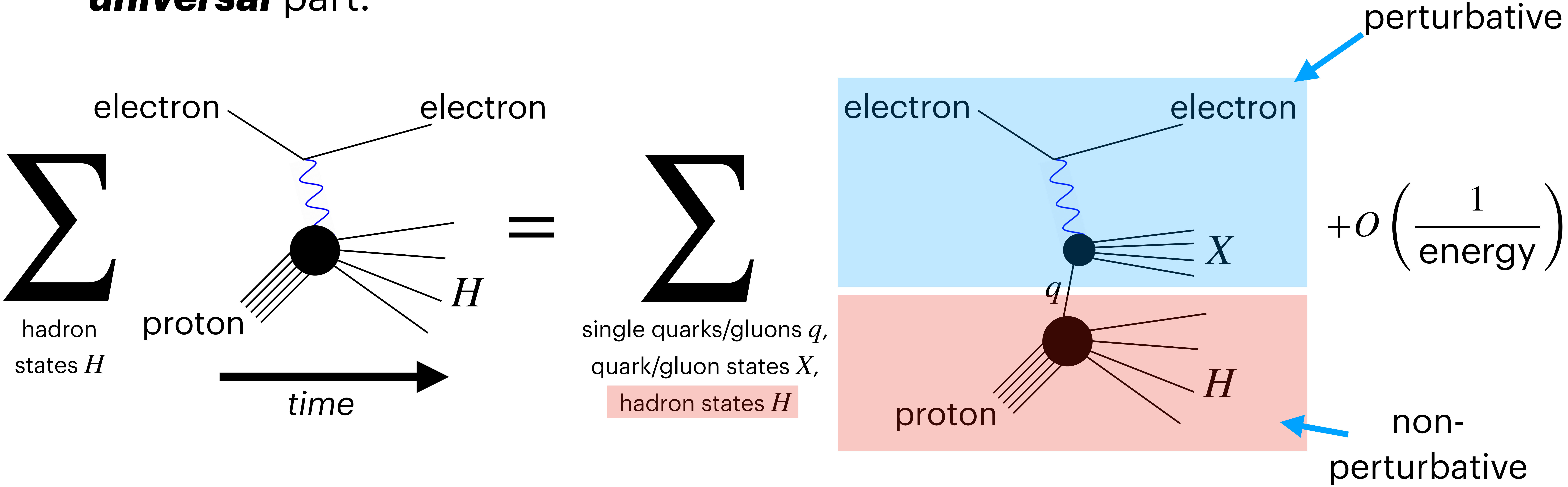
Perturbative QCD for the general reader

- The factorisation theorems separate the physics into a **calculable perturbative part**, and a **non-calculable**, **non-perturbative**, BUT ***universal*** part.



Perturbative QCD for the general reader

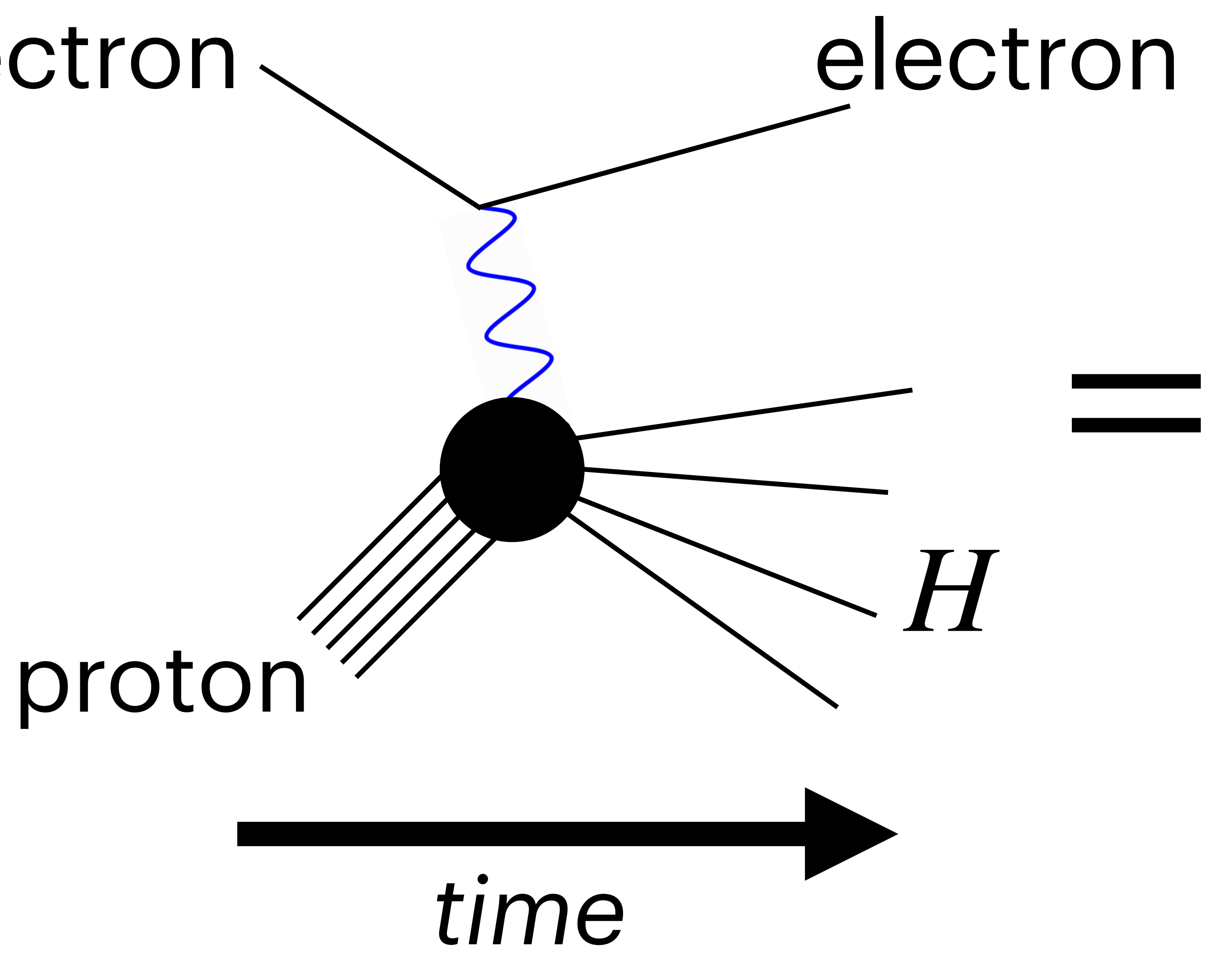
- The factorisation theorems separate the physics into a **calculable perturbative part**, and a **non-calculable**, **non-perturbative**, BUT ***universal*** part.



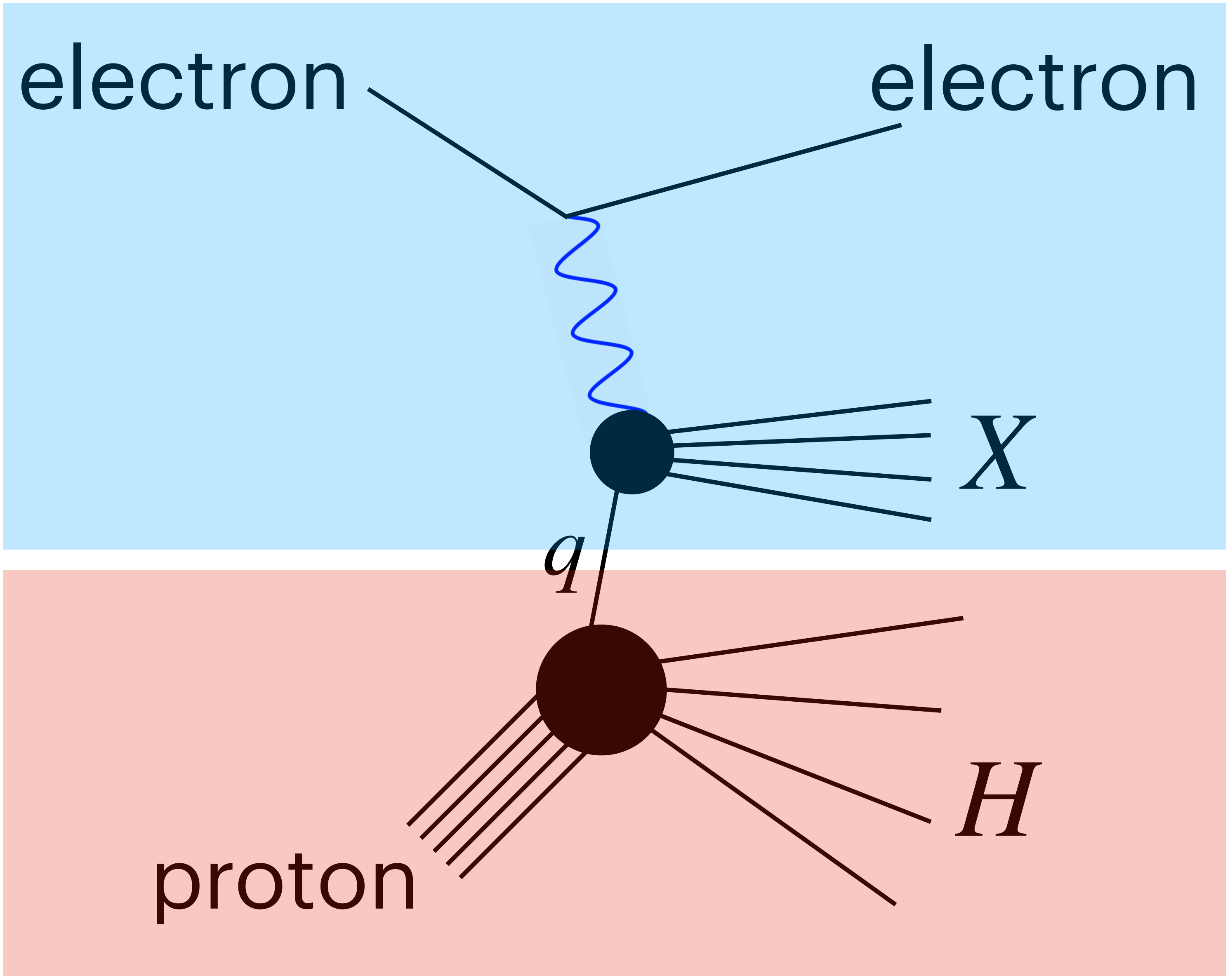
- The universal non-perturbative part is called a **parton distribution function.**

