Modeling F_2 using AI

G. Niculescu (JMU)
S. Brown (VT), I. Niculescu (JMU)

Nucleon and nuclear structure from inclusive measurements JLab, NN, Va

June 21, 2023

What? Why? How?

Disclaimer:

- Even though we had this idea for awhile, working in earnest on this project was a COVID-byproduct.
- … the fact that we're still "at it" can mean one of (several things):
    - value
    - stubbornness
    - long–COVID
- Many collaborators/advisers contributed a great deal to this project. And they done their level–best.
- Misconceptions/mistakes (including starting this in the first place): GN

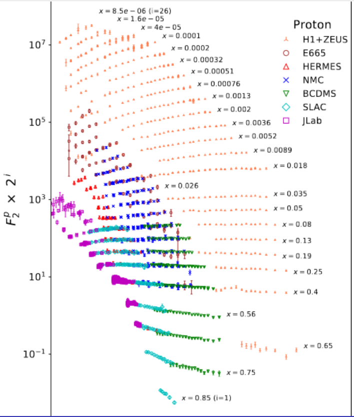
← Segway to the next slide...

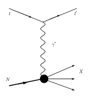
What? Why? How?

What?

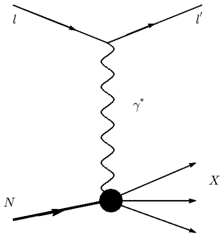
- $e^{+/-}$, γ **beams:** excellent tools for probing the nucleon structure
- **Inclusive electron scattering:** 50+ years of fruitful service to the field...

HighX Workshop, Crete, 2019





Formalism



Define:

$Q^2 = 4EE' \sin^2 \vartheta/2 \qquad x = \frac{Q^2}{2M\nu}$

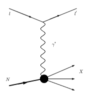
$W^2 = M^2 + 2M\nu - Q^2 \qquad$ (also Z & A)

$$\frac{d^2\sigma}{d\Omega dE'} = \sigma_{Mott}\left(\frac{2}{M}F_1(x,Q^2)\tan^2\frac{\vartheta}{2} + \frac{1}{\nu}F_2(x,Q^2)\right)$$

- F_is connect to pdfs, gpds, etc.
- so studying these is worthwhile.

Why?

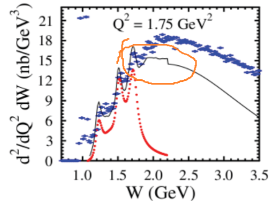
- Large body of data (SLAC, DESY, CERN, JLAB...)
- Several nice models (pdf-based, phenomenological, hybrid)
- Models do a good job of representing the data. Actively maintained.
- So, why bother? **Why, indeed?**



Rationale (I)

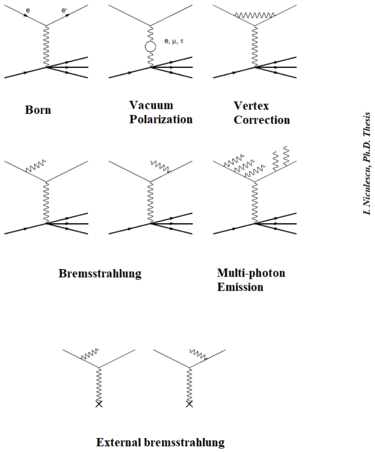
Why?

- Most models have limited kinematic reach.
- Meshing two/more models - problematic.
- ...so  instead of 
- Speed, speed, speed.
- (Audience:) Gabby, CPU cycles are cheap?
- (GN:) Yes, but if you can spend them more fruitfully elsewhere...
- What types of applications* would benefit from a faster model?
- **Good Question...**



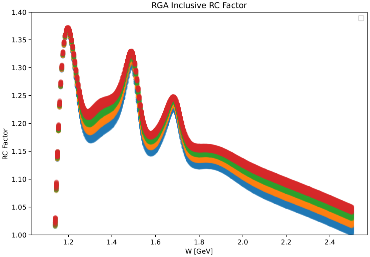
N. Markov - Inclusive electroproduction with CLAS12 and the nucleon structure in the valence quark domain

