Massachusetts Institute of Technology

Connecting nuclear force parameters to nuclear observables via machine learning emulation

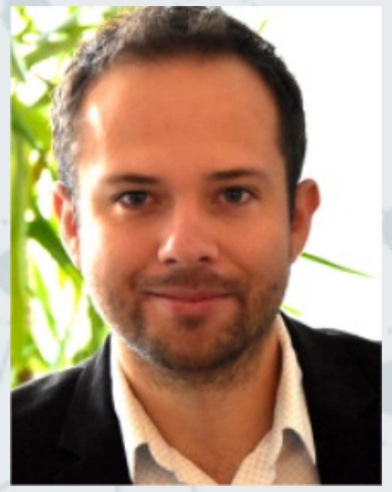
INT26-1: Nuclear Hamiltonians for Advancing Nuclear Physics and Beyond

Antoine Belley

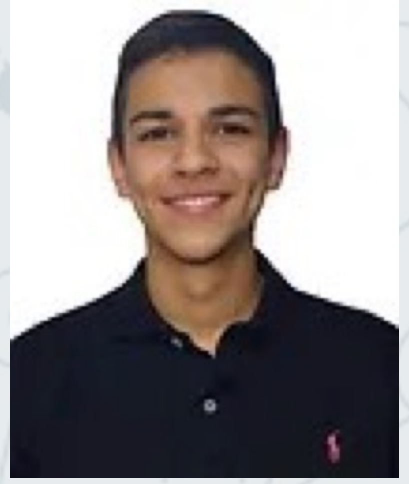
28 April 2026

Acknowledgements

Prof.
Ronald García



Jose Miguel
Muñoz Arias



EMA lab

G. Mondeel

A. Brinson

H. Kakioka

S. Moroch

F. Pastrana

D. Gonzales

J. Hacias

M. Fulghieri

S. Fukaya

M. Flayol

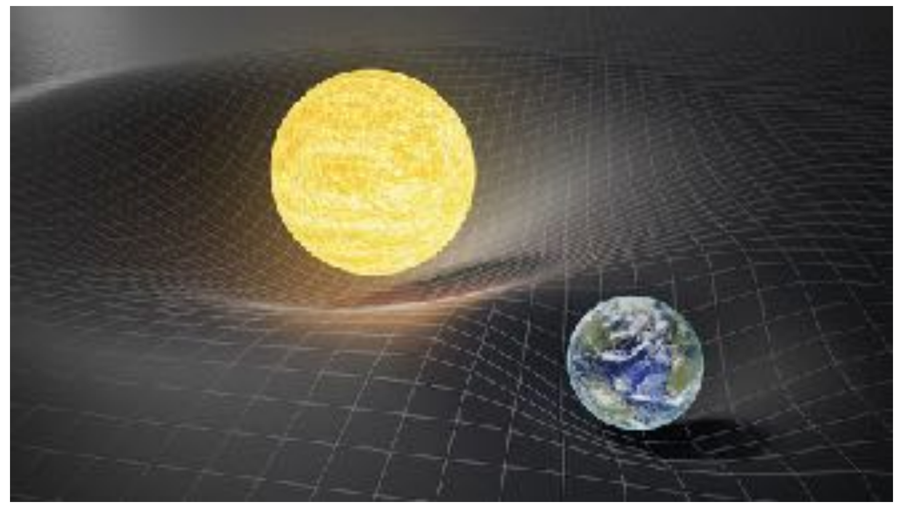
A. Jadbabaie

C. Konig

Shortcomings of the Standard Model

Shortcomings of the Standard Model

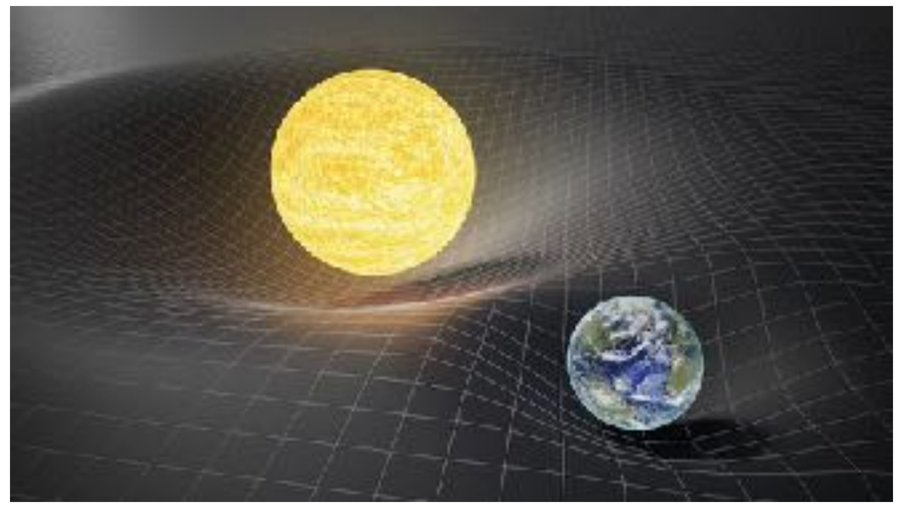
Credit: vchal via Getty Images



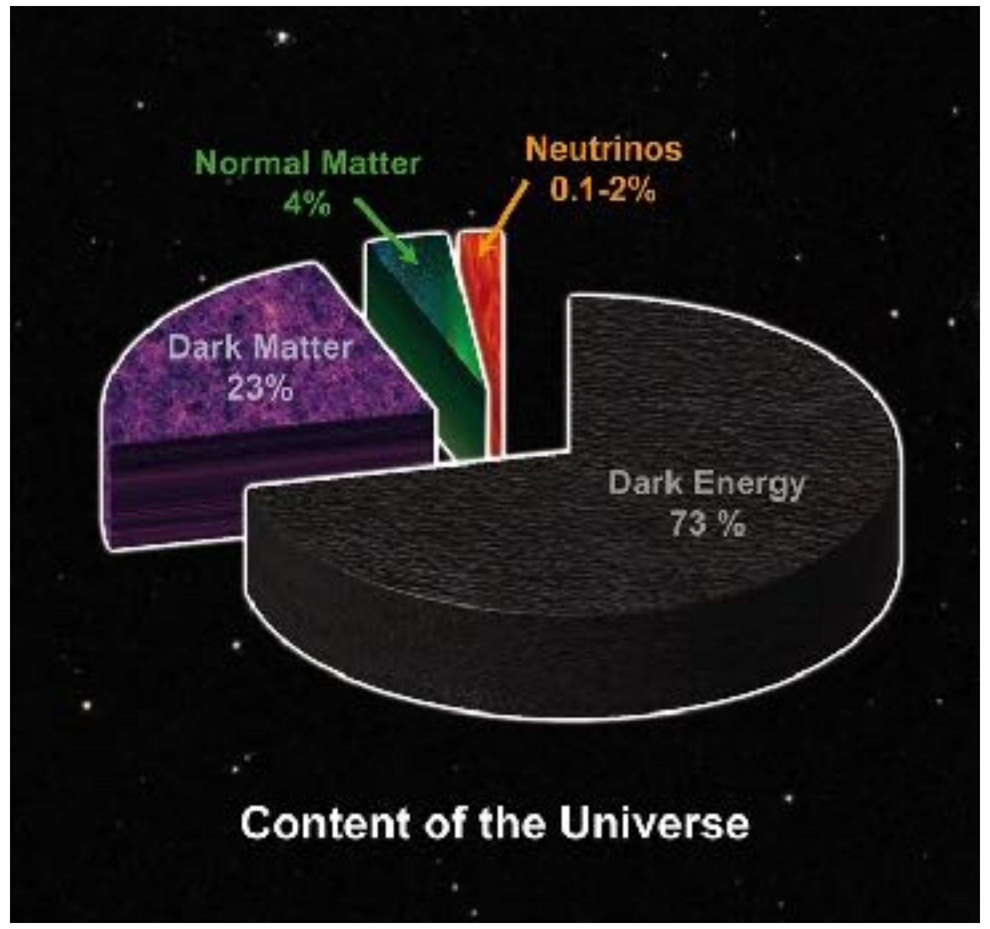
- Fails to explain gravity

Shortcomings of the Standard Model

Credit: vchal via Getty Images



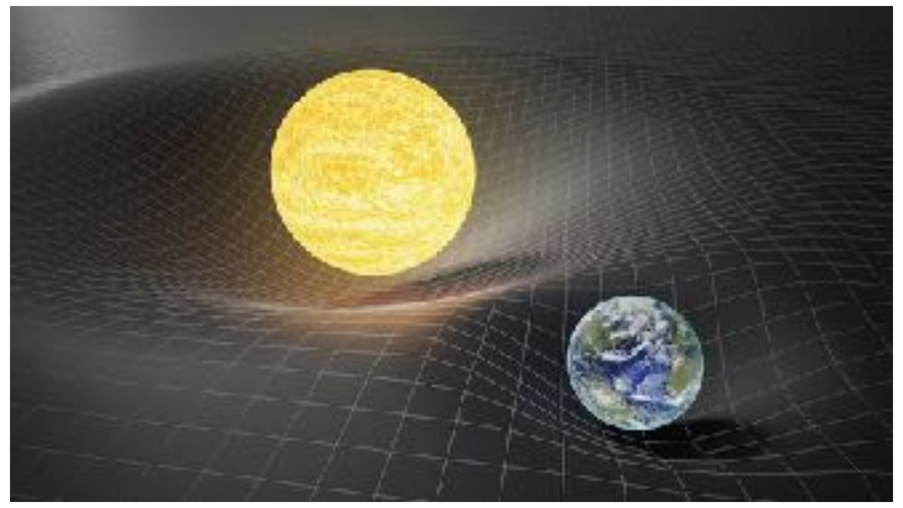
- Fails to explain gravity
- Fails to explain most of the mass and energy of the universe



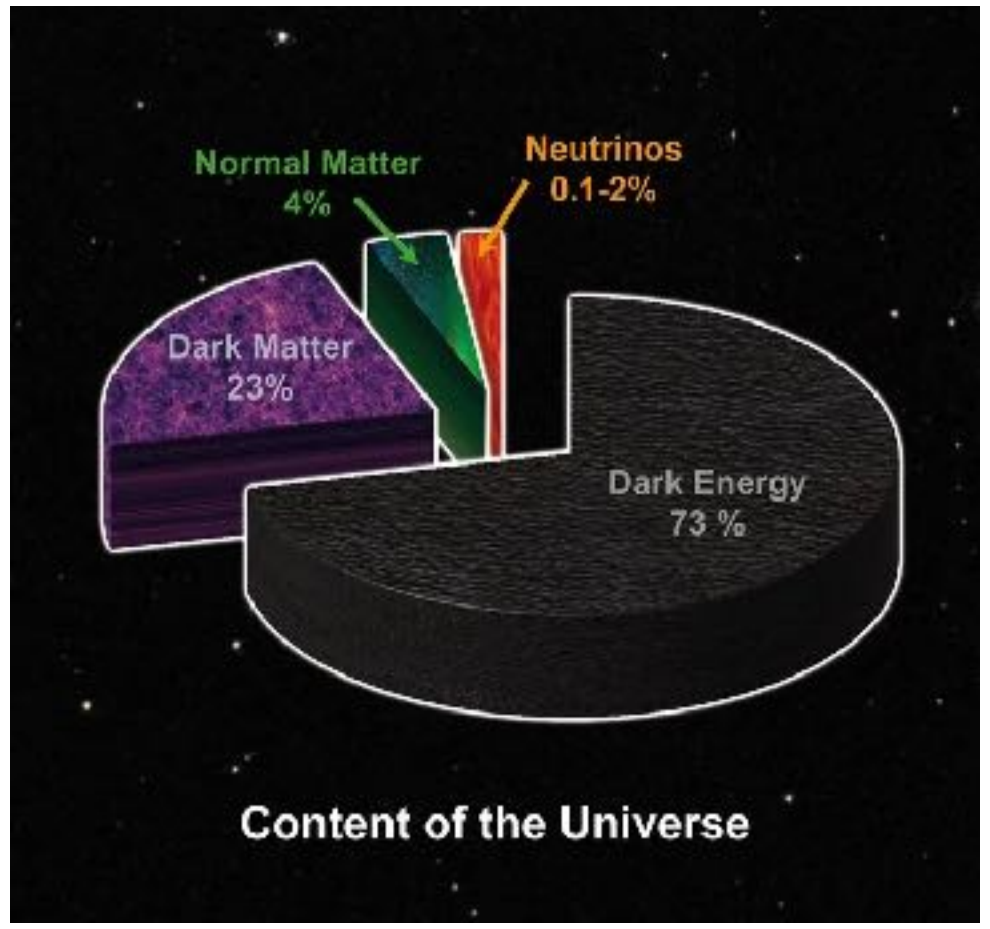
Credit: spacecentre.co.uk

Shortcomings of the Standard Model

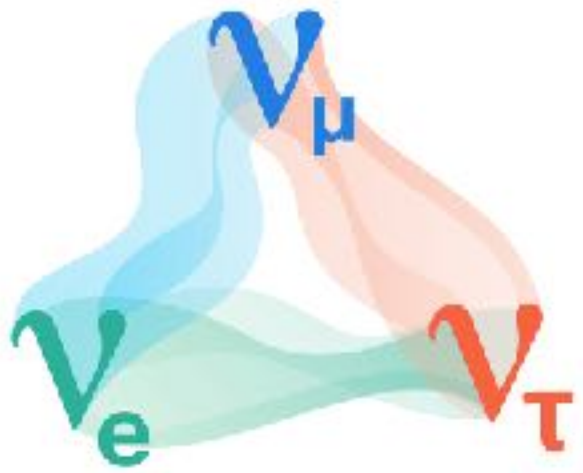
Credit: vchal via Getty Images



- Fails to explain gravity
- Fails to explain most of the mass and energy of the universe
- Fails to explain neutrino mass and oscillation.



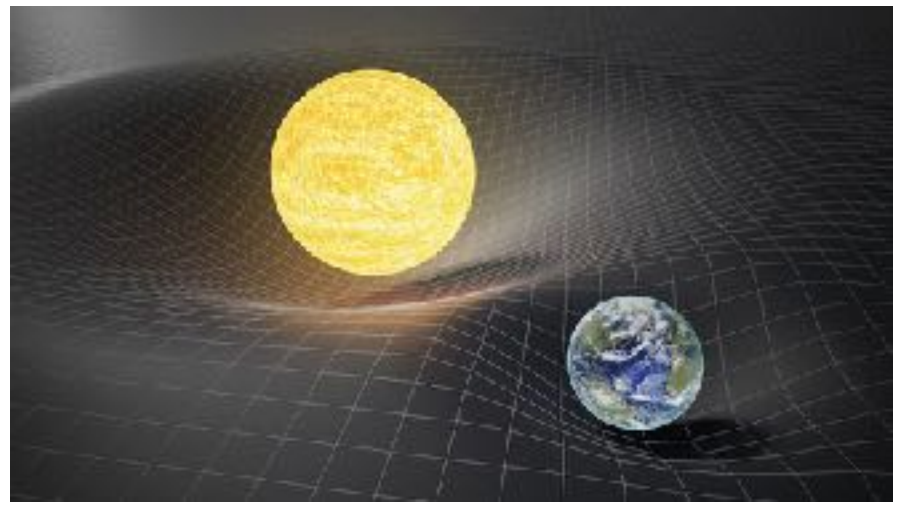
Credit: spacecentre.co.uk



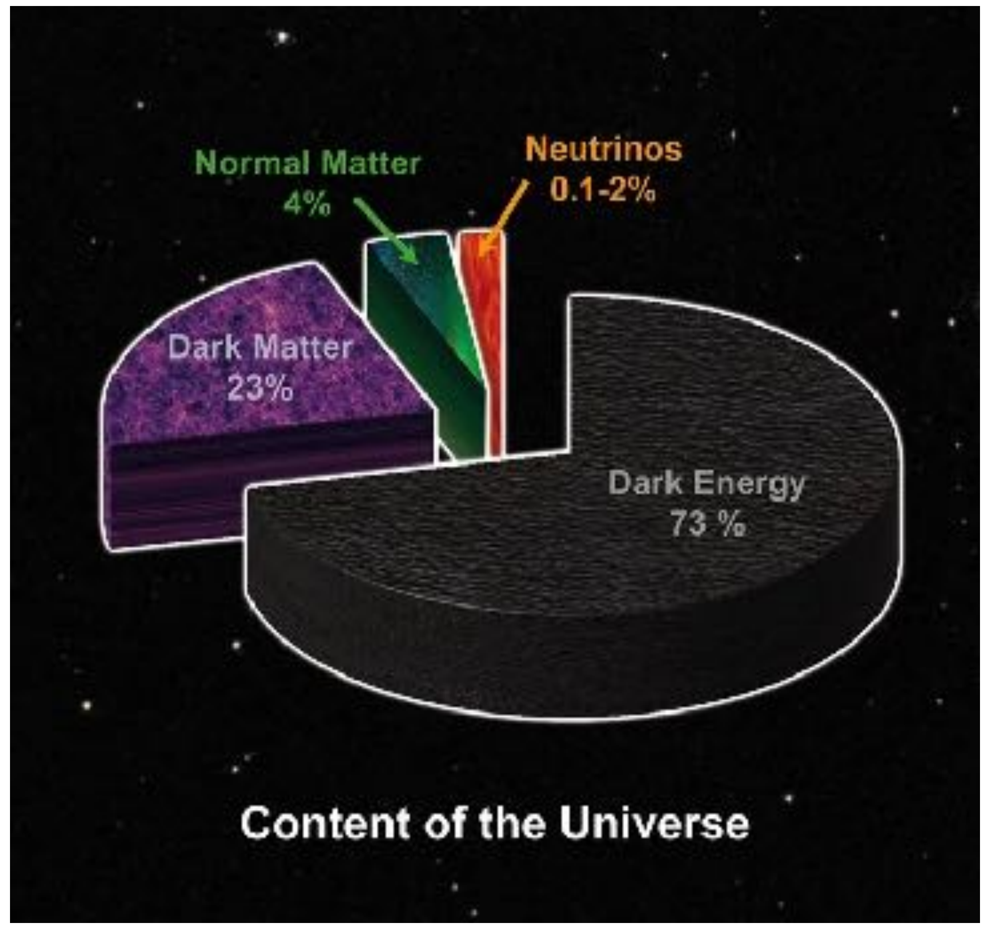
Credit: Sanford Lab/DUNE

Shortcomings of the Standard Model

Credit: vchal via Getty Images

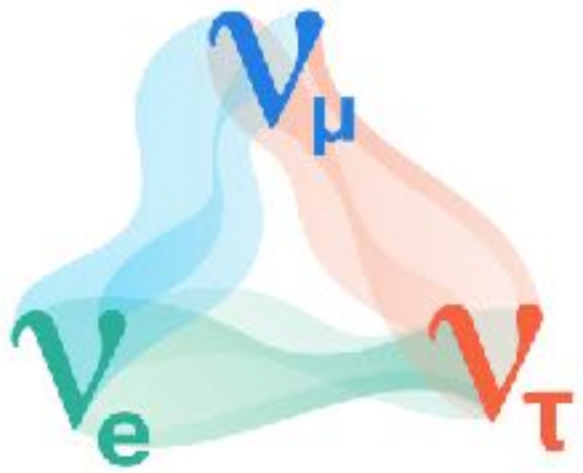


- Fails to explain gravity



Credit: spacecentre.co.uk

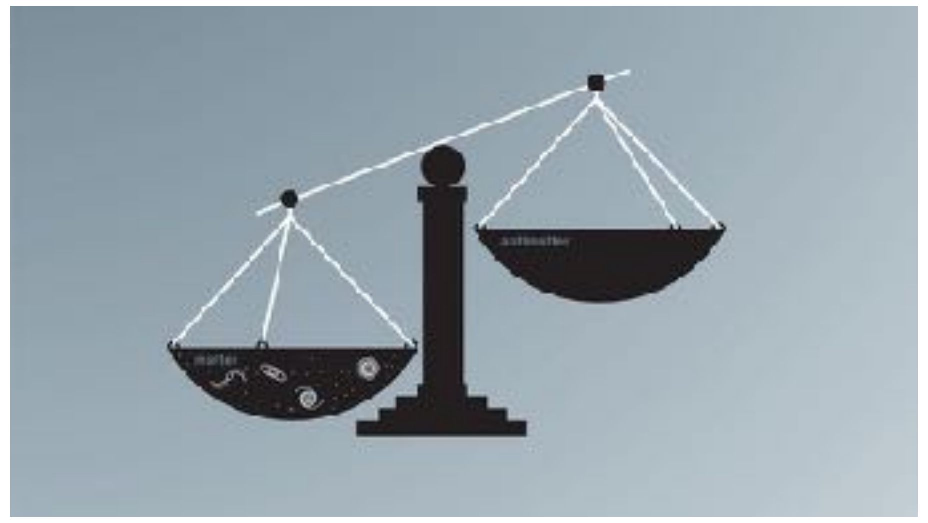
- Fails to explain most of the mass and energy of the universe



Credit: Sanford Lab/DUNE

- Fails to explain neutrino mass and oscillation.

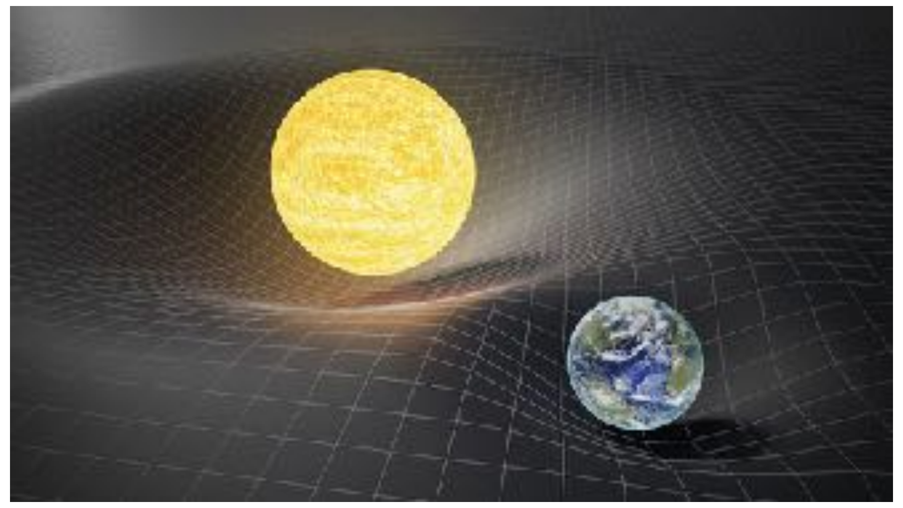
- Fails to explain the matter-antimatter asymmetry.



