High-dimensional model inference with deep-learning: from the LHC to quantum materials

Juan Rojo, VU Amsterdam & Nikhef

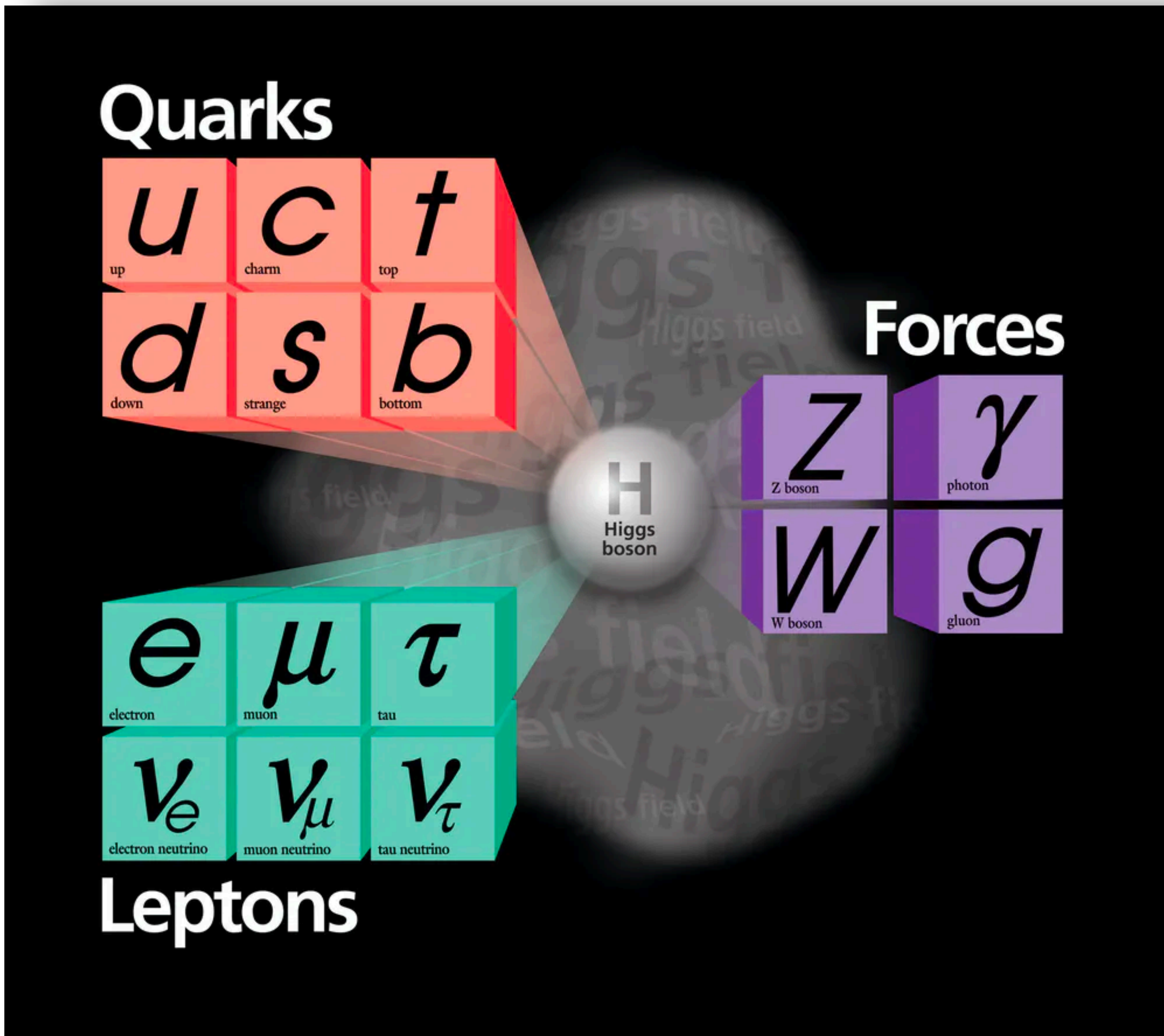
Mini-workshop on AI and Data Modelling @ VU Physics

Amsterdam, 13th March 2024

Scrutinising Proton Structure with AI

The Standard Model

Standard Model of particle physics: hugely succesful, powerful framework describing **elementary particles** and their **interactions**



matter particles

- 6 quarks *(fractional charge)*
- 3 charged leptons (*e.g. electron*)
- 3 neutrinos *(only weak charge)*
- Organised in **3 generations:** identical (?) except for mass

force carriers

- photon (*electromagnetism*)
- gluon (*strong nuclear force*)
- weak bosons (*weak nuclear force*)

Higgs boson

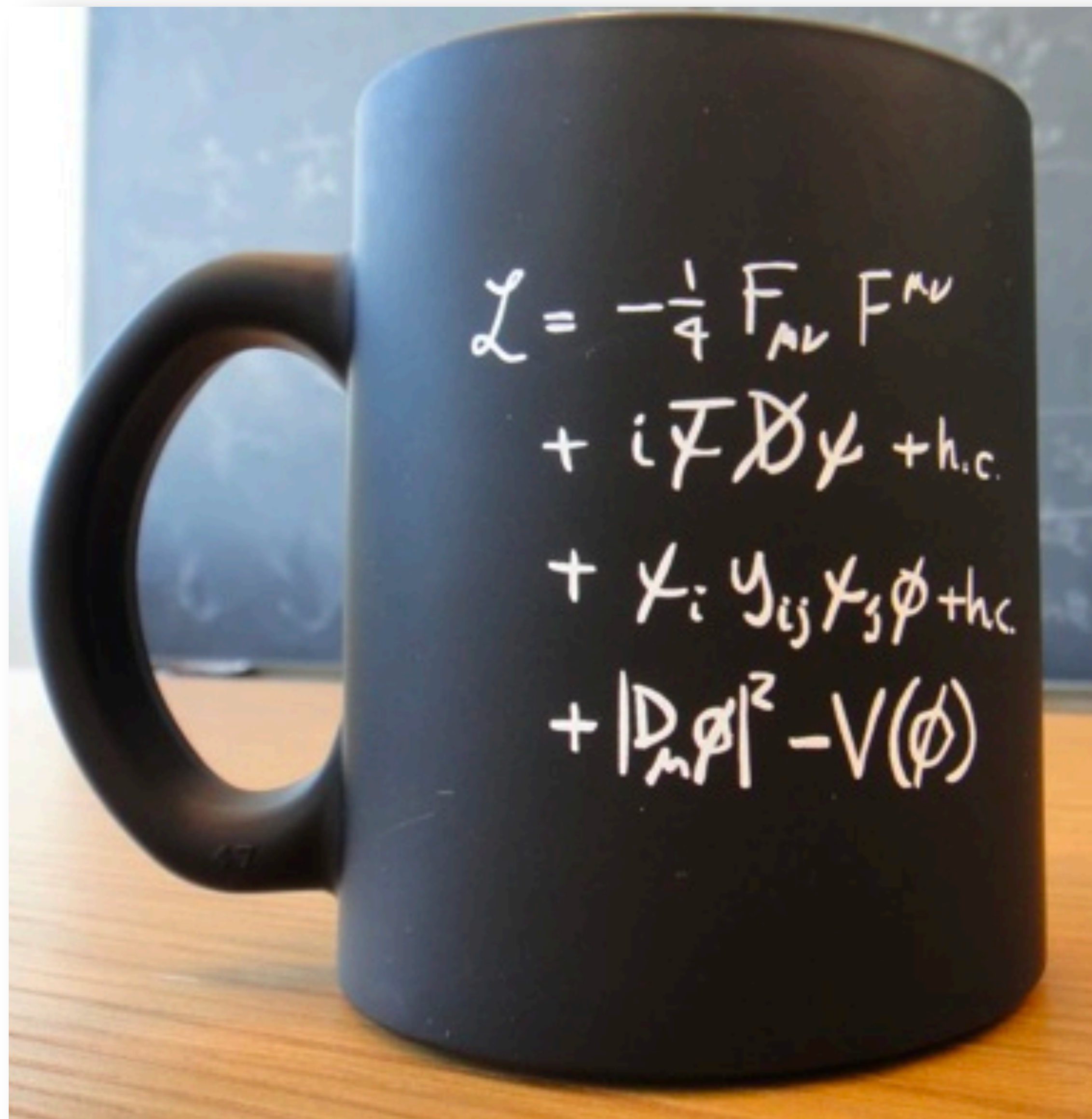
both matter particle and force carrier!

The (incomplete) Standard Model

Standard Model of particle physics: hugely succesful, but leaves many foundational questions unanswered

Origin of particle masses and Higgs force?

Where is all the missing Antimatter?

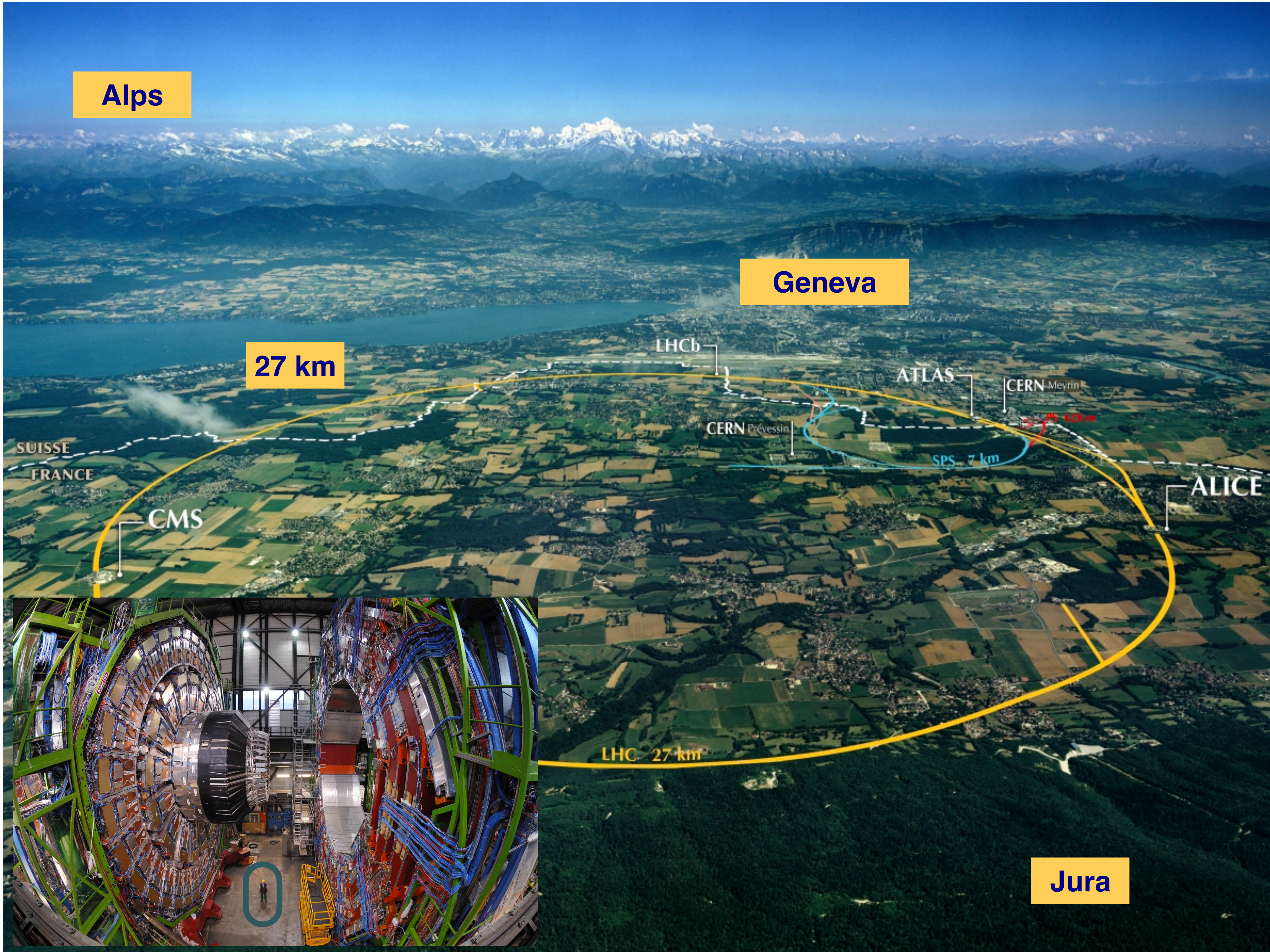


What is Dark Matter?

Quantum Gravity? Inflation?

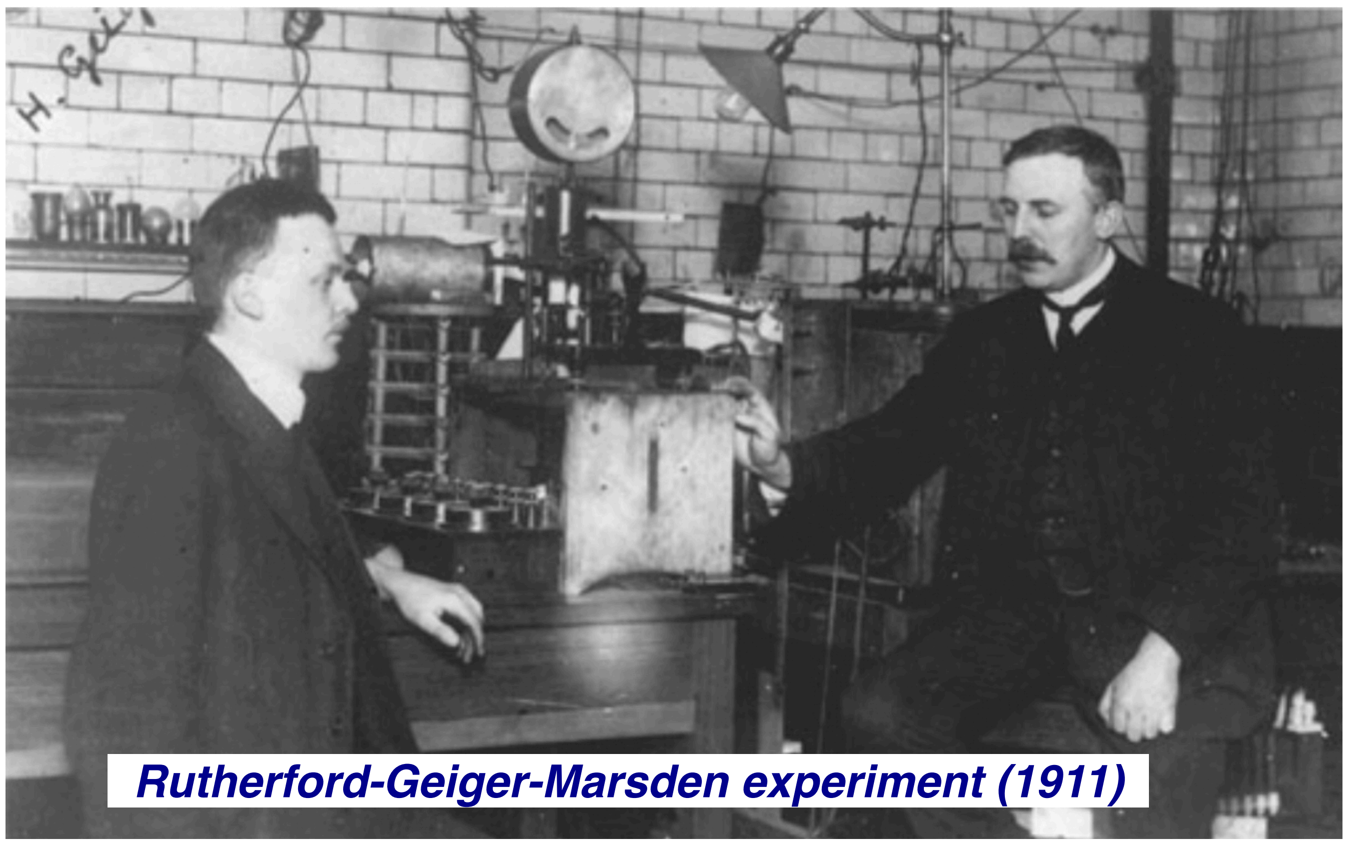
requires new particles and interactions beyond the Standard Model!