Rationale (II)

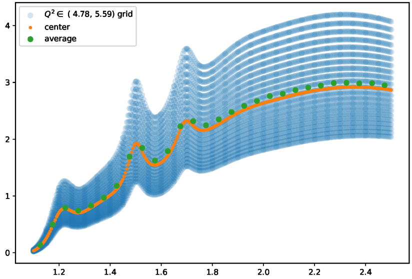


A1: Radiative Corrections

- "as measured by detectors" vs
- "as it happened at the tgt."
- ...the effect can be quite large



Rationale (III)



A2: Bin Centering/Unfolding

- Unfolding is well-beyond the scope of this talk so we'll skip it.
- BC: counting experiments
- Mean Value Theorem:
- $(\exists)\ c \in [a, b]$ so that $f'(c)(b-a) = f(b) - f(a)$
- The whereabouts of c are not (generally) known!

Rationale (IV)

OK! RC, BC, unfolding important. So...

- Where's the AI?
- RC, BC, etc. they all need lots... NO! LOTS of events.
- Said events need event (semi)realistic event generators.
- Existing "artwork" (read "models") are not particularly fast.
- Either for logistic or intrinsic reasons (convolutions, interpolation using large tables, etc.)

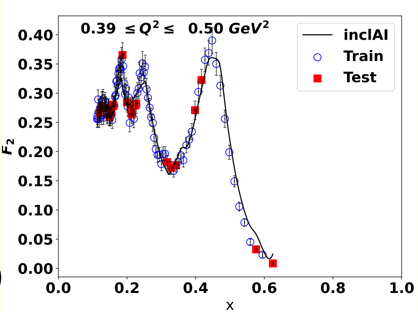
furthermore...

- Even if inclusive scattering is not your game...
- you might still benefit from a super–fast, nimble
- (background?, raw detector rates? etc.)
- ...adaptable/expandable to other reactions, observables...

Enter: inclAI

inclAI

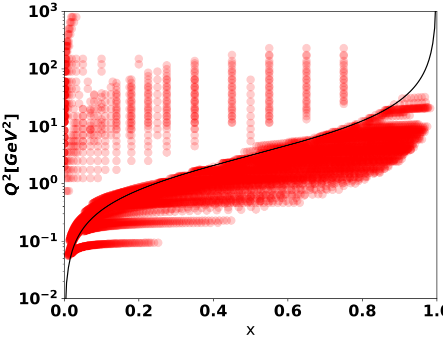
- ML model for F_2.
- no physics assumptions
- spans entire phase space.
- take data uncertainties into account.
- take **Z** and **A** into account
- extensible, customizable.
- fast (for both training and deployment)
- Understandable
- Quantify uncertainty!



inclAI (II)

inclAI: Data Assumptions

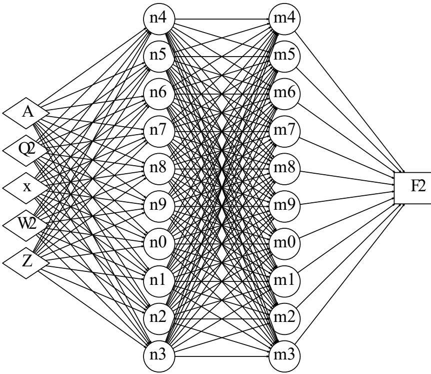
- world data accepted **as is** (no **re–scaling!**)
- uncertainties (stat. & syst.*)
- **features*:** Q^2, x, W^2, Z, A
- **label:** F_2 (...)
- **scaling:** std. & min/max
- ∼12k data points (h & d)
- ∼55k+ for all nuclei
- ∼80% of the work went in curating this data (thanks to all that maintain various databases!)