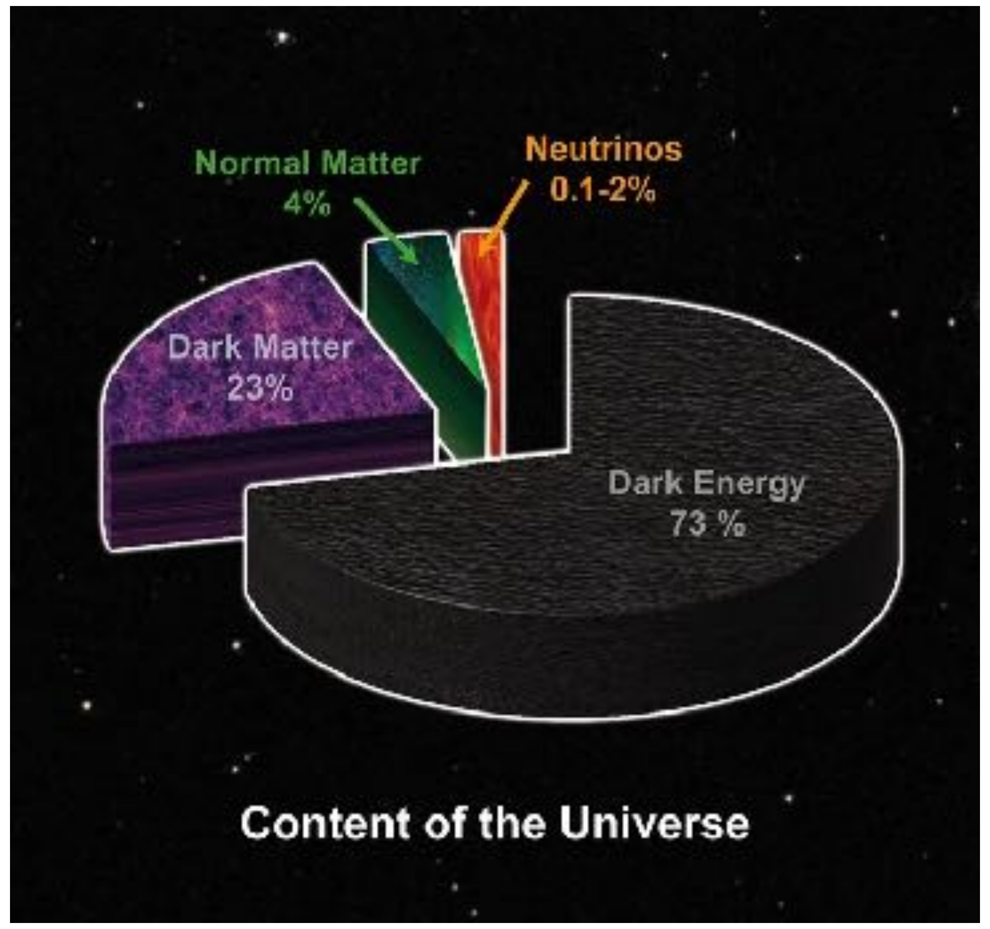
Credit: Symmetry Magazine / Sandbox Studio, Chicago

Shortcomings of the Standard Model

Credit: vchal via Getty Images

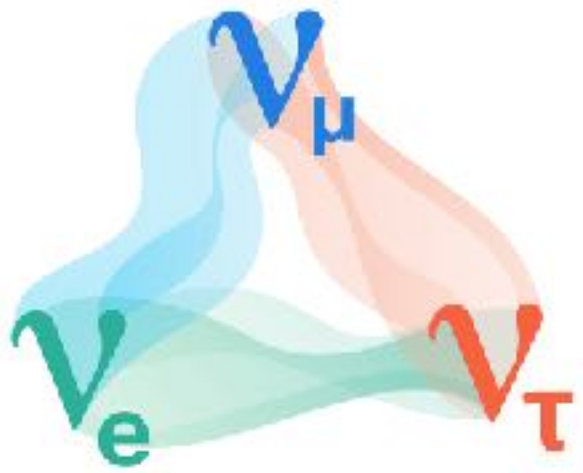


- Fails to explain gravity



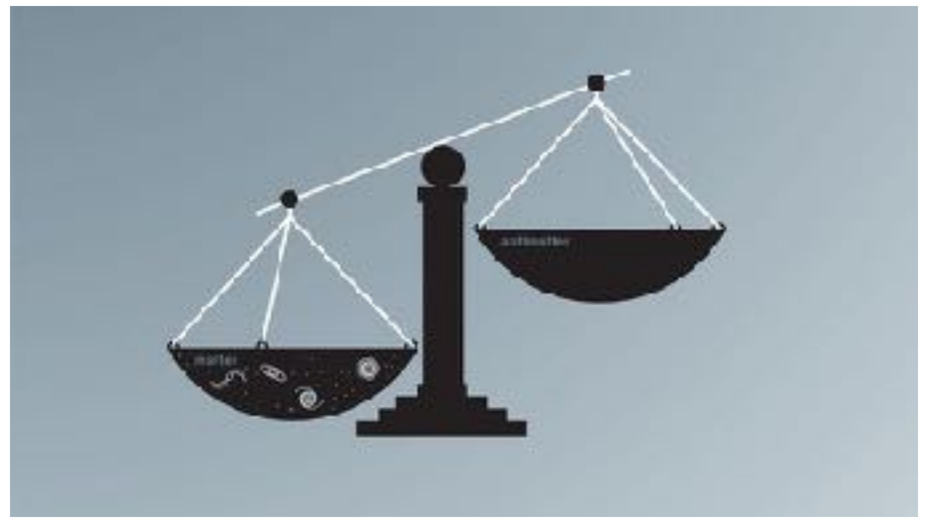
Credit: spacecentre.co.uk

- Fails to explain most of the mass and energy of the universe



Credit: Sanford Lab/DUNE

- Fails to explain neutrino mass and oscillation.



Credit: Symmetry Magazine / Sandbox Studio, Chicago

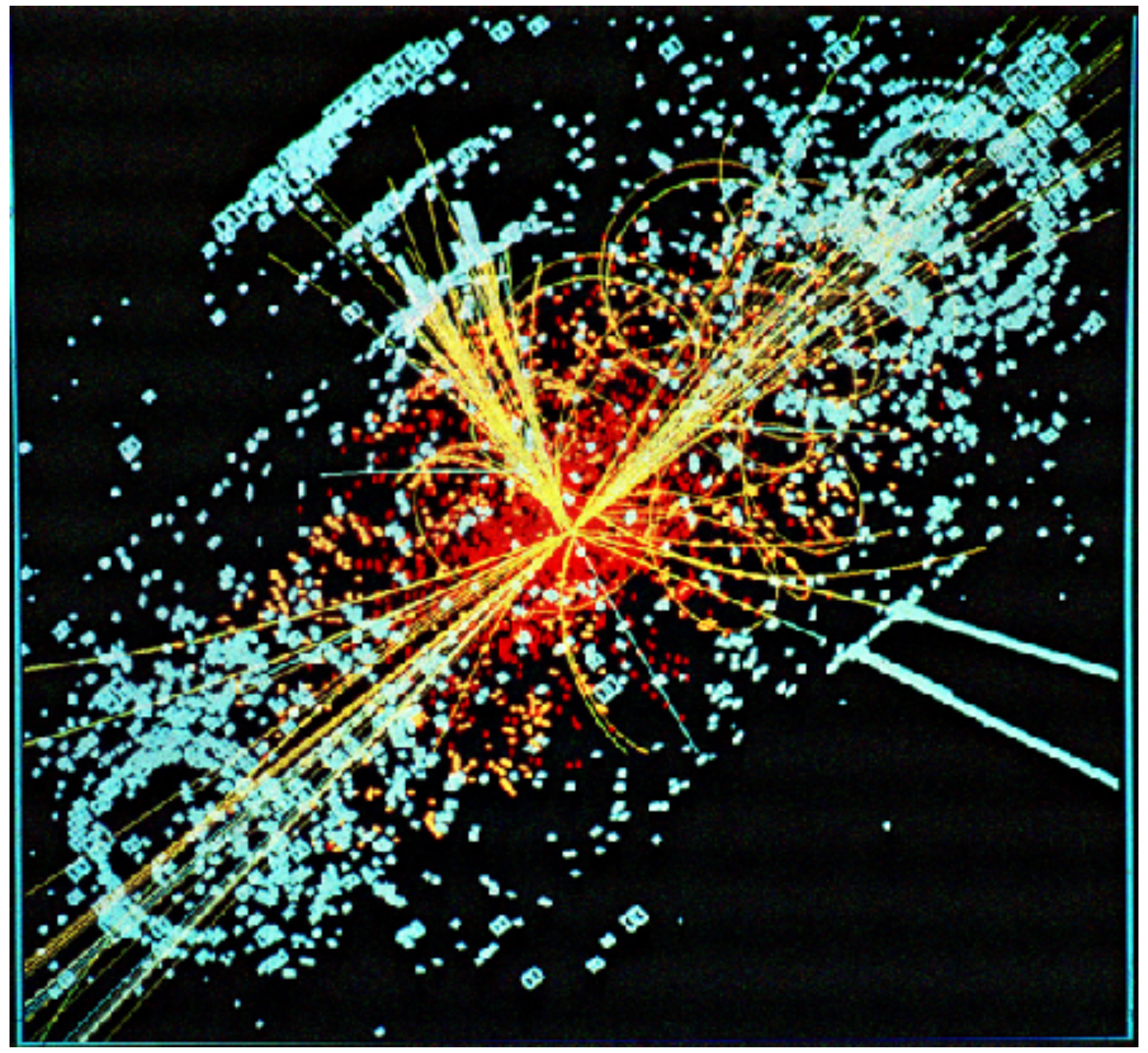
- Fails to explain the matter-antimatter asymmetry.

How should we search beyond the Standard Model to explain these shortcomings?