The Large Hadron Collider

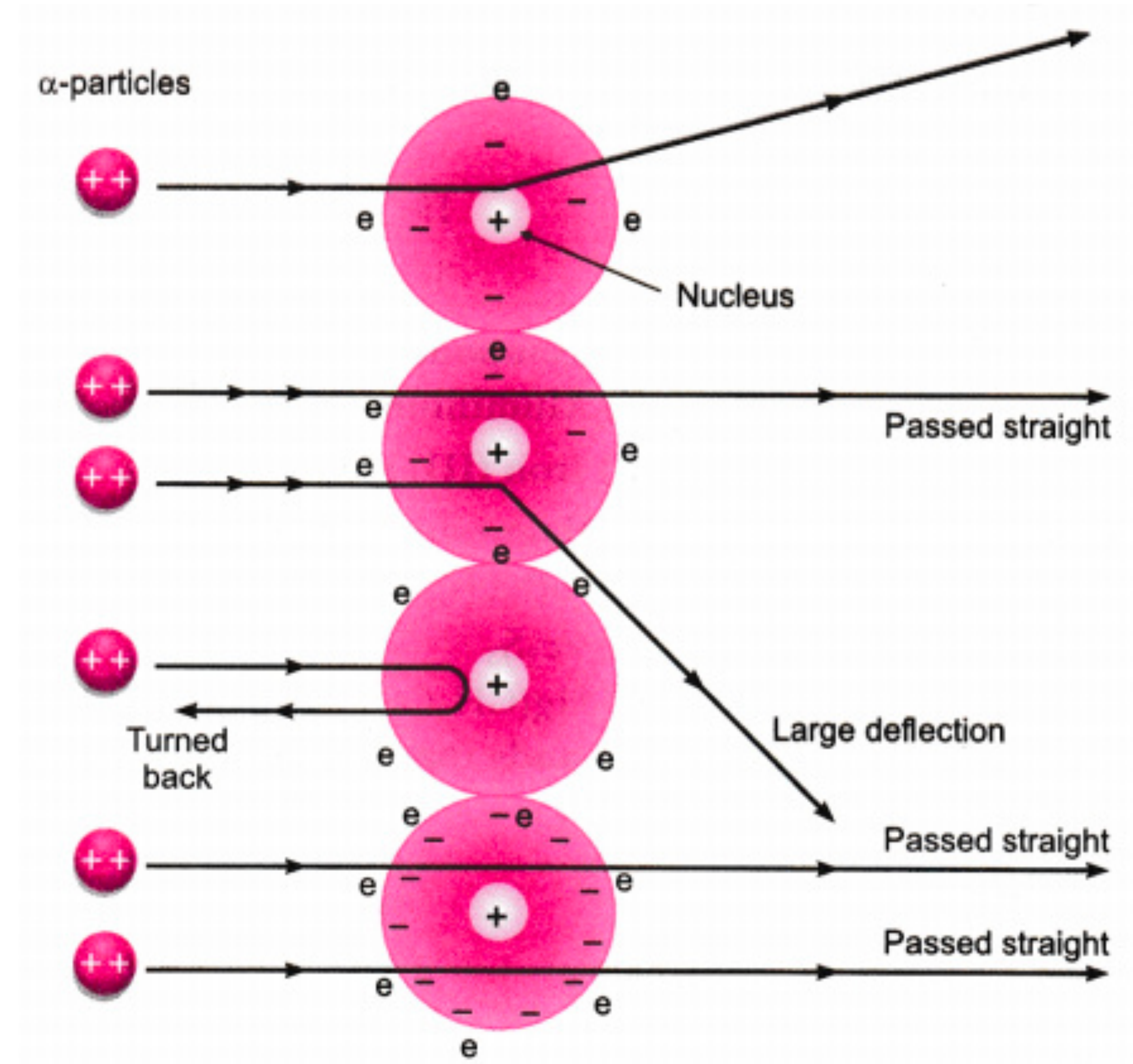


Proton Structure

One may claim that the **nucleon is a rather ``boring" particle**, surely after **one century of studying it**, we know everything about the proton?



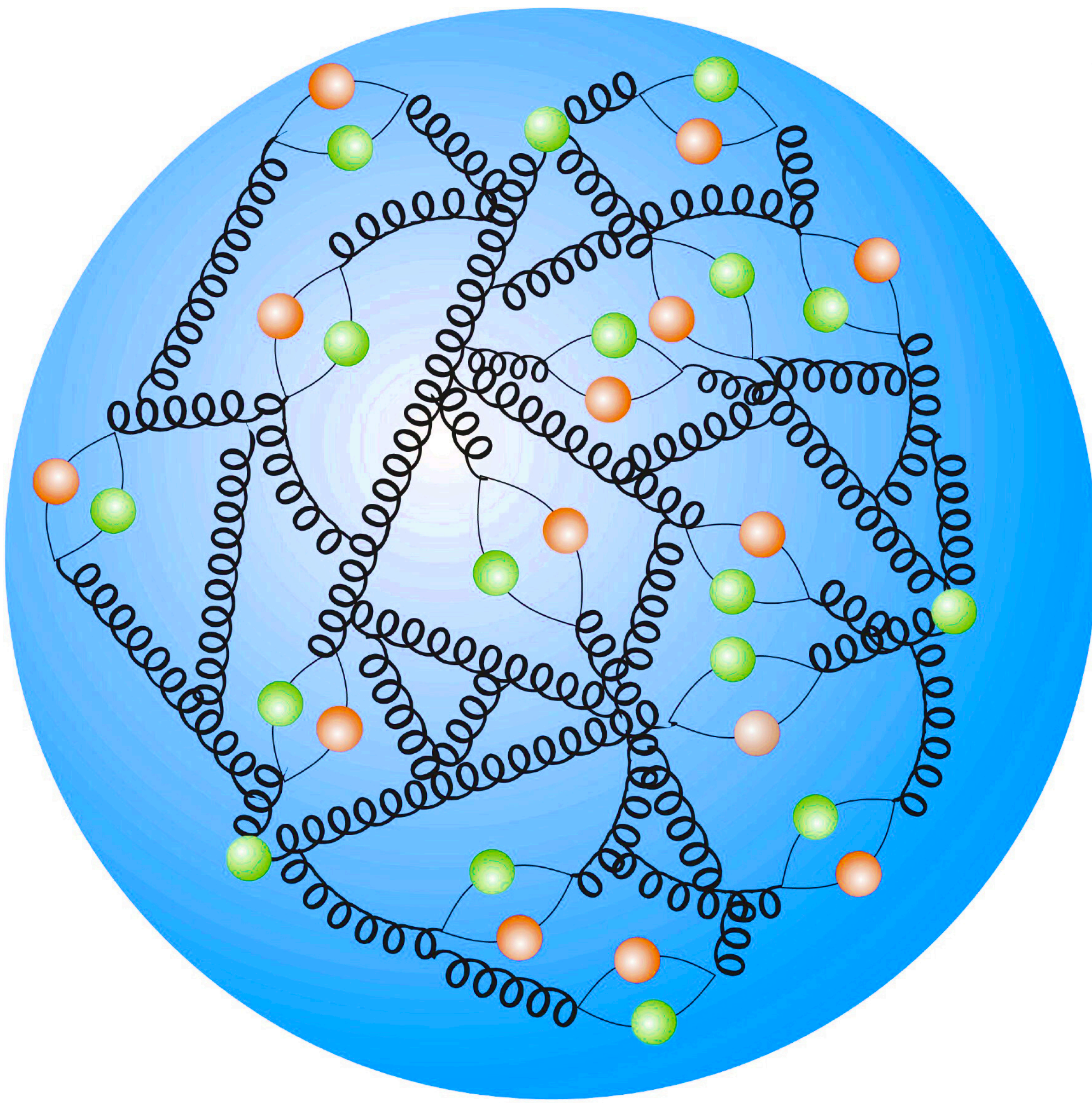
Rutherford-Geiger-Marsden experiment (1911)



nothing farther from reality!

The many faces of the proton

QCD bound state of **quarks** and **gluons**



- **Valence quarks (up and down)** give the proton its quantum numbers (e.g. electric charge)

$$\left|\Psi\right\rangle \approx \left|uud\right\rangle$$

$$Q_p = +1 \qquad Q_u = +2/3 \quad Q_d = -1/3$$

- **Sea quarks (antiup, antidown, strange, …)** arise from quantum fluctuations
- Tightly held together by **gluons,** can only be broken in extremely energetic collisions

Parton Distributions

$$g(x, Q)$$

Energy of hard-scattering reaction**:** inverse of resolution length

Probability of **finding a gluon inside a proton**, carrying a fraction ***x*** of the proton momentum, when probed with energy ***Q***

***x*:** fraction of proton momentum carried by gluon

Dependence on *x* fixed by **non-perturbative QCD dynamics**: extract from experimental data

$$g(x, Q_0, \{a_g\}) = f_g(x, a_g^{(1)}, a_g^{(2)}, \ldots)$$

constrain from data

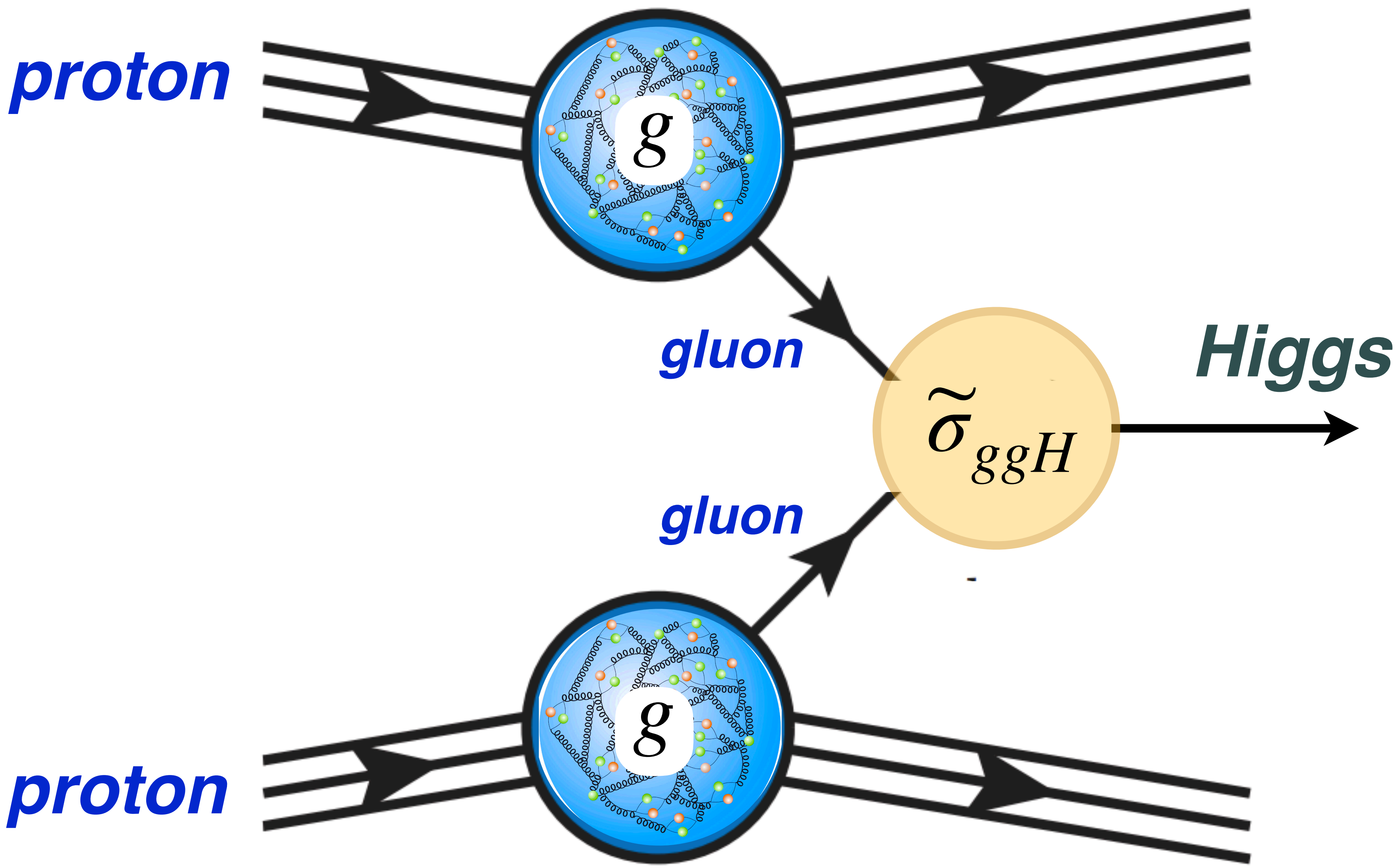
- Dependence with **resolution scale *Q***: DGLAP evolution, computable from first principles

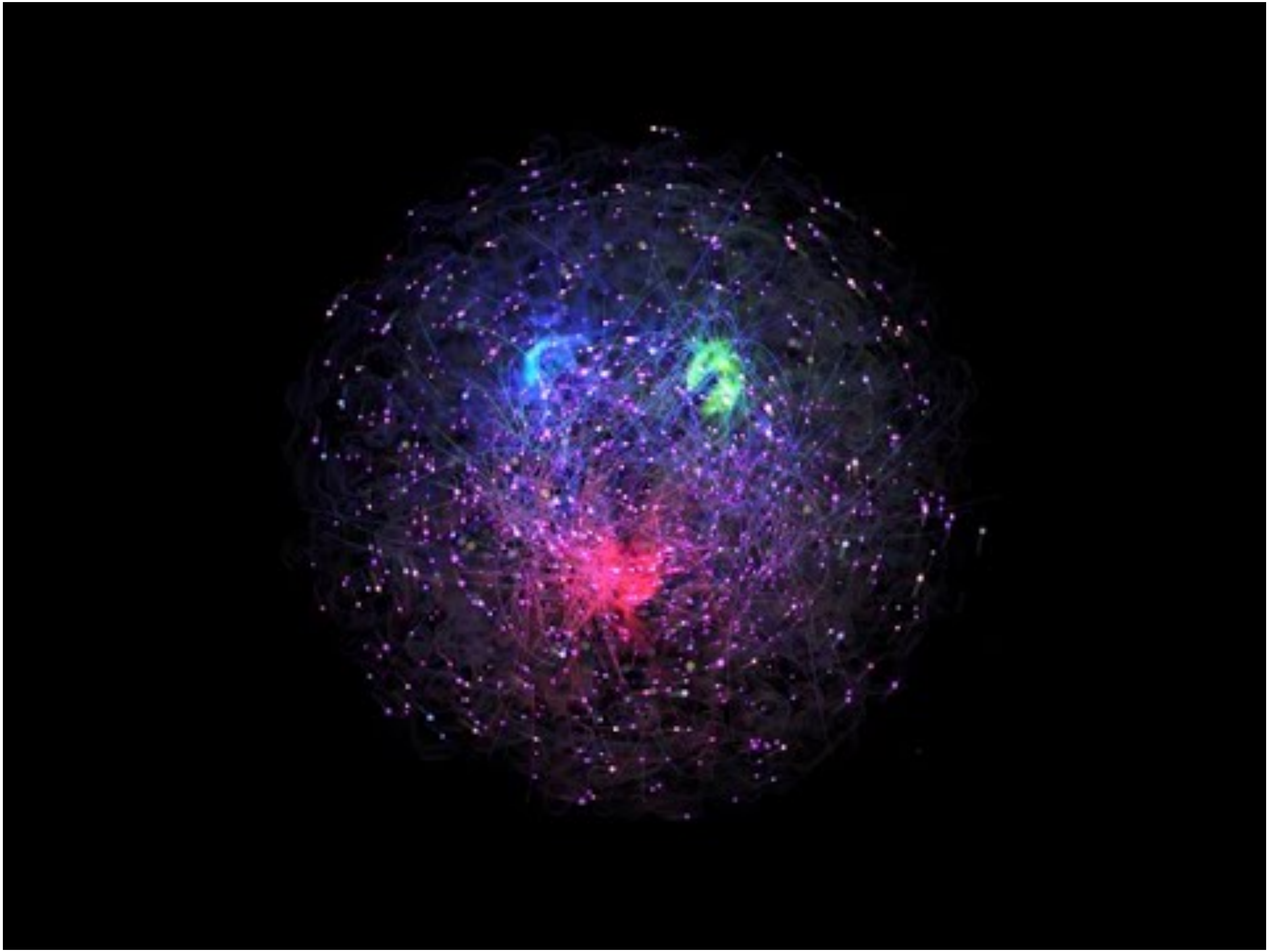
- **Energy conservation** and **quark number conservation** are fixed boundary conditions

Probing Proton Structure

universal *process-dependent*

$$N_{\mathrm{LHC}}(H) \sim \underbrace{g \otimes g}_{\text{Parton Distributions}} \otimes \tilde{\sigma}_{ggH}$$