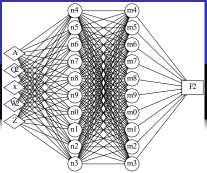
inclAI (III)

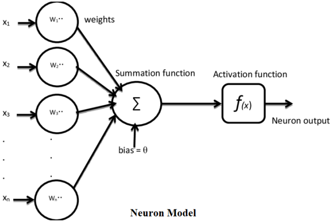
inclAI: ML Assumptions

- fully connected ANN.
- 1...N hidden layers.
- Activation: ReLU & sigmoid
- Additional details:
  - Early stopping
  - LR change on plateau
  - Cold/Hot start
  - Regularization
  - Logging, messaging
  - 80/20 train/test split
  - stratified sampling
- python/keras/tf

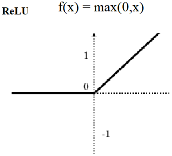


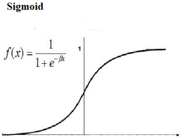
inclAI (IV)





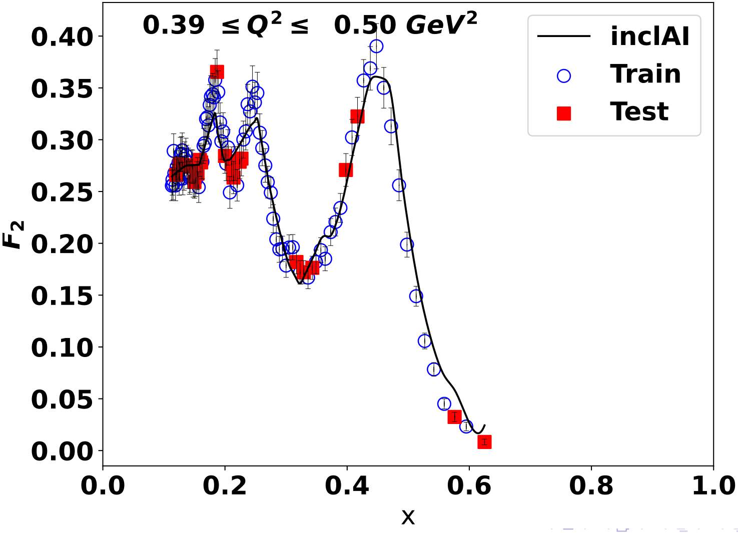
Neuron Model



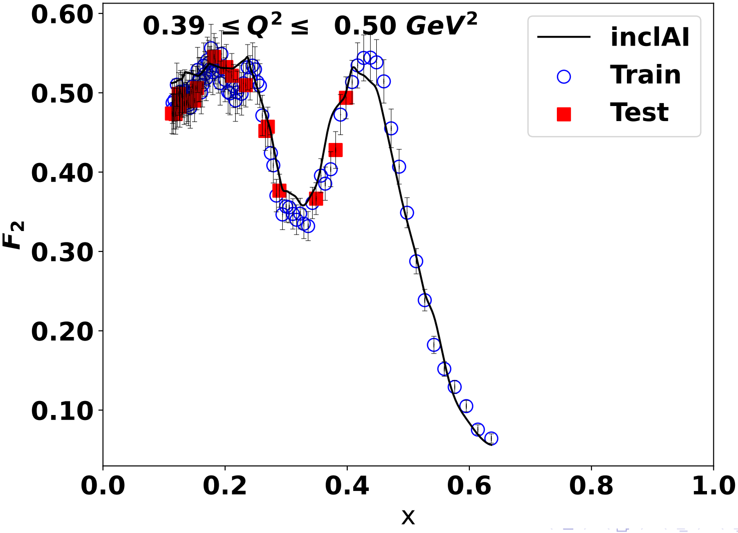


- for each layer "k", at each node "j": $X_{jk} = \sum_i w_{ijk} x_{ijk} + b_{jk}$
- X_{jk} is then fed to the respective activation function, producing the neuron's output. Repeat for all layers and nodes.
- (Audience:) This is so simple. I bet it does not even work!
- (GN:) Well...