Going Beyond

TeV Scale ⟷ < MeV Scale

1. High energy Frontier

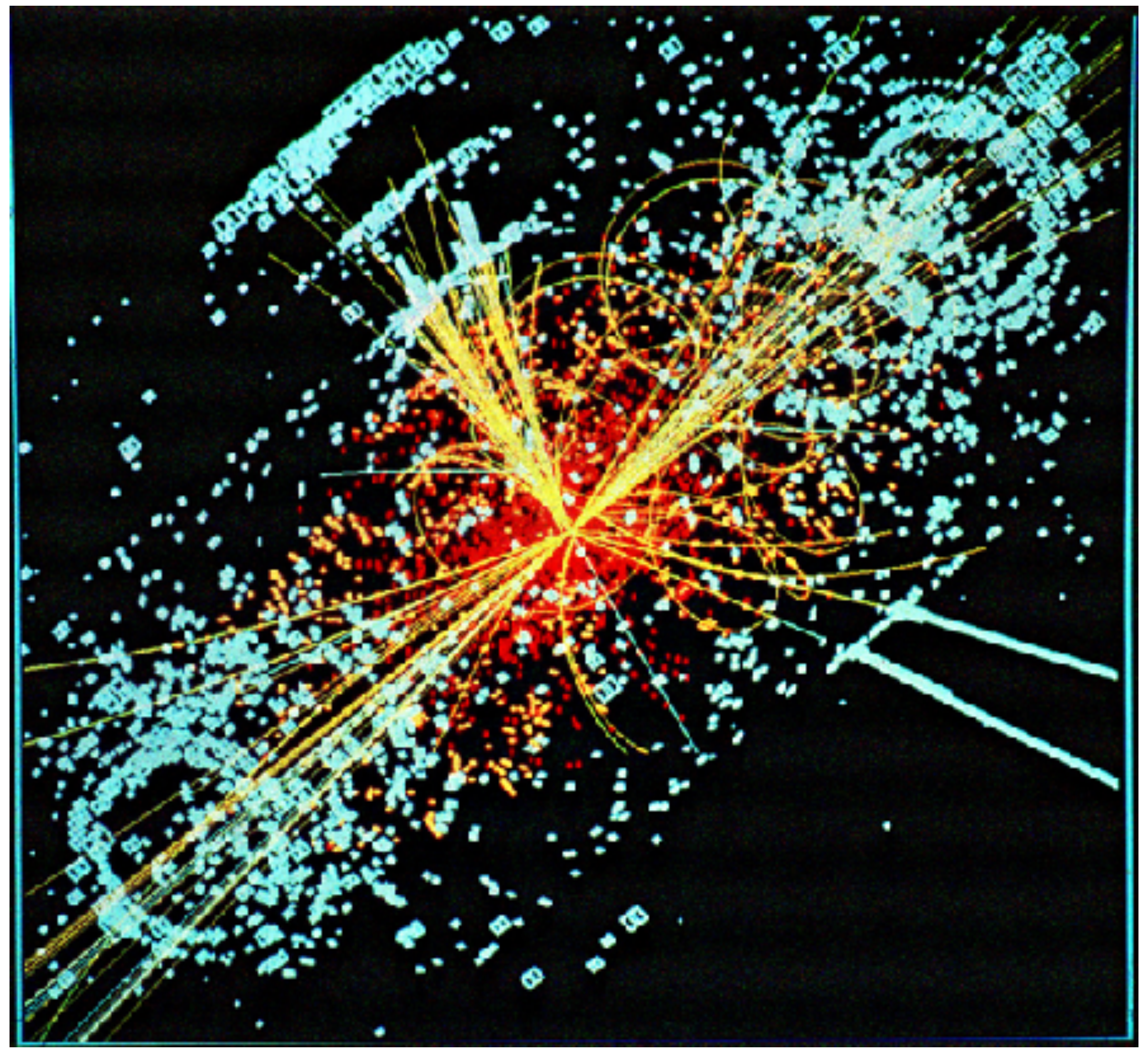


Lucas Taylor / CERN - http://cdsweb.cern.ch/record/628469

TeV Scale ⟷ < MeV Scale

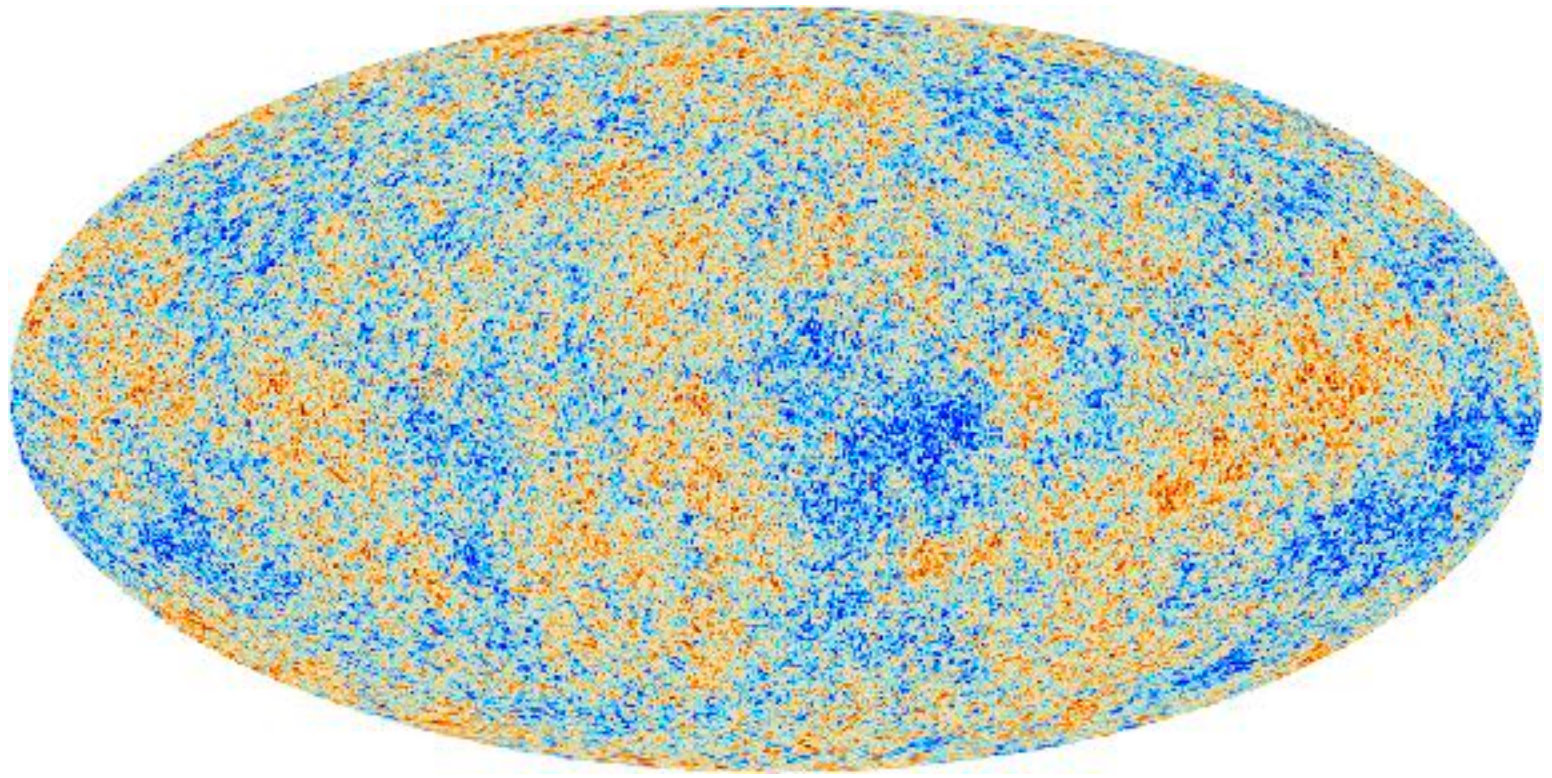
Going Beyond

1. High energy Frontier



Lucas Taylor / CERN - http://cdsweb.cern.ch/record/628469

2. Cosmo+Astro Frontier

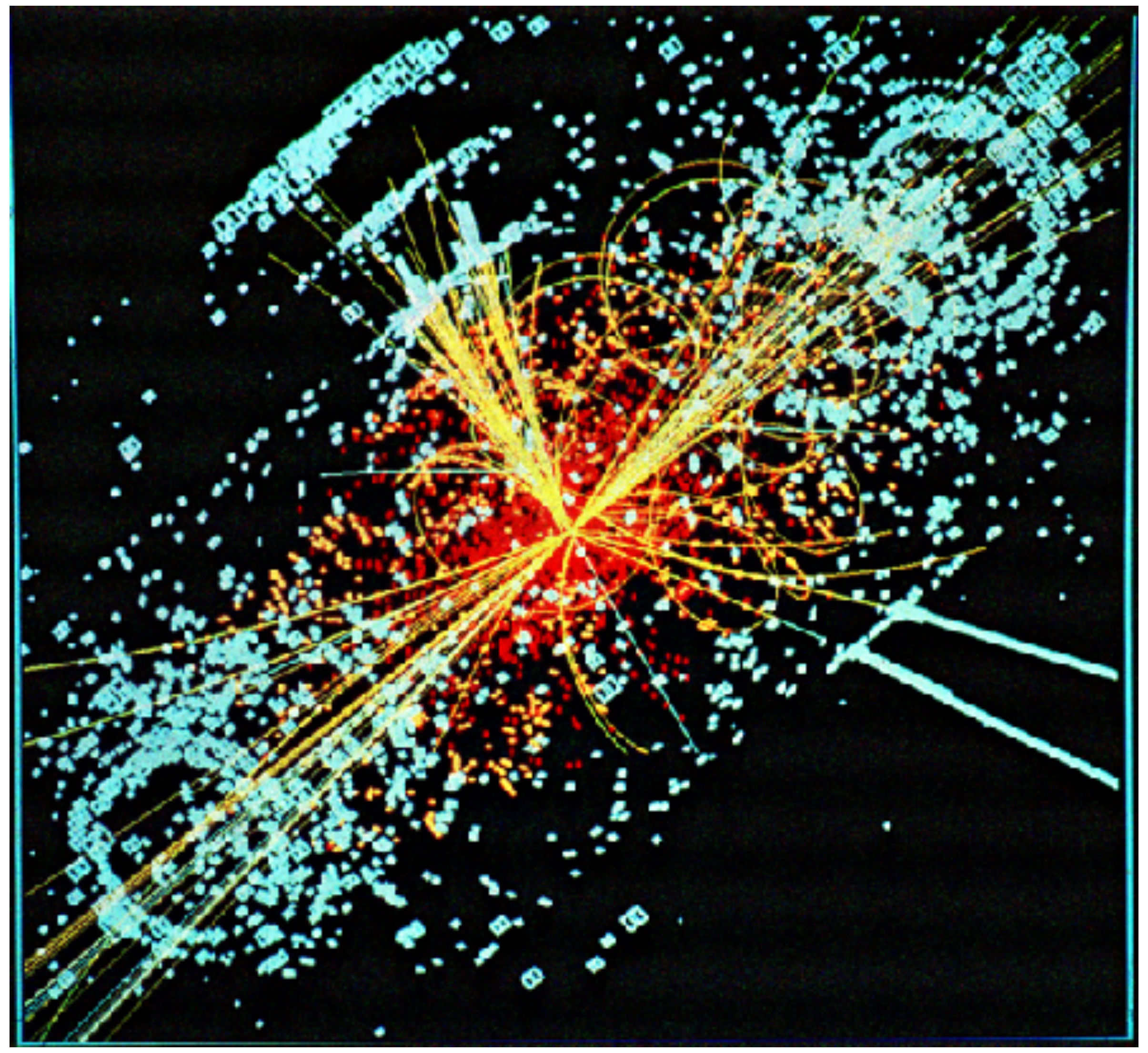


https://www.esa.int/ESA_Multimedia/Images/2013/03/Planck_CMB

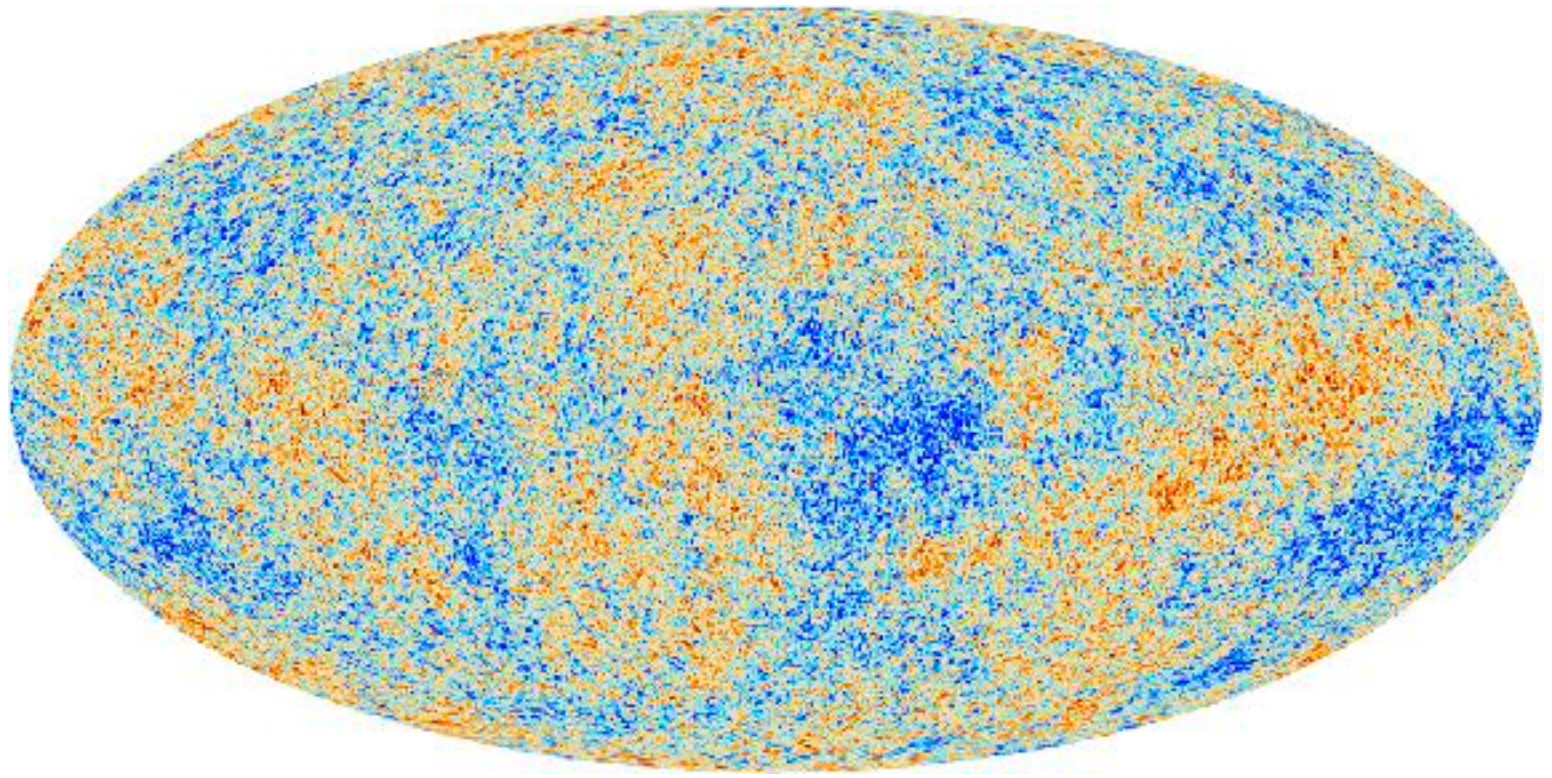
TeV Scale ⟷ < MeV Scale

Going Beyond

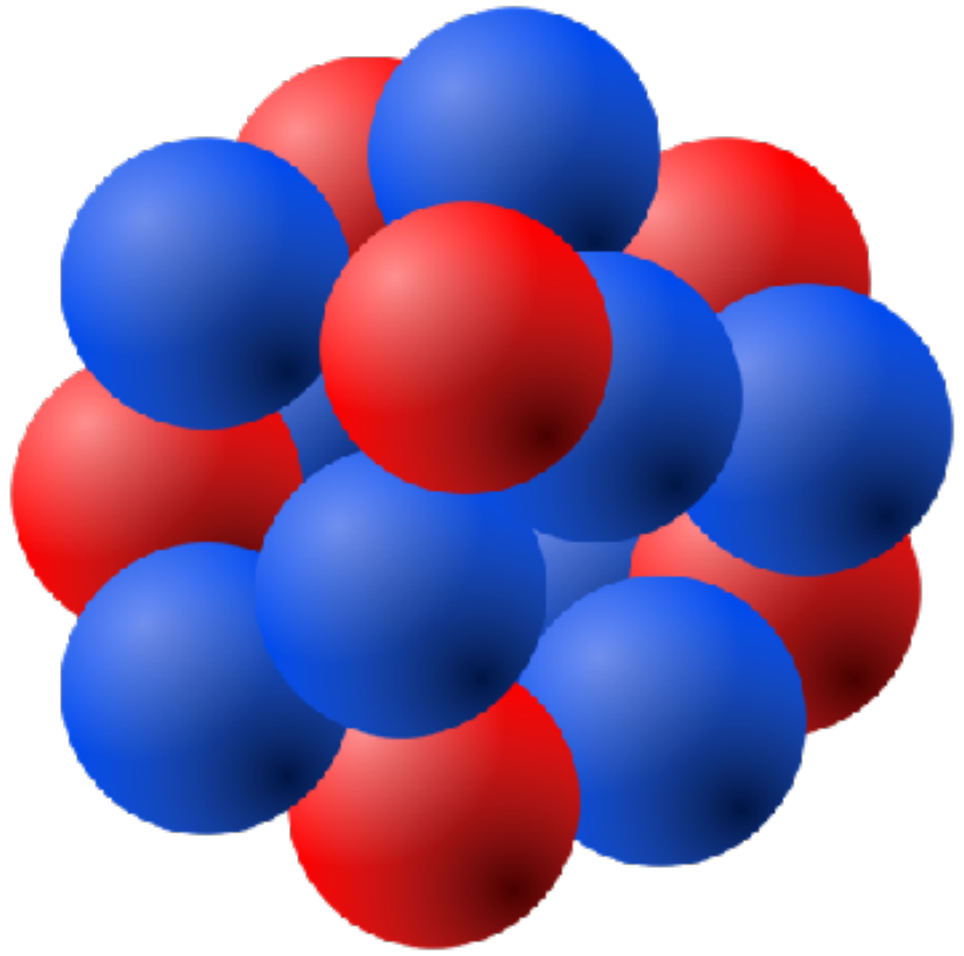
1. High energy Frontier



2. Cosmo+Astro Frontier

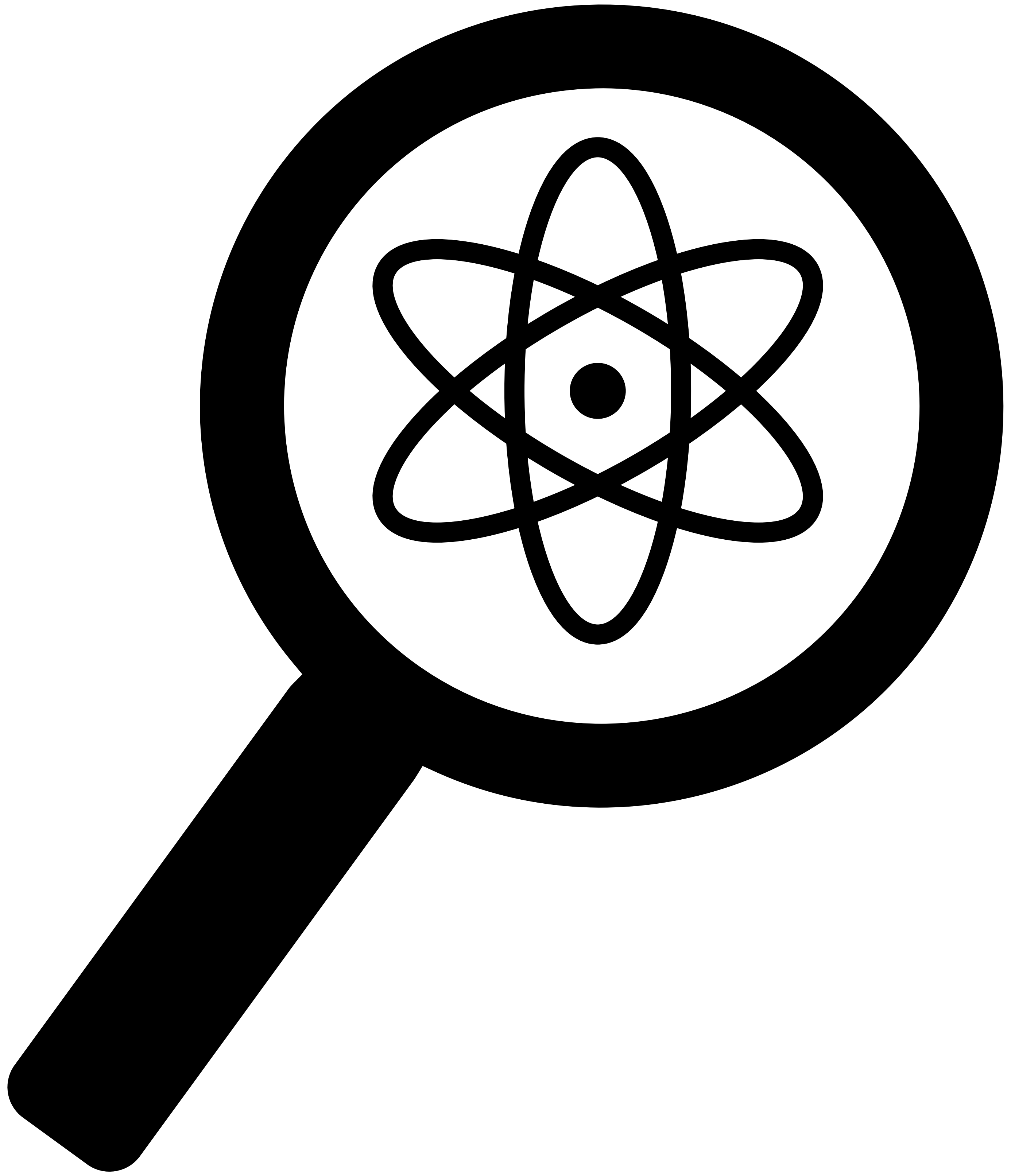


3. Precision frontier

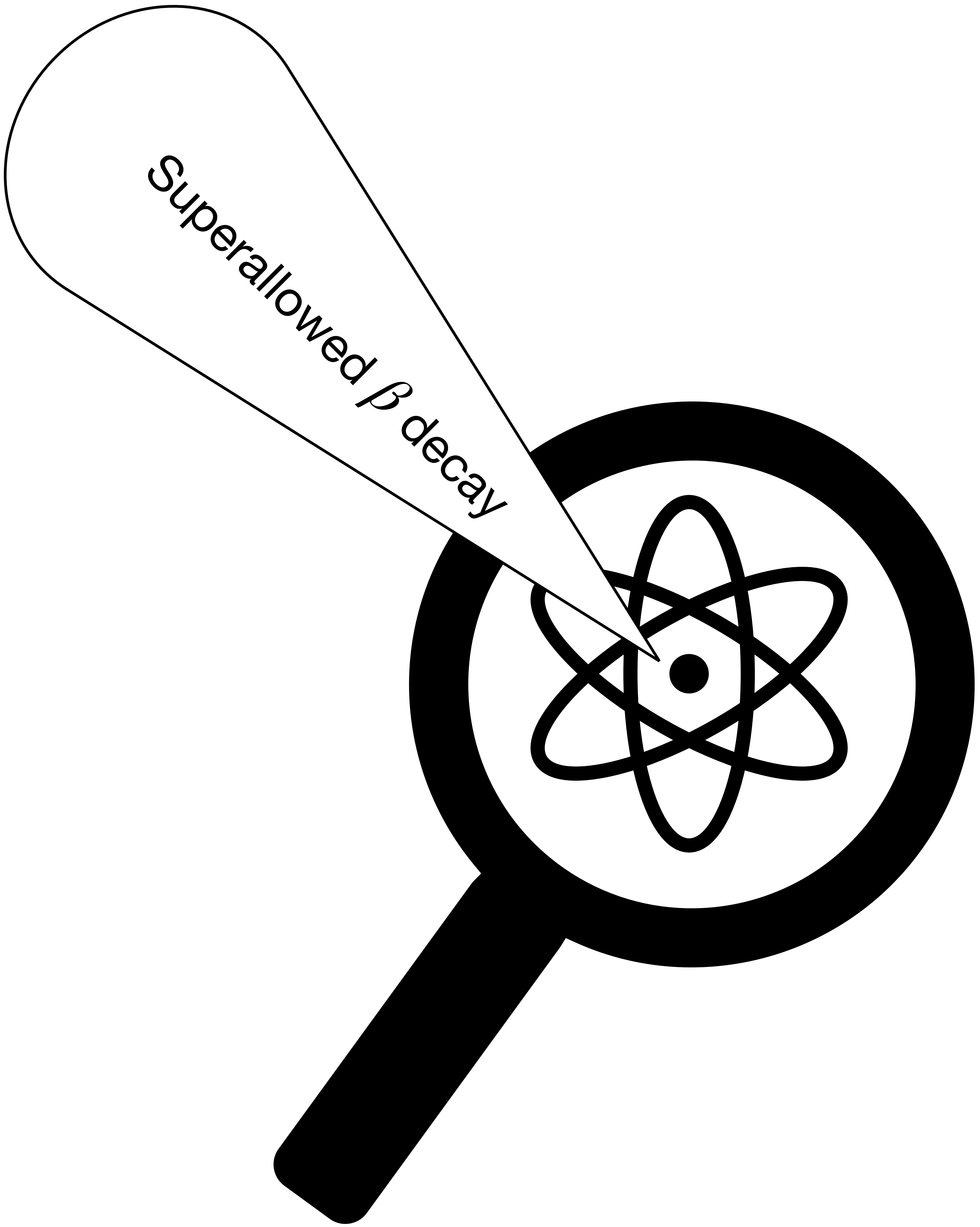


TeV Scale ⟷ < MeV Scale

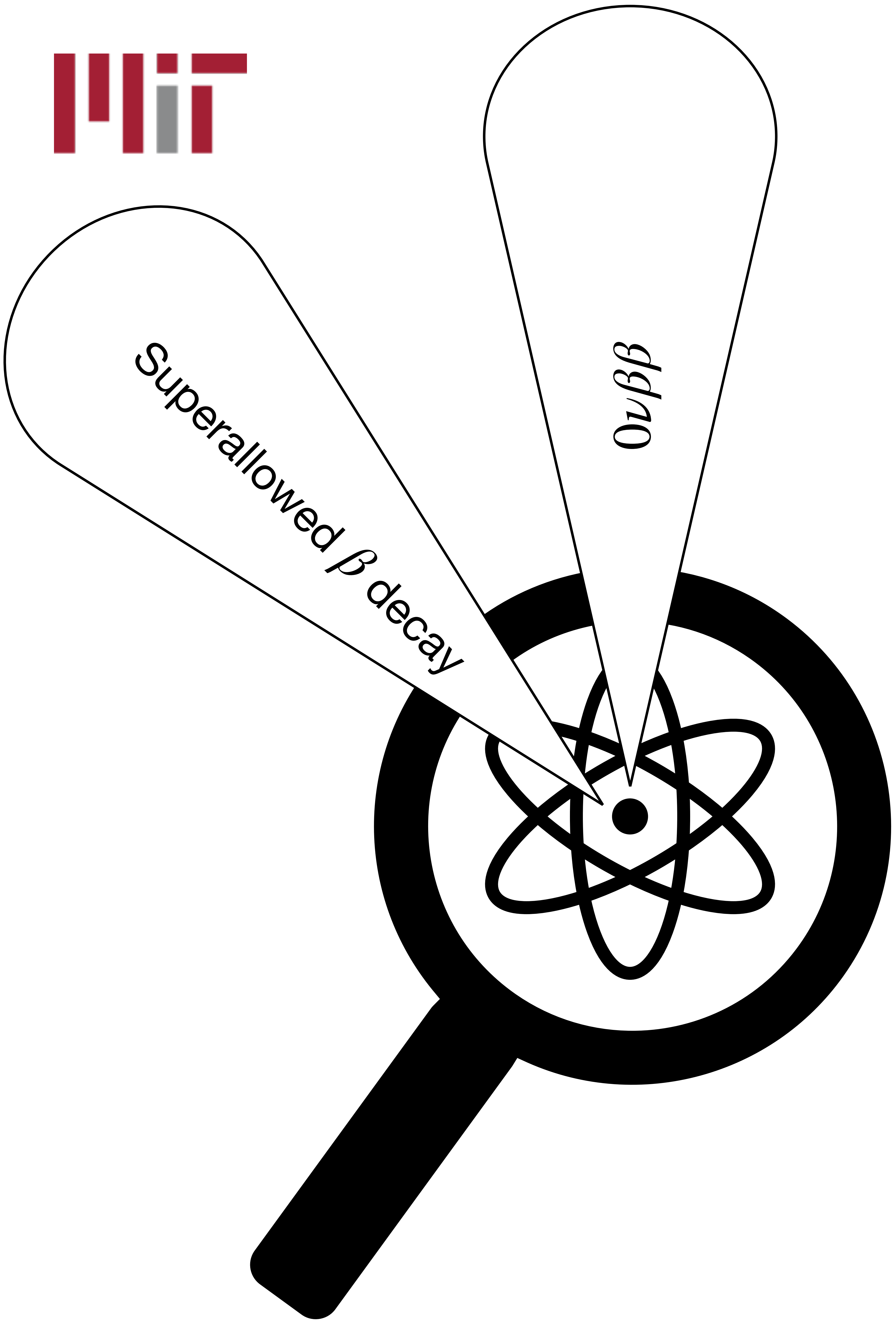
Atomic Nucleus as a Probe