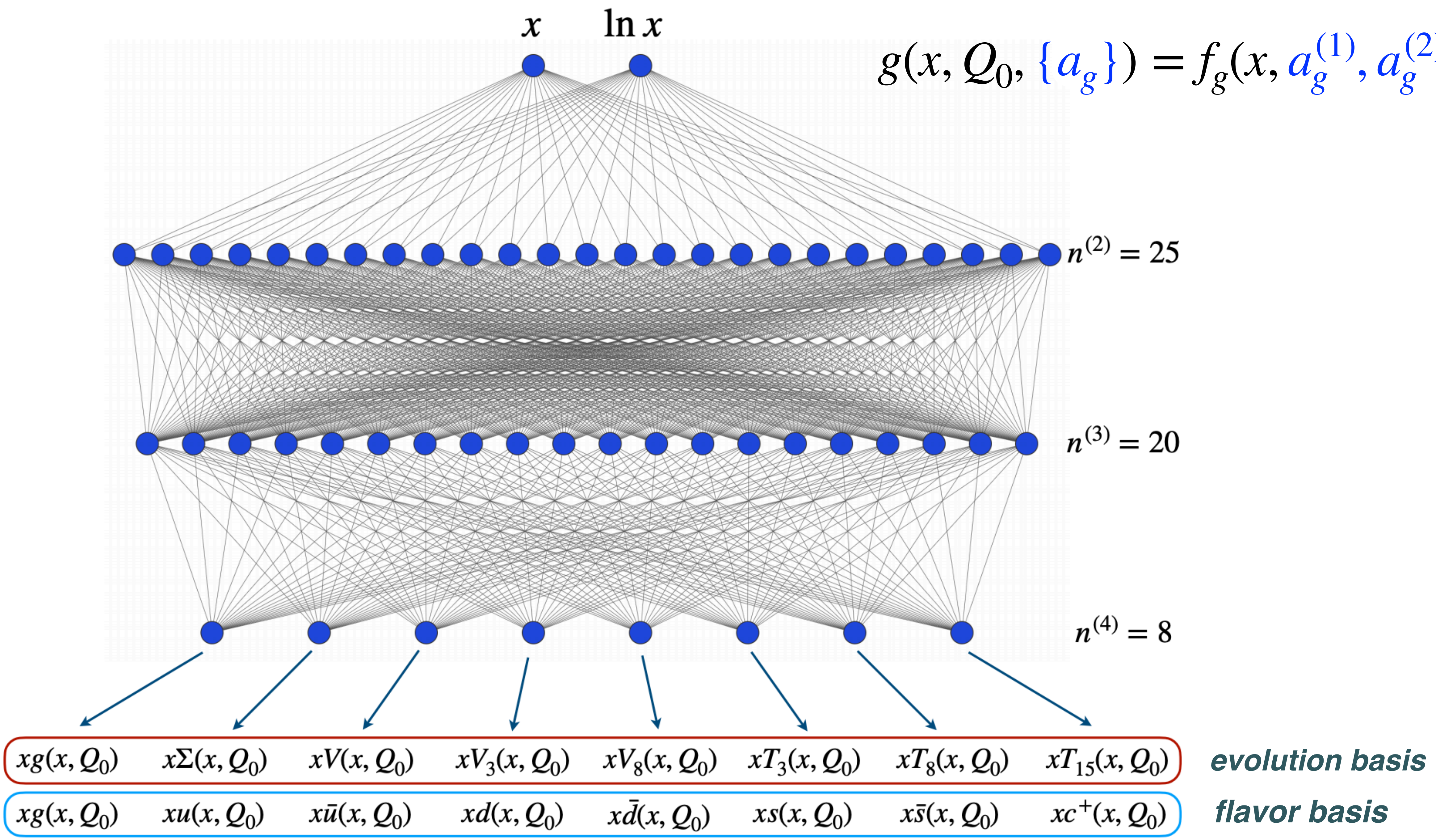


credit: *visualising the proton*, Arts at MIT (https://arts.mit.edu/visualizing-the-proton/)

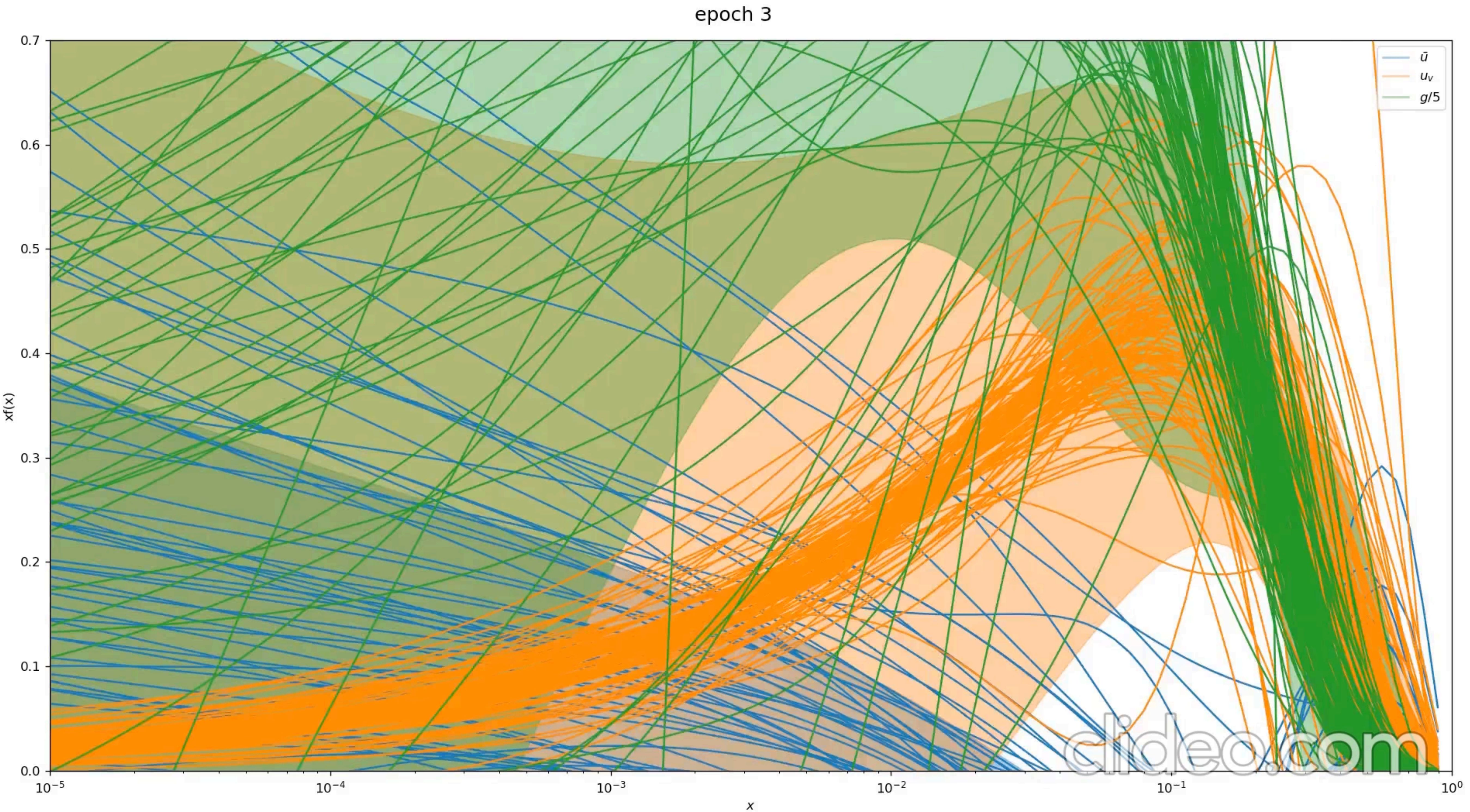
Machine Learning Proton Structure

- ☑ Model-independent PDF parametrisation with neural networks as **universal unbiased interpolants**
- ☑ **Stochastic Gradient Descent** via TensorFlow for neural network training
- ☑ Automated model **hyperparameter optimisation:** NN architecture, minimiser, learning rates …

$$g(x, Q_0, \{a_g\}) = f_g(x, a_g^{(1)}, a_g^{(2)}, \ldots)$$



Machine Learning PDFs



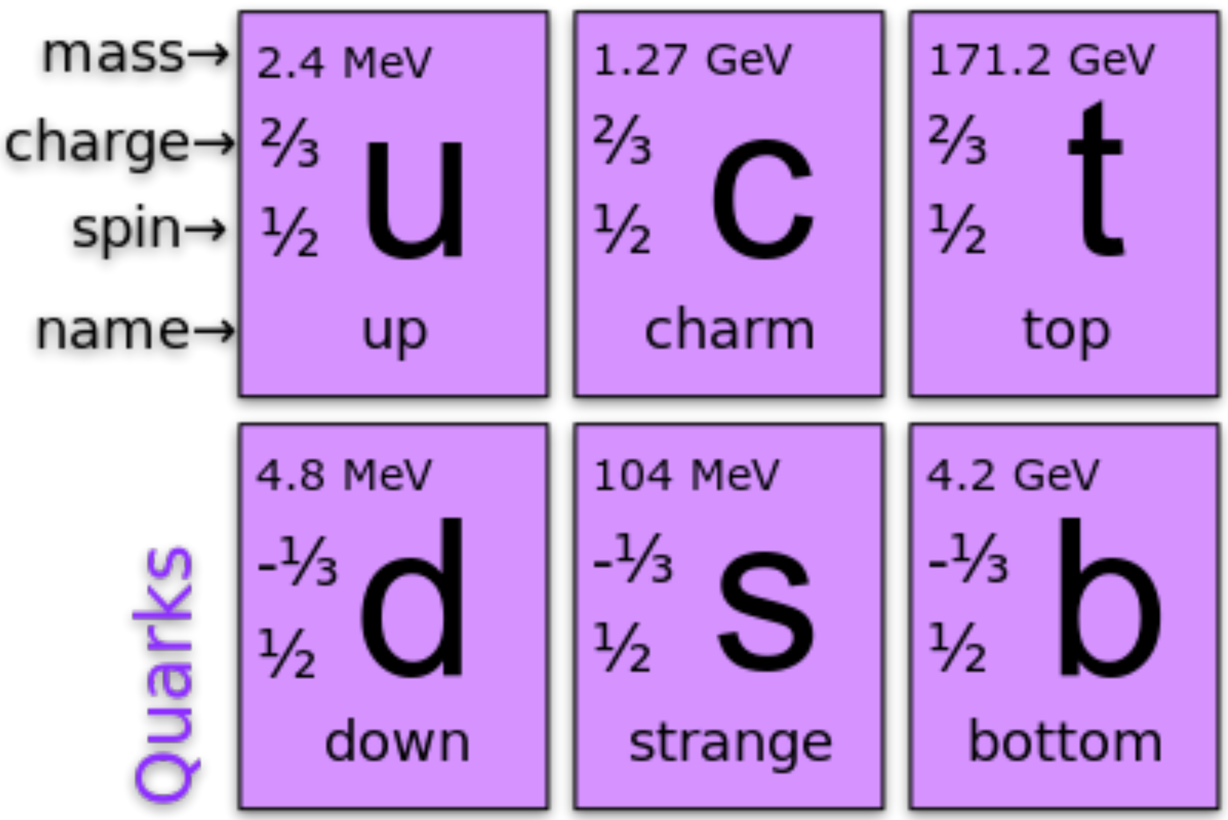
Error estimate based on **Monte Carlo replica method** (band: standard deviation over the MC replicas)

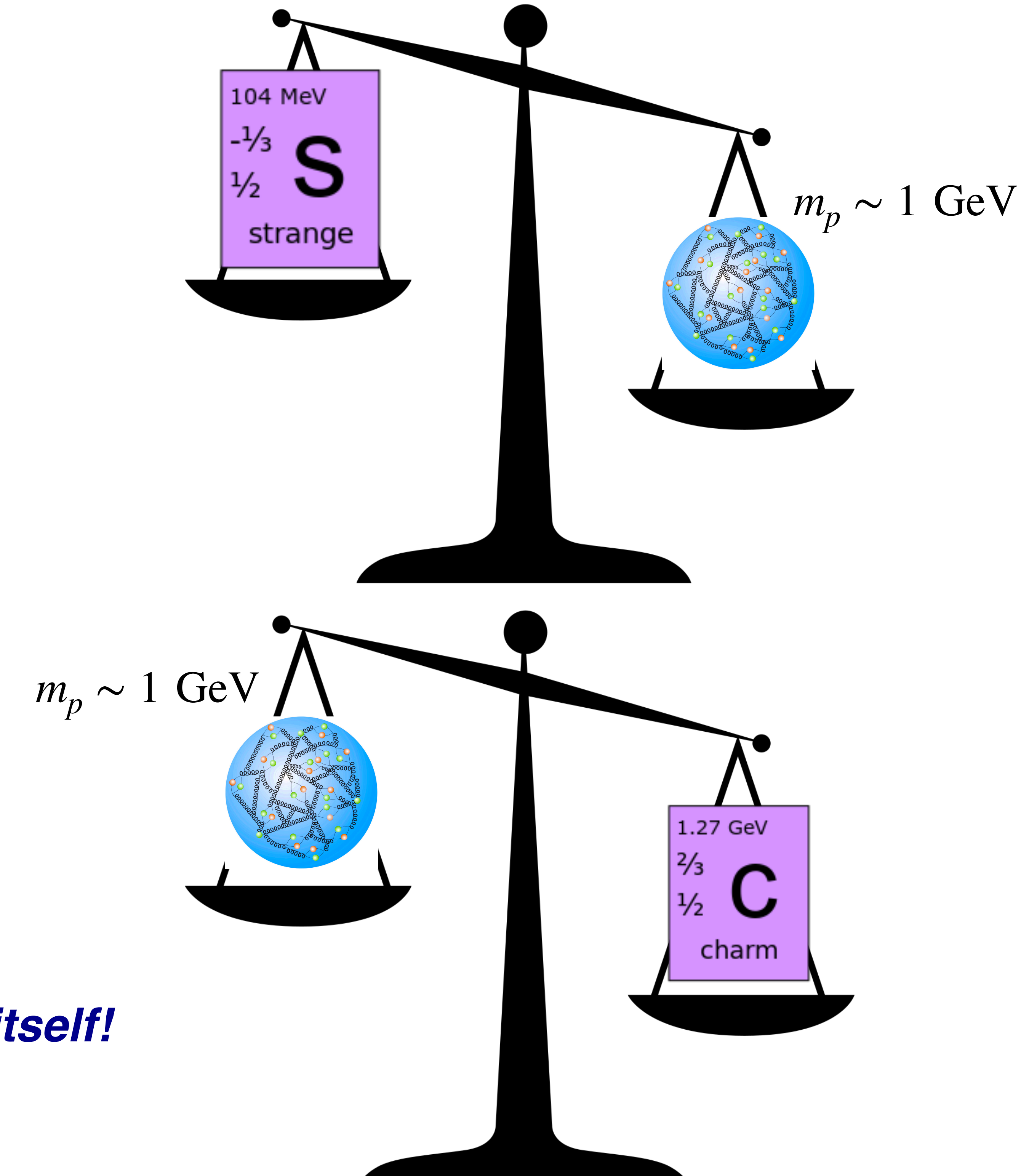
each curve is a separately trained neural network

The charm content of the proton

common assumption: the proton wave function does not contain charm quarks

the proton contains **intrinsic up, down, strange (anti-)quarks** but **no intrinsic charm quarks**





charm quarks heavier than the proton itself!