Perturbative QCD for the general reader

$$\sum_{\text{hadron states } H}$$



$$= \sum_{\substack{\text{single quarks/gluons } q,\\ \text{quark/gluon states } X,\\ \text{hadron states } H}}$$



$$+\,O\left(\frac{1}{\text{energy}}\right)$$

In maths… $$\sigma(x, Q^2) = \sum_{\substack{\text{single quarks/gluons } q,\\ \text{quark/gluon states } X}} \int_x^1 \frac{dy}{y}\, \hat{\sigma}_{eq\to eX}\left(\frac{x}{y}, Q^2\right) f_q(y, Q^2) + O\left(\frac{1}{\text{energy}}\right)$$

Perturbative QCD for the general reader

In maths… $$\sigma(x, Q^2) = \sum_{\substack{\text{single quarks/gluons } q, \\ \text{quark/gluon states } X}} \int_x^1 \frac{dy}{y} \hat{\sigma}_{eq \to eX}\left(\frac{x}{y}, Q^2\right) f_q(y, Q^2) + O\left(\frac{1}{\text{energy}}\right)$$

- Speaking very loosely, the parton distributions capture the probability that a particular quark or gluon will be **ejected by the proton** in a collision.
- We interpret $f_q(x, Q^2)dx$ to be the **number of constituents** of type q carrying a **fraction of the proton's momentum** in the interval $[x, x + dx]$, when the process in which the proton is involved has **energy scale** Q^2.

Parton distributions are *universal*

- The **non-perturbative parton distributions** $f_q(x, Q^2)$ depend on:
  - A **momentum fraction** x - tells us how much of the proton's momentum the ejected quark/gluon carries
  - An **energy scale** Q^2, e.g. energy lost by the proton when ejecting a quark
  - The fact we are colliding **protons** - if we started with a neutron, we would get different PDFs

- They **don't** depend on the fact we are colliding a proton with an electron, so can be used for **other processes**. This is why this approach is useful!

Parton distributions are *universal*

- For example, the **same** parton distributions can **also** be used in the **Drell-Yan process**: the collision of two protons to make an **electron-positron pair**, plus any hadrons.

Parton distributions are *universal*

- For example, the **same** parton distributions can **also** be used in the **Drell-Yan process**: the collision of two protons to make an **electron-positron pair**, plus any hadrons.

Parton distributions *scale*

- Whilst the PDFs are non-perturbative, we can still say something about their Q^2 -dependence. **Renormalisation theory** predicts that PDFs should obey a Callan-Symanzik equation called the **DGLAP equation:**

$$Q^2 \frac{\partial f_q(x, Q^2)}{\partial Q^2} = \sum_{\text{quarks/gluons } q'} \int_x^1 \frac{dy}{y} P_{qq'}\left(\frac{x}{y}\right) f_{q'}(x, Q^2)$$

- The functions (technically distributions) $P_{qq'}$ are called **splitting functions** and can be determined perturbatively.
- This means that **if we know the PDFs for some value of Q^2**, we can **determine them for all values of Q^2**.
- **Only their x-dependence is unknown**.

2. - Fitting parton distributions: A visit to the sausage factory

‘PDFs are like sausages: everyone loves them, but no one really wants to know how they are made.’

- Zahari Kassabov

How to make PDFs...

- *TLDRN*: Fitting PDFs using experimental data is an **ill-posed problem**.
- In short, you have **finite amounts of data** from experiments, but the space of possible PDFs is **infinite-dimensional**. What do we do?
- PDF fitting groups **assume a functional form** for the PDFs at some **initial energy scale**, parametrised by a finite set of parameters. They then obtain the PDF at all energy scales using the **DGLAP equation**.
- Example functional form:

$$f(x, Q_0^2) = Ax^{\alpha}(1-x)^{\beta}\left(1 + ax^{1/2} + bx + cx^{3/2}\right)$$

large and small x behaviour motivated by **Regge theory**

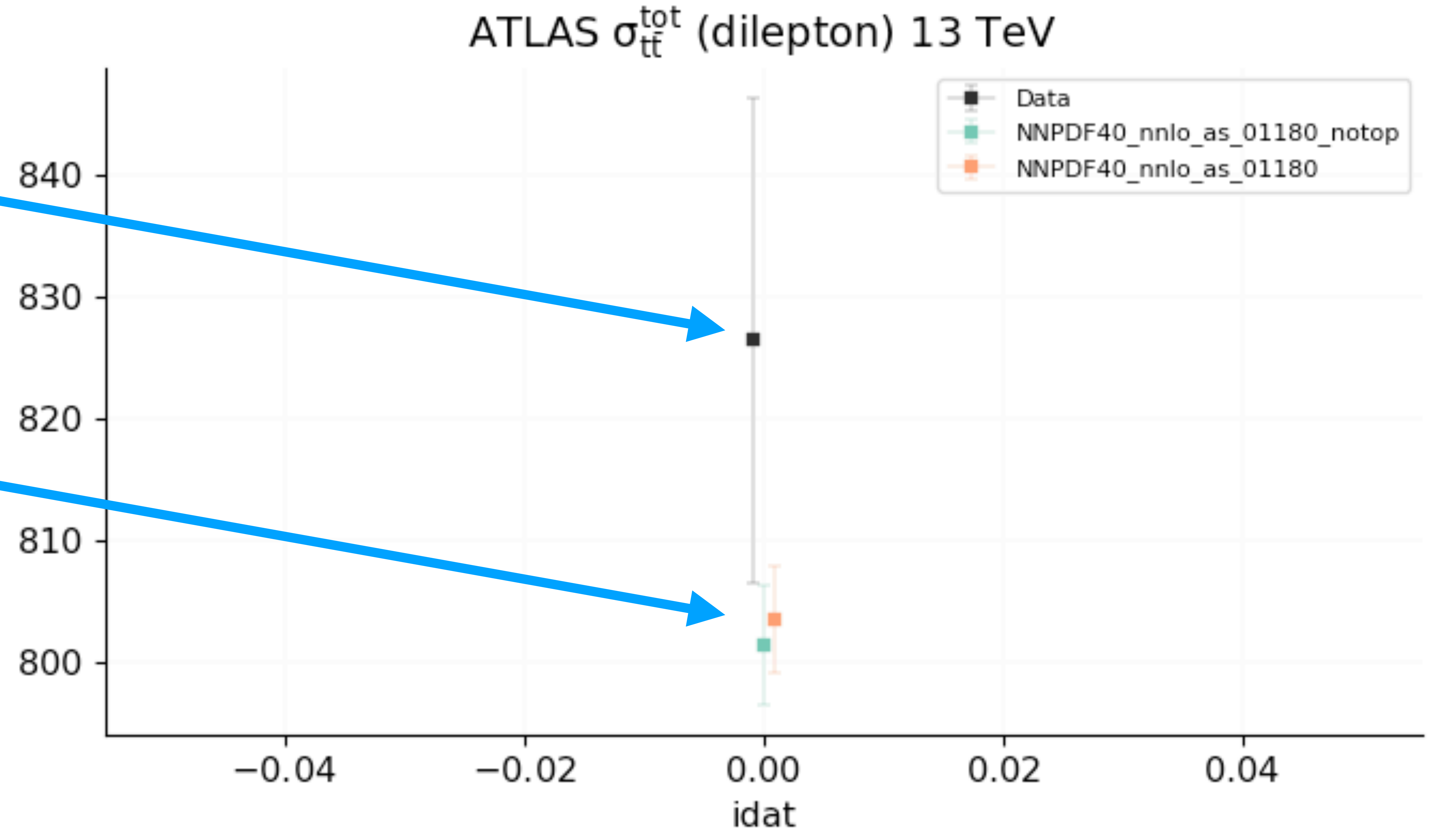
polynomial in $\sqrt{x}$ seems to give nice fit

How to make PDFs...