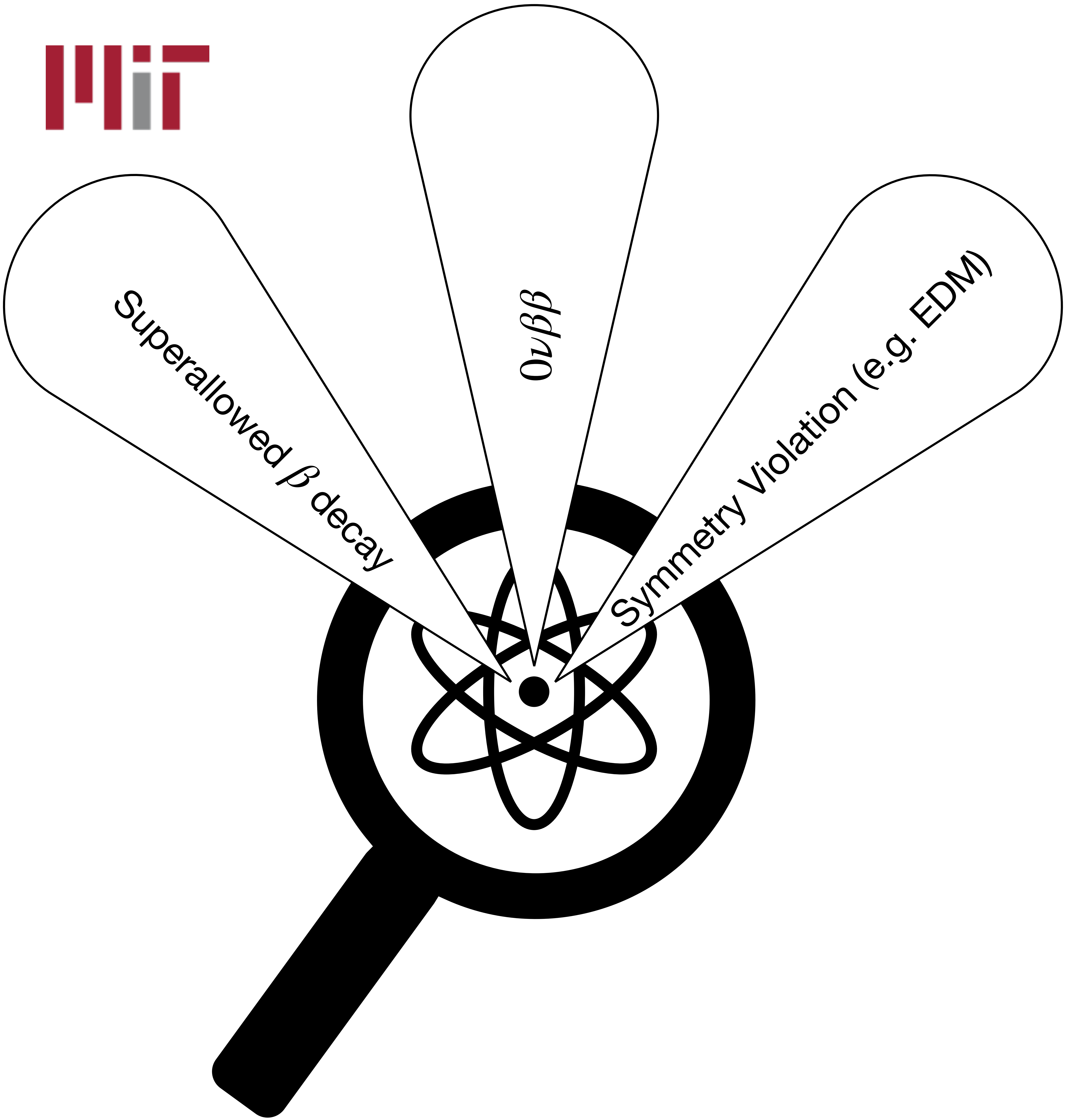
Atomic Nucleus as a Probe



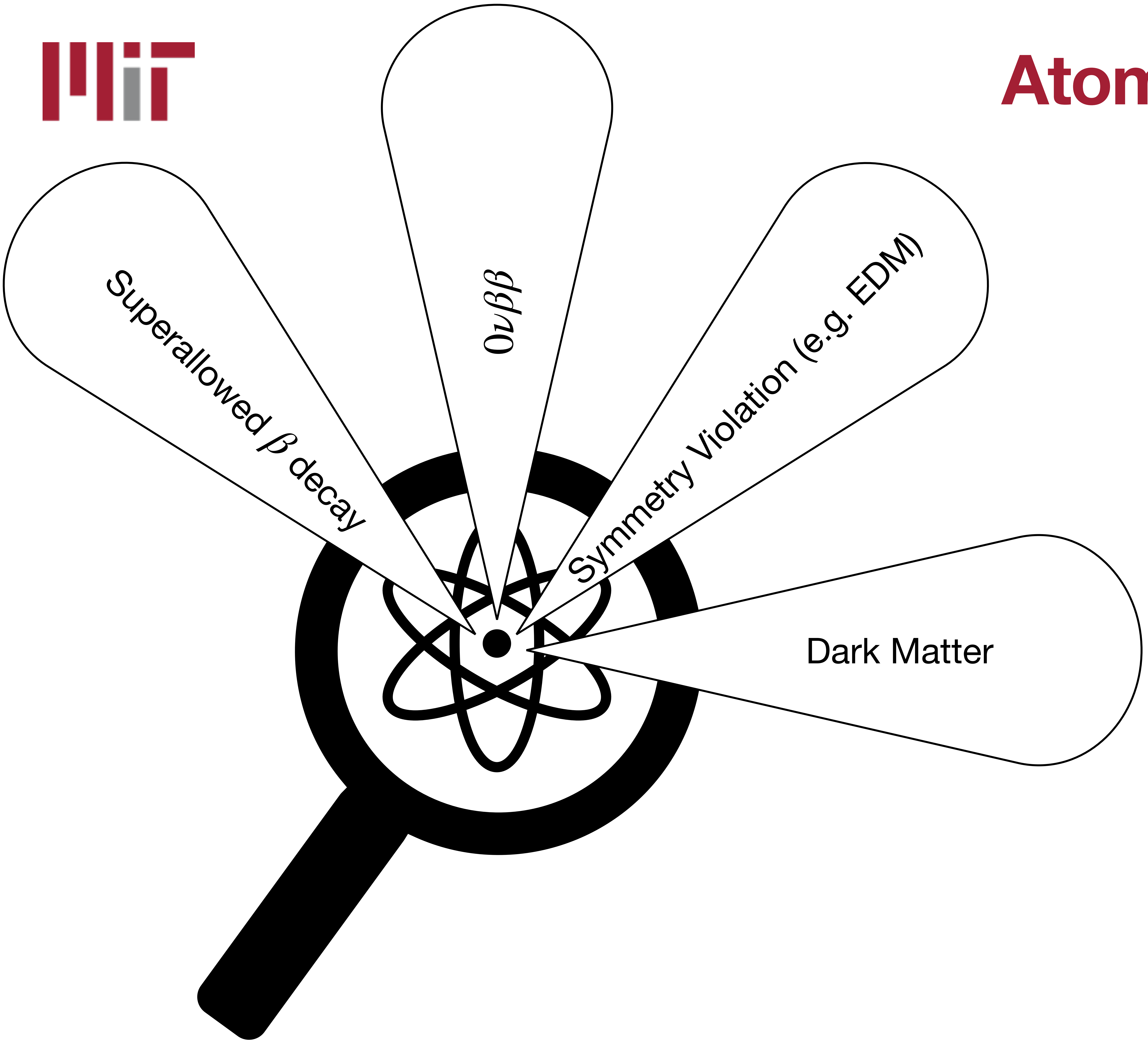
Atomic Nucleus as a Probe



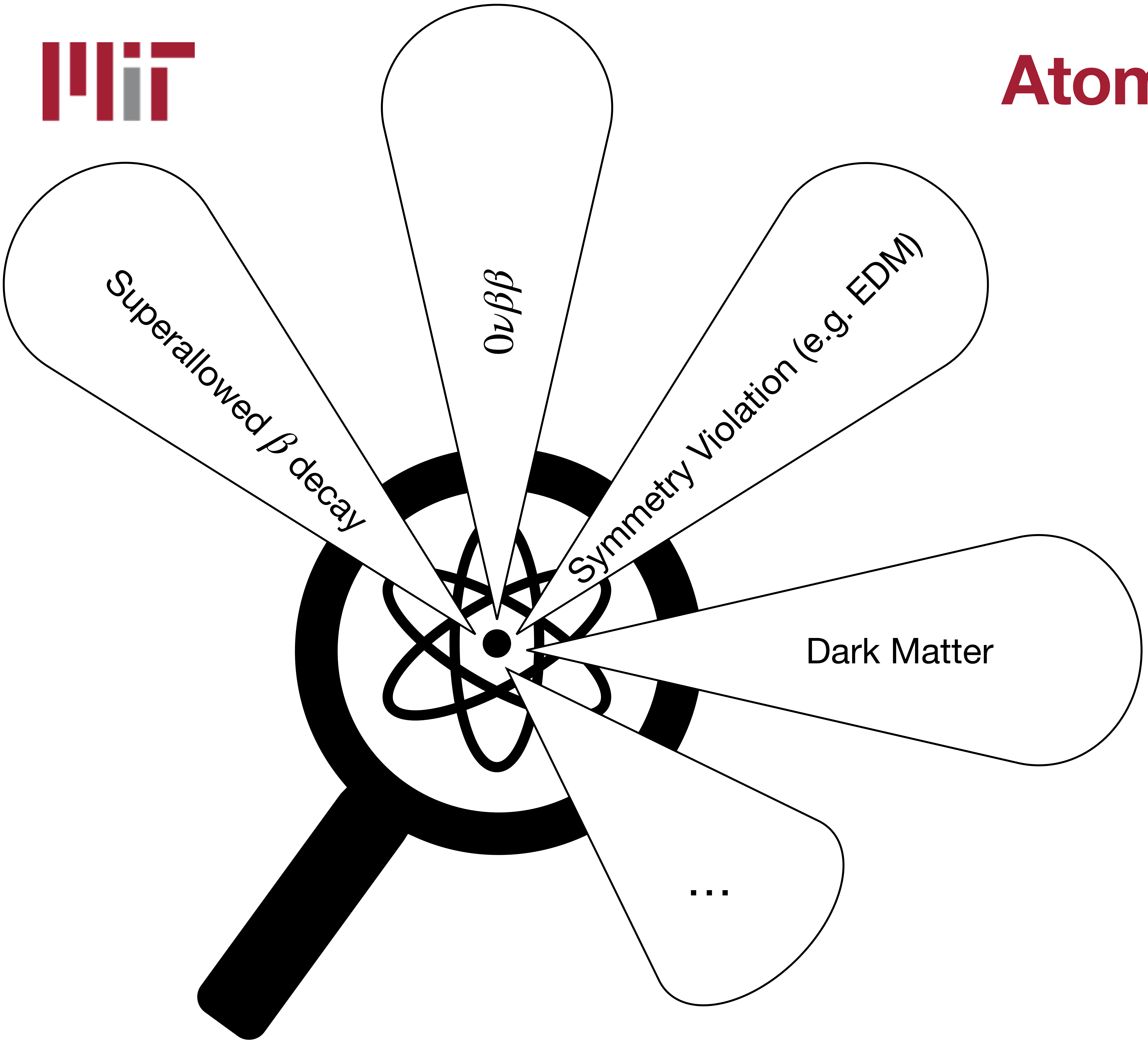
Atomic Nucleus as a Probe



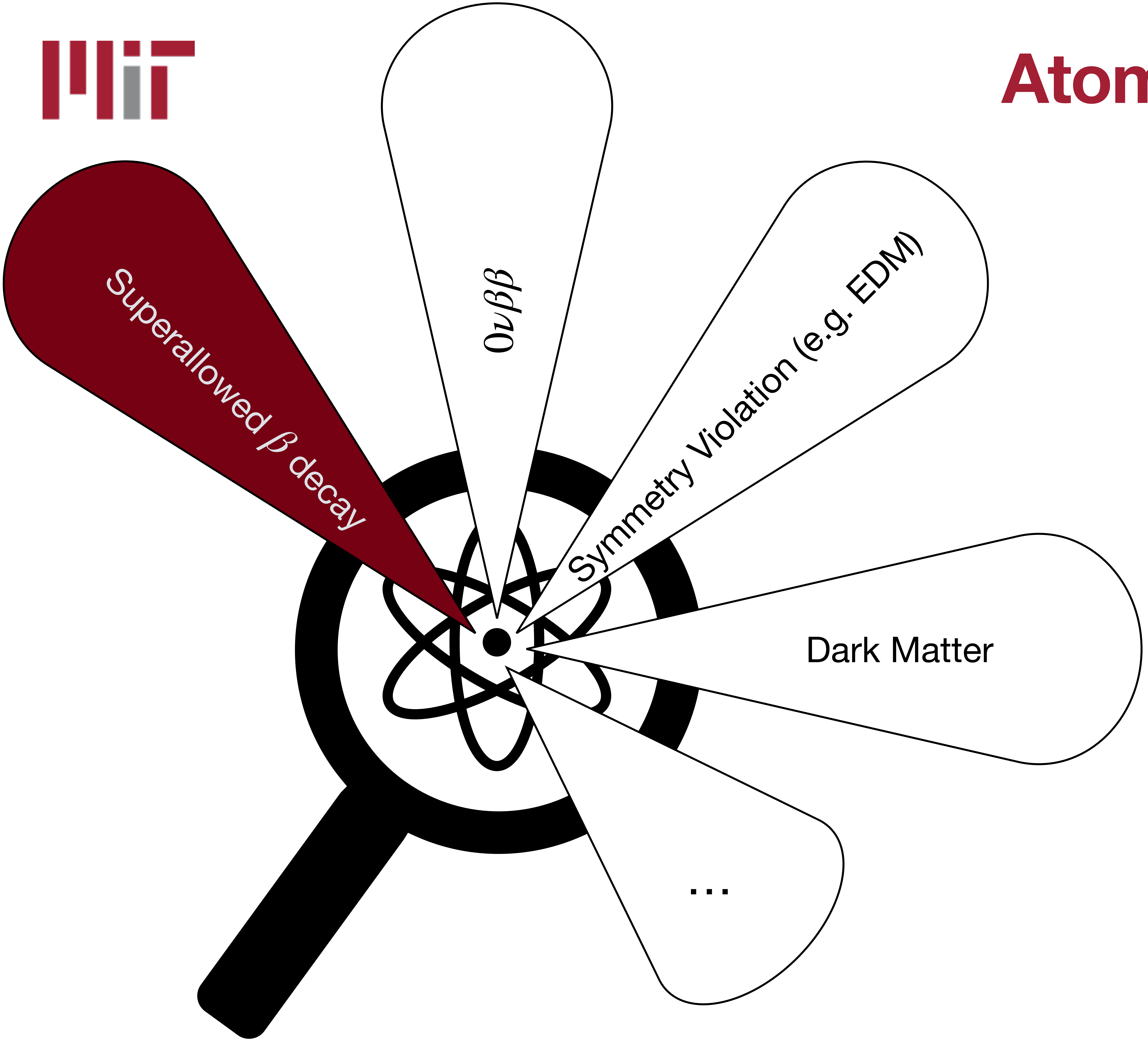
Atomic Nucleus as a Probe



Atomic Nucleus as a Probe

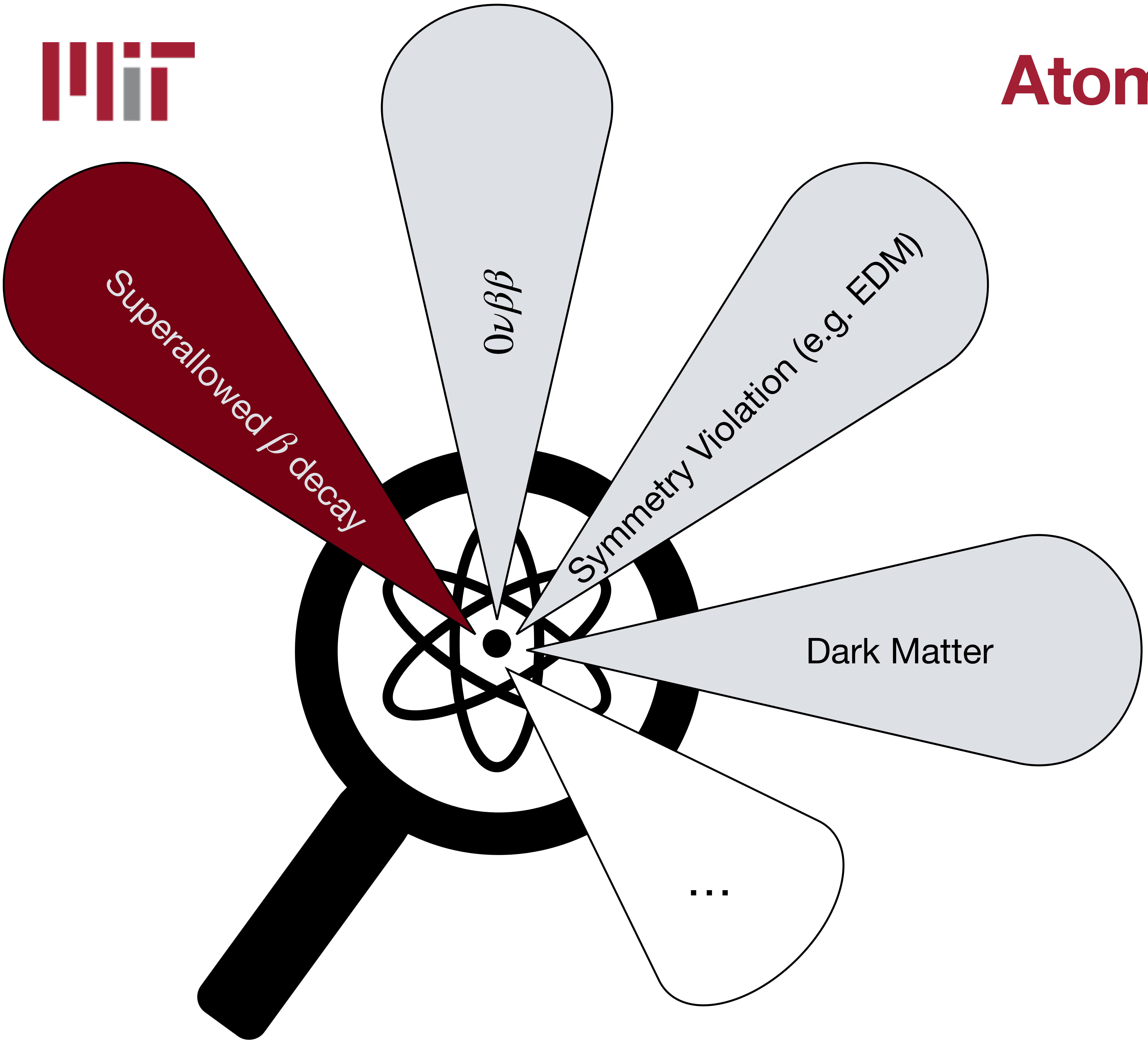


Atomic Nucleus as a Probe



1. Looking at deviation from Standard Model prediction.

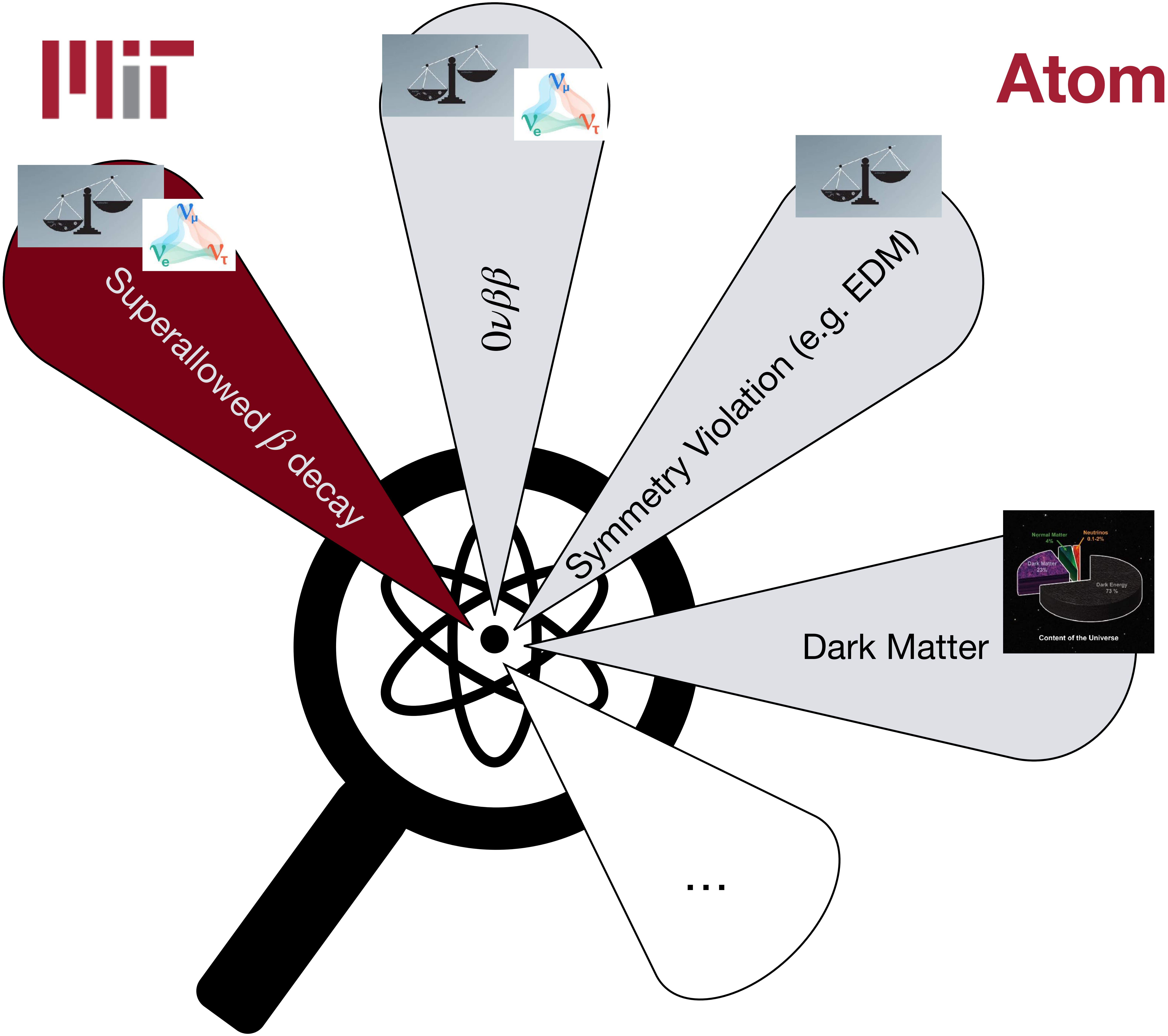
Atomic Nucleus as a Probe



1. Looking at deviation from Standard Model prediction.

2. Search for phenomena not predicted by the Standard Model.

Atomic Nucleus as a Probe



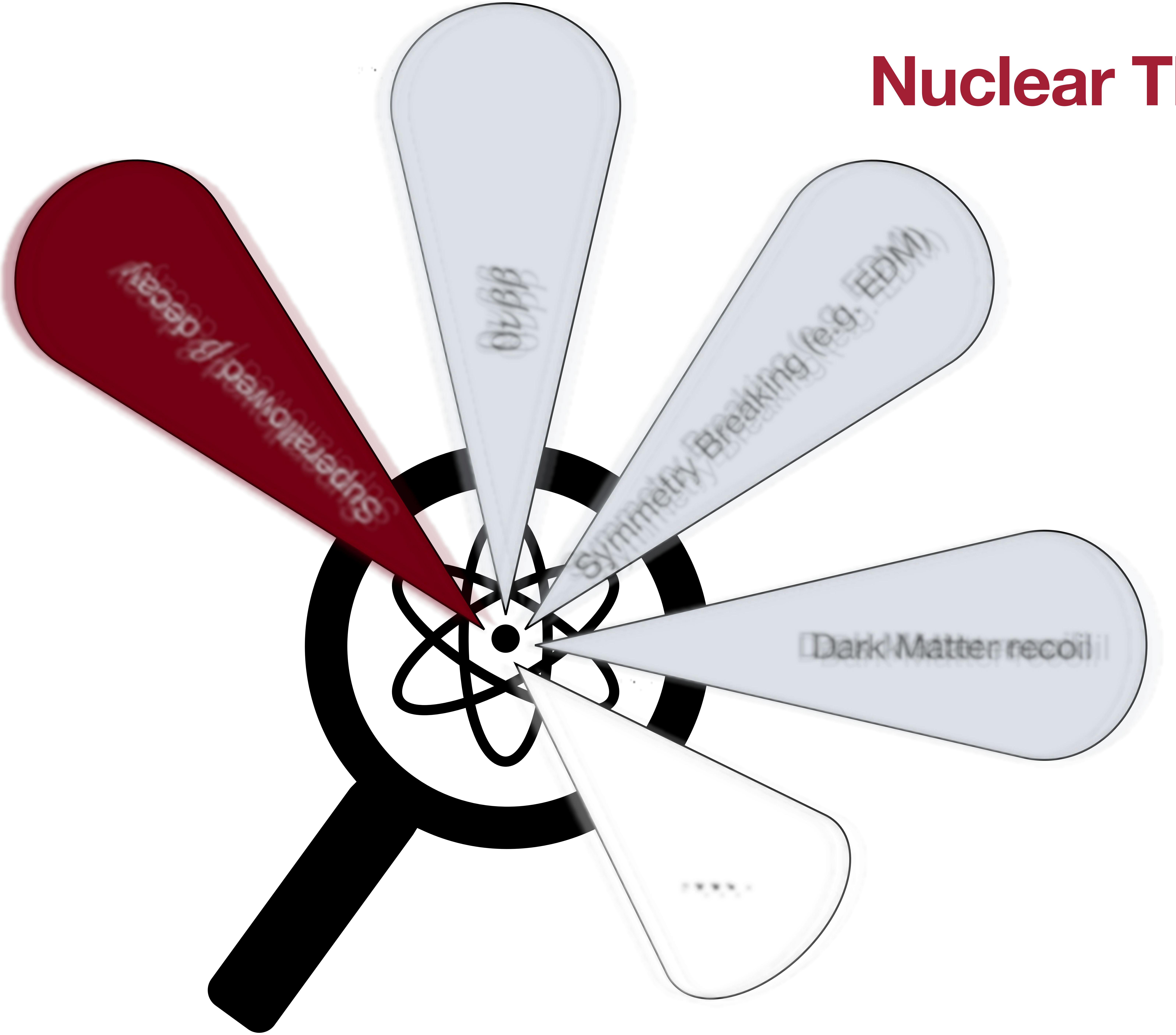
1. Looking at deviation from Standard Model prediction.
2. Search for phenomena not predicted by the Standard Model.