The charm content of the proton

common assumption: the proton wave function does not contain charm quarks

the proton contains **intrinsic up, down, strange (anti-)quarks** but **no intrinsic charm quarks**

It does not need to be so! An **intrinsic charm component** predicted in many models

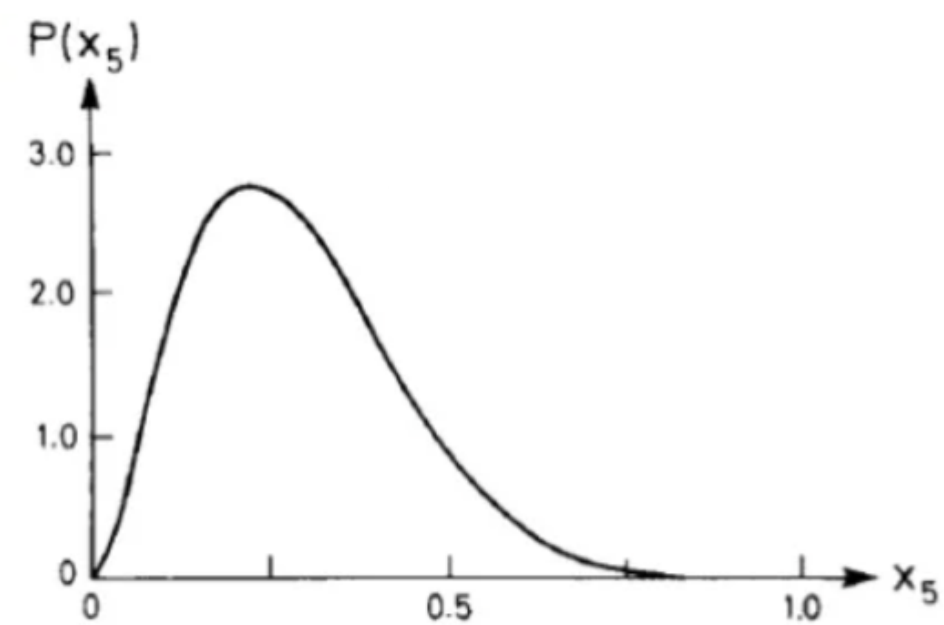
THE INTRINSIC CHARM OF THE PROTON

S.J. BRODSKY [1]
Stanford Linear Accelerator Center, Stanford, California 94305, USA

and

P. HOYER, C. PETERSON and N. SAKAI [2]
NORDITA, Copenhagen, Denmark

Received 22 April 1980

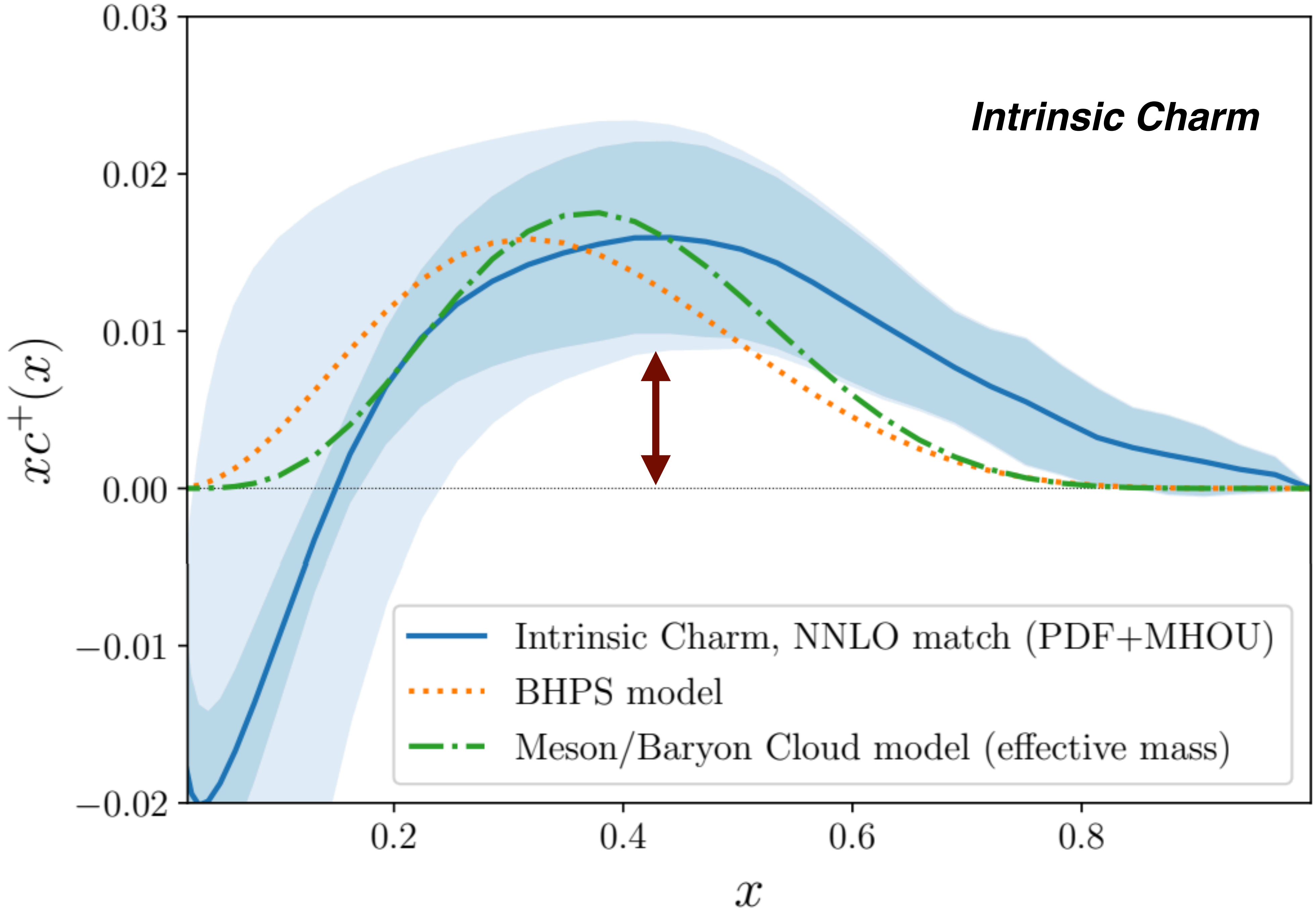


$$|p\rangle = \mathscr{P}_{3q}\,|uud\rangle + \mathscr{P}_{5q}\,|uudc\bar{c}\rangle + \ldots$$

Recent data give unexpectedly large cross-sections for charmed particle production at high x_F in hadron collisions. This may imply that the proton has a non-negligible uudc$\bar{\text{c}}$ Fock component. The interesting consequences of such a hypothesis are explored.

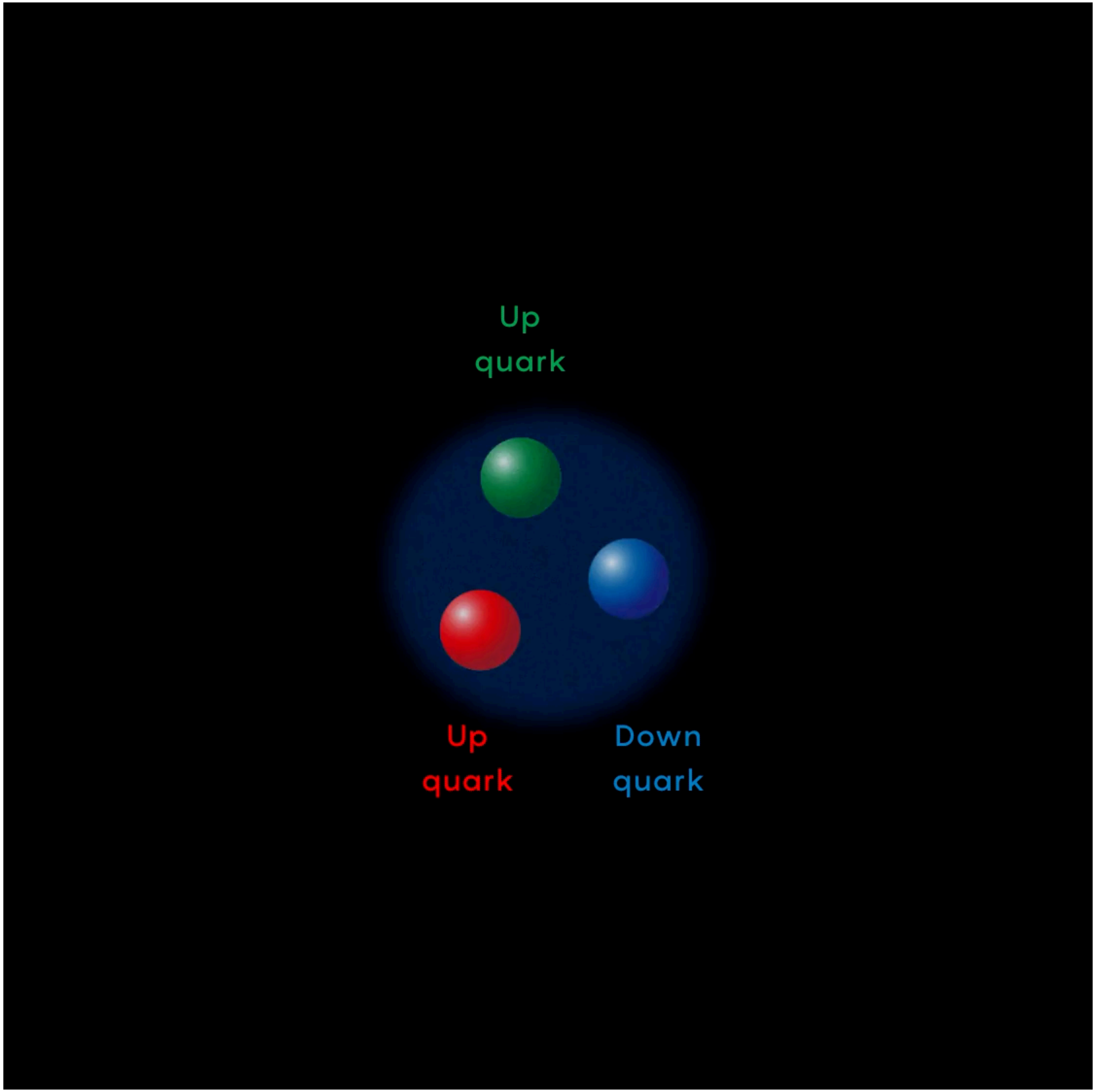
40 years of extensive searches for intrinsic charm: no unambiguous evidence ….

Intrinsic Charm in the Proton



The 3FNS charm PDF displays **non-zero component** peaked at large-*x* which can be identified with **intrinsic charm**

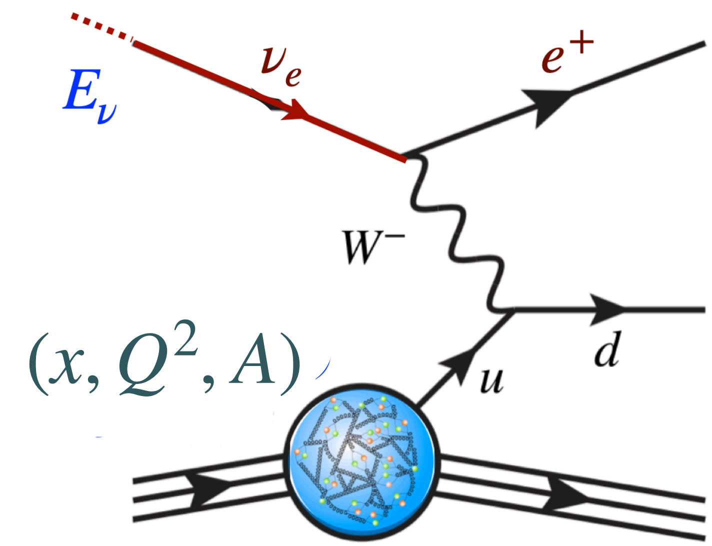
JR et al, Evidence for intrinsic charm quarks in the proton, *Nature* (2023)



credits: https://www.quantamagazine.org/inside-the-proton-the-most-complicated-thing-imaginable-20221019/

Beyond Proton PDFs

- The same approach can be extended to learn **more complex quantities** from the data
- **Neutrino structure functions** require three inputs and six outputs, each with different ranges, variations, asymptotic limits
- Input for the brand-new **collider neutrino program** started recently at CERN with the installation of dedicated detectors

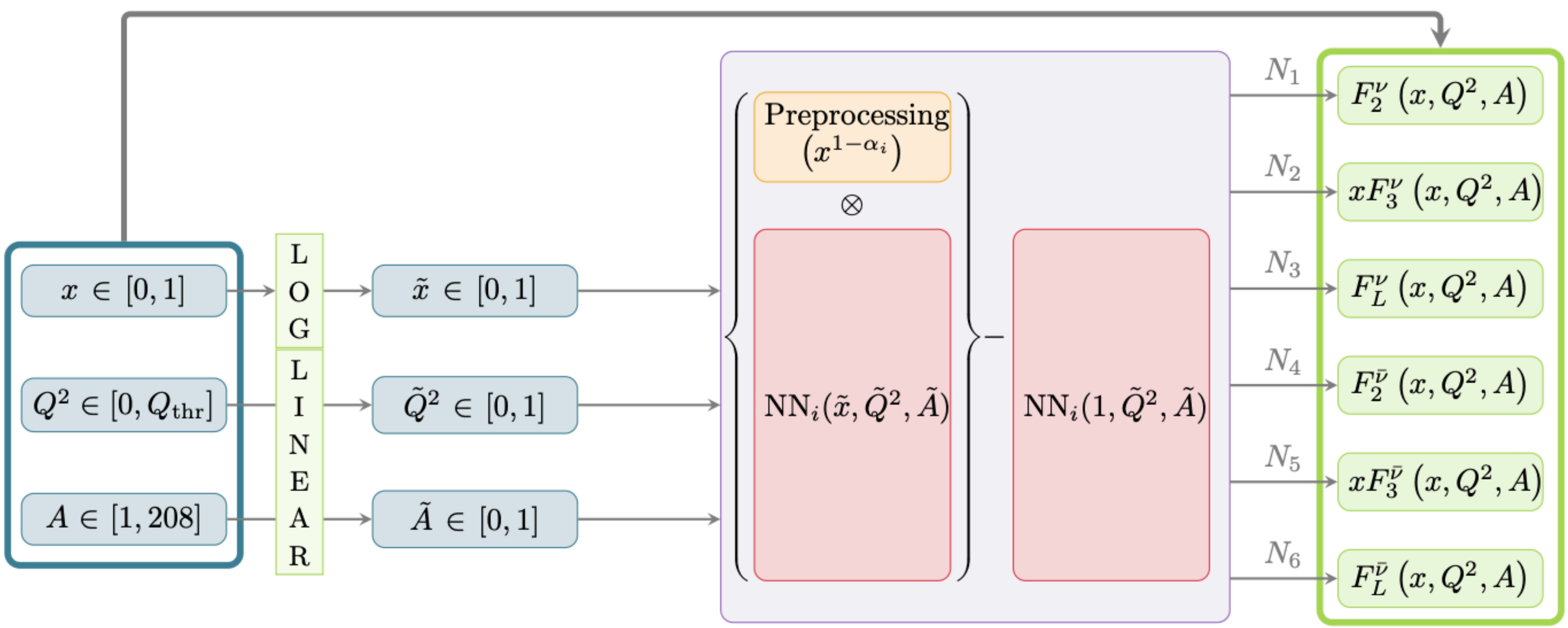


$$\sigma_{\nu p \to e^+ X}(E_\nu) = \tilde{\sigma}_{\nu u \to d} \otimes u(x, Q^2)$$

neutrino-proton scattering rate

partonic cross-section

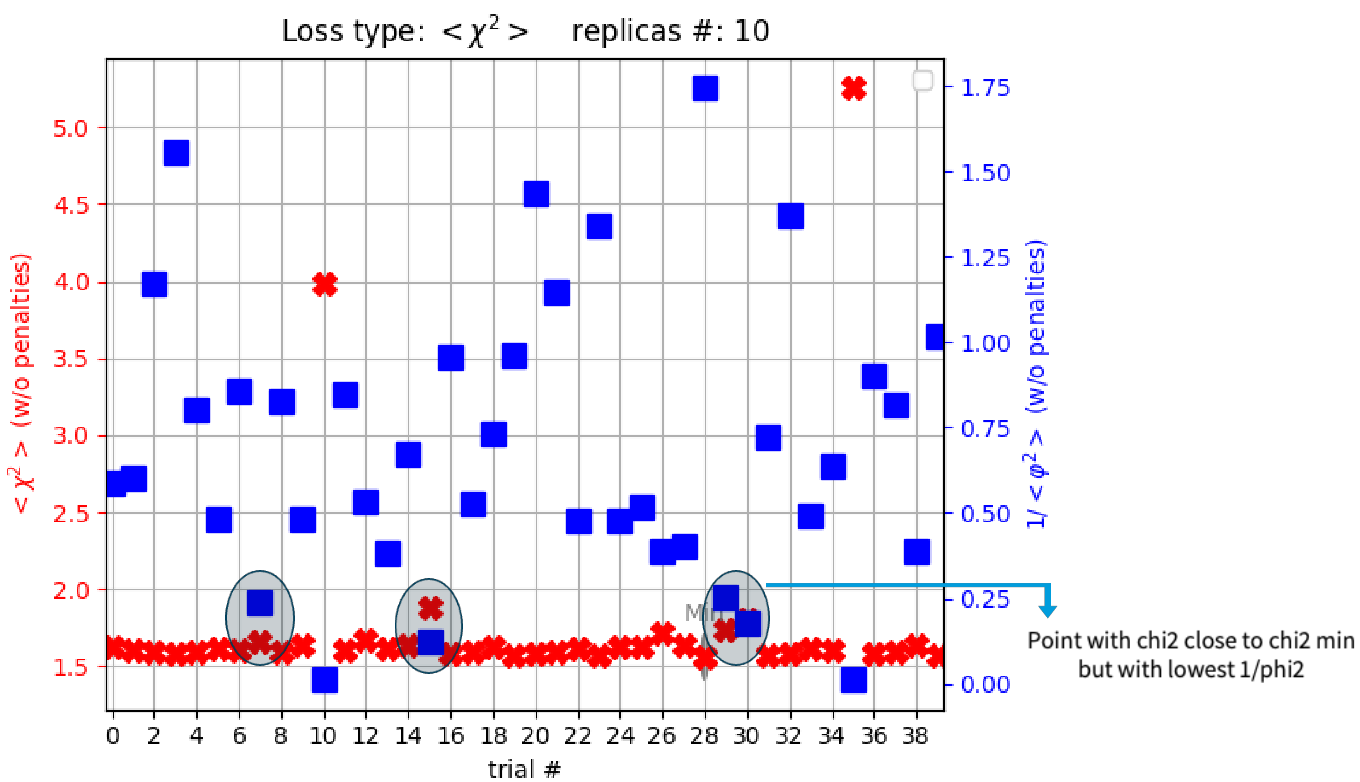
up-quark content int the proton



GPU & Hyperparameter Optimisation

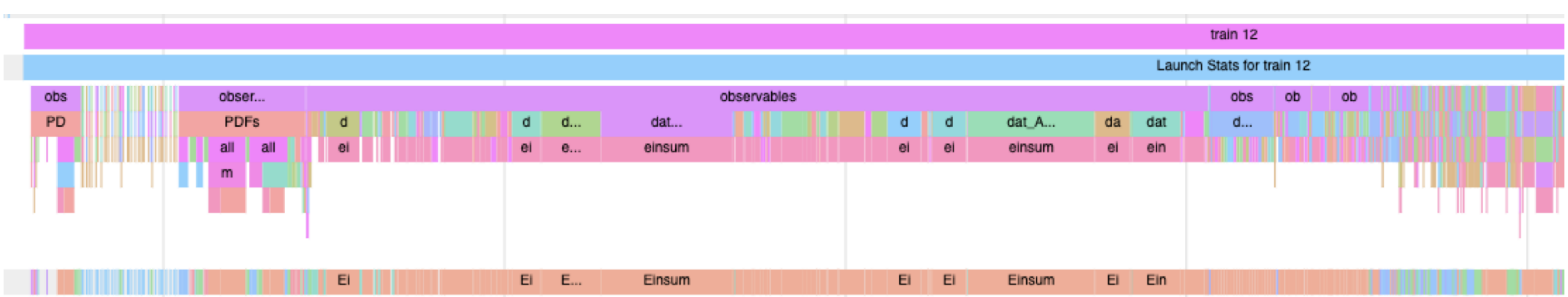
- Deploy NNPDF machinery on GPUs & optimise performance (**factor 200 improvement!**)
- Develop new strategies for **hyperparameter optimisation** based on the full posterior probability distribution, not only on first moment as most approaches

netherlands eScience center



Hyperoptimization with 10 replicas. Statistics: average over replicas and folds

branch	1 replica	100 replicas
master	96	860
replica-axis-first	96	505
multi-dense-layer	112	304
multi-dense + trvl	145	320
fk-refactor	120	90
merge-observable-splits	-	-
avoid-idle-gpu	90?	54