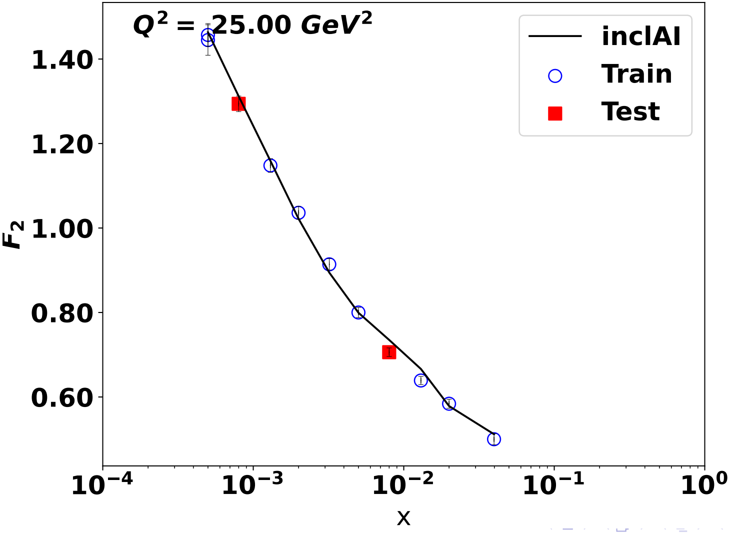
Does it work? (I) ...hydrogen, low Q^2



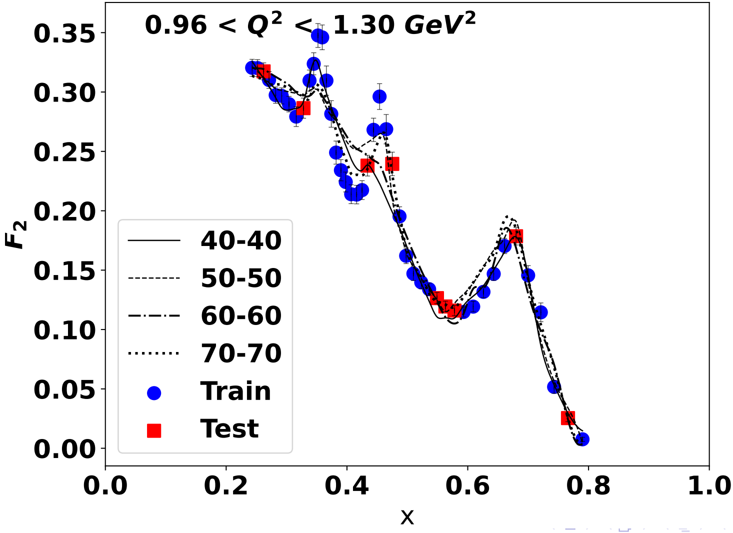
Does it work? (II) ...deuterium, low Q^2



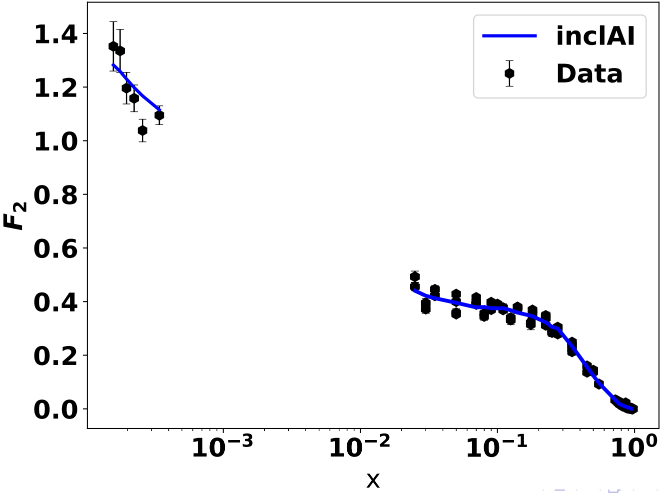
Does it work? (III) ...hydrogen, high Q^2 (log)