The Need for Nuclear Theory

In all cases, nuclear theory inputs are required in order to interpret experimental results:

- **Superallowed β decay:** Corrections from Standard Model δ_{NS} and δ_C.
- $\boldsymbol{0\nu\beta\beta}$**:** Nuclear matrix elements $M^{0\nu}$.
- **Electric Dipole Moment:** The nuclear Schiff Moment.
- **Dark Matter Scattering:** WIMP Scattering structure factor S_A.
- …

Nuclear Theory Blur



Nuclear Theory Challenges

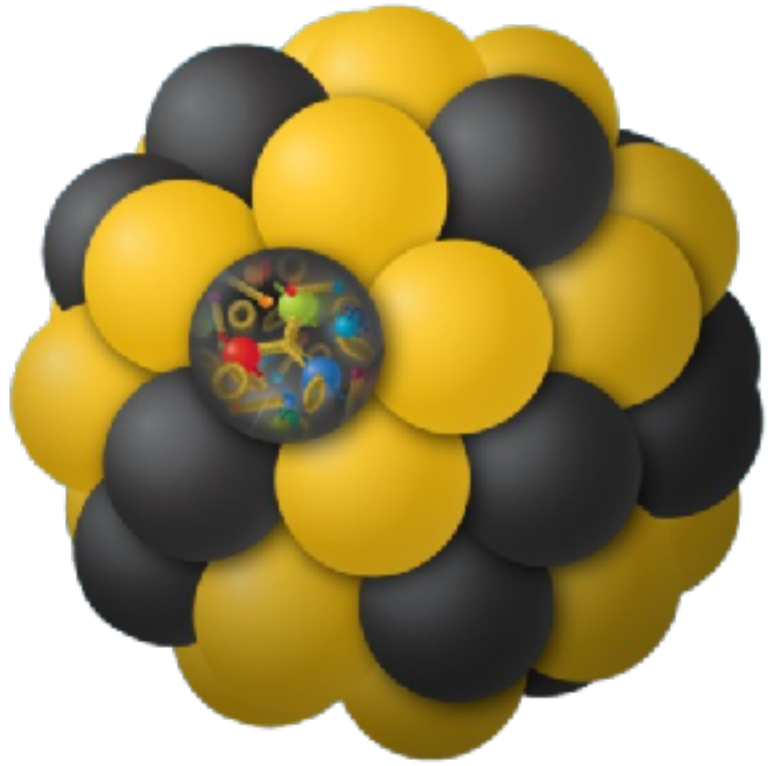
Understanding nuclear structure from microscopic physics

Figures courtesy of J. Muñoz

Nuclear Theory Challenges

Understanding nuclear structure from microscopic physics

Nuclear Interactions

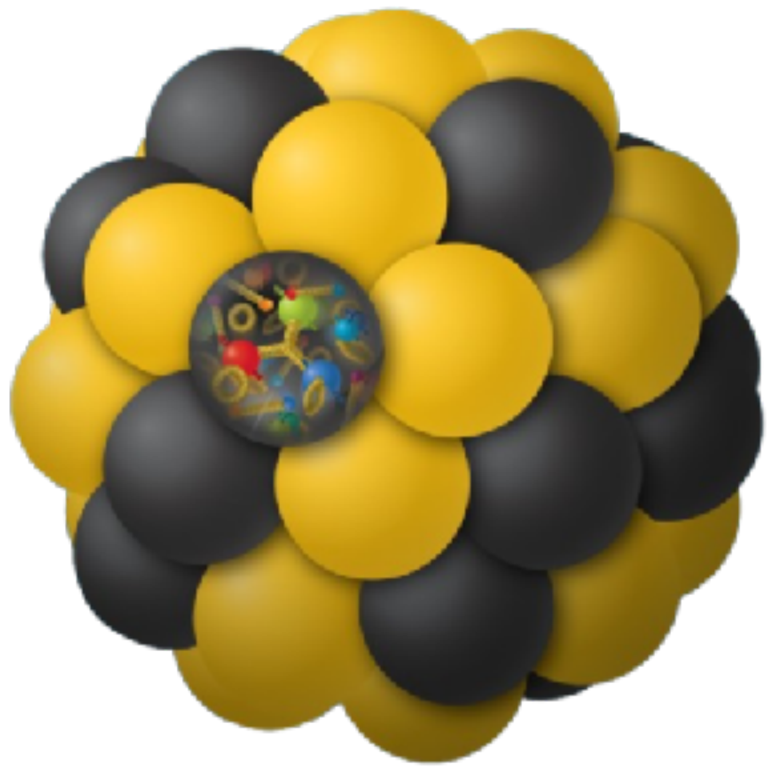


Figures courtesy of J. Muñoz

Nuclear Theory Challenges

Understanding nuclear structure from microscopic physics

Nuclear Interactions



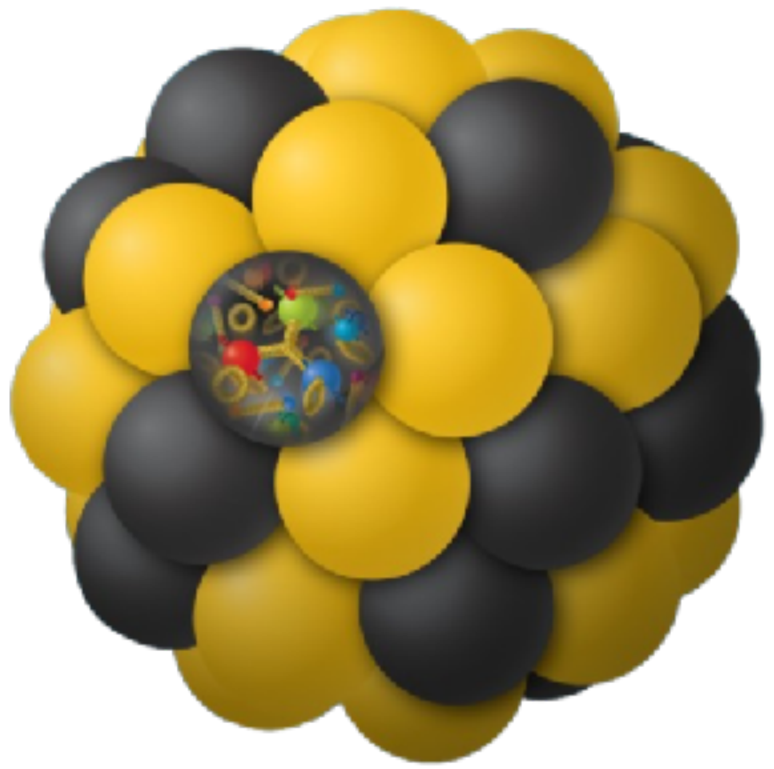
Nucleon-Nucleon force
is not known exactly!

Figures courtesy of J. Muñoz

Nuclear Theory Challenges

Understanding nuclear structure from microscopic physics

Nuclear Interactions



Nucleon-Nucleon force is not known exactly!

Only have parameterizations.

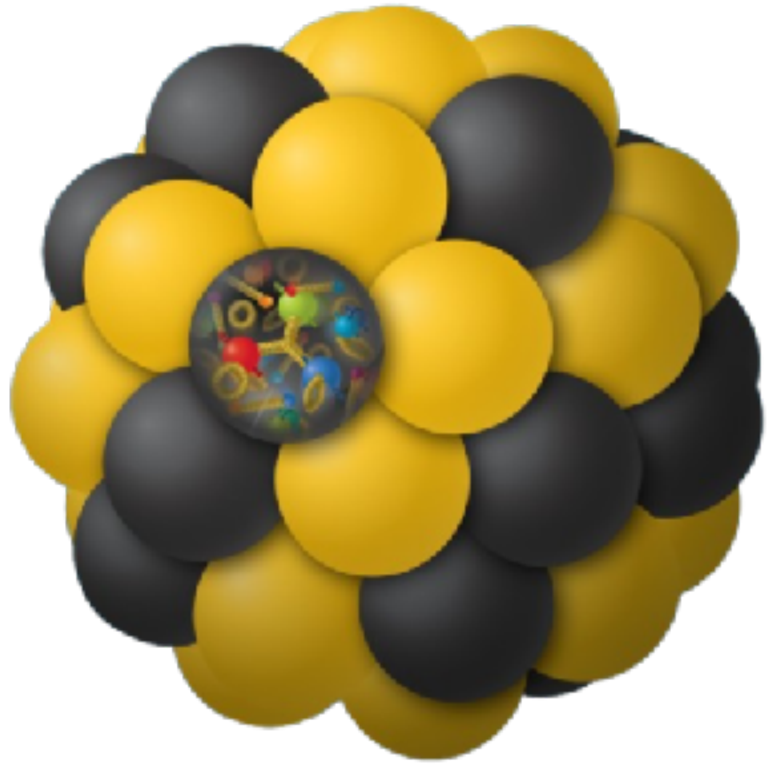
Figures courtesy of J. Muñoz

Nuclear Theory Challenges

Understanding nuclear structure from microscopic physics

Nuclear Interactions

Wave functions



$$H|\Psi\rangle = E|\Psi\rangle$$

Nucleon-Nucleon force
is not known exactly!

Only have
parameterizations.

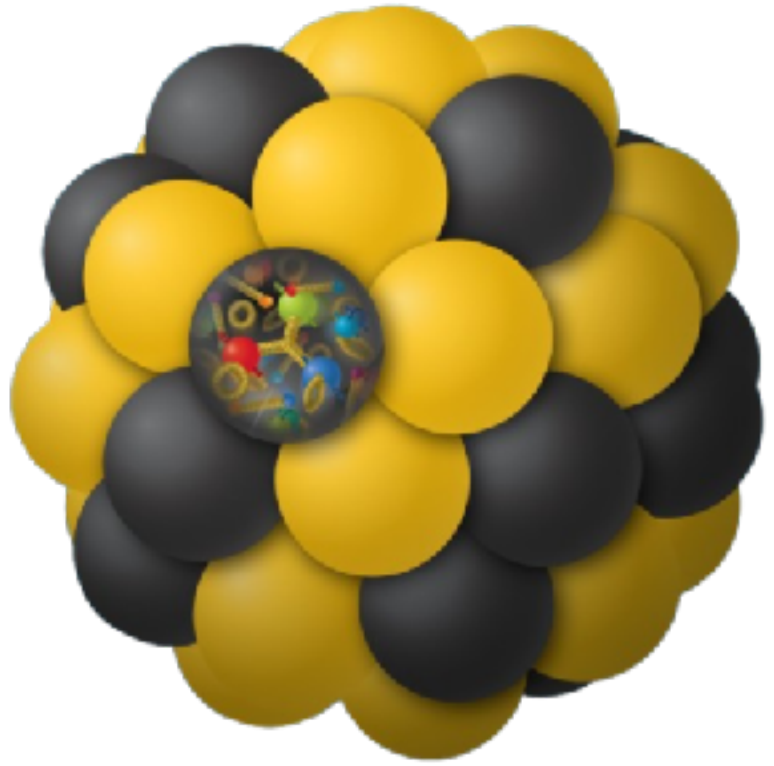
Figures courtesy of J. Muñoz

Nuclear Theory Challenges

Understanding nuclear structure from microscopic physics

Nuclear Interactions

Wave functions



$$H|\Psi\rangle = E|\Psi\rangle$$

Nucleon-Nucleon force is not known exactly!

Only have parameterizations.

Solving the nuclear A-body is computationally expensive…

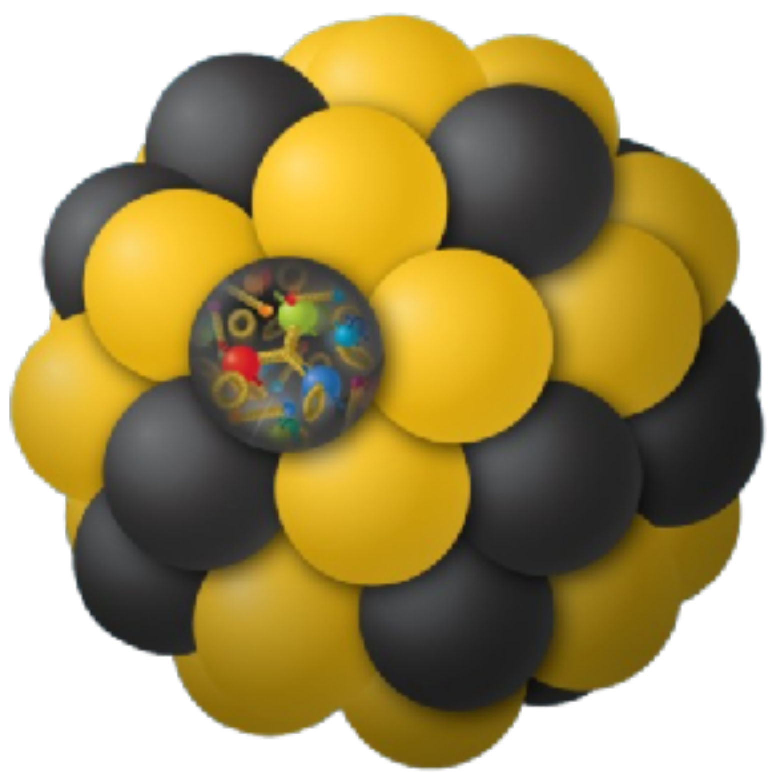
Figures courtesy of J. Muñoz

Nuclear Theory Challenges

Understanding nuclear structure from microscopic physics

Nuclear Interactions

Wave functions



$$H|\Psi\rangle = E|\Psi\rangle$$

Nucleon-Nucleon force is not known exactly!

Only have parameterizations.

Solving the nuclear A-body is computationally expensive…

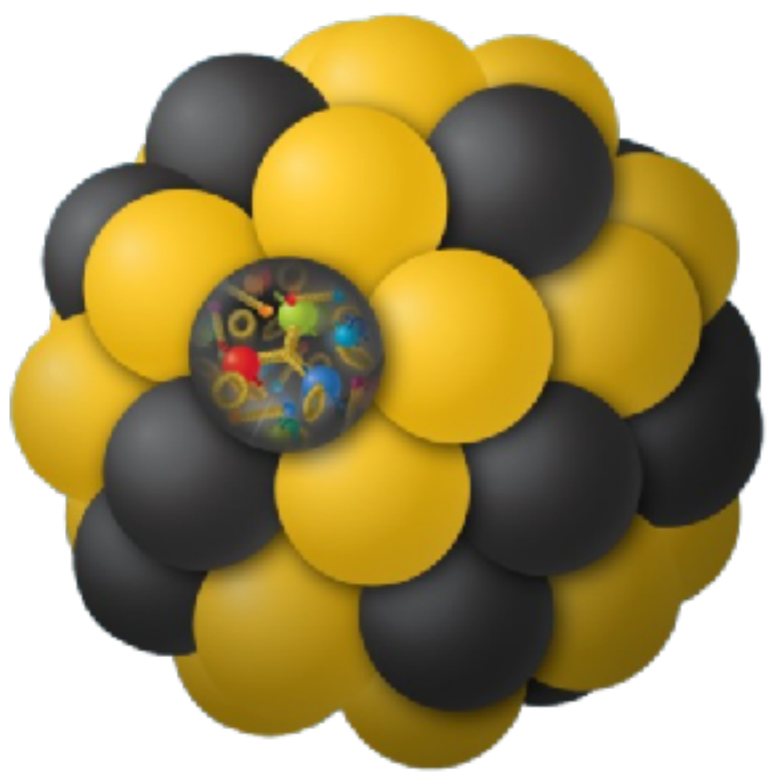
Exact solver ~O(e^A)
Approximate solver ~O(A^6)

Figures courtesy of J. Muñoz

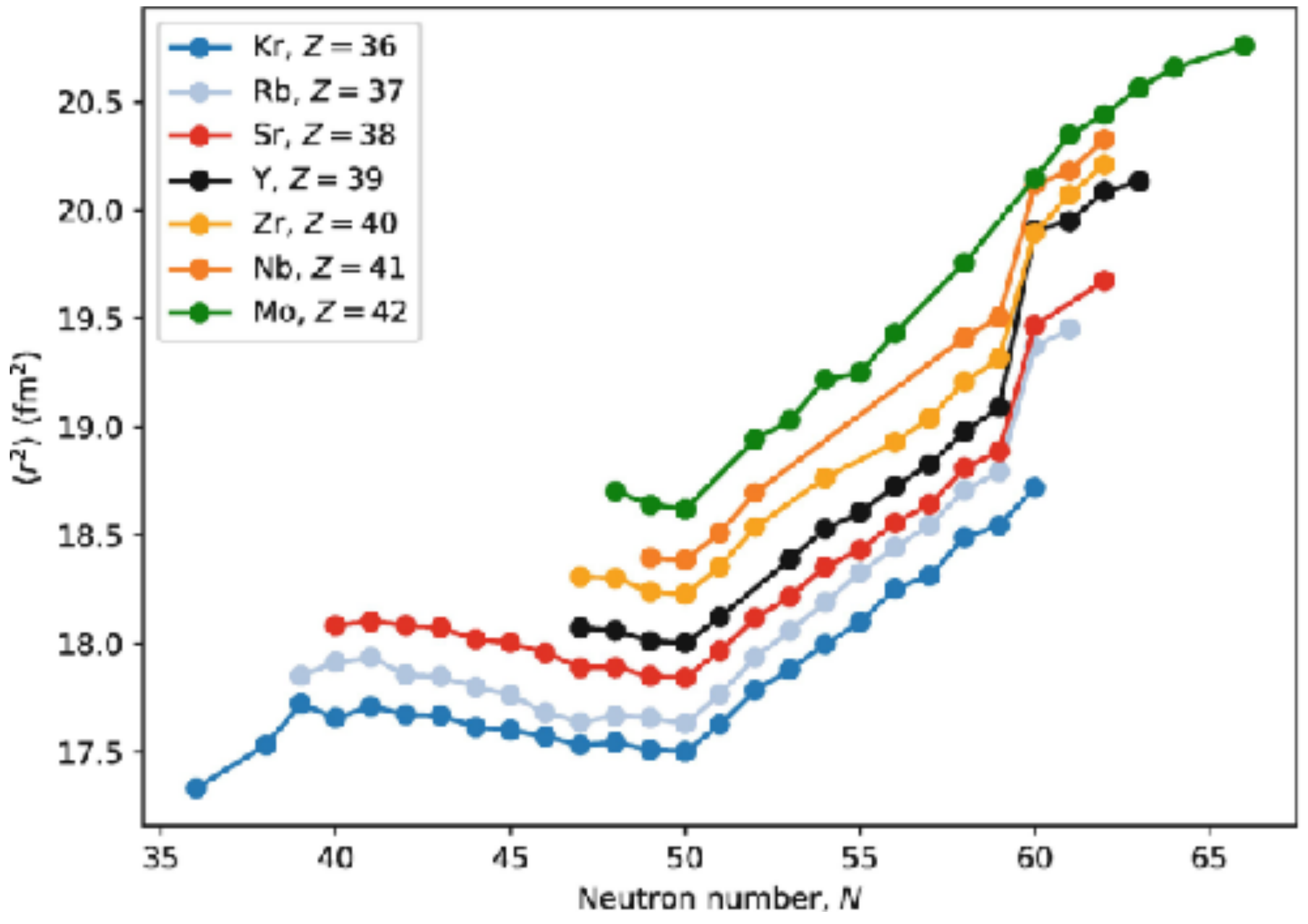
Nuclear Theory Challenges

Understanding nuclear structure from microscopic physics

Nuclear Interactions	Wave functions	Observables



$$H|\Psi\rangle = E|\Psi\rangle$$



$$\langle\Psi|O|\Psi\rangle$$

Nucleon-Nucleon force is not known exactly!

Only have parameterizations.