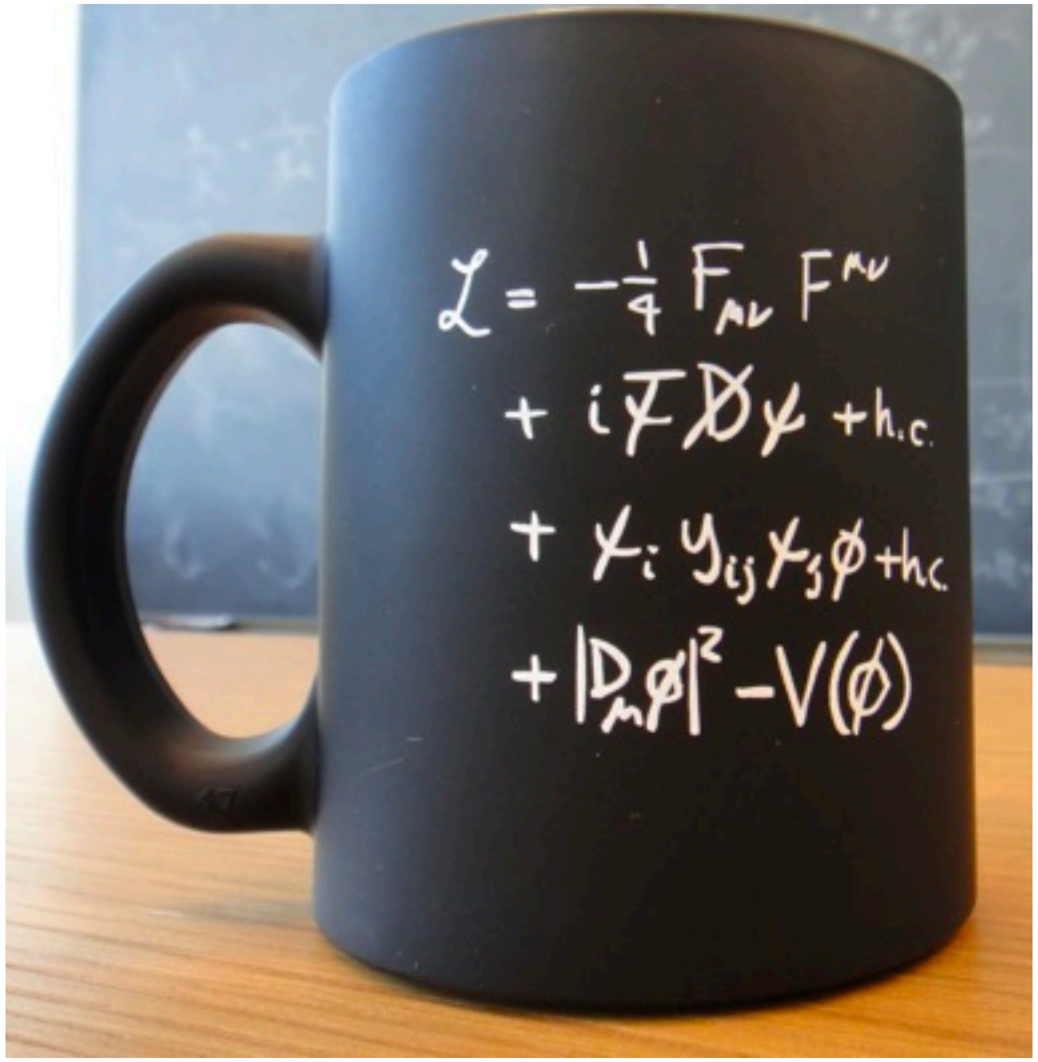
Optimal Observables for New Physics Searches with AI

R. Gomez-Ambrosio, J. ter Hoeve, M. Madigan, J. Rojo, V. Sanz, `*JHEP (2023)*

The Standard Model as an Effective Theory

The Standard Model EFT is defined by:

- **Particle (matter) content**: quarks and leptons
- **Gauge** (local) symmetries and their eventual breaking mechanisms
- **Lorentz** invariance and other global symmetries
- Linearly realised $SU(2)_L$ EW symmetry breaking
- **Validity only up to certain energy scale *Λ***



$$\mathcal{L}_{\mathrm{SMEFT}}(\{c_i\}, \Lambda) = \mathcal{L}_{\mathrm{SM}} + \sum_{d=5}^{\infty} \sum_{i=1}^{N_d} c_i^{(d)} \frac{\mathcal{O}_i^{(d)}}{\Lambda^{d-4}}$$

EFT coupling constants, to be determined from **data**

All possible operators of **mass-dimension *d*** consistent with above requirements

The Standard Model EFT

- The number of SMEFT operators is large: **59 non-redundant operators at dimension 6** for one fermion generation, **2499 operators** without any flavour assumption
- A global SMEFT analysis needs to explore a **huge complicated parameter space**

pure bosonic

X^3		$X^2\varphi^2$	
Q_G	$f^{ABC} G_\mu^{A\nu} G_\nu^{B\rho} G_\rho^{C\mu}$	$Q_{\varphi G}$	$\varphi^\dagger \varphi G^A_{\mu\nu} G^{A\mu\nu}$
$Q_{\widetilde{G}}$	$f^{ABC} \widetilde{G}_\mu^{A\nu} G_\nu^{B\rho} G_\rho^{C\mu}$	$Q_{\varphi B}$	$\varphi^\dagger \varphi B_{\mu\nu} B^{\mu\nu}$
Q_W	$\varepsilon^{IJK} W_\mu^{I\nu} W_\nu^{J\rho} W_\rho^{K\mu}$	$Q_{\varphi W}$	$\varphi^\dagger \varphi W^I_{\mu\nu} W^{I\mu\nu}$
$Q_{\widetilde{W}}$	$\varepsilon^{IJK} \widetilde{W}_\mu^{I\nu} W_\nu^{J\rho} W_\rho^{K\mu}$	$Q_{\varphi WB}$	$\varphi^\dagger \tau^I \varphi W^I_{\mu\nu} B^{\mu\nu}$
φ^6		$Q_{\varphi\widetilde{G}}$	$\varphi^\dagger \varphi \widetilde{G}^A_{\mu\nu} G^{A\mu\nu}$
Q_φ	$(\varphi^\dagger \varphi)^3$	$Q_{\varphi\widetilde{B}}$	$\varphi^\dagger \varphi \widetilde{B}_{\mu\nu} B^{\mu\nu}$
$\varphi^4 D^2$		$Q_{\varphi\widetilde{W}}$	$\varphi^\dagger \varphi \widetilde{W}^I_{\mu\nu} W^{I\mu\nu}$
$Q_{\varphi\Box}$	$(\varphi^\dagger \varphi)\Box(\varphi^\dagger \varphi)$	$Q_{\varphi\widetilde{W}B}$	$\varphi^\dagger \tau^I \varphi \widetilde{W}^I_{\mu\nu} B^{\mu\nu}$
$Q_{\varphi D}$	$(\varphi^\dagger D^\mu \varphi)^* (\varphi^\dagger D_\mu \varphi)$		

four-fermion operators

bosonic-fermionic

$\psi^2\varphi^3$		$\psi^2\varphi^2 D$	
$Q_{u\varphi}$	$(\varphi^\dagger \varphi)(\bar{q} u \widetilde{\varphi})$	$Q^{(1)}_{\varphi\ell}$	$(\varphi^\dagger i \overleftrightarrow{D}_\mu \varphi)(\bar{\ell}\gamma^\mu \ell)$
$Q_{d\varphi}$	$(\varphi^\dagger \varphi)(\bar{q} d \varphi)$	$Q^{(3)}_{\varphi\ell}$	$(\varphi^\dagger i \overleftrightarrow{D}^I_\mu \varphi)(\bar{\ell}\tau^I\gamma^\mu \ell)$
$Q_{e\varphi}$	$(\varphi^\dagger \varphi)(\bar{\ell} e \varphi)$	$Q_{\varphi e}$	$(\varphi^\dagger i \overleftrightarrow{D}_\mu \varphi)(\bar{e}\gamma^\mu e)$
$\psi^2 X\varphi$		$Q^{(1)}_{\varphi q}$	$(\varphi^\dagger i \overleftrightarrow{D}_\mu \varphi)(\bar{q}\gamma^\mu q)$
Q_{eW}	$(\bar{\ell}\sigma^{\mu\nu} e)\tau^I \varphi W^I_{\mu\nu}$	$Q^{(3)}_{\varphi q}$	$(\varphi^\dagger i \overleftrightarrow{D}^I_\mu \varphi)(\bar{q}\tau^I\gamma^\mu q)$
Q_{eB}	$(\bar{\ell}\sigma^{\mu\nu} e)\varphi B_{\mu\nu}$	$Q_{\varphi u}$	$(\varphi^\dagger i \overleftrightarrow{D}_\mu \varphi)(\bar{u}\gamma^\mu u)$
Q_{uG}	$(\bar{q}\sigma^{\mu\nu} T^A u)\widetilde{\varphi} G^A_{\mu\nu}$	$Q_{\varphi d}$	$(\varphi^\dagger i \overleftrightarrow{D}_\mu \varphi)(\bar{d}\gamma^\mu d)$
Q_{uW}	$(\bar{q}\sigma^{\mu\nu} u)\tau^I \widetilde{\varphi} W^I_{\mu\nu}$	$Q_{\varphi ud}$	$(\widetilde{\varphi}^\dagger i D_\mu \varphi)(\bar{u}\gamma^\mu d)$

$(\bar{L}L)(\bar{L}L)$		$(\bar{L}L)(\bar{R}R)$	
$Q_{\ell\ell}$	$(\bar{\ell}\gamma_\mu \ell)(\bar{\ell}\gamma^\mu \ell)$	$Q_{\ell e}$	$(\bar{\ell}\gamma_\mu \ell)(\bar{e}\gamma^\mu e)$
$Q^{(1)}_{qq}$	$(\bar{q}\gamma_\mu q)(\bar{q}\gamma^\mu q)$	$Q_{\ell u}$	$(\bar{\ell}\gamma_\mu \ell)(\bar{u}\gamma^\mu u)$
$Q^{(3)}_{qq}$	$(\bar{q}\gamma_\mu \tau^I q)(\bar{q}\gamma^\mu \tau^I q)$	$Q_{\ell d}$	$(\bar{\ell}\gamma_\mu \ell)(\bar{d}\gamma^\mu d)$
$Q^{(1)}_{\ell q}$	$(\bar{\ell}\gamma_\mu \ell)(\bar{q}\gamma^\mu q)$	Q_{qe}	$(\bar{q}\gamma_\mu q)(\bar{e}\gamma^\mu e)$
$Q^{(3)}_{\ell q}$	$(\bar{\ell}\gamma_\mu \tau^I \ell)(\bar{q}\gamma^\mu \tau^I q)$	$Q^{(1)}_{qu}$	$(\bar{q}\gamma_\mu q)(\bar{u}\gamma^\mu u)$

Statistically optimal observables for EFTs

Which kind of measurement is **most sensitive to SMEFT operators?**

Difficult question to answer in general since SMEFT-sensitive measurements can be:

- Inclusive or (1,2,3, …)-differential (in which specific variables?)
- Binned (choice of binning?) or unbinned
- Unfolded at parton level, at particle level, or at detector level

relevant to many other extractions of SM & BSM parameters from data

Our approach:

deploy **unbinned multivariate measurements** to determine the best sensitivity that a given process can have on SMEFT operators by means of **machine learning techniques**

Gaussian likelihood

$$\mathcal{L}(\boldsymbol{n};\boldsymbol{\nu}(\boldsymbol{c})) = \prod_{i=1}^{N_b} \exp\left[-\frac{1}{2}\frac{(\underbrace{n_i}_{\text{observed event counts}} - \underbrace{\nu_i(\boldsymbol{c})}_{\text{predicted event counts}})^2}{\nu_i(\boldsymbol{c})}\right]$$

retains full information on event-by-event kinematics

Unbinned multivariate likelihood

$$\mathcal{L}(\boldsymbol{c}) = \frac{\nu_{\text{tot}}(\boldsymbol{c})^{N_{\text{ev}}}}{N_{\text{ev}}!} e^{-\nu_{\text{tot}}(\boldsymbol{c})} \underbrace{\prod_{i=1}^{N_{\text{ev}}}}_{\text{sum over events}} f_\sigma(\underbrace{\boldsymbol{x}_i}_{\text{event kinematics}}, \boldsymbol{c}) \qquad \underbrace{f_\sigma(\boldsymbol{x}, \boldsymbol{c})}_{\text{event probability}} = \frac{1}{\sigma_{\text{fid}}(\boldsymbol{c})} \frac{d\sigma(\boldsymbol{x}, \boldsymbol{c})}{d\boldsymbol{x}} \; ,$$

Statistically optimal observables for EFTs

Which kind of measurement is **most sensitive to SMEFT operators?**

Difficult question to answer in general since SMEFT-sensitive measurements can be:

- Inclusive or (1,2,3, …)-differential (in which specific variables?)
- Binned (choice of binning?) or unbinned
- Unfolded at parton level, at particle level, or at detector level

relevant to many other extractions of SM & BSM parameters from data

Our approach:

deploy **unbinned multivariate measurements** to determine the best sensitivity that a given process can have on SMEFT operators by means of **machine learning techniques**

Challenges:

- Parameter inference requires knowledge of the likelihood for **any value of the EFT coefficients**
- Evaluation of likelihood functions **computationally costly due to high dimensionality** both of the space of kinematic features $\boldsymbol{x}$ and of EFT parameters $\boldsymbol{c}$

Solution:

- Neural networks as **universal unbiased interpolants** to parametrise high-dimensional likelihoods

Statistically optimal observables from ML

the dependence of the cross-section on **kinematic variables and all EFT coefficients**

$$r_\sigma(\boldsymbol{x}, \boldsymbol{c}) \equiv \frac{f_\sigma(\boldsymbol{x}, \boldsymbol{c})}{f_\sigma(\boldsymbol{x}, \boldsymbol{0})} = 1 + \sum_{j=1}^{n_{\rm eft}} r_\sigma^{(j)}(\boldsymbol{x}) c_j + \sum_{j=1}^{n_{\rm eft}} \sum_{k\geq j}^{n_{\rm eft}} r_\sigma^{(j,k)}(\boldsymbol{x}) c_j c_k$$

parametrised with **neural networks** trained to Monte Carlo simulations & benchmarked with exact calculations

$$\hat{r}_\sigma(\boldsymbol{x}, \boldsymbol{c}) = 1 + \sum_{j=1}^{n_{\rm eft}} \mathrm{NN}^{(j)}(\boldsymbol{x}) c_j + \sum_{j=1}^{n_{\rm eft}} \sum_{k\geq j}^{n_{\rm eft}} \mathrm{NN}^{(j,k)}(\boldsymbol{x}) c_j c_k$$

extendable to **arbitrary number** of kinematic variables and EFT coefficients: training can be parallelised