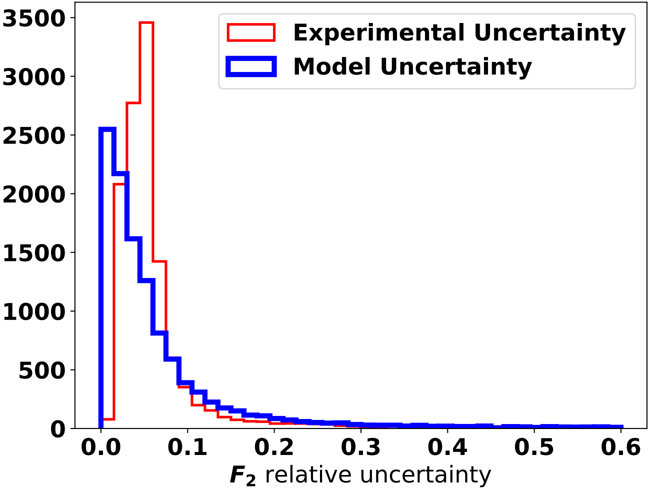
Does it work? (IV) customizable...



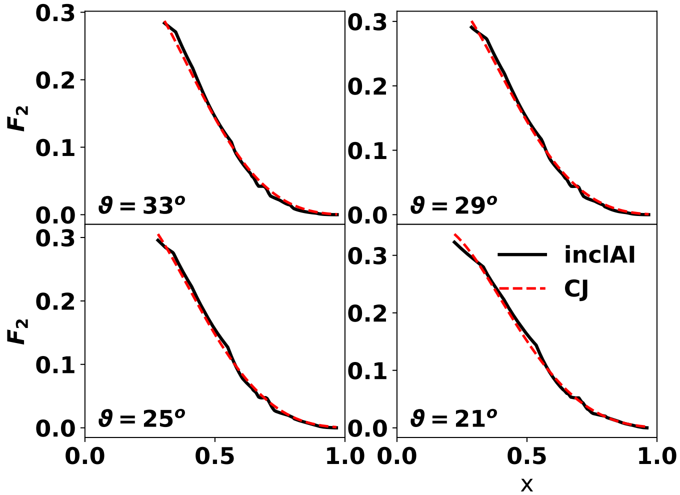
Does it work? (V) ...kinematic range ($7 \leq Q^2 \leq 13$)



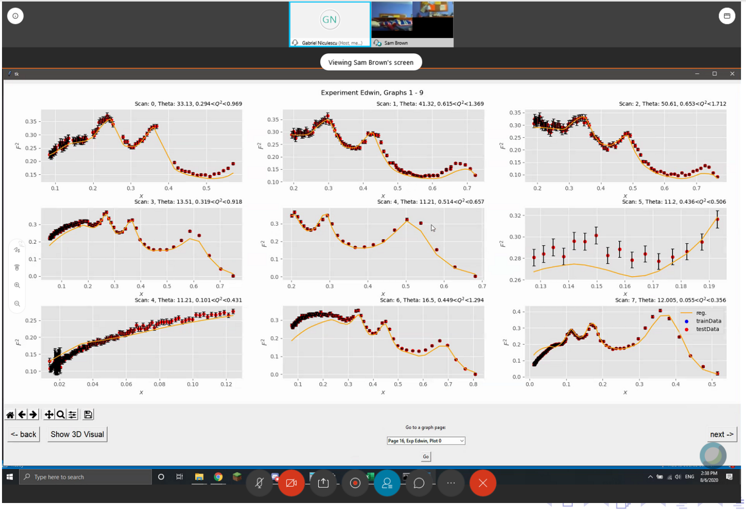
Does it work? (VI) ...data & model uncertainties



Does it work? (VII) ...vs existing artwork



Does it work? (VIII) ...predicting a whole exp. Live shot off of a BlueJeans session!



inclAI H & D summary

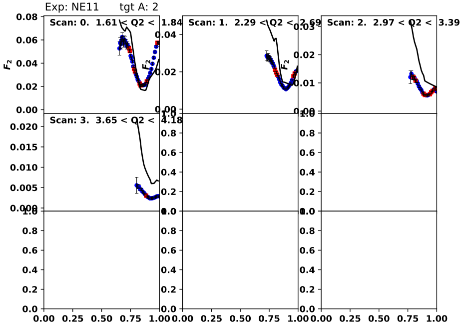
Machine learning representation of the F_2 structure function over all charted Q^2 and x range