Solving the nuclear A-body is computationally expensive…

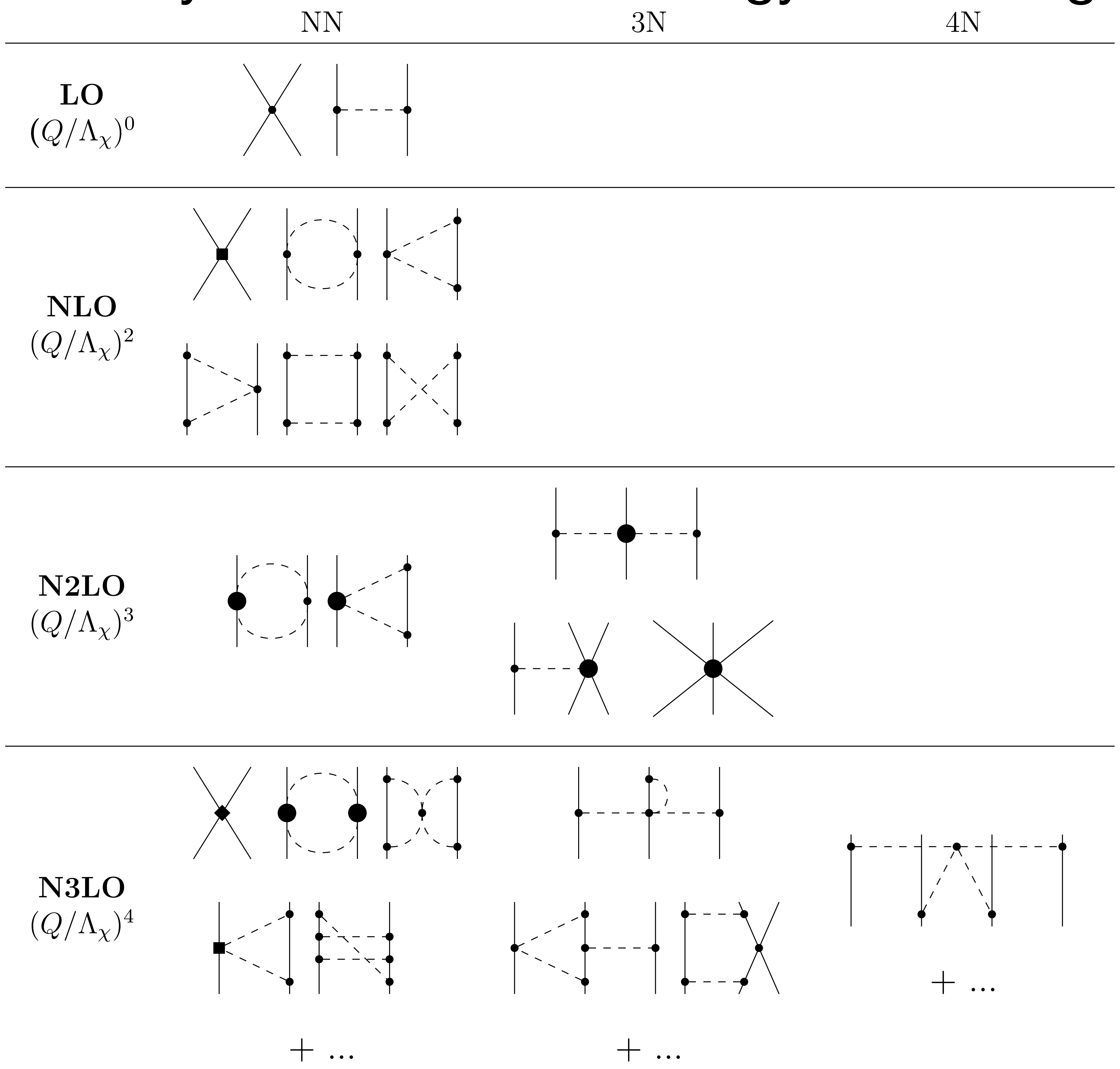
Exact solver ~O(e^A)
Approximate solver ~O(A^6)

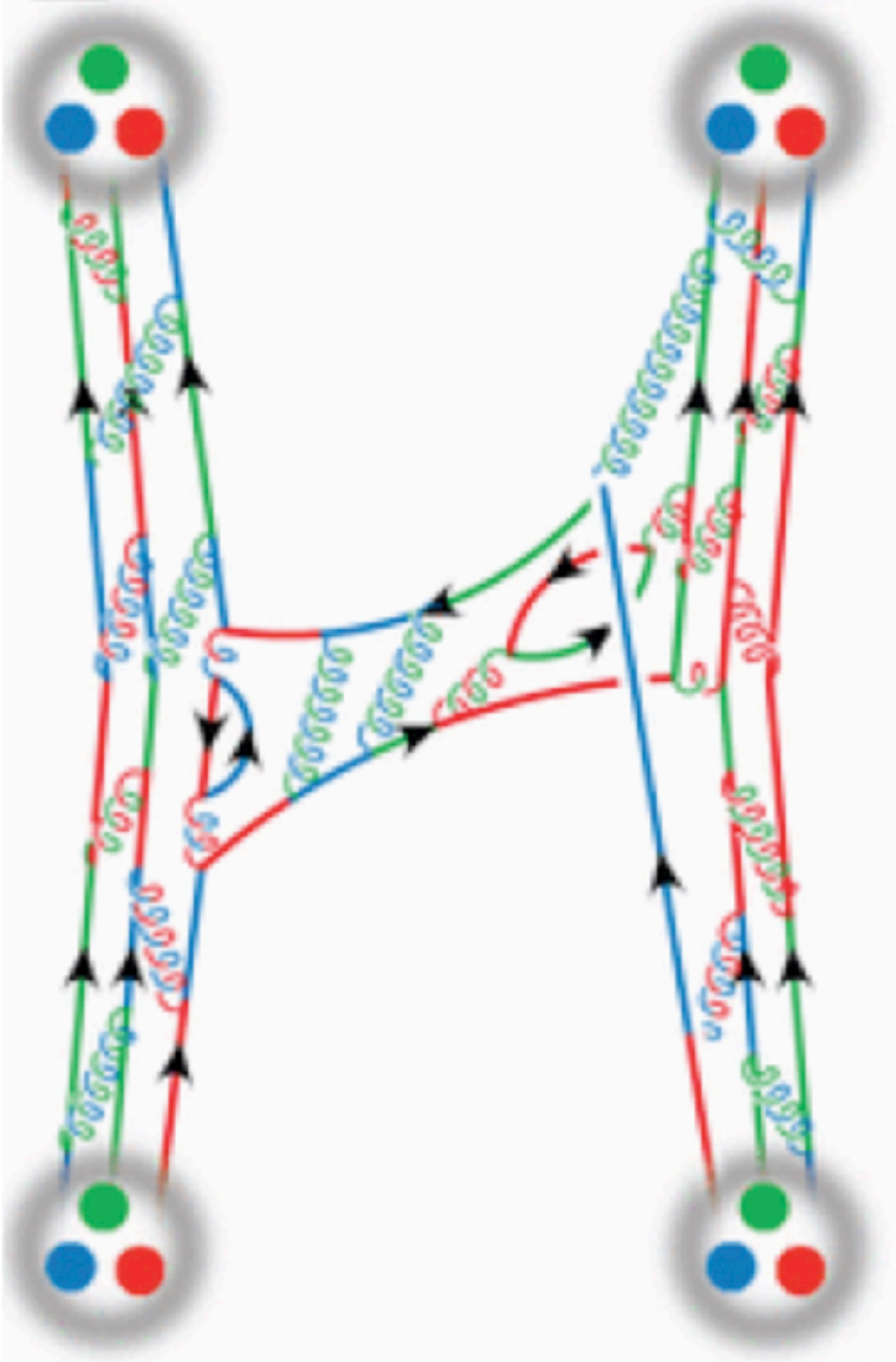
Figures courtesy of J. Muñoz

χ-EFT

Expansion order by order of the nuclear forces

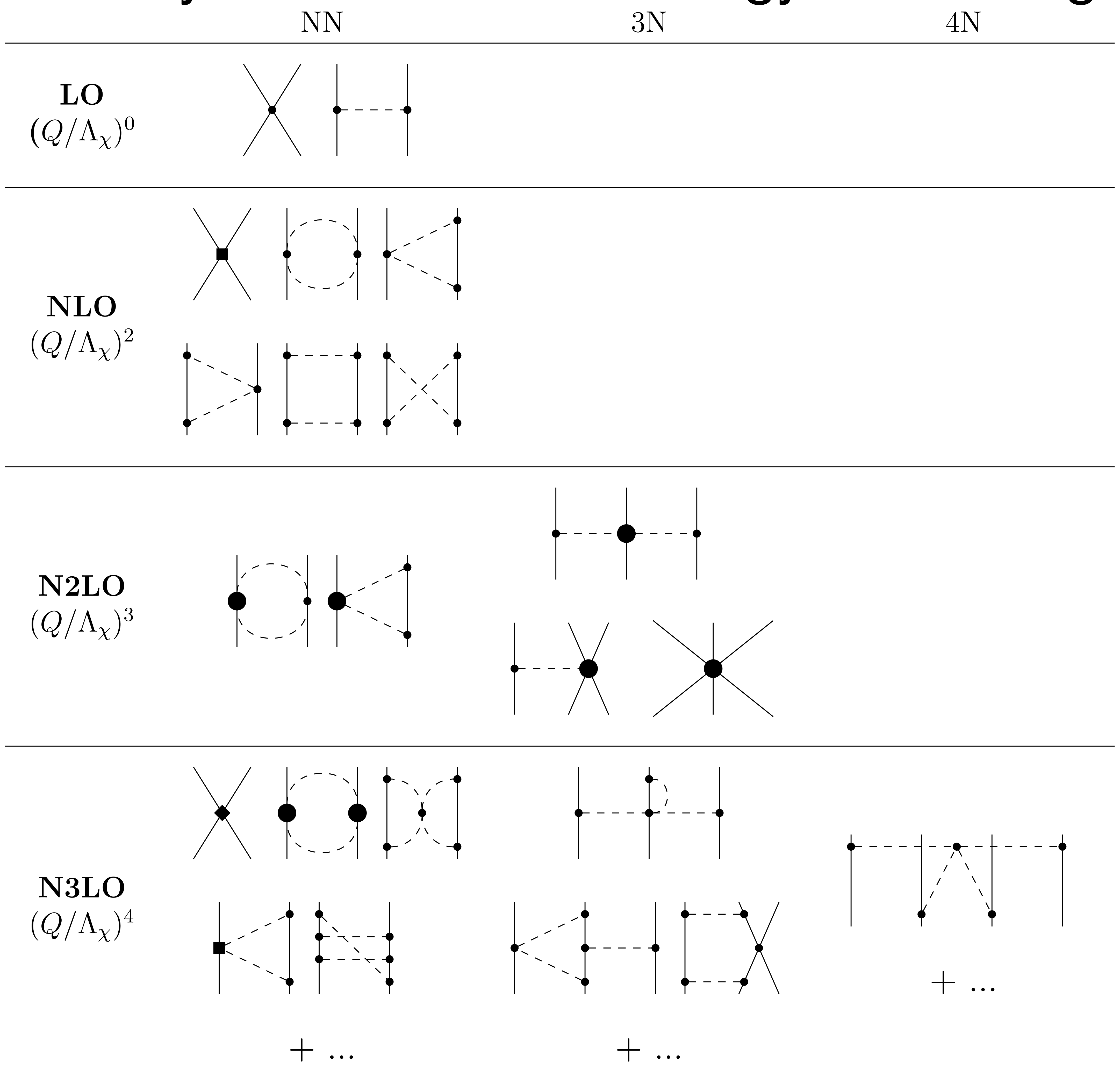
Reproduces symmetries of low-energy QCD using nucleons as fields and pions as force carriers.





Expansion order by order of the nuclear forces

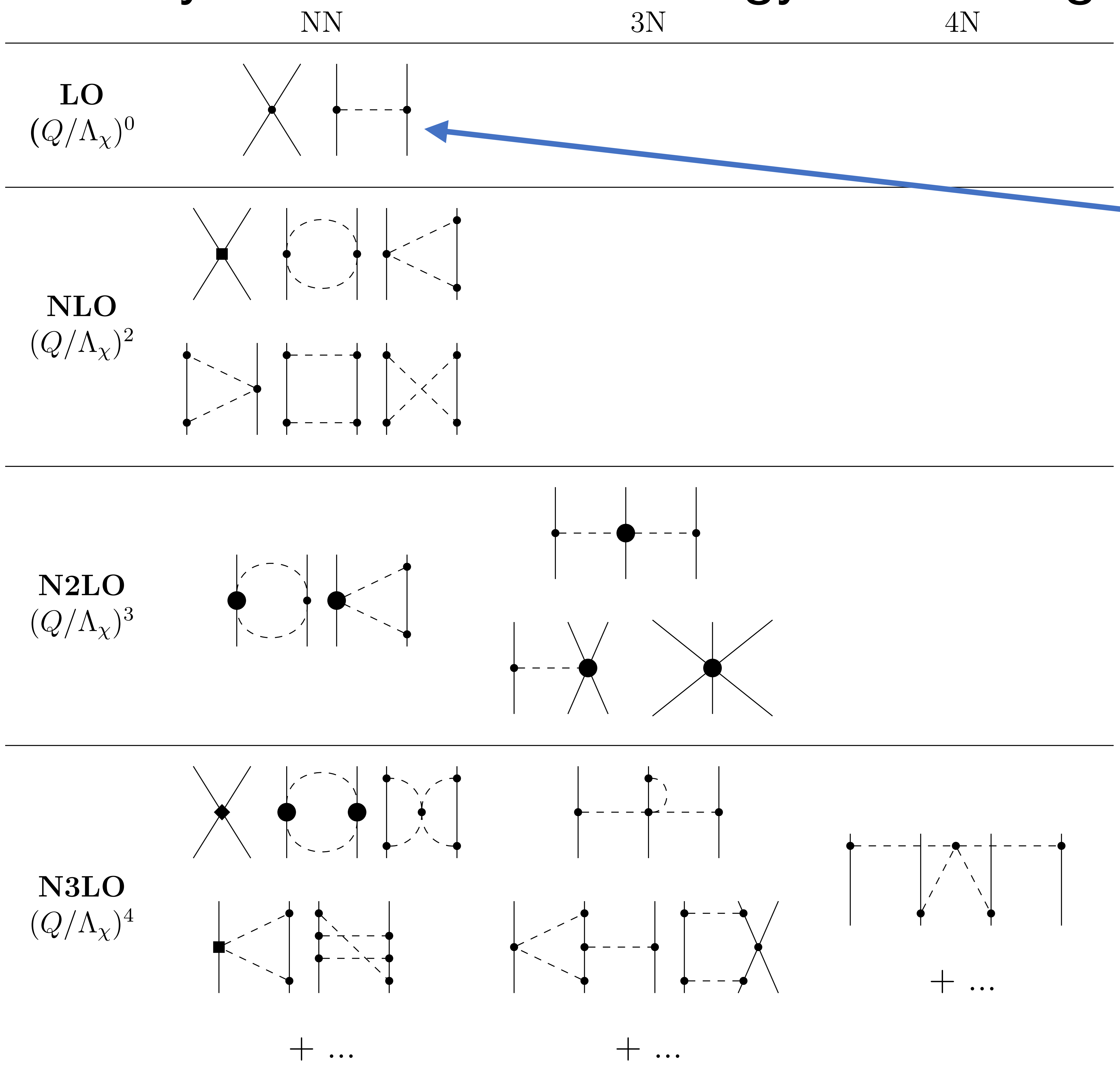
Reproduces symmetries of low-energy QCD using nucleons as fields and pions as force carriers.



χ-EFT

Expansion order by order of the nuclear forces

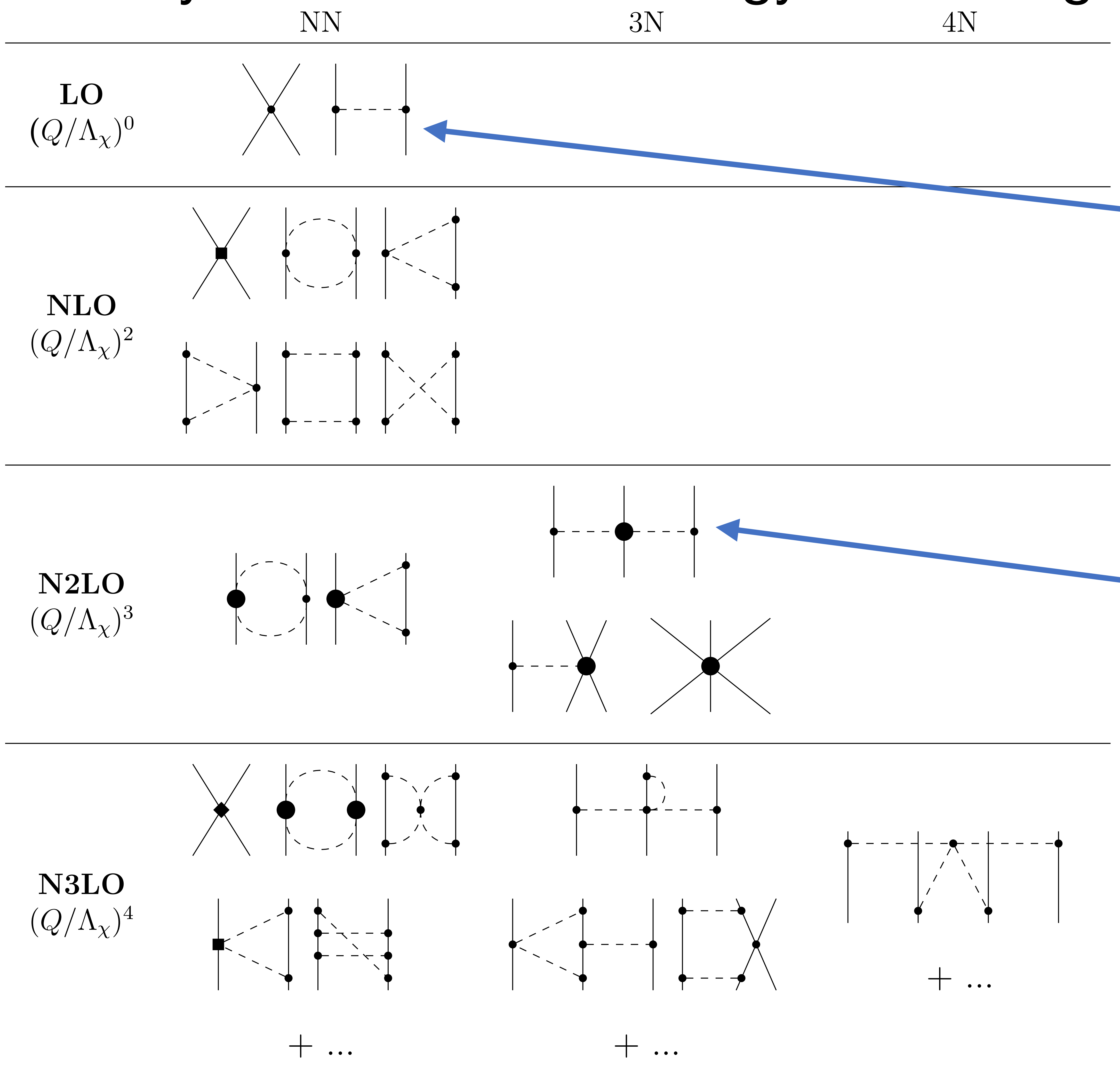
Reproduces symmetries of low-energy QCD using nucleons as fields and pions as force carriers.



The different low energy coupling constants (LECs) are fitted to few-nucleon data to absorb the effect of higher order terms χ-EFT

Expansion order by order of the nuclear forces

Reproduces symmetries of low-energy QCD using nucleons as fields and pions as force carriers.



The different low energy coupling constants (LECs) are fitted to few-nucleon data to absorb the effect of higher order terms

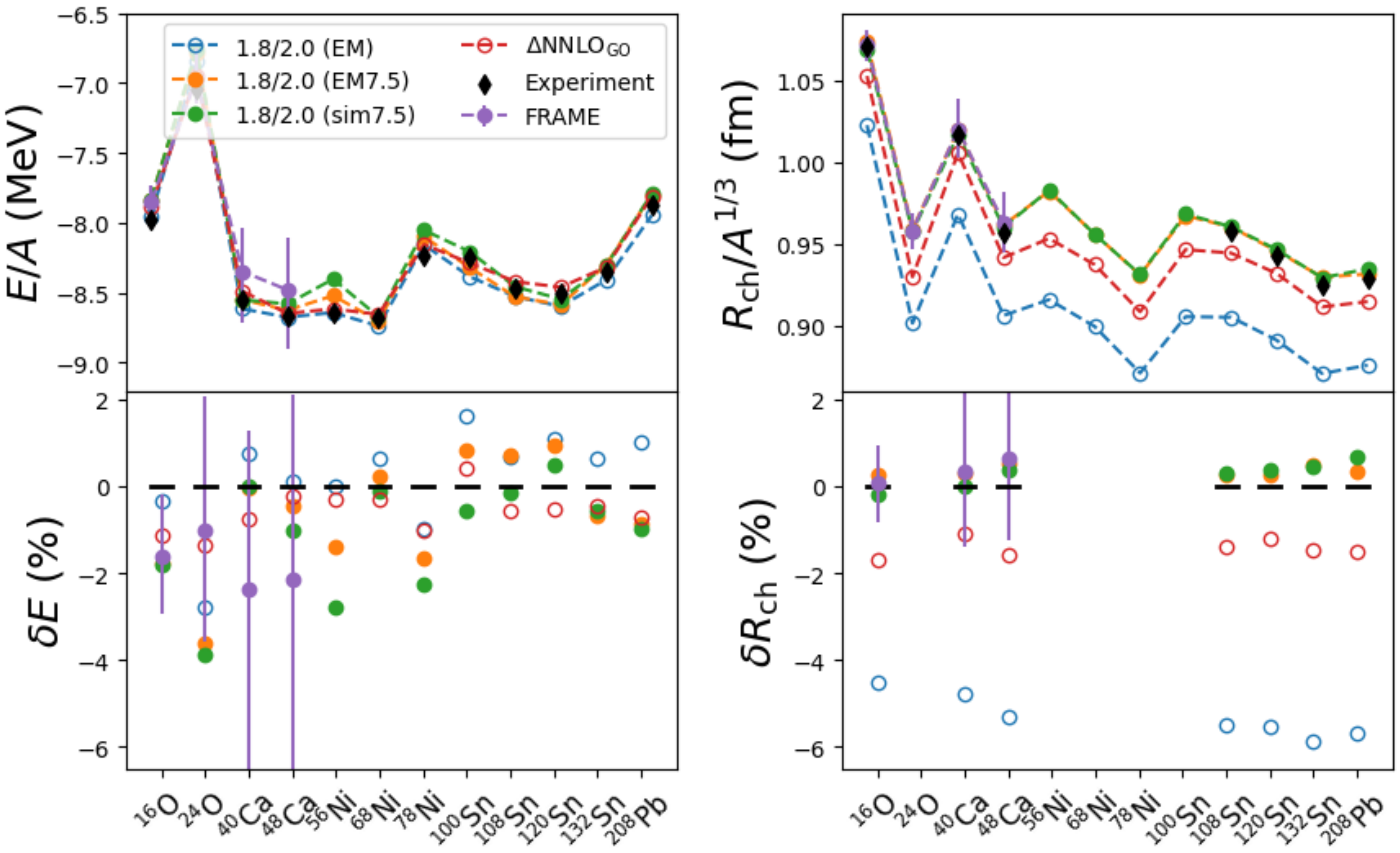
Three- (and higher-)body forces needed

Issues with Hamiltonians

- Large variance for many observables.