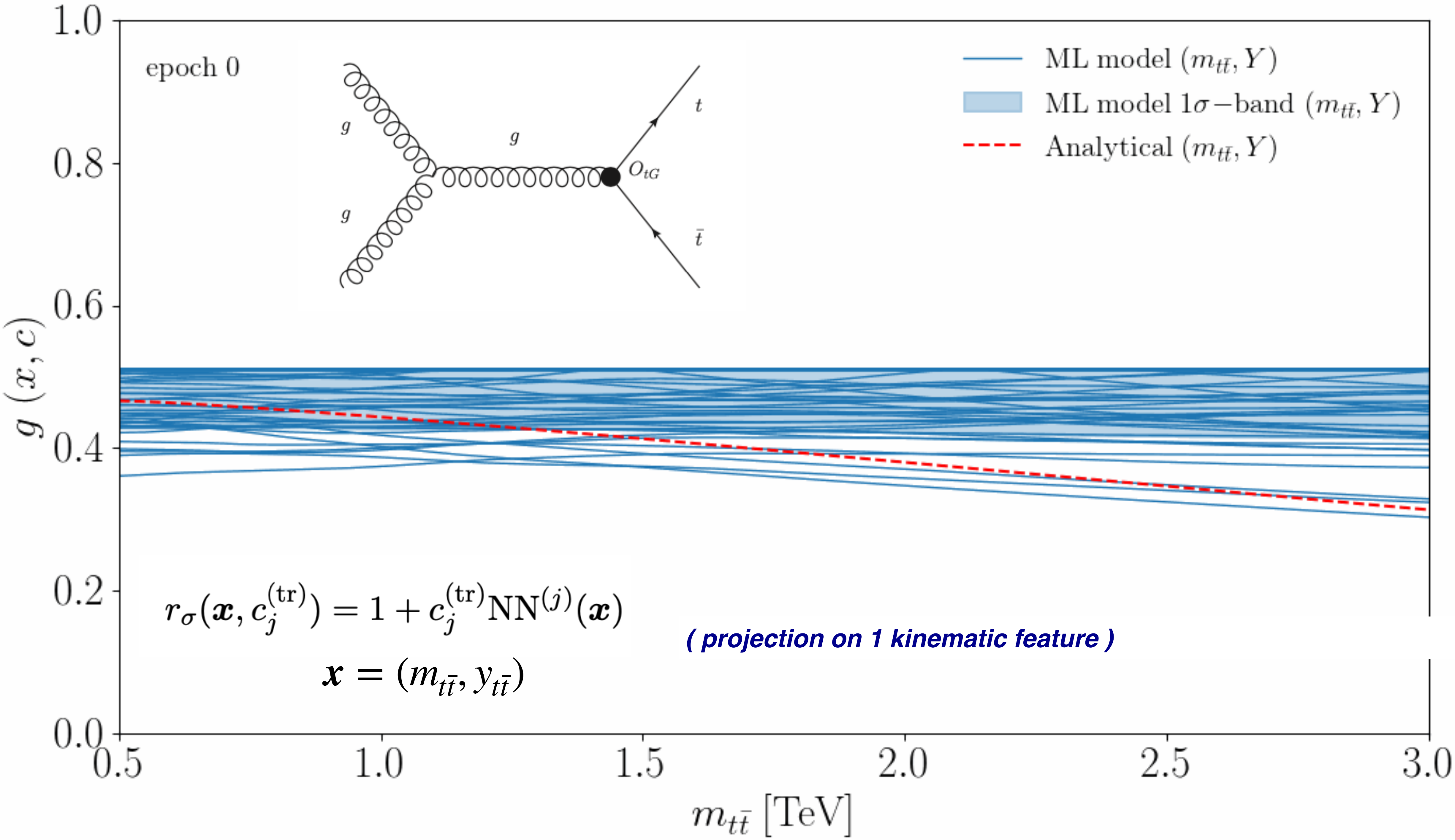
methodological uncertainties (e.g. finite training samples) assess with the **replica method**

$$\hat{r}_\sigma^{(i)}(\boldsymbol{x}, \boldsymbol{c}) \equiv 1 + \sum_{j=1}^{n_{\rm eft}} \mathrm{NN}_i^{(j)}(\boldsymbol{x}) c_j + \sum_{j=1}^{n_{\rm eft}} \sum_{k\geq j}^{n_{\rm eft}} \mathrm{NN}_i^{(j,k)}(\boldsymbol{x}) c_j c_k \, , \qquad i = 1, \ldots, N_{\rm rep}$$

each replica trained to an independent set of MC events

representation of the probability distribution in the space of ML models

Neural network training



$$r_\sigma(\boldsymbol{x}, c_j^{(\mathrm{tr})}) = 1 + c_j^{(\mathrm{tr})}\mathrm{NN}^{(j)}(\boldsymbol{x})$$

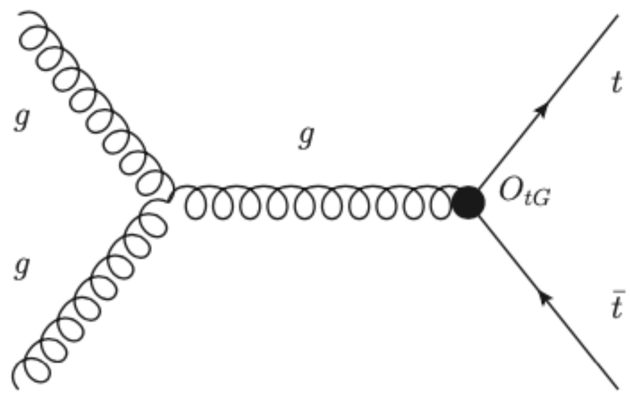
$$\boldsymbol{x} = (m_{t\bar{t}}, y_{t\bar{t}})$$

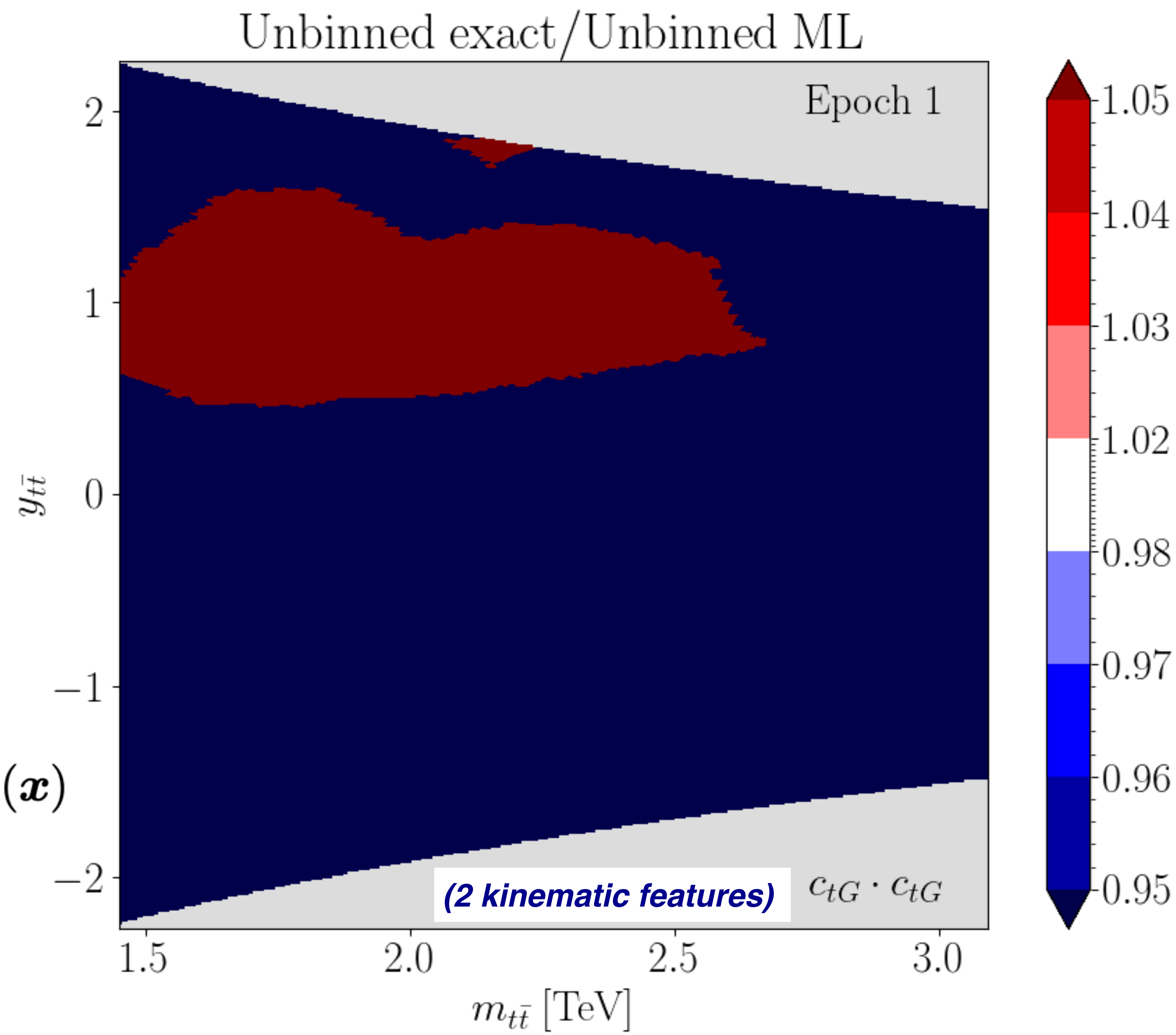
(projection on 1 kinematic feature)

NN training by minimising cross-entropy loss function

$$L[g(\boldsymbol{x}, \boldsymbol{c})] = -\sigma_{\mathrm{fid}}(\boldsymbol{c}) \sum_{i=1}^{N_{\mathrm{ev}}} \log(1 - g(\boldsymbol{x}_i, \boldsymbol{c})) - \sigma_{\mathrm{fid}}(\boldsymbol{0}) \sum_{j=1}^{N_{\mathrm{ev}}} \log g(\boldsymbol{x}_j, \boldsymbol{c}) \qquad g = (1 + r_\sigma)^{-1}$$

Neural network training





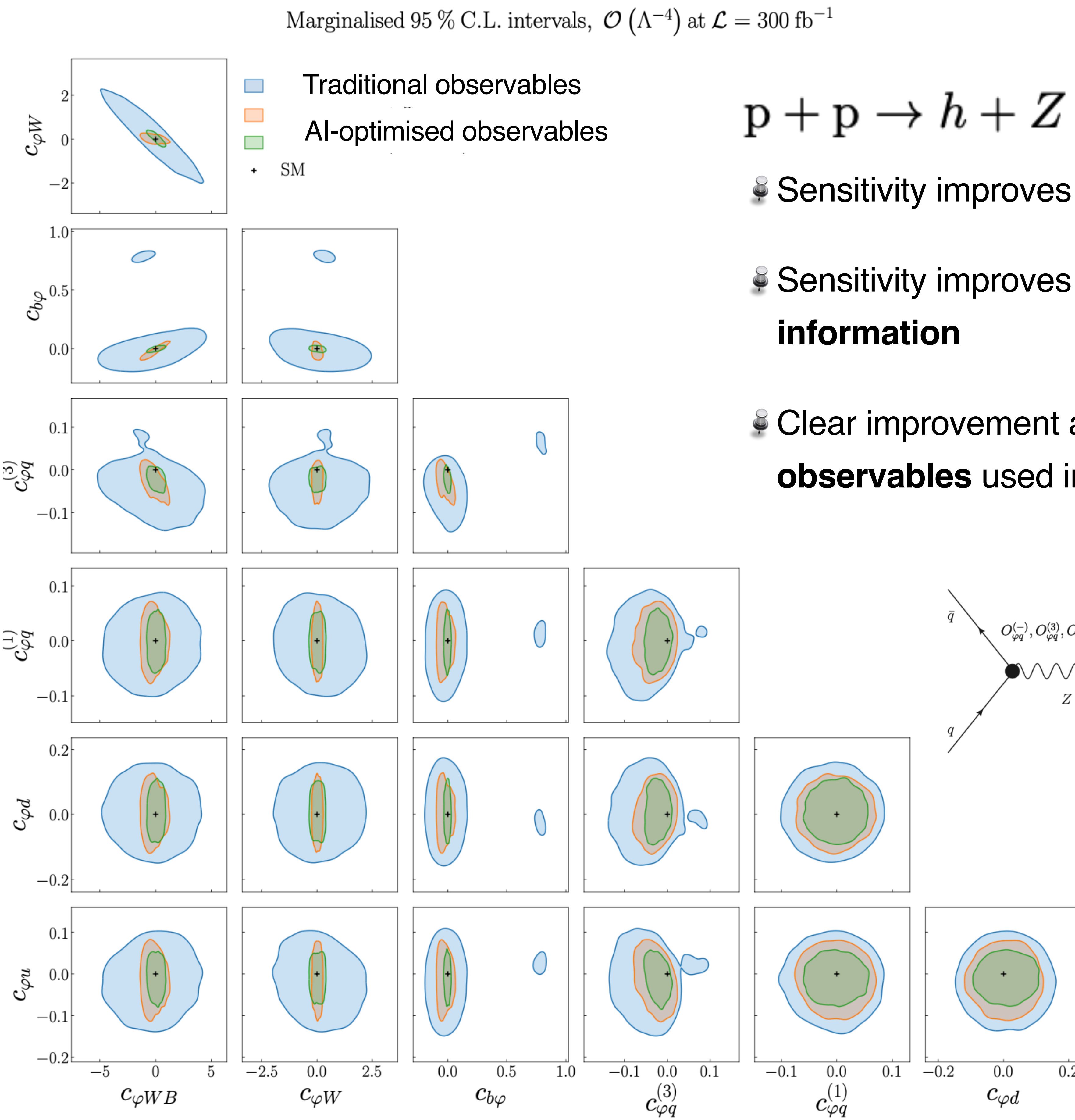
$$r_\sigma(\boldsymbol{x}, c_j^{(\mathrm{tr})}, c_k^{(\mathrm{tr})}) = 1 + c_j^{(\mathrm{tr})} c_k^{(\mathrm{tr})} \mathrm{NN}^{(j,k)}(\boldsymbol{x})$$

$$\boldsymbol{x} = (m_{t\bar{t}}, y_{t\bar{t}})$$

NN training by minimising cross-entropy loss function

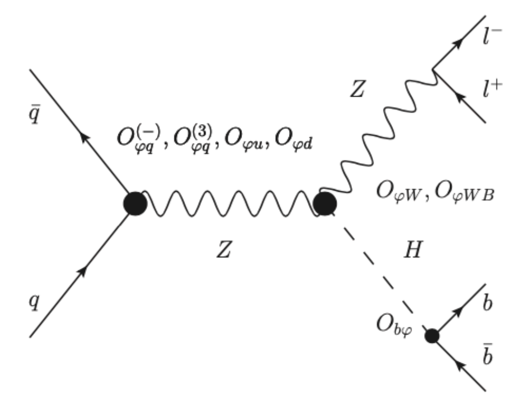
$$L[g(\boldsymbol{x}, \boldsymbol{c})] = -\sigma_{\mathrm{fid}}(\boldsymbol{c}) \sum_{i=1}^{N_{\mathrm{ev}}} \log(1 - g(\boldsymbol{x}_i, \boldsymbol{c})) - \sigma_{\mathrm{fid}}(\boldsymbol{0}) \sum_{j=1}^{N_{\mathrm{ev}}} \log g(\boldsymbol{x}_j, \boldsymbol{c}) \qquad g = (1 + r_\sigma)^{-1}$$

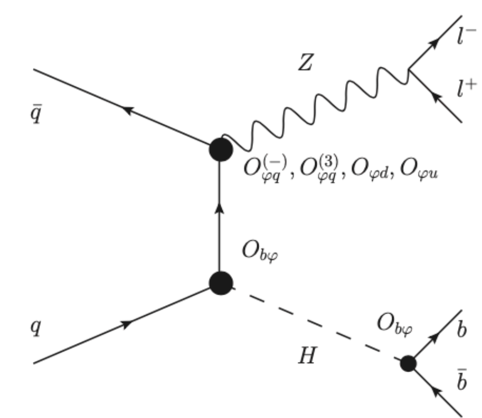
Results: Higgs+Z production



$$\mathrm{p} + \mathrm{p} \to h + Z \to b + \bar{b} + \ell^+ + \ell^-$$

- Sensitivity improves in **unbinned analysis**
- Sensitivity improves when **using all kinematic information**
- Clear improvement as compared to **traditional observables** used in EFT fits