- Once we have selected a functional form, we find the parameters which **best describe experimental data**.

data point

predictions for two different choices of PDF



How to make PDFs…

- This is usually done by minimising the χ^2-statistic, which measures the **goodness of fit** of our model:

$$\chi^2 = (\text{data} - \text{theory})^T \text{covariance}^{-1} (\text{data} - \text{theory})$$

vector of data points

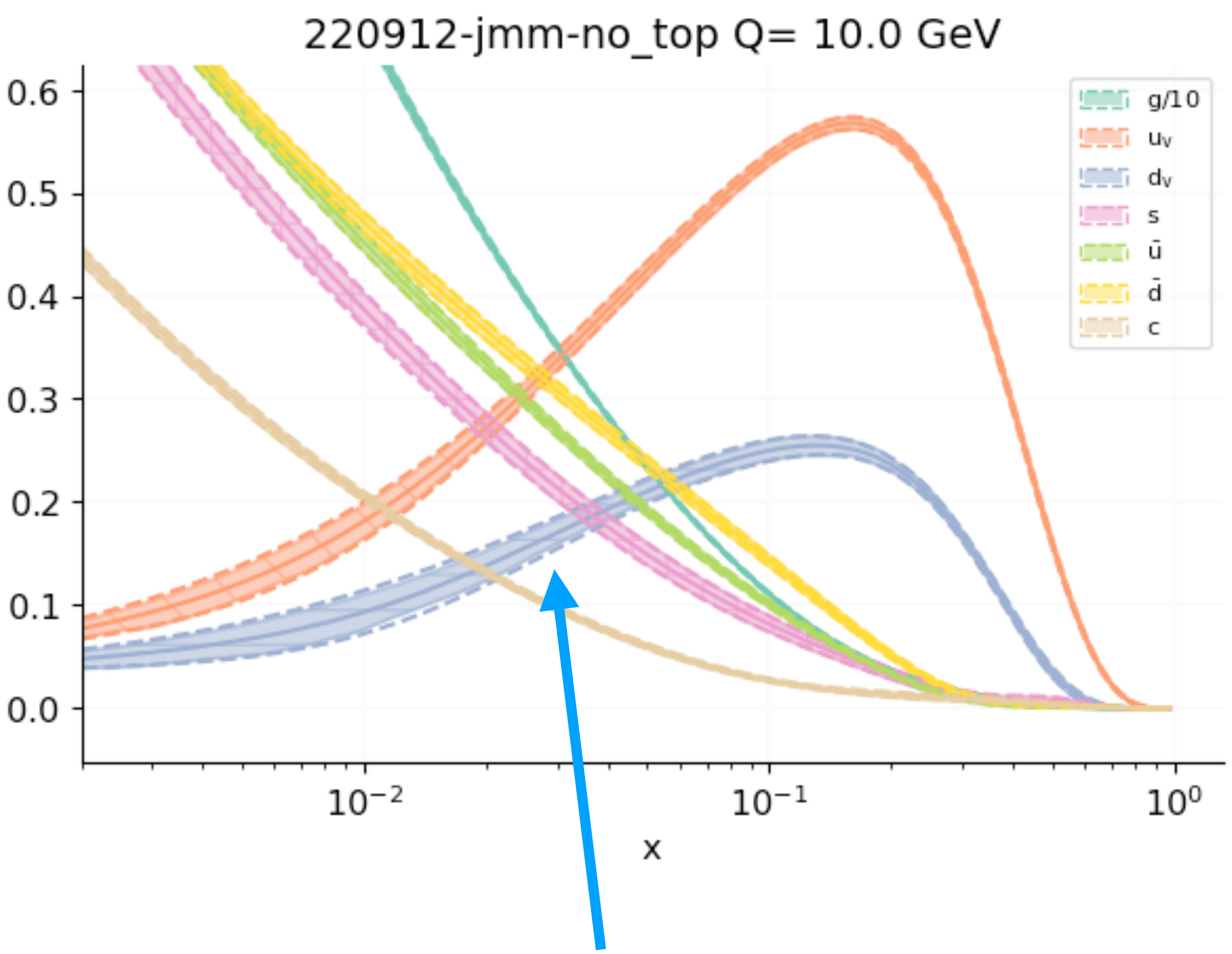
vector of theory predictions

experimental covariance matrix

- *General idea*: we want **theory to be close to data**, but if the data is **more uncertain**, we don’t require such precise agreement.

How to make PDFs...

- It's not good enough to find the PDF parameters which give just the **central data values** because experimental data comes with **uncertainty**. We must also **propagate errors** properly too.
- This can be handled using **Monte Carlo error propagation**. We create 100 different copies of **Monte Carlo pseudodata**, generated as a **multivariate Gaussian distribution** around the central data, then find the best-fit PDF parameters for each of the 100 copies.
- We can then take **envelopes** to get uncertainties from the resulting **PDF ensemble**.



PDFs with error bands

The choice of functional form

- The choice of functional form that we have suggested so far is:

$$f(x, Q_0^2) = Ax^{\alpha}(1-x)^{\beta}\left(1 + ax^{1/2} + bx + cx^{3/2}\right)$$

- This seems a bit arbitrary though! To try to remove as much bias as possible, another possible choice is to parametrise the PDFs using a **neural network** instead:

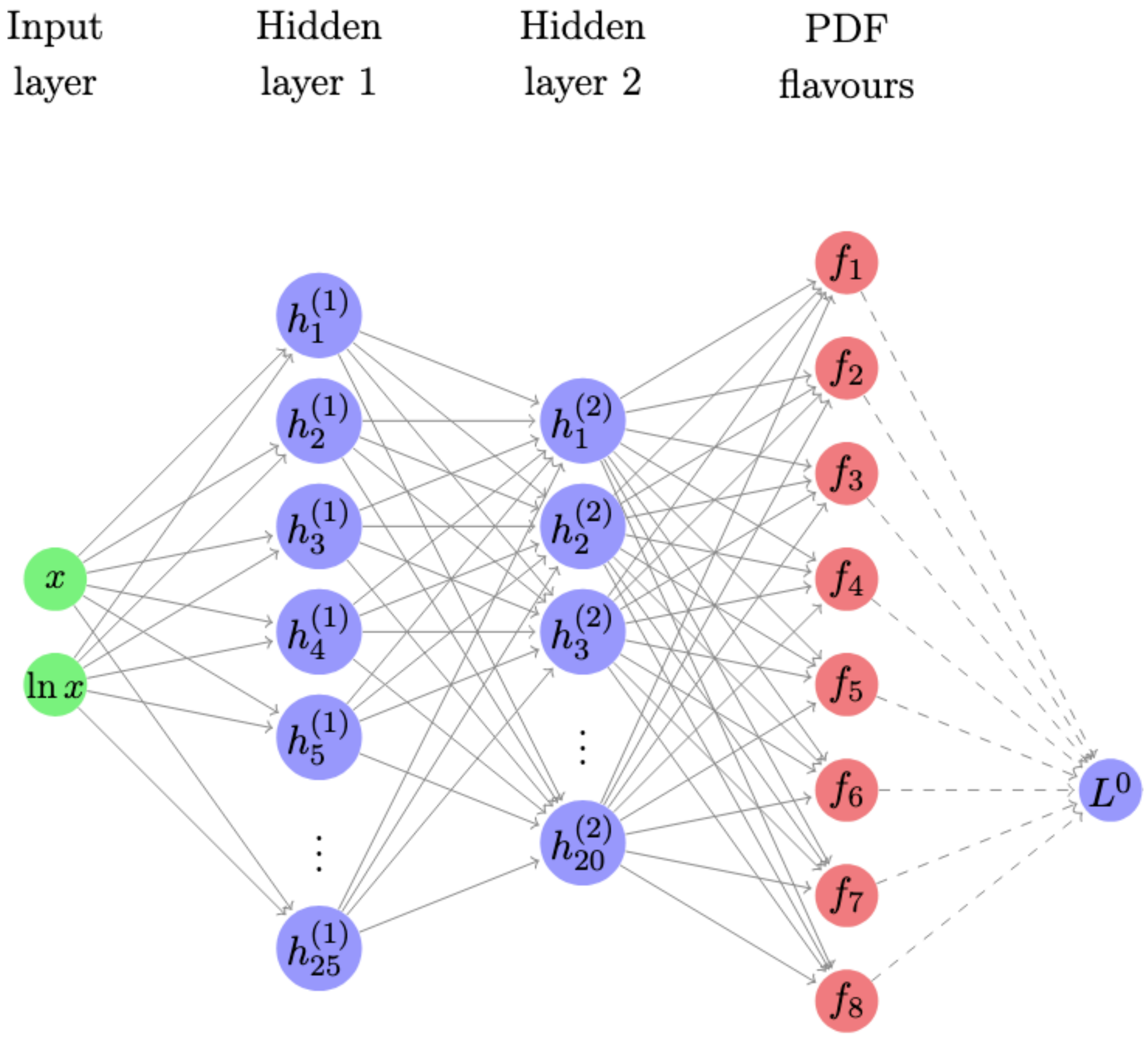
$$f(x, Q_0^2) = Ax^{\alpha}(1-x)^{\beta}\mathrm{NN}(x, \omega)$$

- Here, $\mathrm{NN}(x, \omega)$ is a **neural network** which takes in x as an argument, and has network parameters ω.

The choice of functional form

$$f(x, Q_0^2) = Ax^{\alpha}(1-x)^{\beta}\mathrm{NN}(x, \omega)$$

- The neural network parametrisation is used by the **NNPDF collaboration**, whose fitting code is publicly available (and I use regularly!).



So what do PDFs look like, and why?