Issues with Hamiltonians

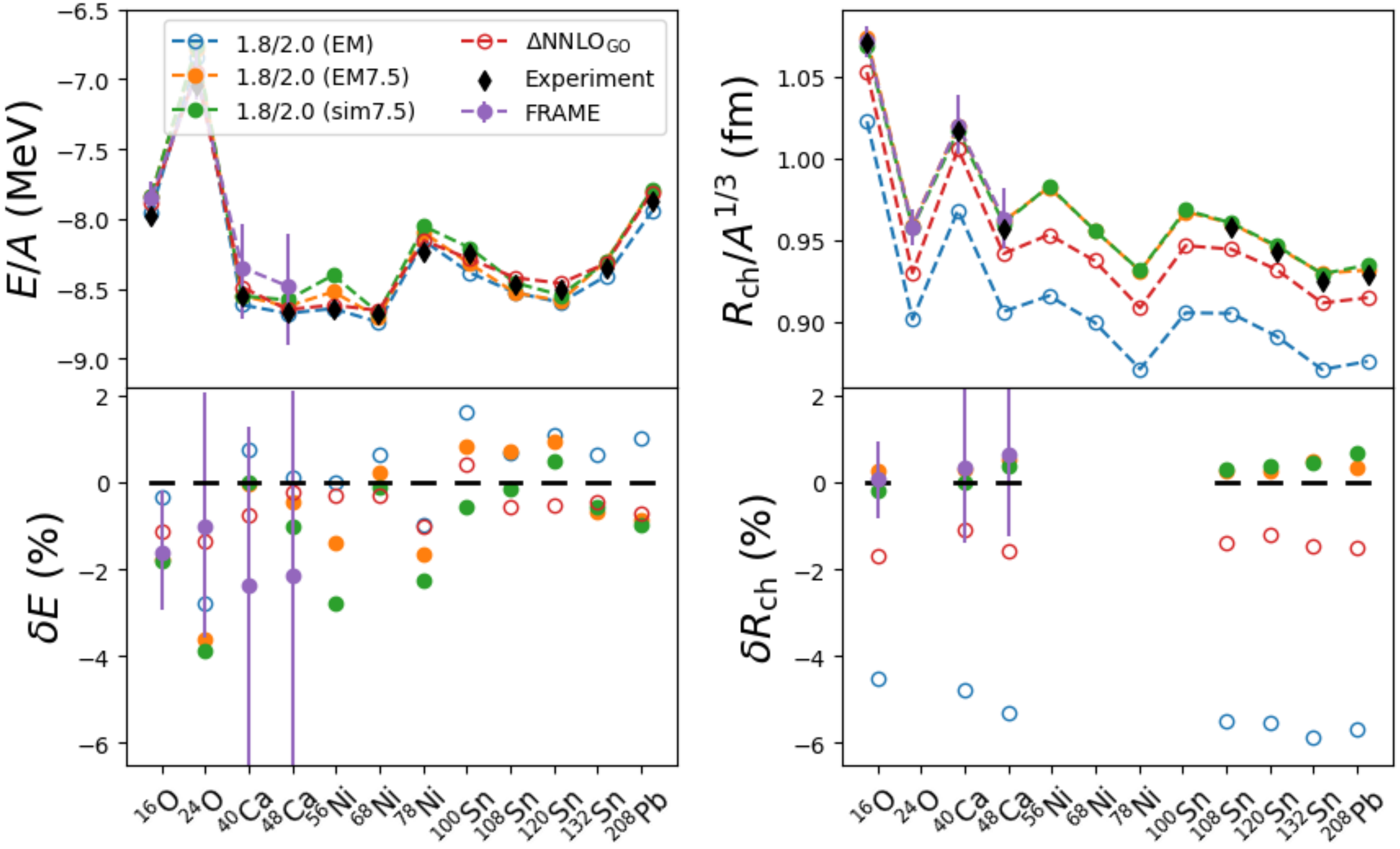
- Large variance for many observables.

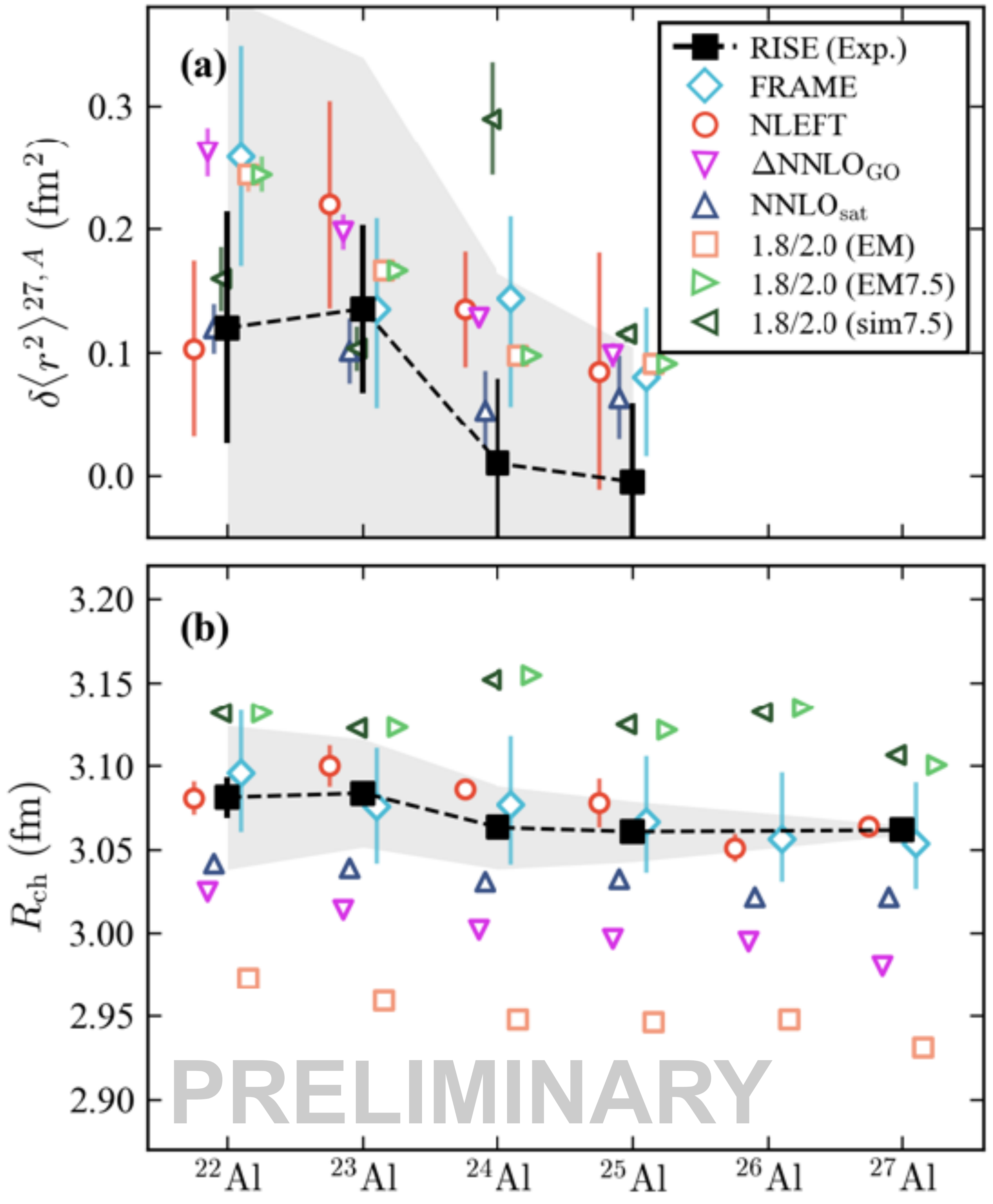


Arthuis, et al., arXiv:2401.06675 (2024)

Issues with Hamiltonians

- Large variance for many observables.





Arthuis, et al., arXiv:2401.06675 (2024)

Issues with Hamiltonians

Issues with Hamiltonians

- Large variance for many observables.
- Little understanding of how choices of parametrization affect results across nuclei chart

Issues with Hamiltonians

- Large variance for many observables.
- Little understanding of how choices of parametrization affect results across nuclei chart
- Lack of unified description:

Issues with Hamiltonians

- Large variance for many observables.
- Little understanding of how choices of parametrization affect results across nuclei chart
- Lack of unified description:
    - Cutoffs

Issues with Hamiltonians

- Large variance for many observables.
- Little understanding of how choices of parametrization affect results across nuclei chart
- Lack of unified description:
  - Cutoffs
  - EFT orders

Issues with Hamiltonians

- Large variance for many observables.
- Little understanding of how choices of parametrization affect results across nuclei chart
- Lack of unified description:
    - Cutoffs
    - EFT orders
    - SRG

Issues with Hamiltonians

- Large variance for many observables.
- Little understanding of how choices of parametrization affect results across nuclei chart
- Lack of unified description:
    - Cutoffs
    - EFT orders
    - SRG
    - Power Counting

Issues with Hamiltonians

- Large variance for many observables.
- Little understanding of how choices of parametrization affect results across nuclei chart
- Lack of unified description:
  - Cutoffs
  - EFT orders
  - SRG
  - Power Counting
  - LEC fitting strategy

Issues with Hamiltonians

- Large variance for many observables.
- Little understanding of how choices of parametrization affect results across nuclei chart
- Lack of unified description:
    - Cutoffs
    - EFT orders
    - SRG
    - Power Counting
    - LEC fitting strategy
    - Terms included

Issues with Hamiltonians

- Large variance for many observables.
- Little understanding of how choices of parametrization affect results across nuclei chart
- Lack of unified description:
    - Cutoffs
    - EFT orders
    - SRG
    - Power Counting
    - LEC fitting strategy
    - Terms included
    - …

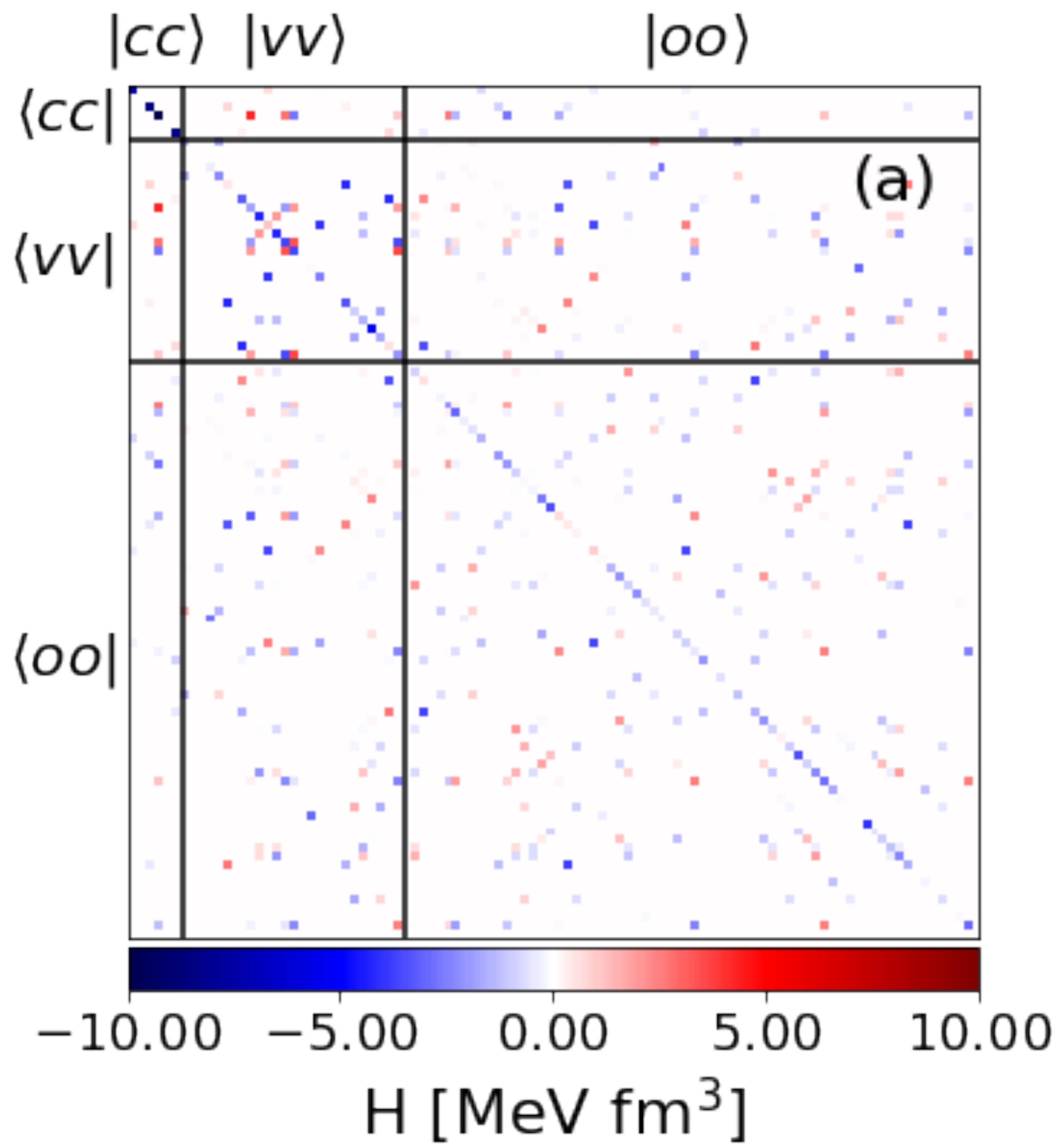
Issues with Hamiltonians

- Large variance for many observables.
- Little understanding of how choices of parametrization affect results across nuclei chart
- Lack of unified description:
  - Cutoffs
  - EFT orders
  - SRG
  - Power Counting
  - LEC fitting strategy
  - Terms included
  - …

Hard/impossible to consider all interactions together to assign a robust uncertainty

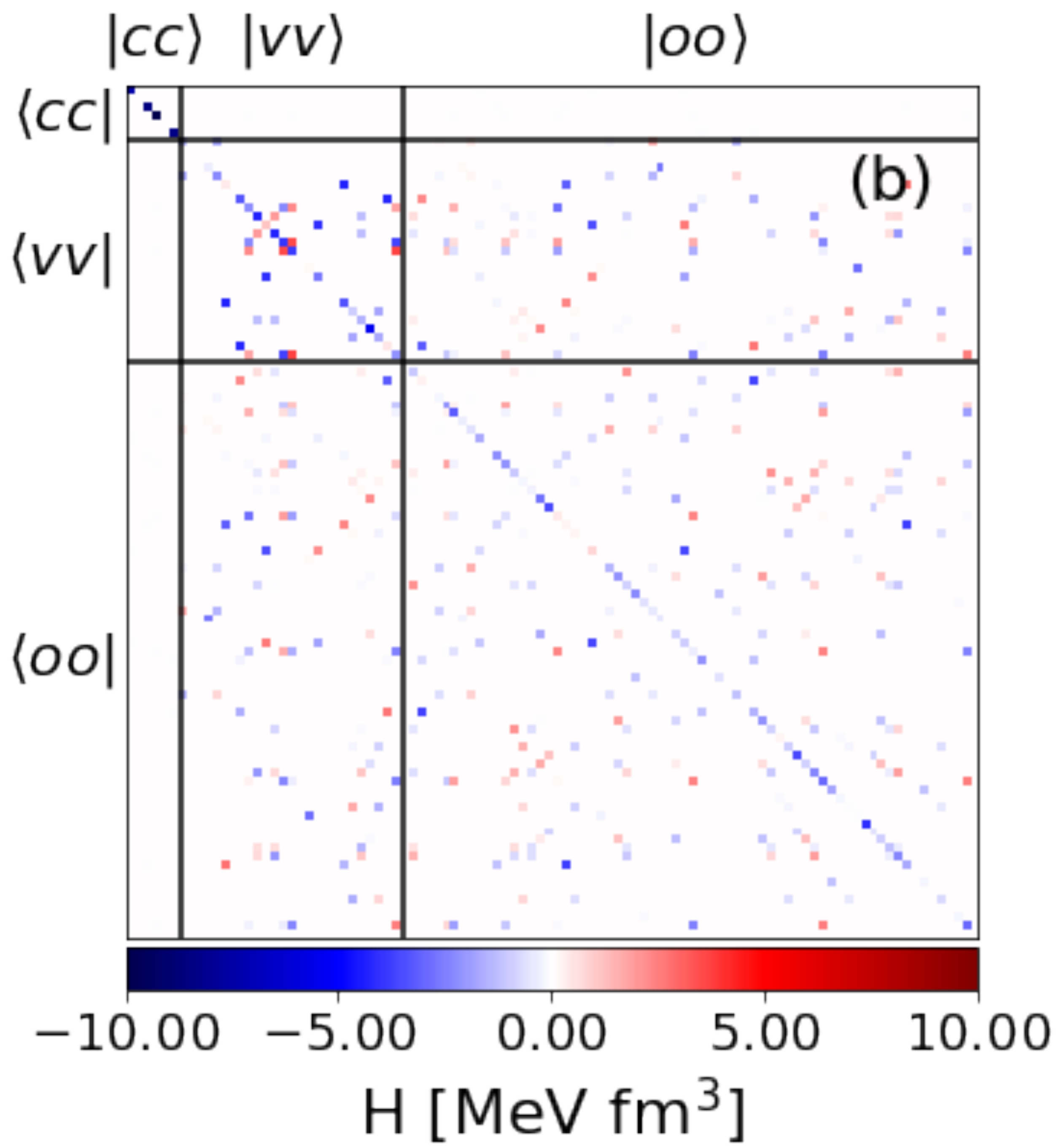
The VS-IMSRG

Valence-Space In Medium Similarity Renormalization Group



Bare Hamiltonian

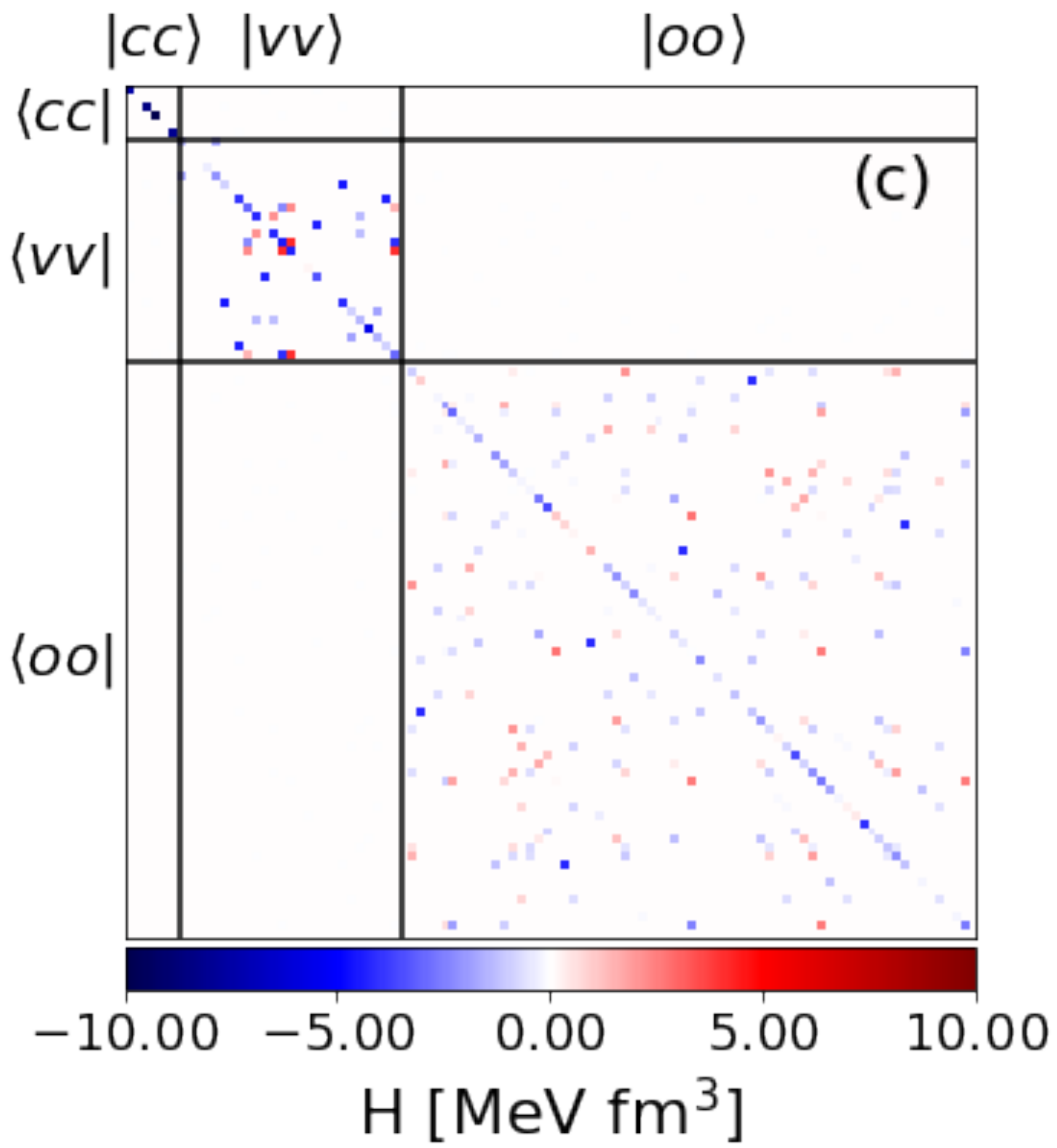
$$\hat{H}(0)$$



Core is decoupled

$$\hat{H}(s) = e^{\Omega_c(s)}\hat{H}(0)e^{-\Omega_c(s)}$$

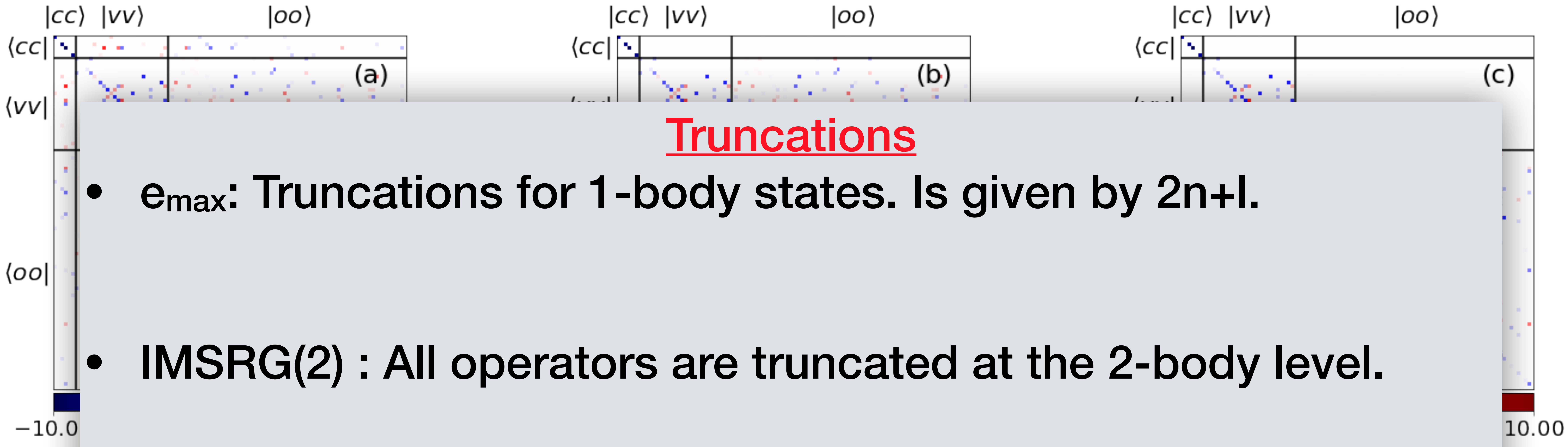
$$\hat{H}_c = e^{\Omega_c(\infty)}\hat{H}(0)e^{-\Omega_c(\infty)}$$



Valence-space is decoupled

$$\hat{H}(s) = e^{\Omega_v(s)}\hat{H}_c e^{-\Omega_v(s)}$$

Valence-Space In Medium Similarity Renormalization Group



Truncations

- e_{max}: Truncations for 1-body states. Is given by 2n+l.
- IMSRG(2) : All operators are truncated at the 2-body level.