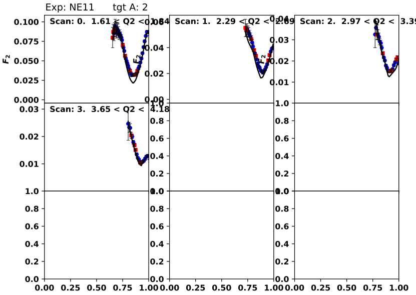
S. Brown, G. Niculescu, and I. Niculescu
Phys. Rev. C **104**, 064321 – Published 23 December 2021

- hydrogen and deuterium results published in 2021
- precision comparable w/ the data uncertainties
- Speed: 10-100x faster than existing artwork

- Good!
- Now, onward to extensions, adaptation, current (and future) work
- ... in other words: "emerging capabilities"

inclAI as “anomaly detector”

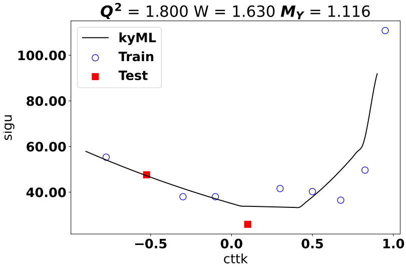


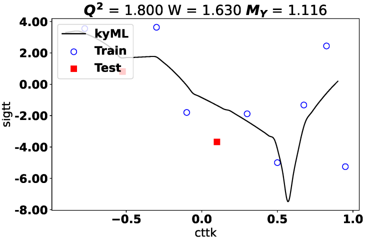


Finding “problems” in existing databases...

kyML

Modeling $e + p \rightarrow e' + K^+ + \Lambda/\Sigma^0$ reduced cross–sections





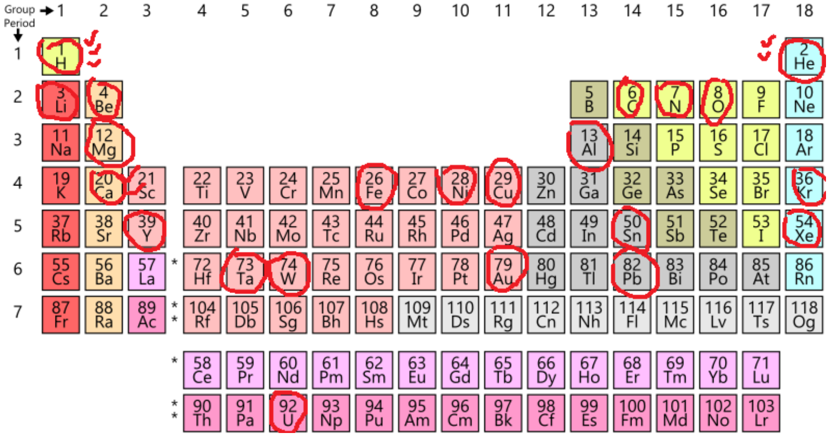
inclAI extension to nuclei

inclAI strikes back (and at higher Z!)

- inclAI had target **Z A** as features *ab initio*
- ... with the obvious goal of extending the model to nuclei.
- This presented a few new challenges:
  - Finding/reading the data! (m data sources, n different formats, $n > m$ (!!))
    Thank you to all the maintainers of these databases/websites!!!
  - omG! (some of) this data has quasi–elastic!
  - Some of the data comes as ratios wrt another nucleus (usually deuterium).
  - Add a few (more) columns to our DF (year, type of obs, secondary Z and A).
  - Devise a way of handling ratios (HINT: existing F_2^D artwork does not work above $x = 1$).
  - Switch from F_2 "per nucleon" to F_2 "absolute".
  - Revise (a little) the way we plot things.

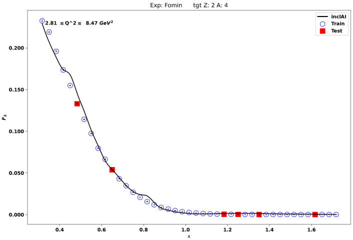
Apologies! there's wayyyy to much text here.