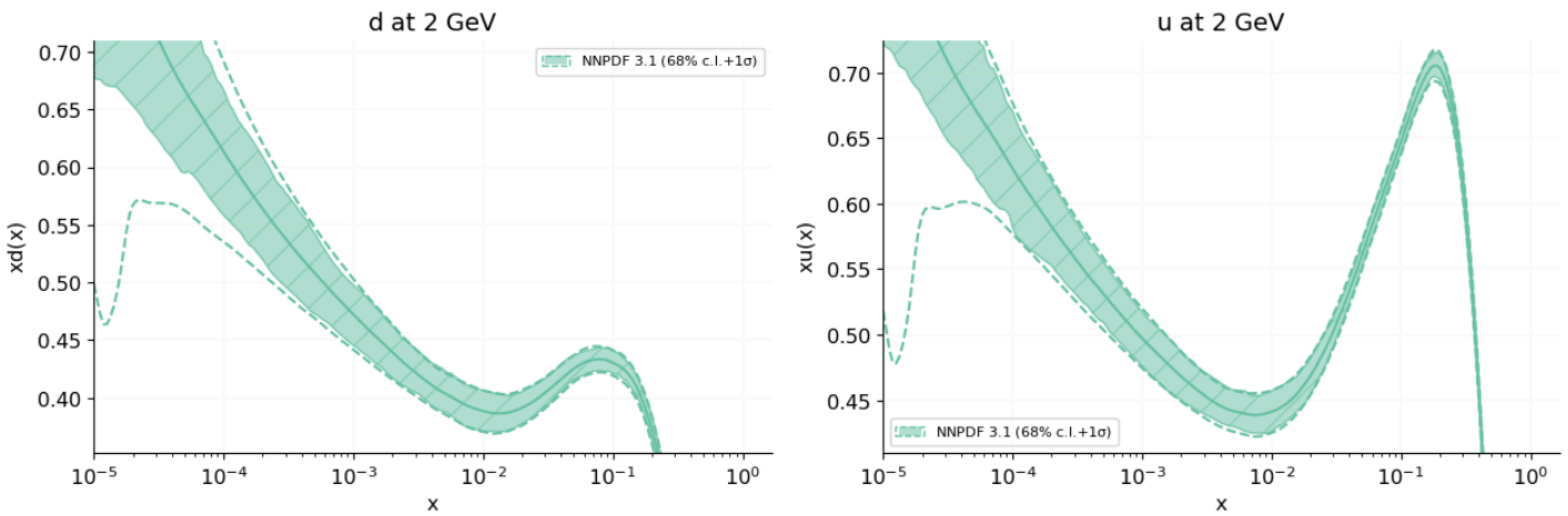
- Now we have described how to obtain PDFs, let's look at some examples!

So what do PDFs look like, and why?

- Now we have described how to obtain PDFs, let's look at some examples!

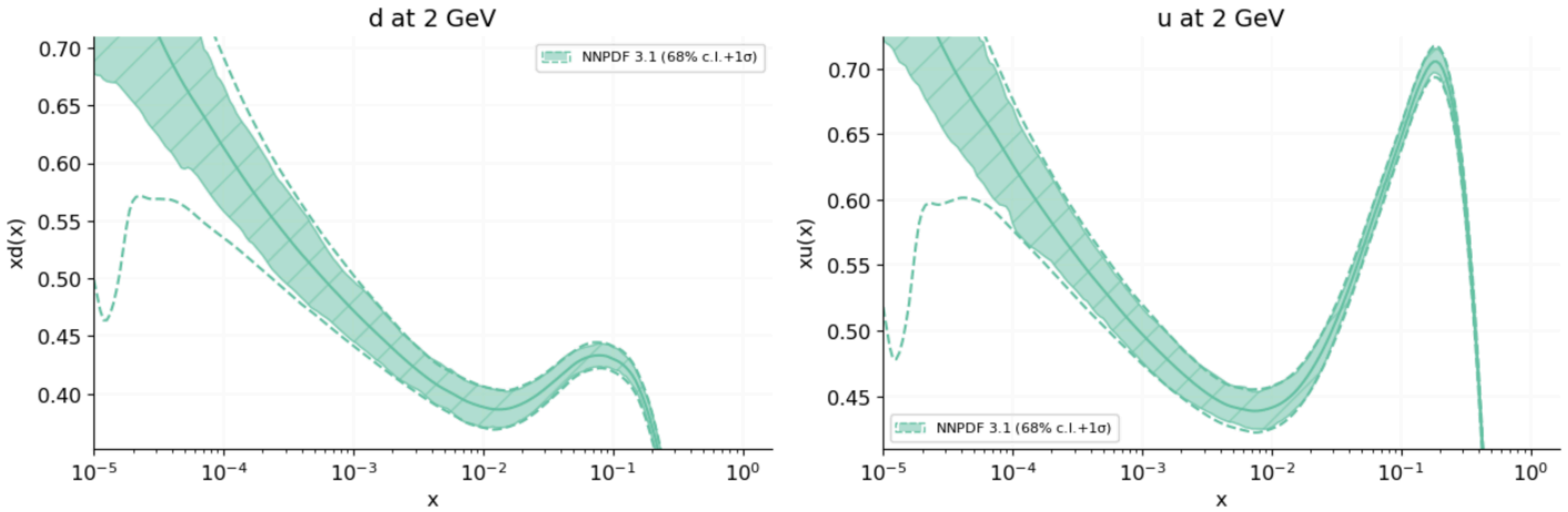
So what do PDFs look like, and why?

- Now we have described how to obtain PDFs, let’s look at some examples!



So what do PDFs look like, and why?

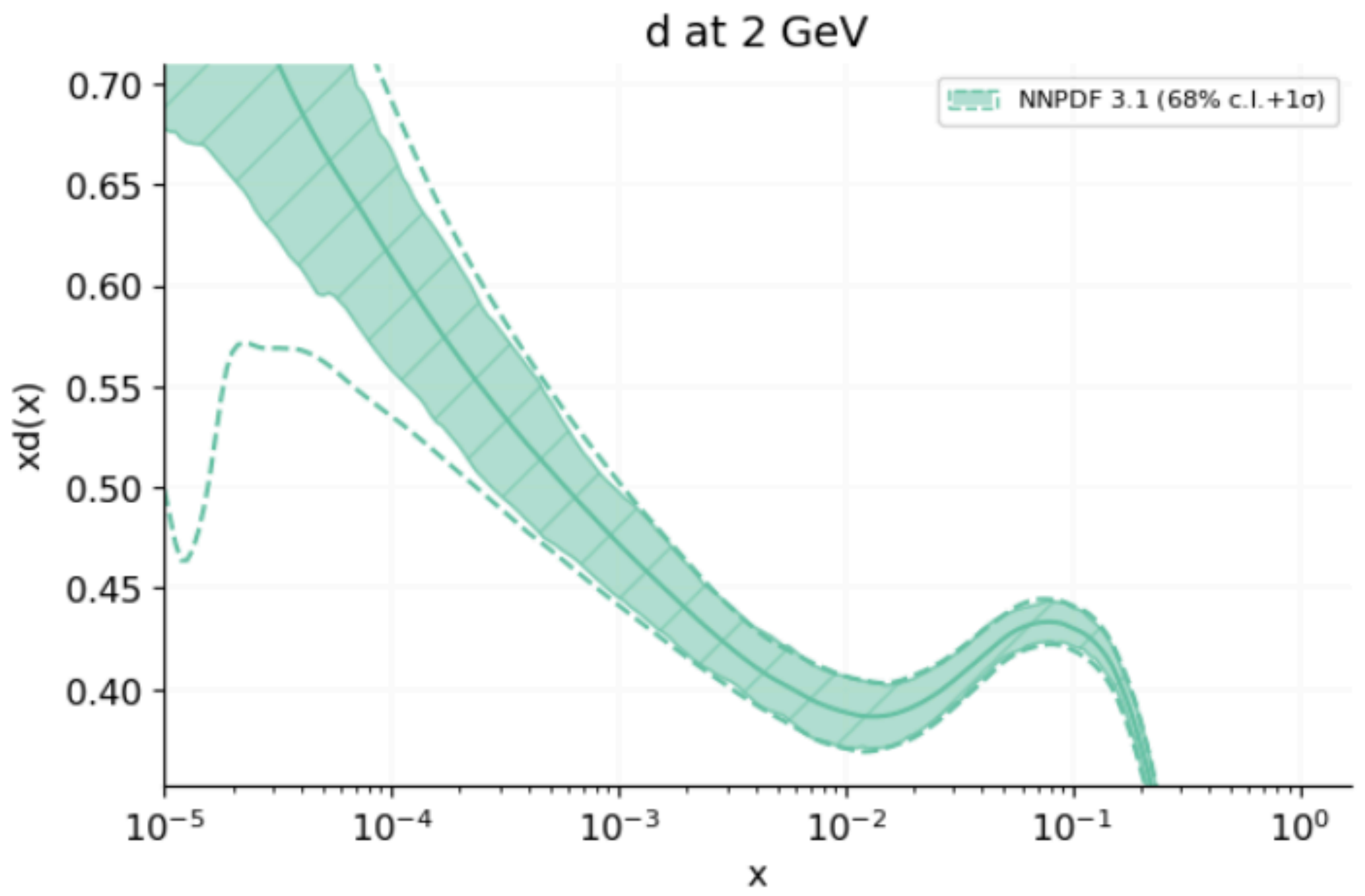
- Now we have described how to obtain PDFs, let's look at some examples!