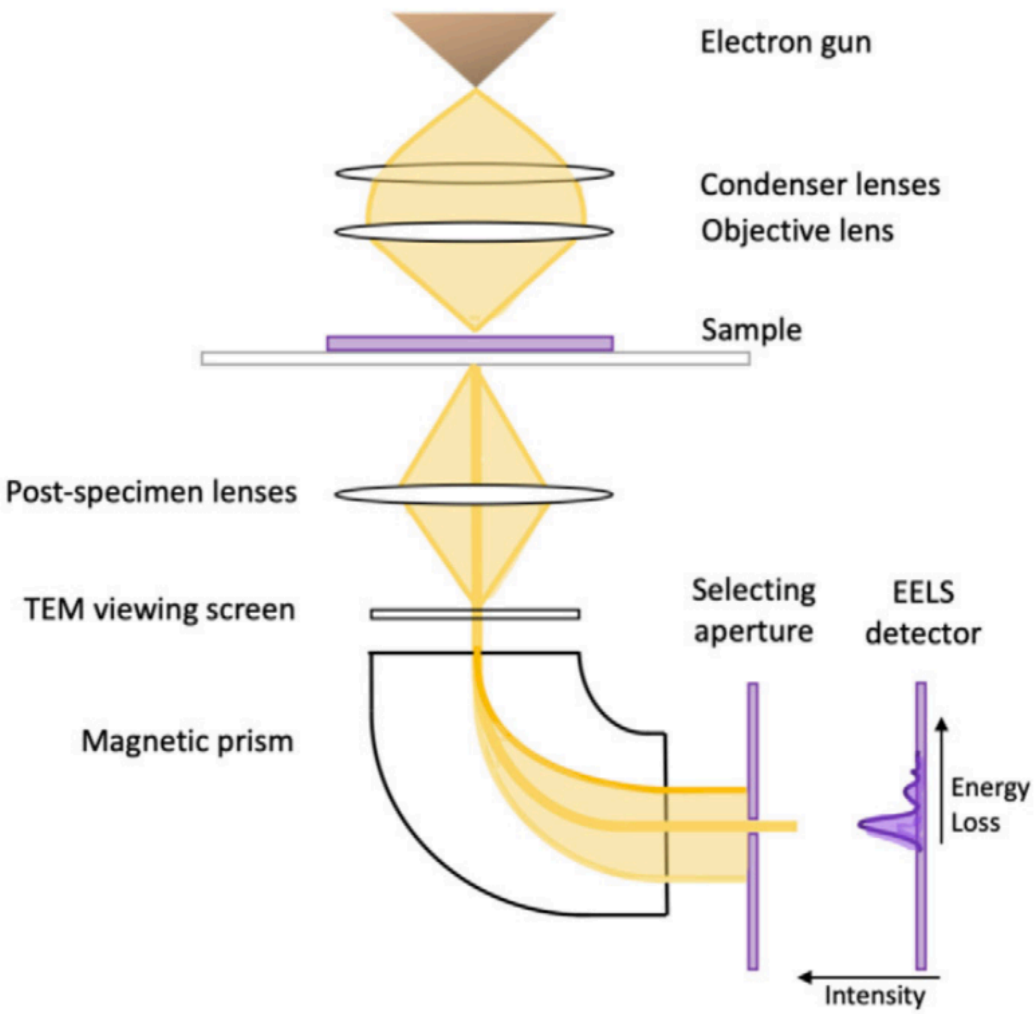


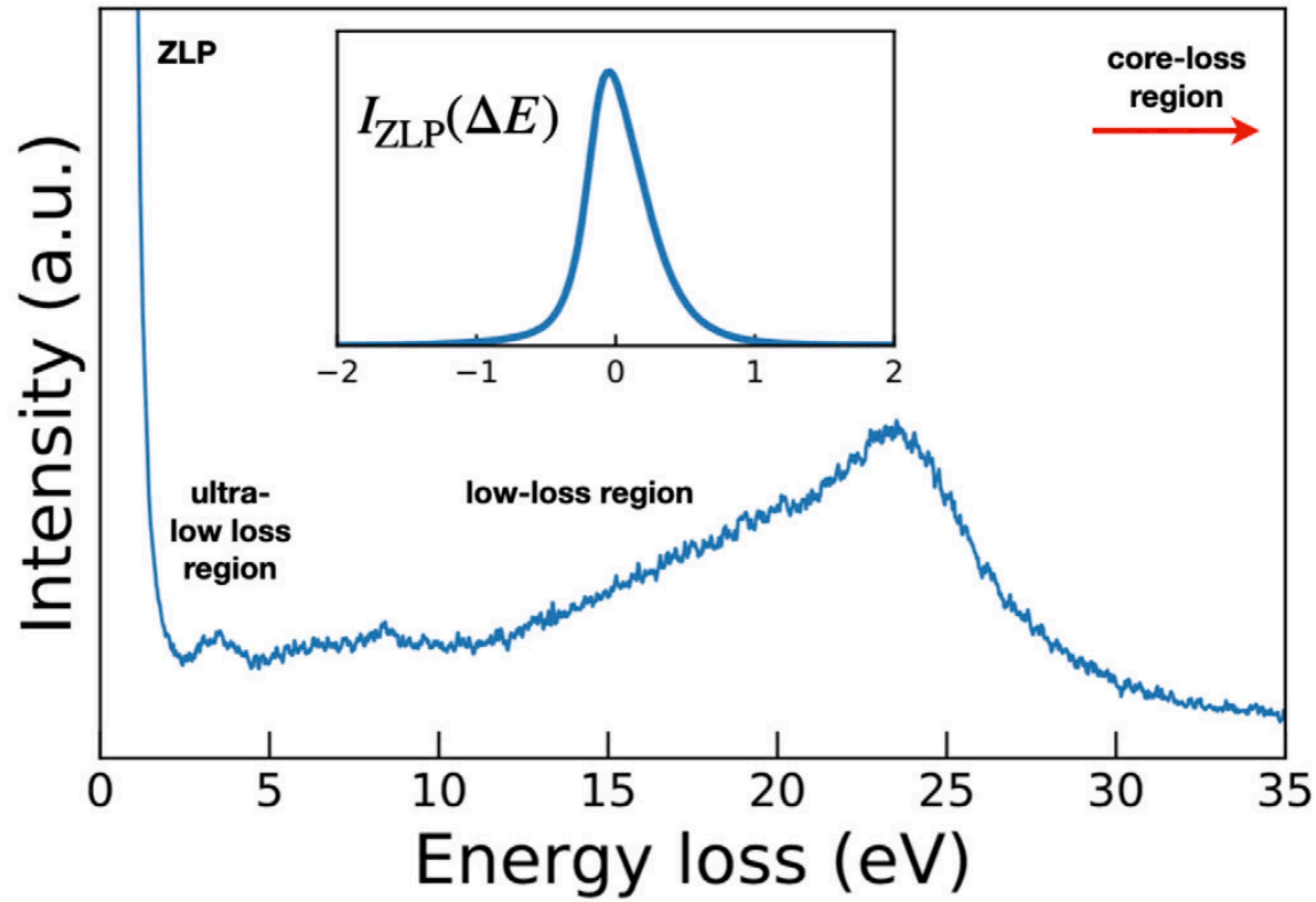
Quantum Materials under the AI-Boosted Microscope

- Roest, van Heijst, Maduro, **JR**, Conesa-Boj, Charting the low-loss region in electron energy loss spectroscopy with machine learning *Ultramicroscopy* (2021)
- van Heijst, Mukai, Okunishi, Hashiguchi, Maduro, Roest, **JR**, Conesa-Boj, Illuminating the Electronic Properties of WS_2 Polytypism with Electron Microscopy*, Annalen der Physiek* (2021)
- Brokkelkamp, ter Hoeve, Brokkelkamp van Heijst, ter Hoeve, Maduro, Davydof, Kryluyk, **JR**, Conesa-Boj, Spatially resolved band gap and dielectric function in two-dimensional materials from electron energy loss spectroscopy*, Journal of Physical Chemistry A* (2022)
- Stijn van der Lippe, Abel Brokkelkamp, **JR**, Sonia Conesa-Boj, Localized Exciton Anatomy and BandGap Energy Modulation in 1D MoS_2 Nanostructures, *Advanced Functional Materials* (2023)
- La, Brokkelkamp, van der Lippe, Ter Hoeve, **JR**, Conesa-Boj, Edge-induced excitations in Bi2Te3 from spatially-resolved electron energy-gain spectroscopy, *Ultramicroscopy* (2023).

Electron Energy Loss Spectroscopy

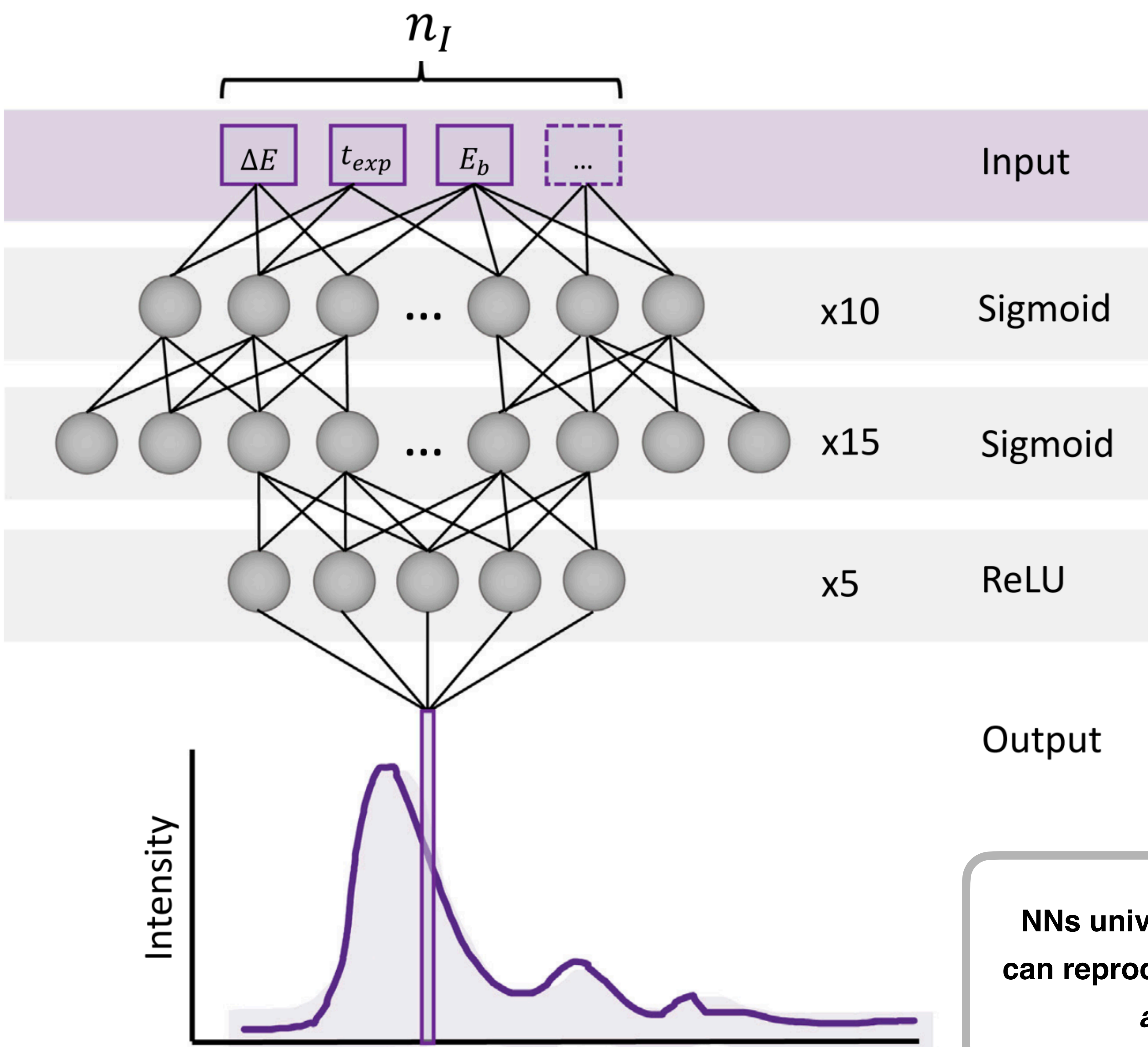
EELS: monitor **energy losses** suffered by the electrons from a Transmission Electron Microscope (TEM) beam upon **interaction with the sample**





- **Challenge**: EELS measurements affected by **huge background** (Zero-Loss Peak) at low-energy losses from elastic scatterings: complicates interpretation of material properties!
- **Solution:** parametrise backgrounds from data using **Deep Neural Networks** and **Monte Carlo sampling** to remove them in a model-independent manner

A ML model for EELS backgrounds



energy loss, beam energy, exposure time, …

hidden layers

intensity of the Zero-Loss Peak background

NNs universal unbiased interpolants: can reproduce any physical law without *a priori* assumptions

The Monte Carlo replica method

- Generate Monte Carlo **replicas of the original data points** with multi-Gaussian distribution with central values and covariance matrices taken from the input measurements

$$I^{(\mathrm{art})(k)}_{\mathrm{ZLP},i} = I^{(\mathrm{exp})}_{\mathrm{ZLP},i} + r_i^{(\mathrm{stat},k)}\sigma_i^{(\mathrm{stat})} + \sum_{j=1}^{n_{\mathrm{sys}}} r_{i,j}^{(\mathrm{sys},k)}\sigma_{i,j}^{(\mathrm{sys})}\,, \quad \forall i\,, \quad k=1,\ldots,N_{\mathrm{rep}}\,,$$

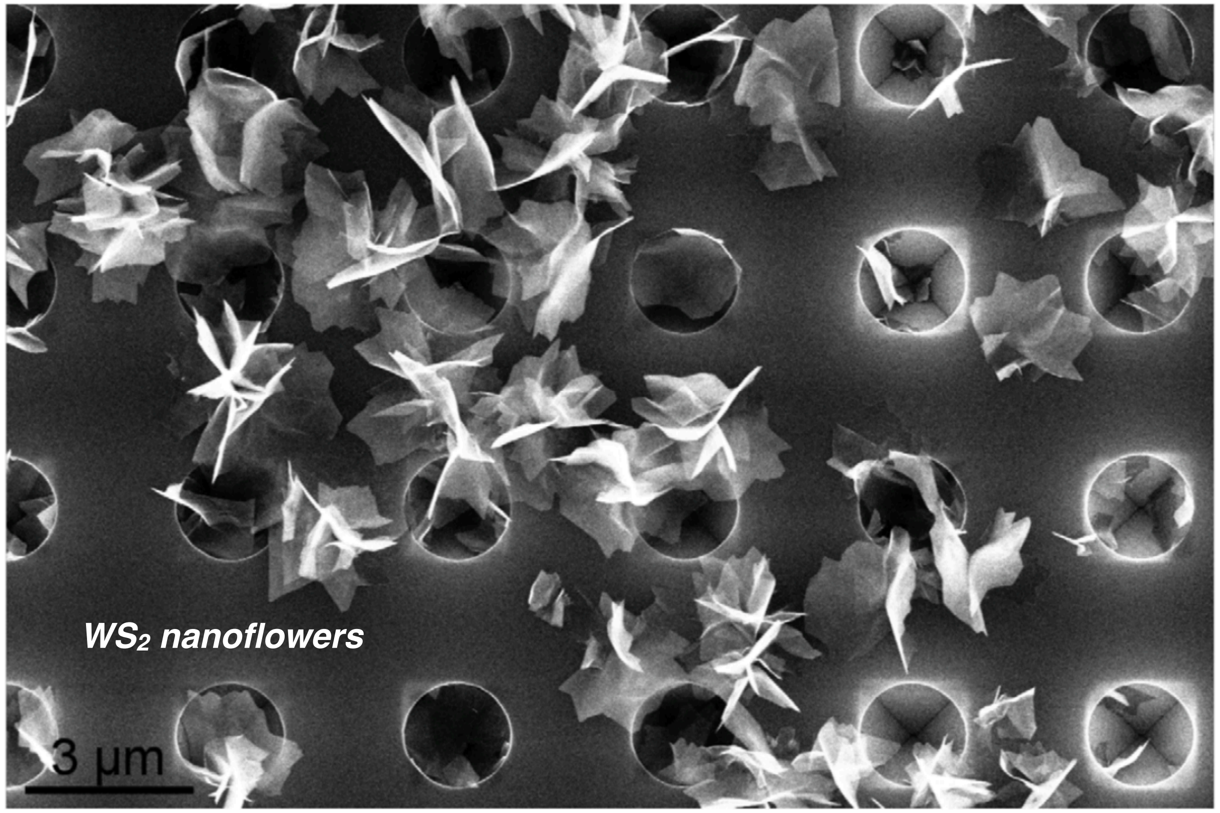
- Train a NN model on each replica from the **minimisation of the log-likelihood**