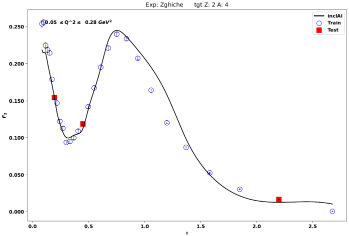
Data used for training

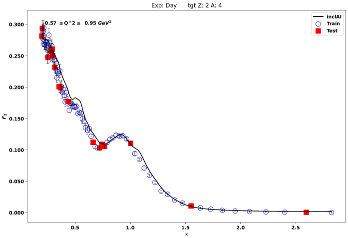


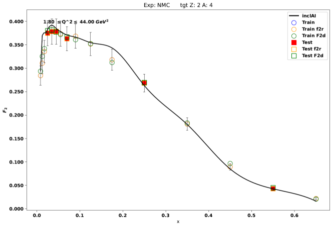
${}^4_2He$

This, and following pages are PRELIMINARY!!

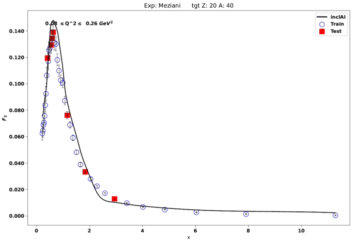


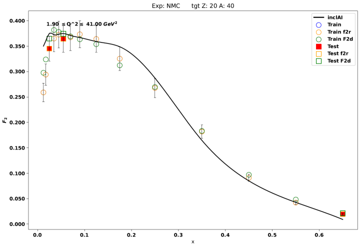






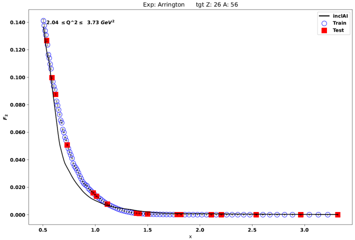
$^{40}_{20}Ca$

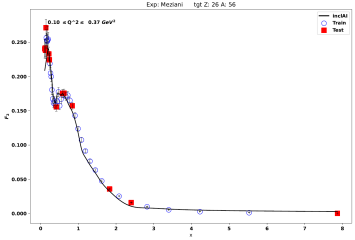

Exp: Meziani tgt Z: 20 A: 40
0.08 ≤ Q^2 ≤ 0.26 GeV^2
inclAI
Train
Test
F_2
x

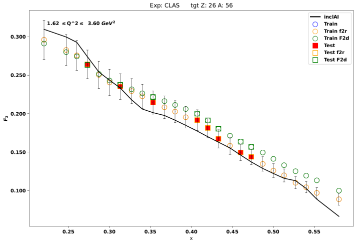

Exp: NMC tgt Z: 20 A: 40
1.90 ≤ Q^2 ≤ 41.00 GeV^2
inclAI
Train
Train f2r
Train F2d
Test
Test f2r
Test F2d
F_2
x

$^{56}_{26}Fe$

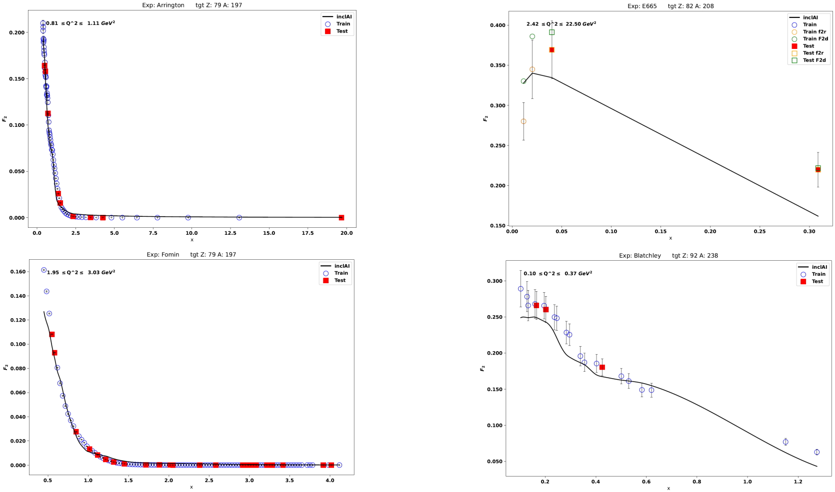
As we said: work in progress!!