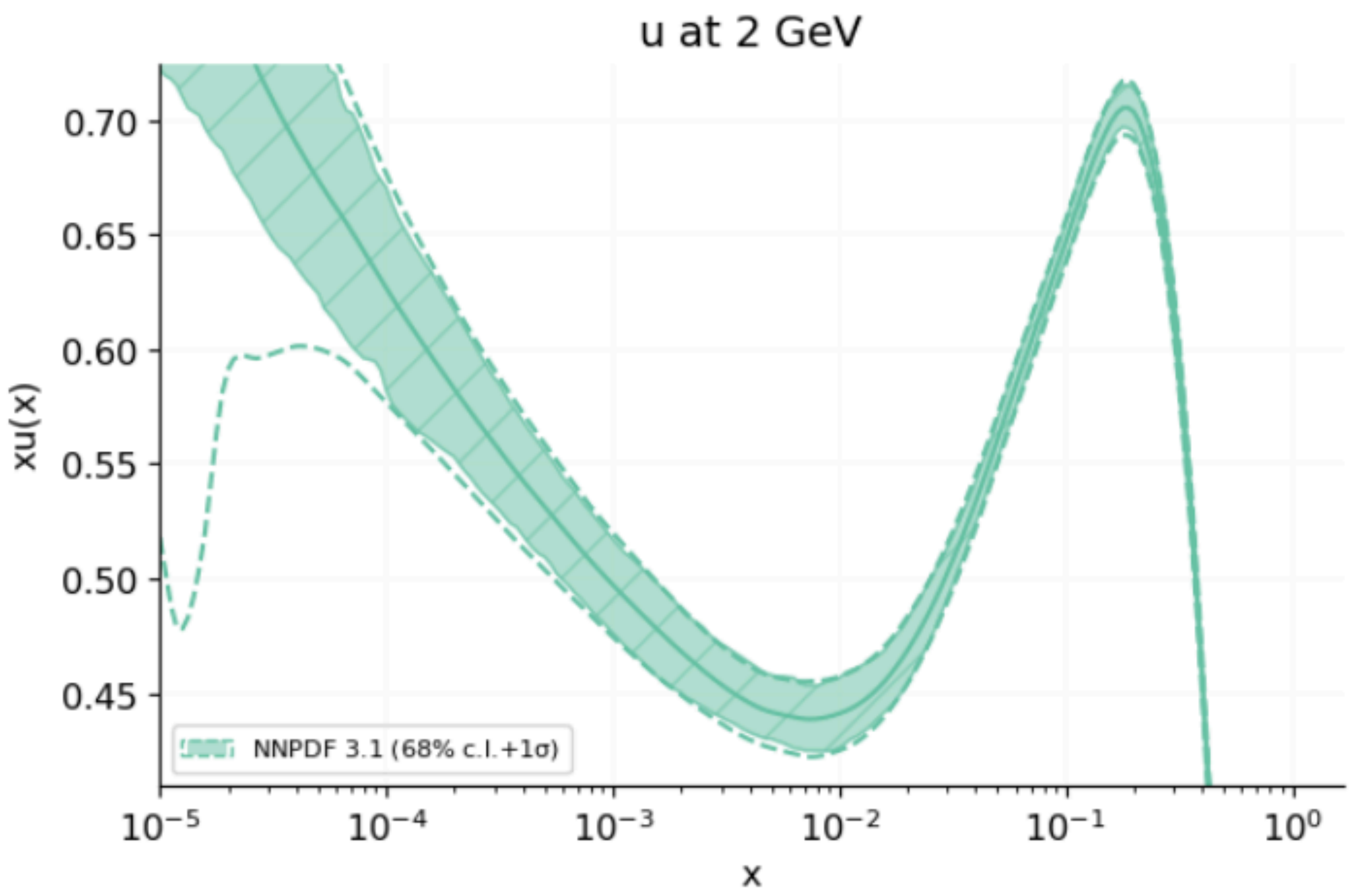


- If we think of the proton as '**two up quarks and two down quarks**', we naively expect the up, down distributions to be delta functions peaked at $x = 1/3$.

So what do PDFs look like, and why?

- Now we have described how to obtain PDFs, let's look at some examples!

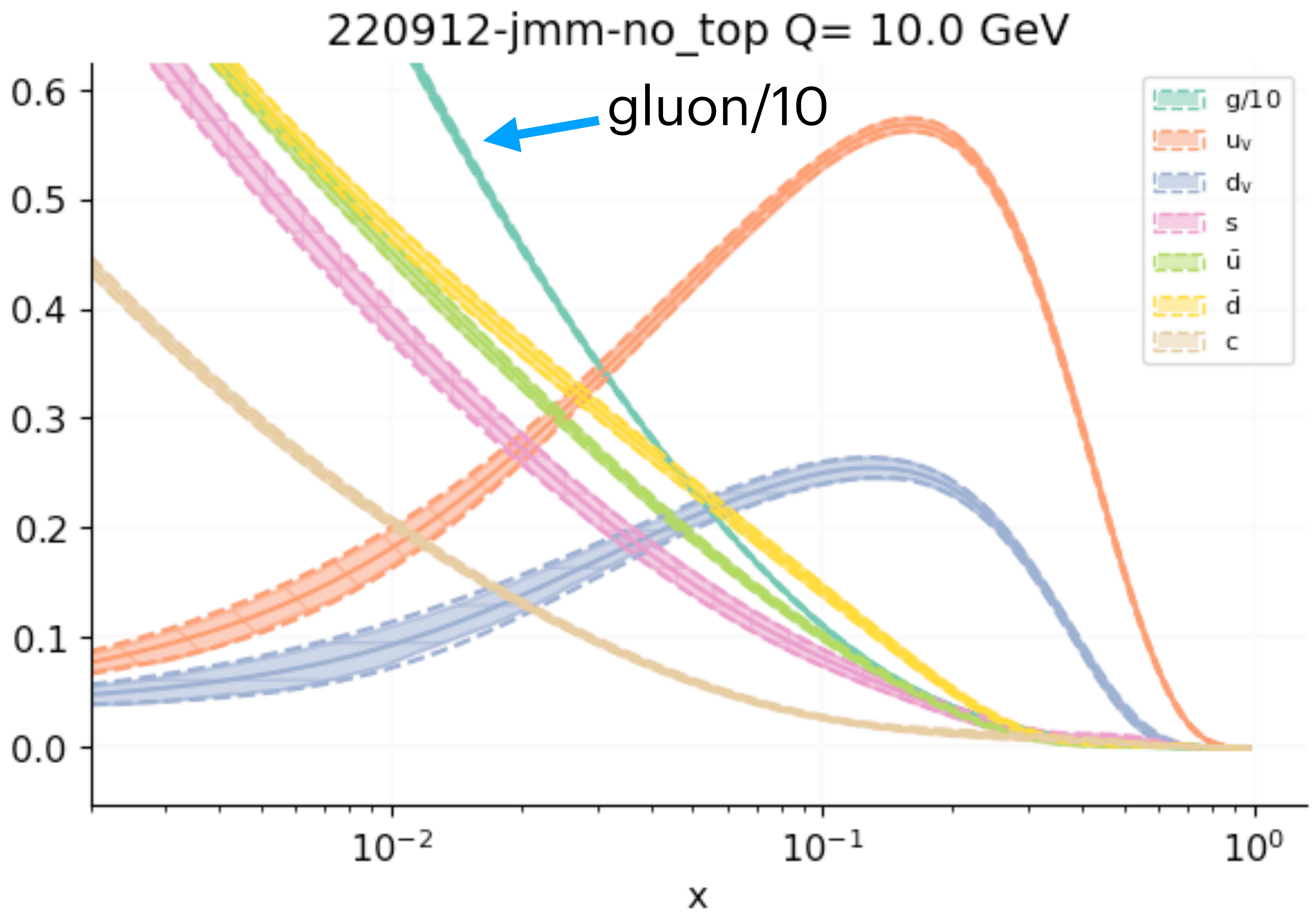




- In reality, we see that **quantum fluctuations** result in the creation of up/anti-up and down/anti-down pairs with small momentum fractions, which cause the distributions to **increase at small** x.

So what do PDFs look like, and why?

- Most flavours only arise **virtually** inside the proton, so we don't get the peaked behaviour for other species of quark.



- One flavour features much more heavily than others: **gluons**.
- In fact, the momentum due to the gluons accounts for nearly **1/3 of all momentum of a proton**!

3. - Beyond the standard proton

The Standard Model is *incomplete*...

- Whilst the Standard Model has been **extremely successful**, it is known to be incomplete. There are lots of things it does not describe:
  - *Gravity*
  - *Dark matter*
  - *Neutrino masses*
  - *many more obscure things...*

- People working to extend the Standard Model to account for these phenomena are said to be working on **Beyond the Standard Model physics** (BSM).

So how do we fix the Standard Model?

- For example, to **include dark matter** in the Standard Model, we might **hypothesise new particles** and add them in. The Standard Model Lagrangian density is augmented to:

$$\mathscr{L}_{\text{new}} = \mathscr{L}_{\text{SM}} + \mathscr{L}_{\text{dark matter}}$$

So how do we fix the Standard Model?

- For example, to **include dark matter** in the Standard Model, we might **hypothesise new particles** and add them in. The Standard Model Lagrangian density is augmented to:

$$\mathscr{L}_{\text{new}} = \mathscr{L}_{\text{SM}} + \mathscr{L}_{\text{dark matter}}$$

- We could then **try to produce the new particles directly** (*direct detection*), or **fit existing data using this theory to see if we get a better fit** (*indirect detection*).

So how do we fix the Standard Model?

- For example, to **include dark matter** in the Standard Model, we might **hypothesise new particles** and add them in. The Standard Model Lagrangian density is augmented to:

$$\mathscr{L}_{\text{new}} = \mathscr{L}_{\text{SM}} + \mathscr{L}_{\text{dark matter}}$$

- We could then **try to produce the new particles directly** (*direct detection*), or **fit existing data using this theory to see if we get a better fit** (*indirect detection*).
- However, there are **thousands** of possibilities, so just guessing particles seems a bit like **stabbing in the dark**!

So how do we fix the Standard Model?