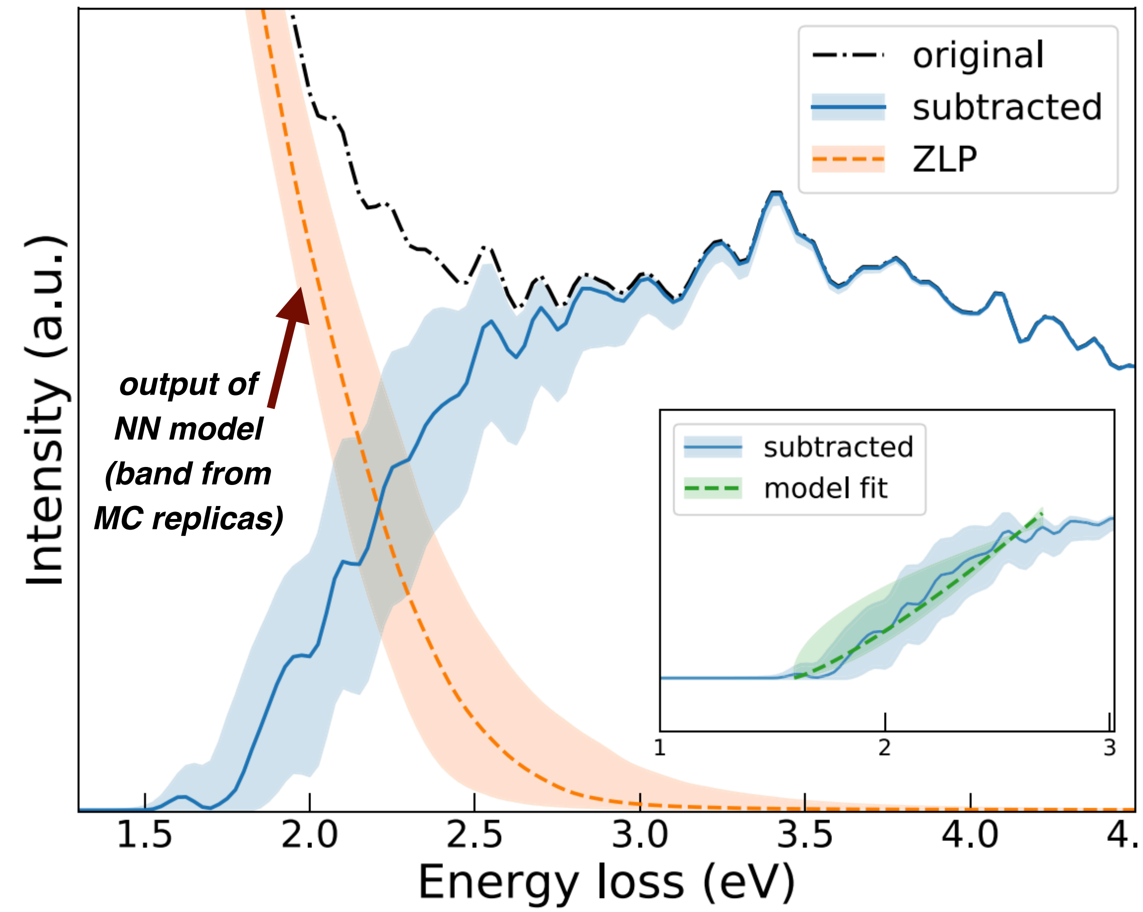
$$E^{(k)}\left(\{\theta^{(k)}\}\right) = \frac{1}{n_{\mathrm{dat}}}\sum_{i=1}^{n_{\mathrm{dat}}}\left(\frac{I^{(\mathrm{art})(k)}_{\mathrm{ZLP},i} - I^{(\mathrm{mod})}_{\mathrm{ZLP},i}\left(\{\theta^{(k)}\}\right)}{\sigma_i^{(\mathrm{exp})}}\right)^2\,,$$

- We end up with a sampling of the **probability density in the space of NN models,** from which we can compute e.g. the variance of the predicted ZLP intensity for arbitrary inputs

$$\sigma^{(\mathrm{mod})}_{I_{\mathrm{ZLP}}}(\{z_1\}) = \left(\frac{1}{N_{\mathrm{rep}}-1}\sum_{k=1}^{N_{\mathrm{rep}}}\left(I^{(\mathrm{mod})(\mathrm{k})}_{\mathrm{ZLP}} - \left\langle I^{(\mathrm{mod})}_{\mathrm{ZLP}}\right\rangle\right)\right)^{1/2}$$

Band gap extraction in polytypic WS_2





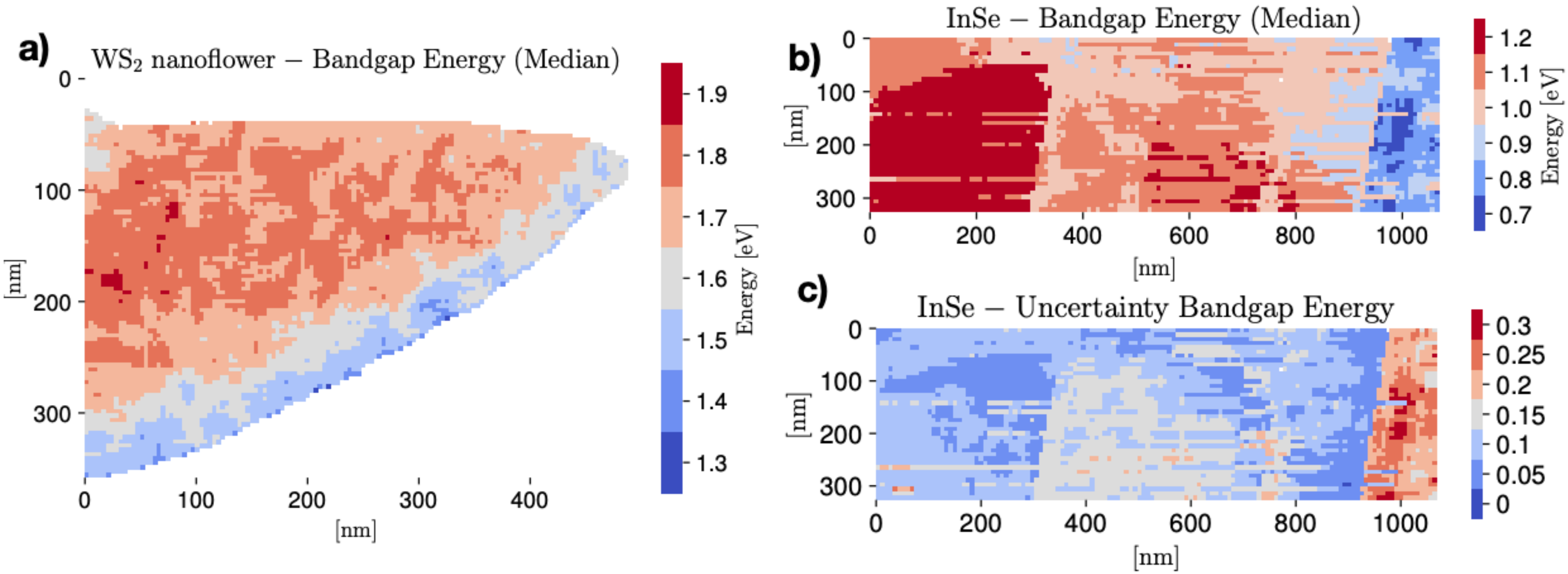
- Apply to **nanoflowers** composed by 2H/3R **polytypic WS_2**
- First extraction of band gap in this material from fit to subtracted EEL spectra

$$I_{\rm inel}(\Delta E) \simeq A\,(\Delta E - E_{\rm BG})^b \qquad E_{\rm BG} = 1.6^{+0.3}_{-0.2}\,{\rm eV}\,, \quad b = 1.3^{+0.3}_{-0.7}\,.$$

- ML-subtracted spectra make possible mapping **exciton transitions** down to 1.5 eV

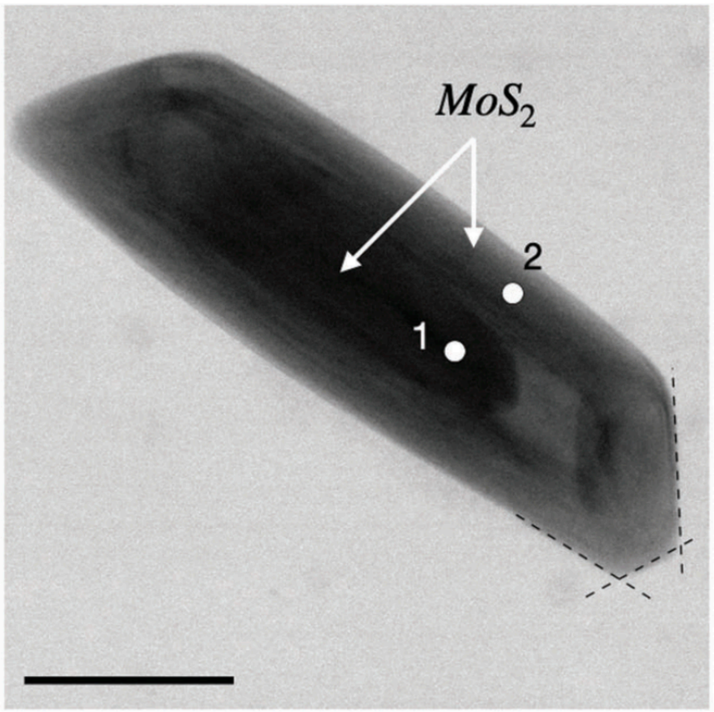
consistent with ab-initio DFT calculations

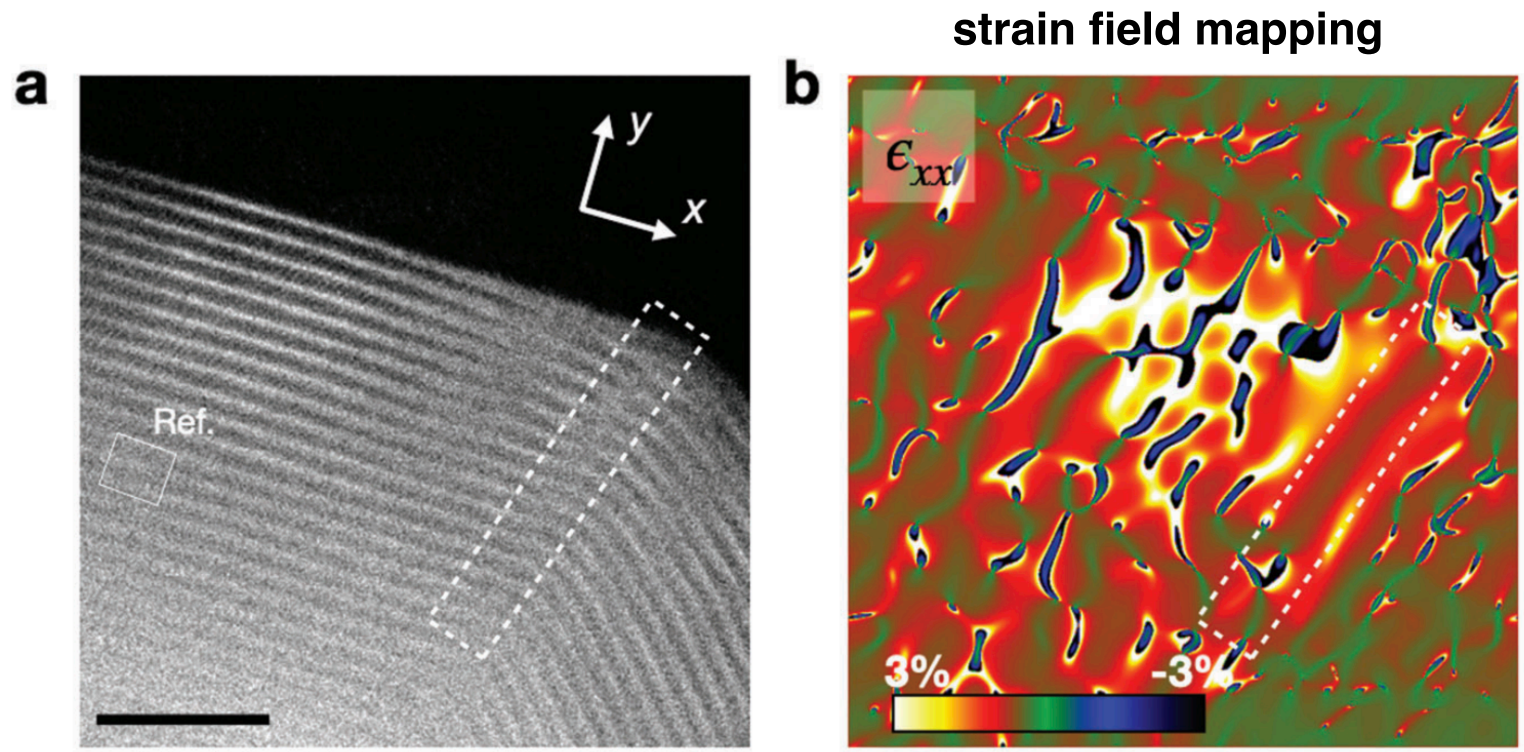
ML analysis of spectral images

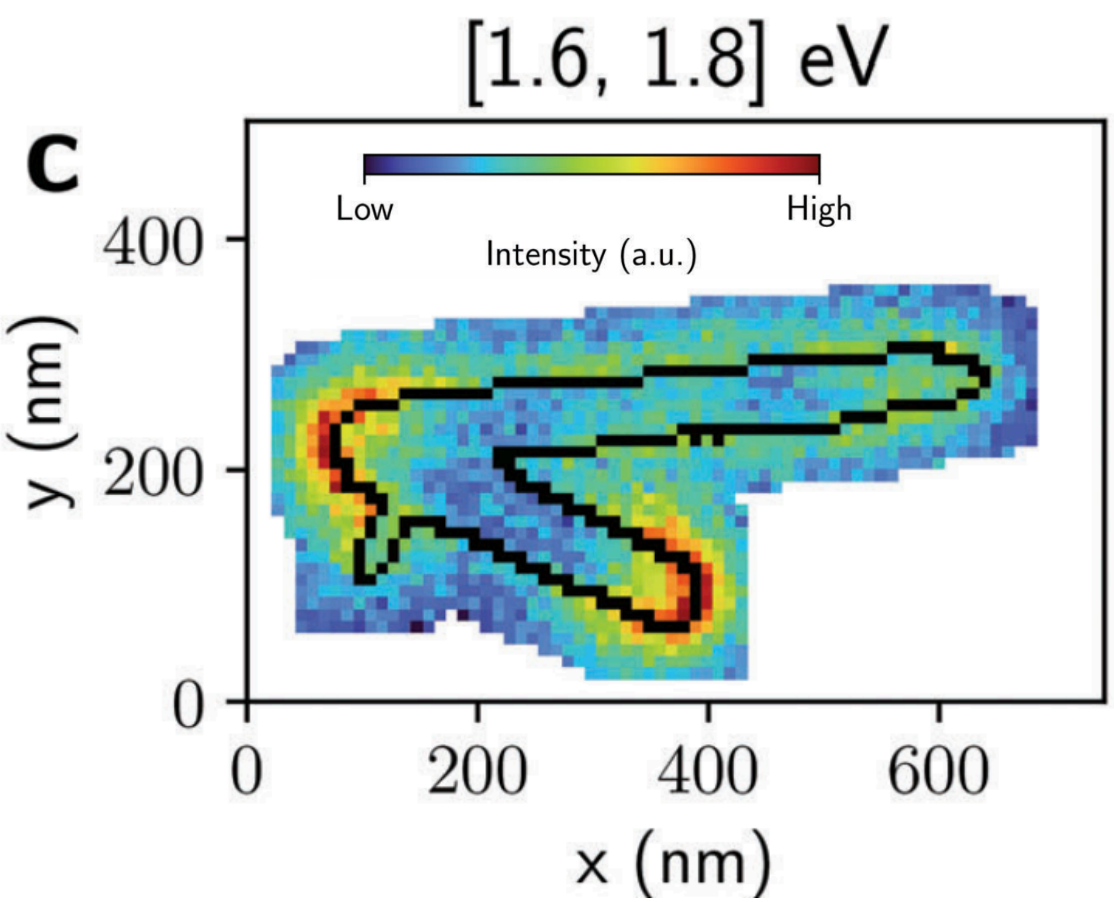


- EELS spectral image contains up to **O(10^5) individual spectra**
- Use **unsupervised learning** (*K-means clustering*) to identify clusters of pixels with comparable sample thickness and combine them for the **(supervised) NN training**
- Simultaneous determination of physical properties across the **whole nanostructure** with their **uncertainties:** thickness, band gap, position and width of plasmonic and excitonic resonances,…

Exciton Anatomy in 1D-MoS_2







- ☑ Direct correlation of **strain fields, band gap modulation, and exciton localisation** in 1D-MoS_2 nanostructures with different morphologies
- ☑ Demonstrate that **excitons are localised** in regions with large strain (e.g. bends, tips)
- ☑ Implications e.g. for single photon emitters for quantum communication
- ☑ Same methods could be also applied to other types of **spectroscopy techniques**

Summary and outlook

- Machine learning makes possible **identifying patterns in the data** whereby one can efficiently solve problems which are difficult of intractable with traditional approaches

- Powerful to parametrise in an unbiased way **high-dimensional functions** and **infer them from the data**, while ensuring **faithful uncertainty estimates**

- **Enable discoveries** such as intrinsic charm quarks in the proton & make possible to **optimise the sensitivity of searches** for interesting phenomena hidden in the data

- Our technology is portable to many other problems, as demonstrated for their applicability to data analysis in **electron microscopy of quantum materials**

- All codes are **open source** and extensively documented, and have benefitted from contributions as well from BSc and MSc students in our groups