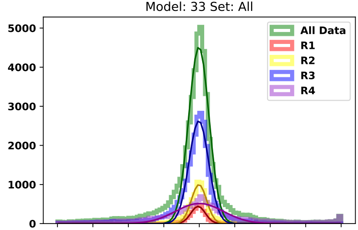


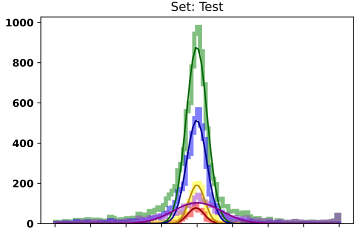
How about some heavier nuclei?

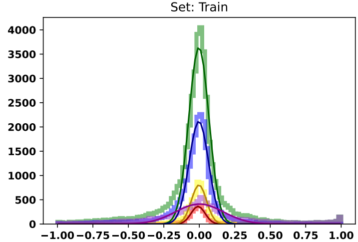


Gauging (in)success

How do you know you won (or lost) the game?



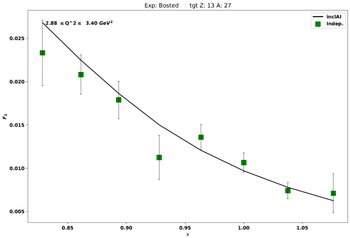


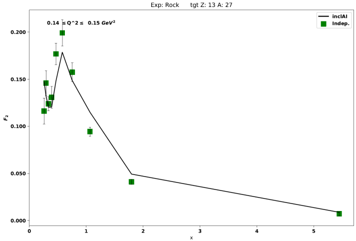


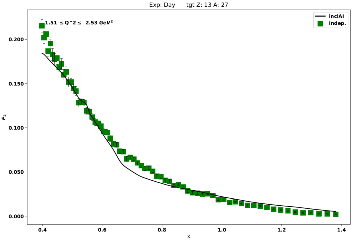
Reg.	x_{min}	x_{max}	σ_{DATA}	res_{fit}
All	0.0e+00	1.0e+04	8.52e-02	7.39e-02
R0	1.0e-03	1.0e-01	5.59e-02	5.78e-02
R1	1.0e-01	3.0e-01	4.41e-02	5.56e-02
R2	3.0e-01	1.0e+00	5.95e-02	7.08e-02
R3	1.0e+00	1.0e+04	1.62e-01	1.70e-01

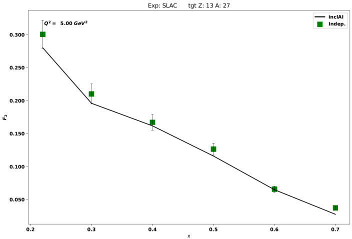
How about...

...leaving a nucleus out of the training and trying to predict its F2 data afterward?







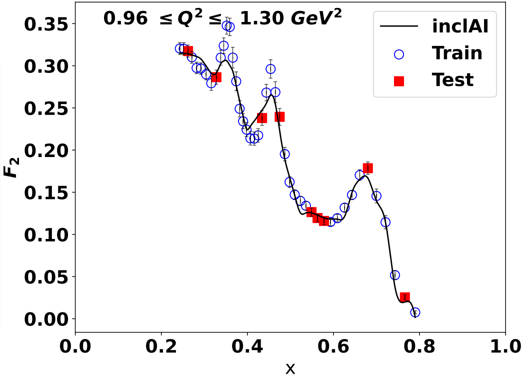


Quo Vadis?

To do: finish/publish the work on nuclei.
Start phase III of the project.

Hopefully I convinced you that inclAI...

- ML F_2 representation.
- Flexible, adaptable.
- 10-100x+ speed improvement.
- ideally suited for RC, BC...



THANK YOU!