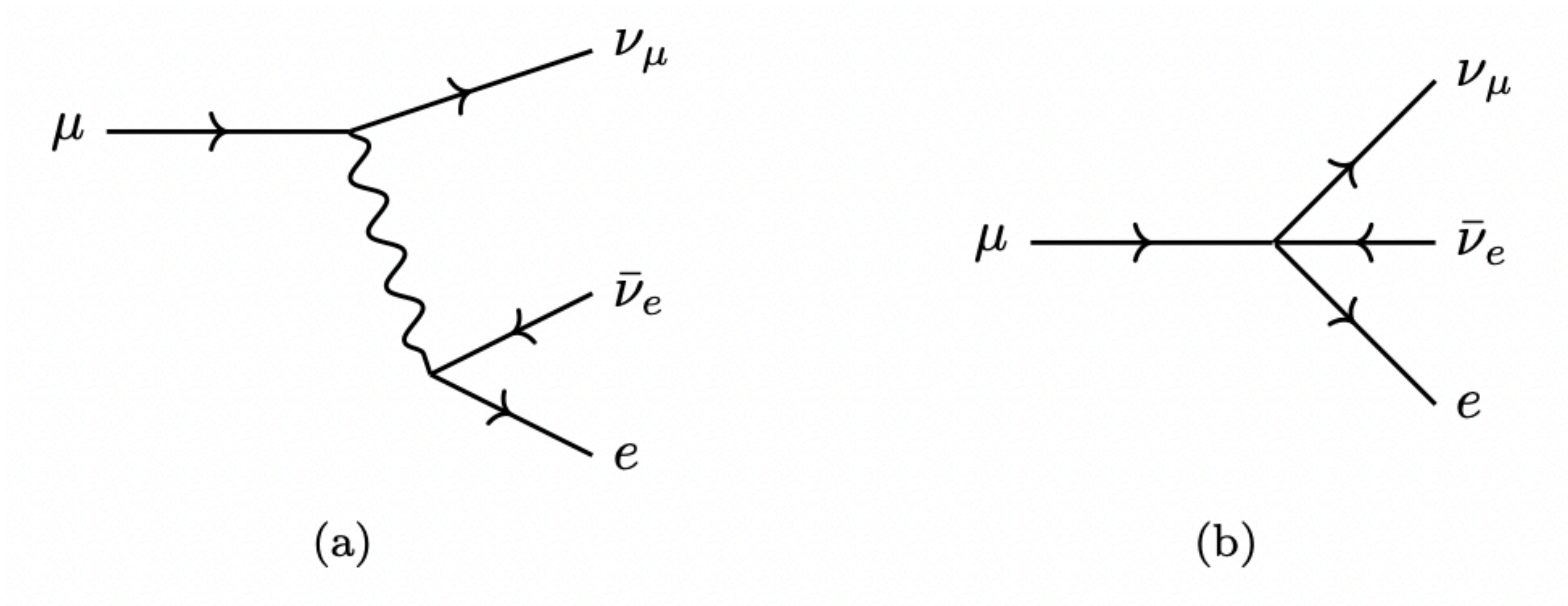
- For example, to **include dark matter** in the Standard Model, we might **hypothesise new particles** and add them in. The Standard Model Lagrangian density is augmented to:

$$\mathscr{L}_{\text{new}} = \mathscr{L}_{\text{SM}} + \mathscr{L}_{\text{dark matter}}$$

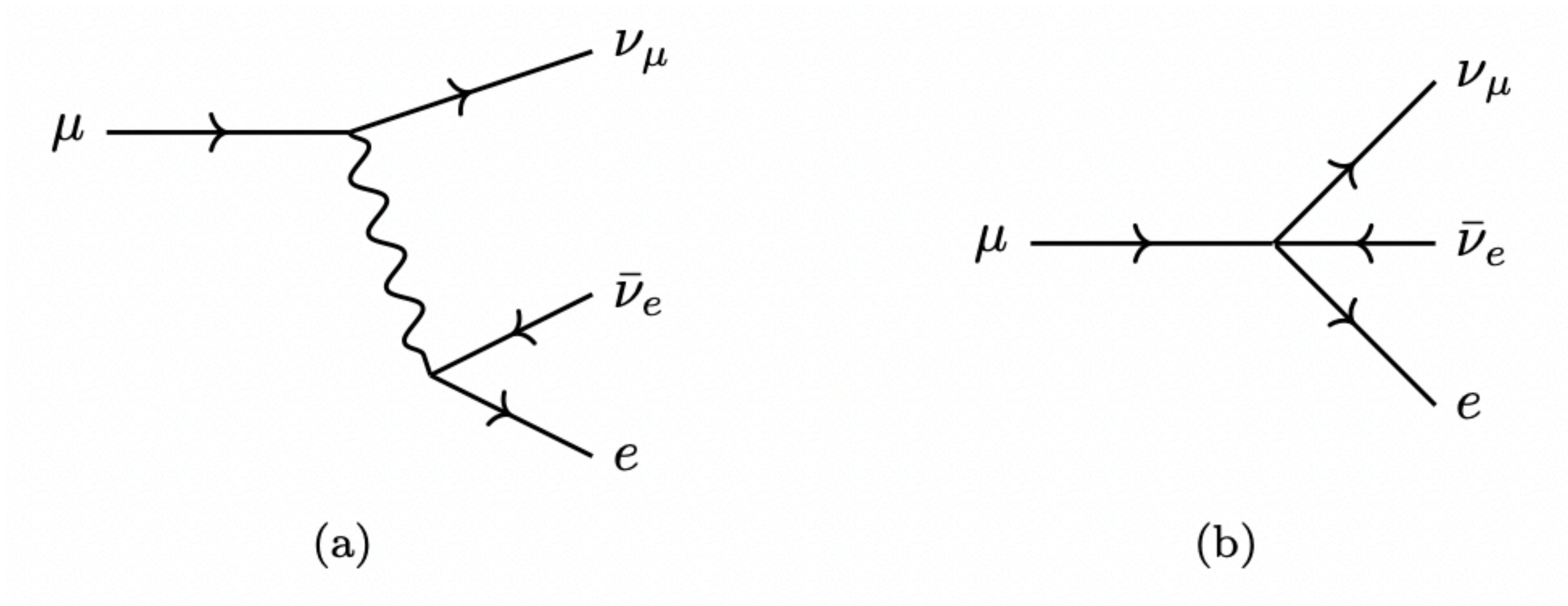
- We could then **try to produce the new particles directly** (*direct detection*), or **fit existing data using this theory to see if we get a better fit** (*indirect detection*).
- However, there are **thousands** of possibilities, so just guessing particles seems a bit like **stabbing in the dark**!
- Some models are **more motivated** than others, but it would be nice to have a more general approach...

Enter the SMEFT...

- Fortunately, the language of **effective field theory** exists to help us tackle this problem.
- *Idea:* at **low energies** we can't distinguish between a **particle being exchanged**, or an **interaction between multiple particles**.

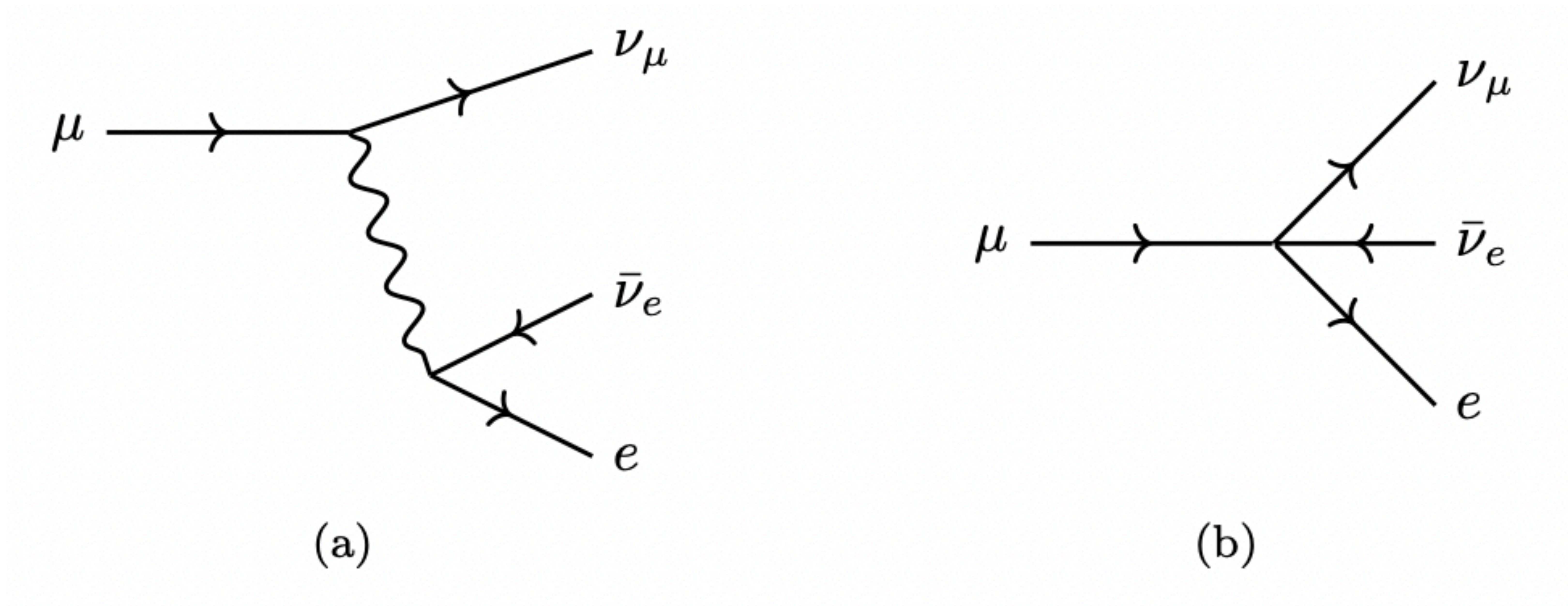


Enter the SMEFT...



- For example, in **muon decay**, the final decay products are two neutrinos and an electron, and the decay is mediated by a W-boson.
- But if we didn't know the W-boson existed, we would think that there was a **direct interaction** between muons, neutrinos and electrons.

Enter the SMEFT...



- It can be shown that four-point interactions, like those in (b), are actually forbidden in a fundamental quantum field theoretic description of Nature - they are '**non-renormalisable**'.
- In particular, if we saw the process (b) without knowing the existence of the W-boson, we could **infer its existence**!

Enter the SMEFT...

- This is the idea of the **Standard Model effective field theory (SMEFT)**. We add to the SM Lagrangian density all possible **non-renormalisable interactions** between the **SM particles**.
- Roughly speaking, they can be organised by the number of particles participating in the interaction:

$$\mathscr{L}_{\text{SMEFT}} = \mathscr{L}_{\text{SM}} + \mathscr{L}_{4\text{-point}} + \mathscr{L}_{5\text{-point}} + \cdots$$

- Looking at the smallest number of particles first, the **interaction strengths** in $\mathscr{L}_{4-\text{point}}$ are unknown, but can be **found by precise fits to data**. If we see non-zero values, it means there **must be new particles**.

Enter the SMEFT...

$$\mathscr{L}_{\text{SMEFT}} = \mathscr{L}_{\text{SM}} + \mathscr{L}_{4\text{-point}} + \mathscr{L}_{5\text{-point}} + \cdots$$

- Unfortunately, there are **2499 different interactions** in $\mathscr{L}_{\text{SMEFT}}$, so this is a lot of work! At the moment, people can only fit subsets of the interactions at a time.
- Various fitting groups **just fit** the interactions strengths, for example the **SMEFiT collaboration**, and the **FitMaker collaboration**.
- This can be problematic if **data involving protons** is used in the fits because of **PDFs**...

Joint PDF-SMEFT fits?

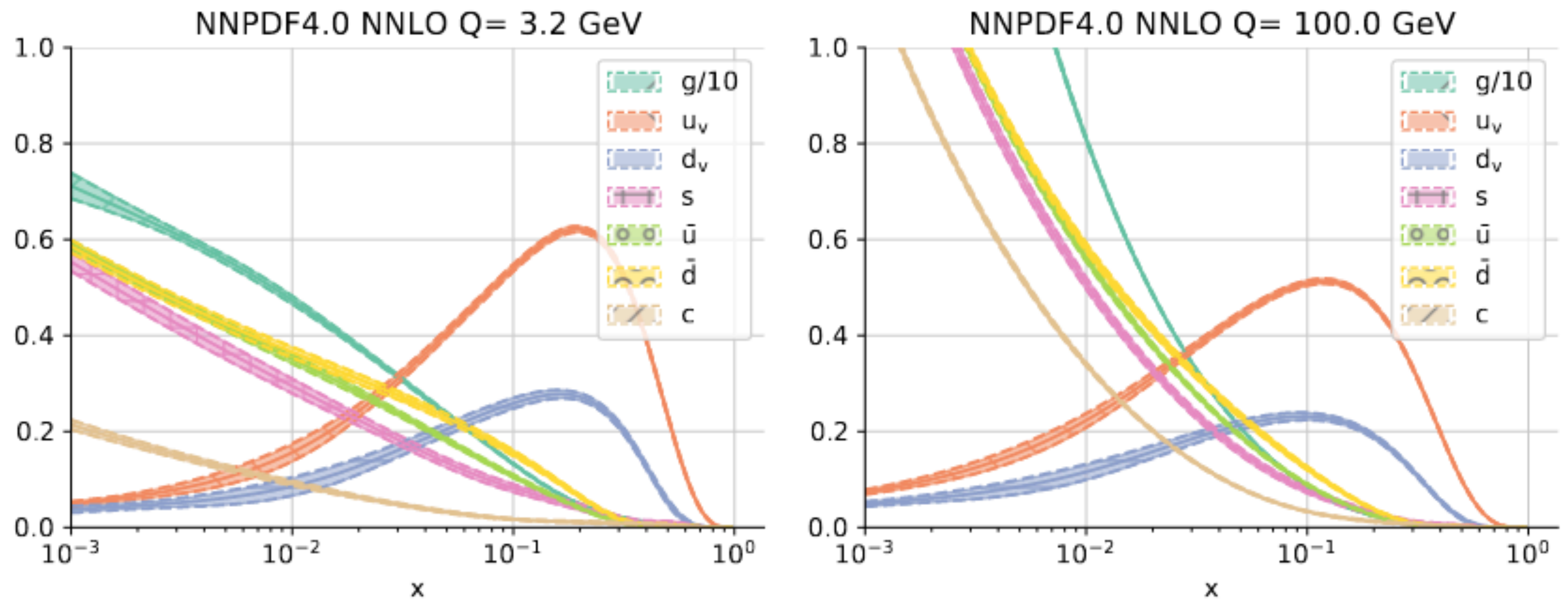
- Usually, people **fit the SMEFT parameters and PDFs separately**:

PDF parameter fits

- Fix SMEFT parameters (usually to zero), $c = \bar{c}$:

$$\sigma(\bar{c}, \theta) = \hat{\sigma}(\bar{c}) \otimes \text{PDF}(\theta)$$

- Optimal PDF parameters θ^* then have an **implicit dependence** on initial SMEFT parameter choice: $\text{PDF}(\theta^*) \equiv \text{PDF}(\theta^*(\bar{c}))$.
- E.g. NNPDF4.0 fit, Ball et al., 2109.02653.

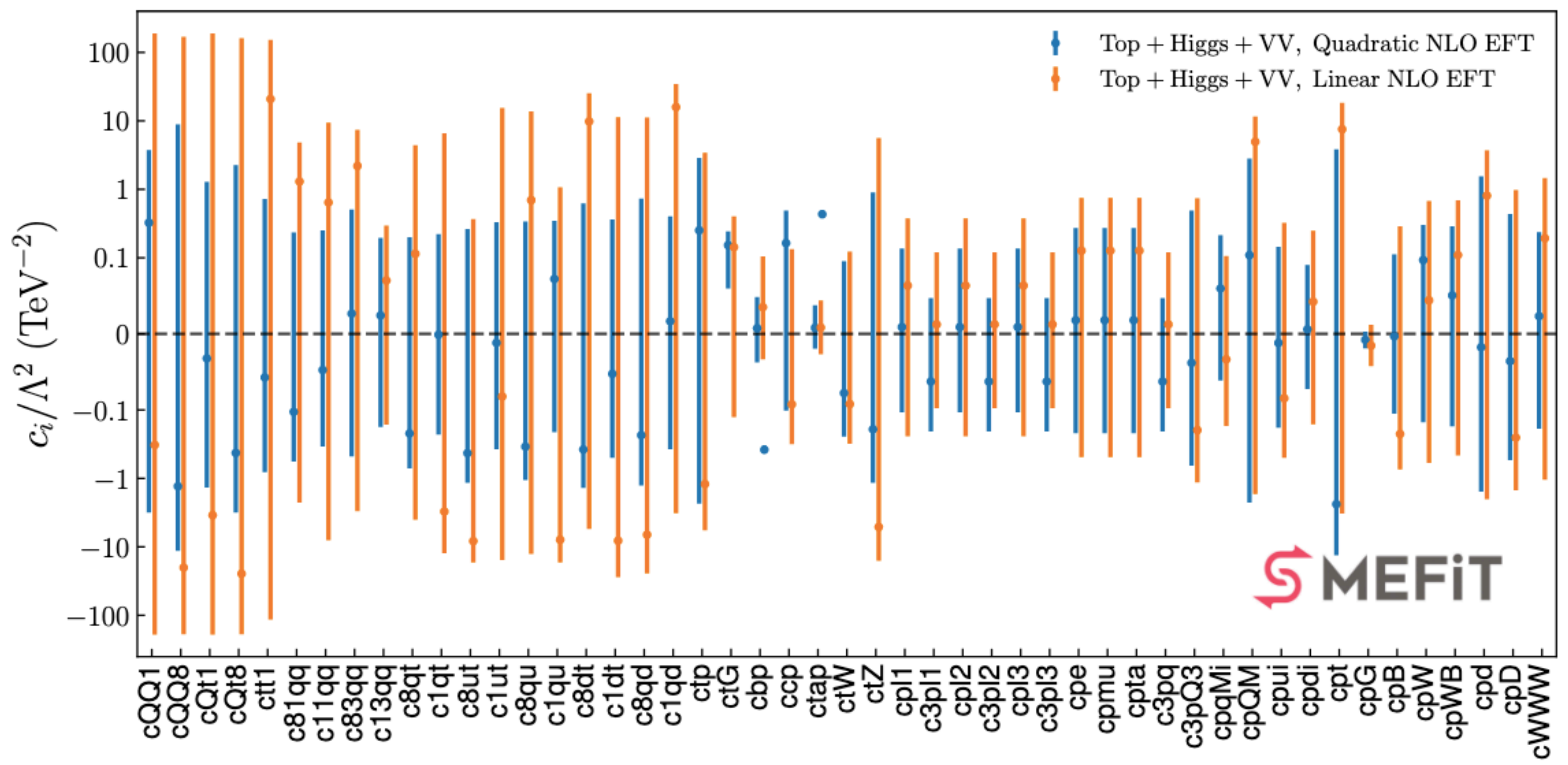


SMEFT parameter fits

- Fix PDF parameters $\theta = \bar{\theta}$:

$$\sigma(c, \bar{\theta}) = \hat{\sigma}(c) \otimes \text{PDF}(\bar{\theta})$$

- Optimal SMEFT parameters c^* then have an **implicit dependence** on PDF choice: $c^* = c^*(\bar{\theta})$.
- E.g. SMEFiT, Ethier et al., 2105.00006.



Fitting PDFs and physical parameters

- **This could lead to inconsistencies.**

PDF parameter fits

$$\text{PDF}(\theta^*) \equiv \text{PDF}(\theta^*(\bar{c}))$$

- Fitted PDFs can depend implicitly on fixed SMEFT parameters used in the fit.

SMEFT parameter fits

$$c^* \equiv c^*(\bar{\theta})$$

- Bounds on SMEFT parameters can depend implicitly on the fixed PDF set used in the fit.

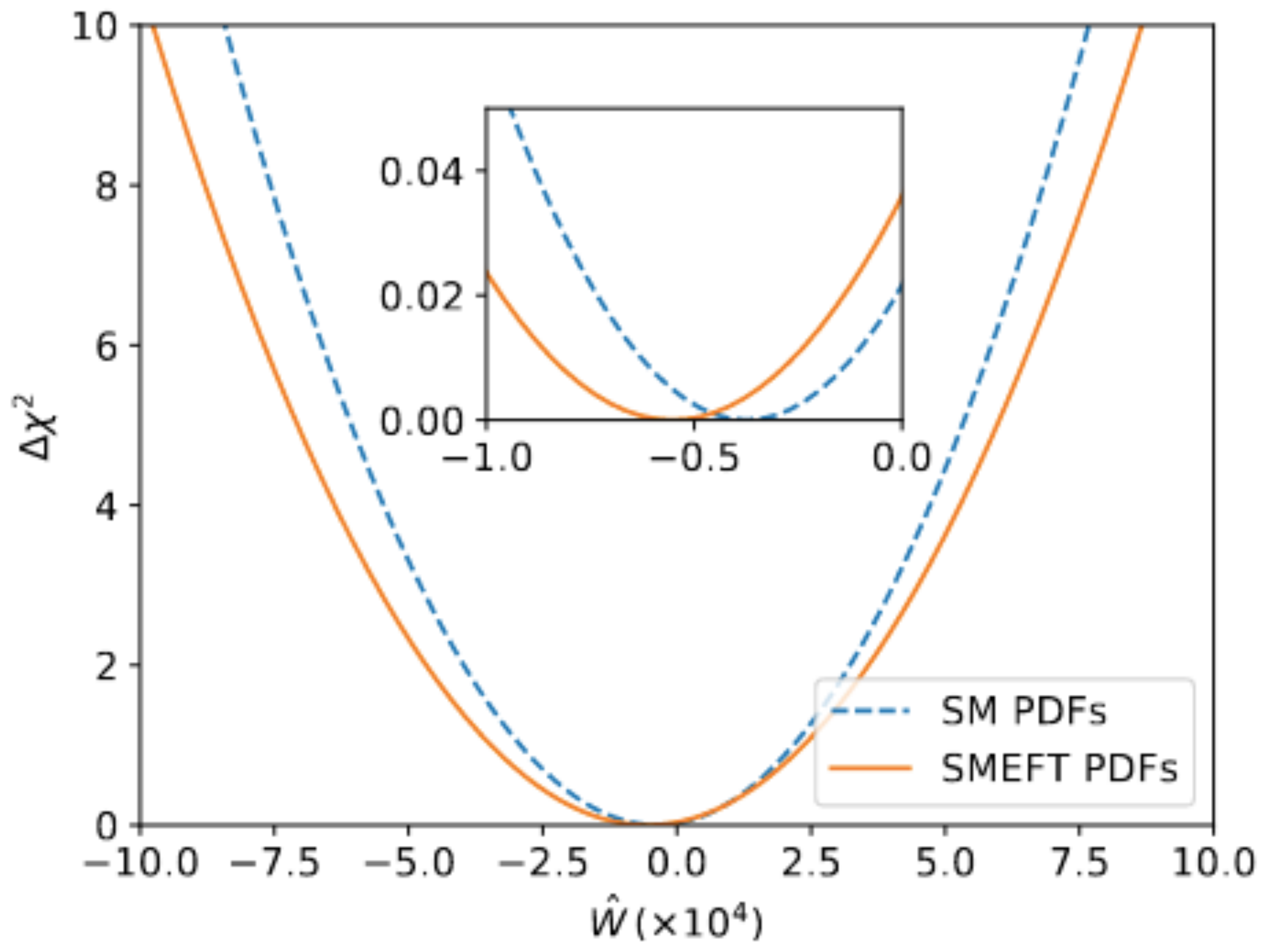
- In particular, if we fit PDFs **assuming all SMEFT interactions are zero**, but then **use those PDFs in a fit of SMEFT interactions**, our resulting bounds **might be misleading**. The same applies to SM parameters.
- In the case of BSM models, we could even **miss New Physics**, or **see New Physics that isn't really there**!

Key question for remainder of talk:

To what extent do bounds on SMEFT parameters change if they are fitted simultaneously with PDF parameters? Is a consistent treatment important?

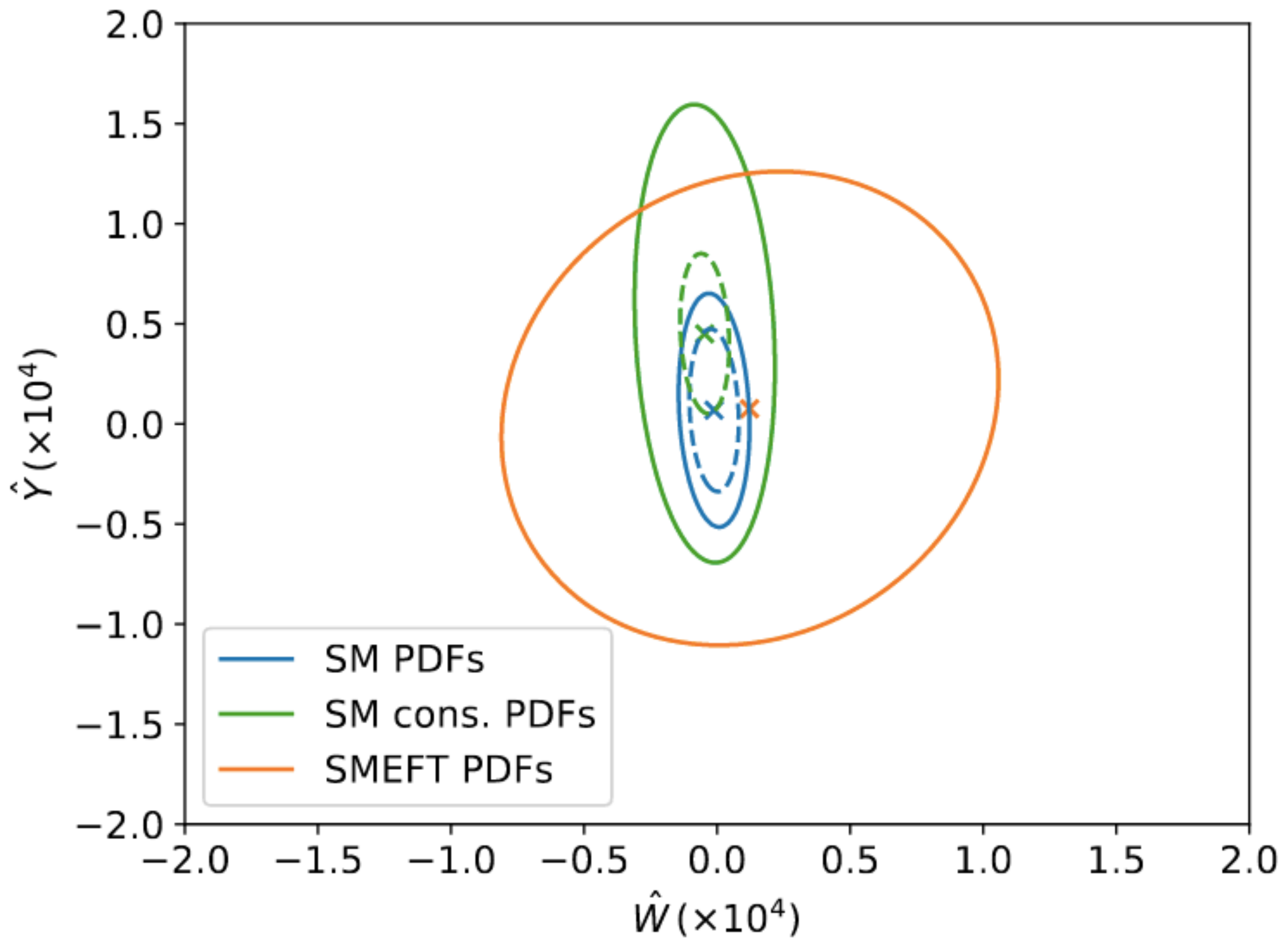
Parton distributions in the SMEFT from high-energy Drell-Yan tails

- In particular, in the paper 2104.02723 from the PBSP team (+ Greljo, Rojo), we find that in the context of the **oblique W, Y parameters**, a simultaneous fit of PDFs and the SMEFT parameters using **current data** has a **small impact on the bounds**.
- The methodology used is similar to the **'scan' methodology**; we simply take the χ^2 of a PDF fit at each **benchmark point** in Wilson coefficient space to **construct bounds**.



Parton distributions in the SMEFT from high-energy Drell-Yan tails

- On the other hand, when we use **projected HL-LHC data**, the impact of a simultaneous fit versus a fixed PDF fit becomes **enormous**!
- Without a simultaneous fit, we find that the size of the bounds is **significantly underestimated** - this could lead to claims of discovering New Physics when it **isn't necessarily there**.



4. - Conclusions

Conclusions

- The Standard Model of particle physics has proven **robust to all challenges so far**, but remains **incomplete**. We can search for New Physics is an organised way using the **Standard Model effective field theory**.

- One of the key ingredients of collider predictions, namely **PDFs**, must be obtained from **global fits to data**.

- Assuming that there is **no interplay** between **PDF fitting** and **fits of the SMEFT interaction strengths** can result in **misleading bounds**.

Thanks for listening!
Questions?