Bare Hamiltonian

$$\hat{H}(0)$$

Core is decoupled

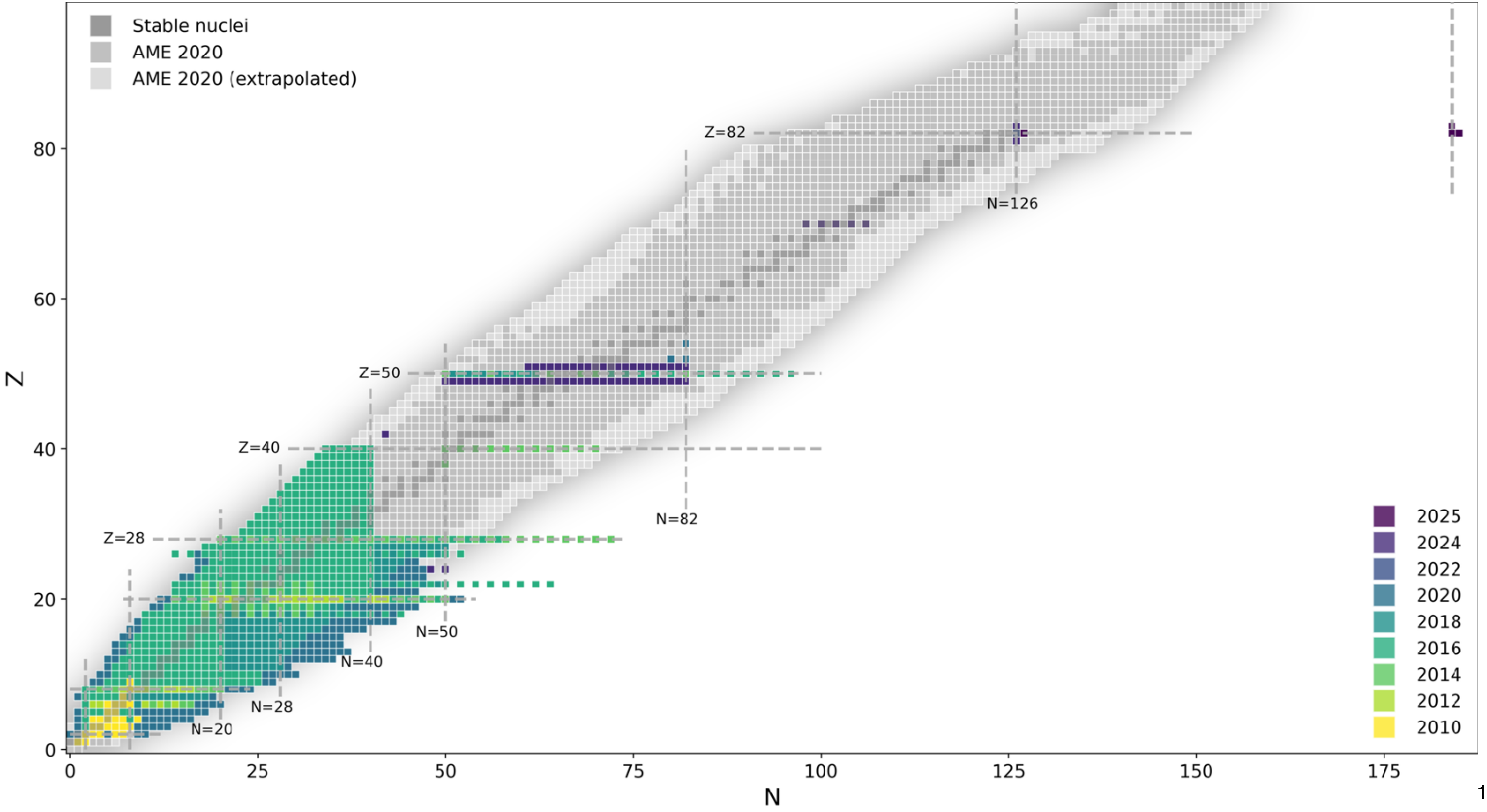
$$\hat{H}(s) = e^{\Omega_c(s)}\hat{H}(0)e^{-\Omega_c(s)}$$

$$\hat{H}_c = e^{\Omega_c(\infty)}\hat{H}(0)e^{-\Omega_c(\infty)}$$

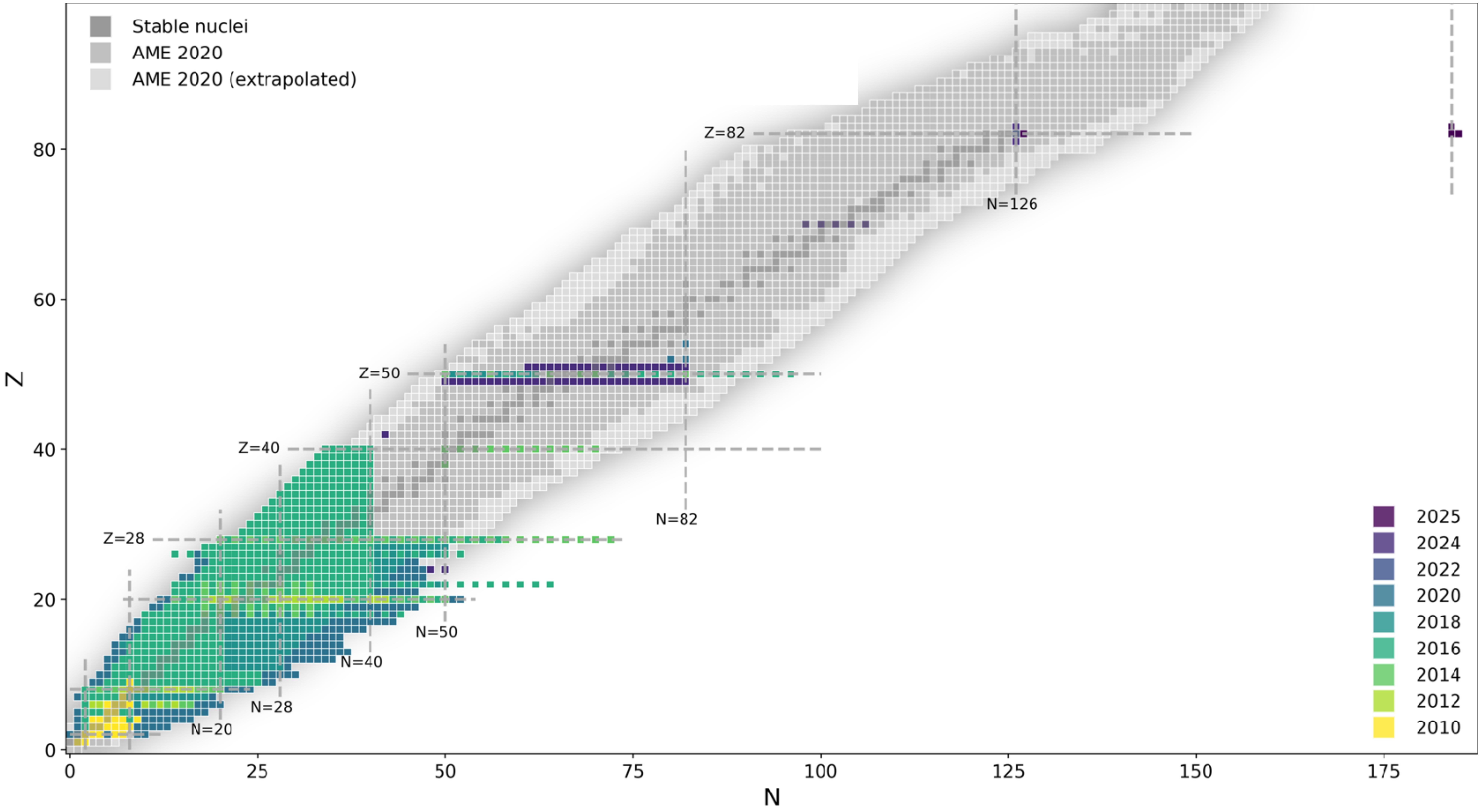
Valence-space is decoupled

$$\hat{H}(s) = e^{\Omega_v(s)}\hat{H}_c e^{-\Omega_v(s)}$$

Ab initio Revolution

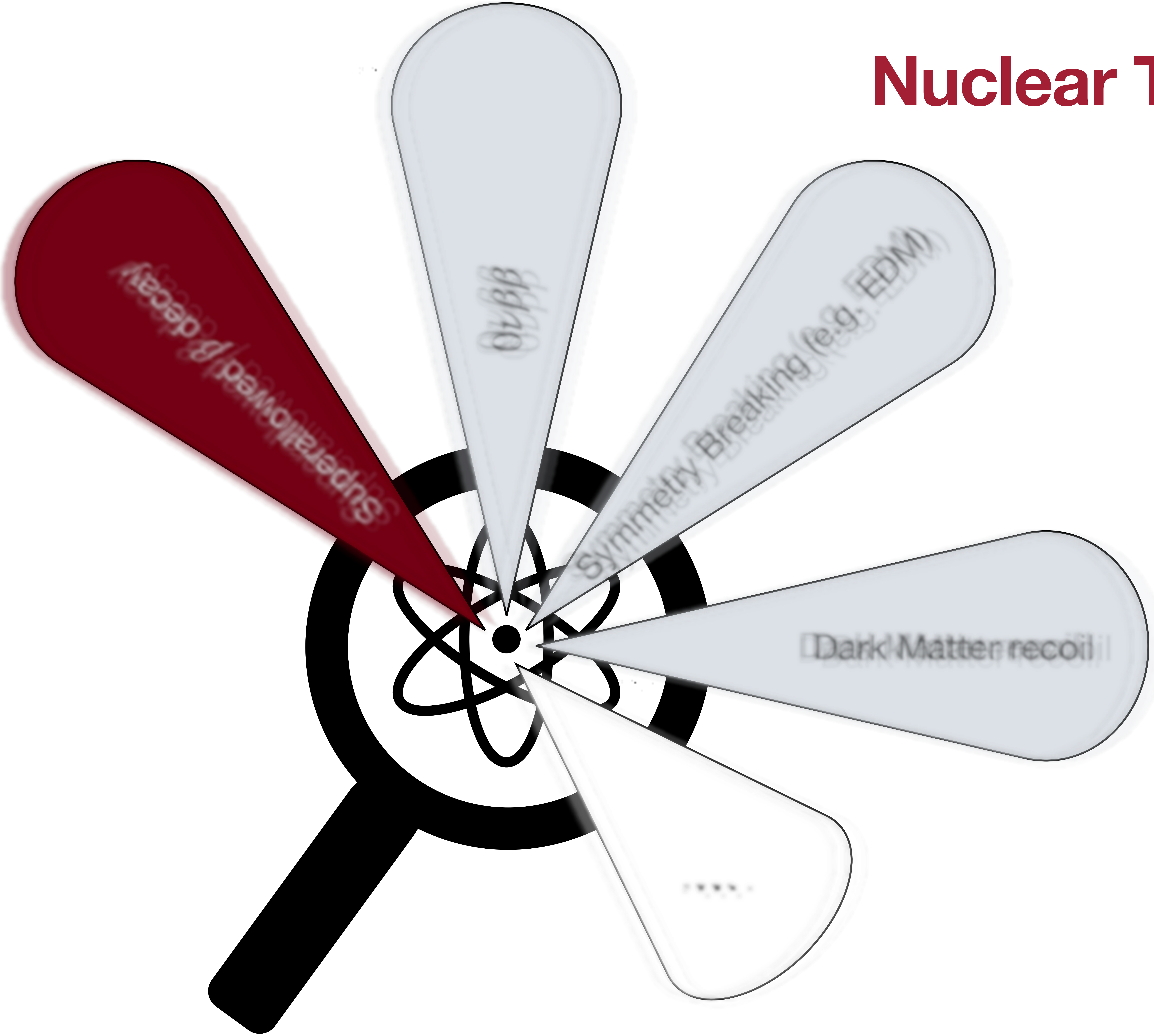


Ab initio Revolution

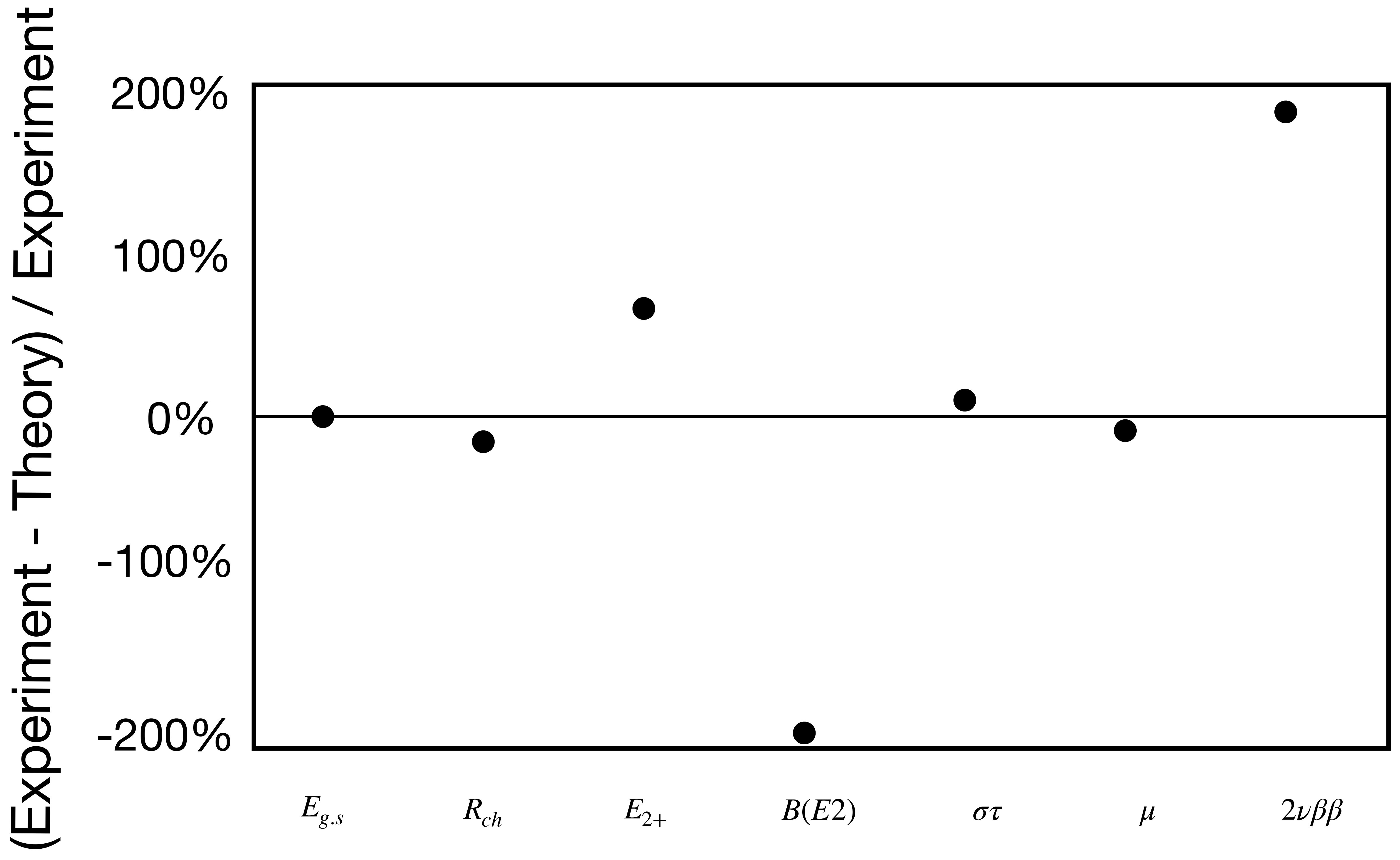


Python version on github!

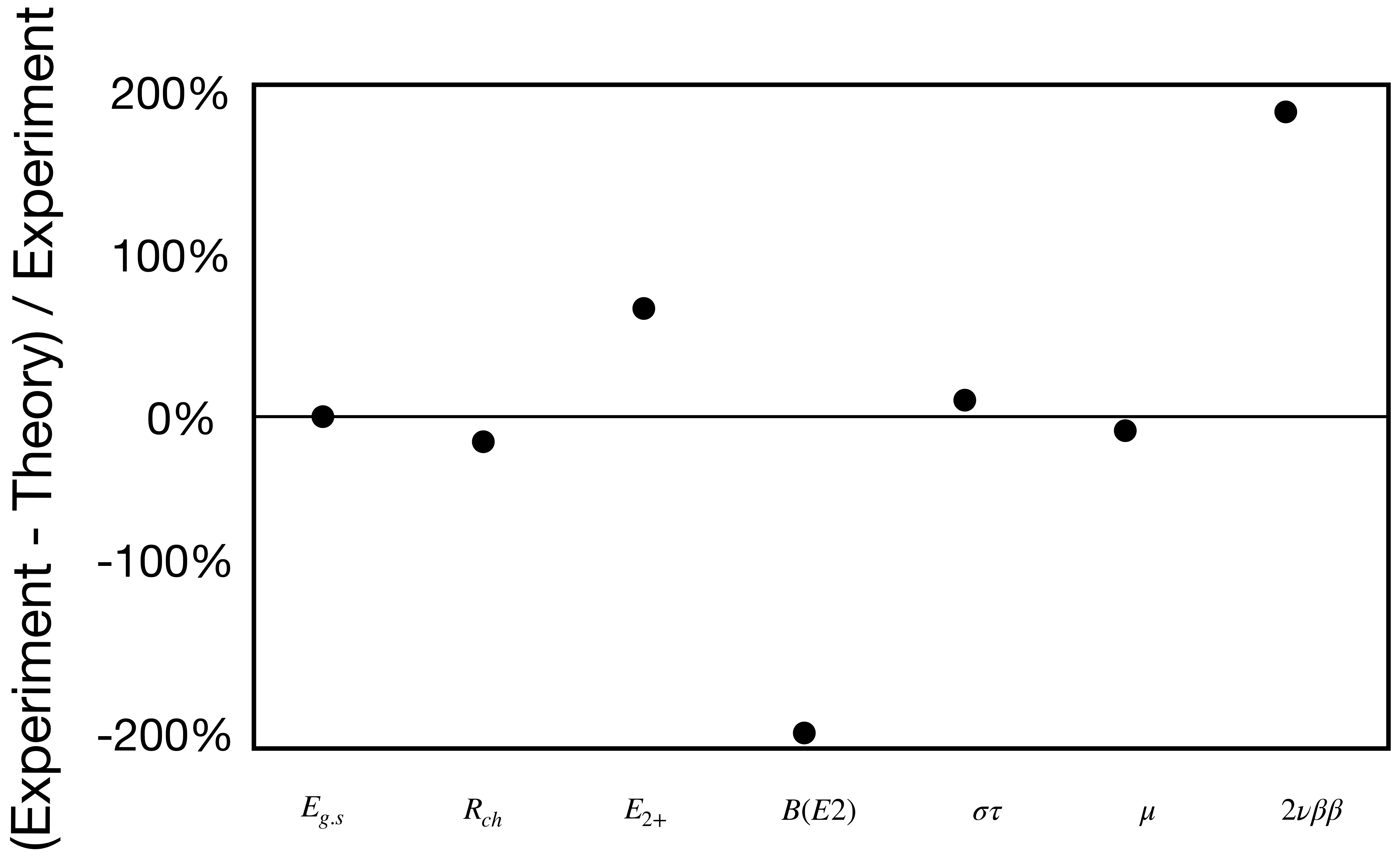
Nuclear Theory Blur



Using other observables

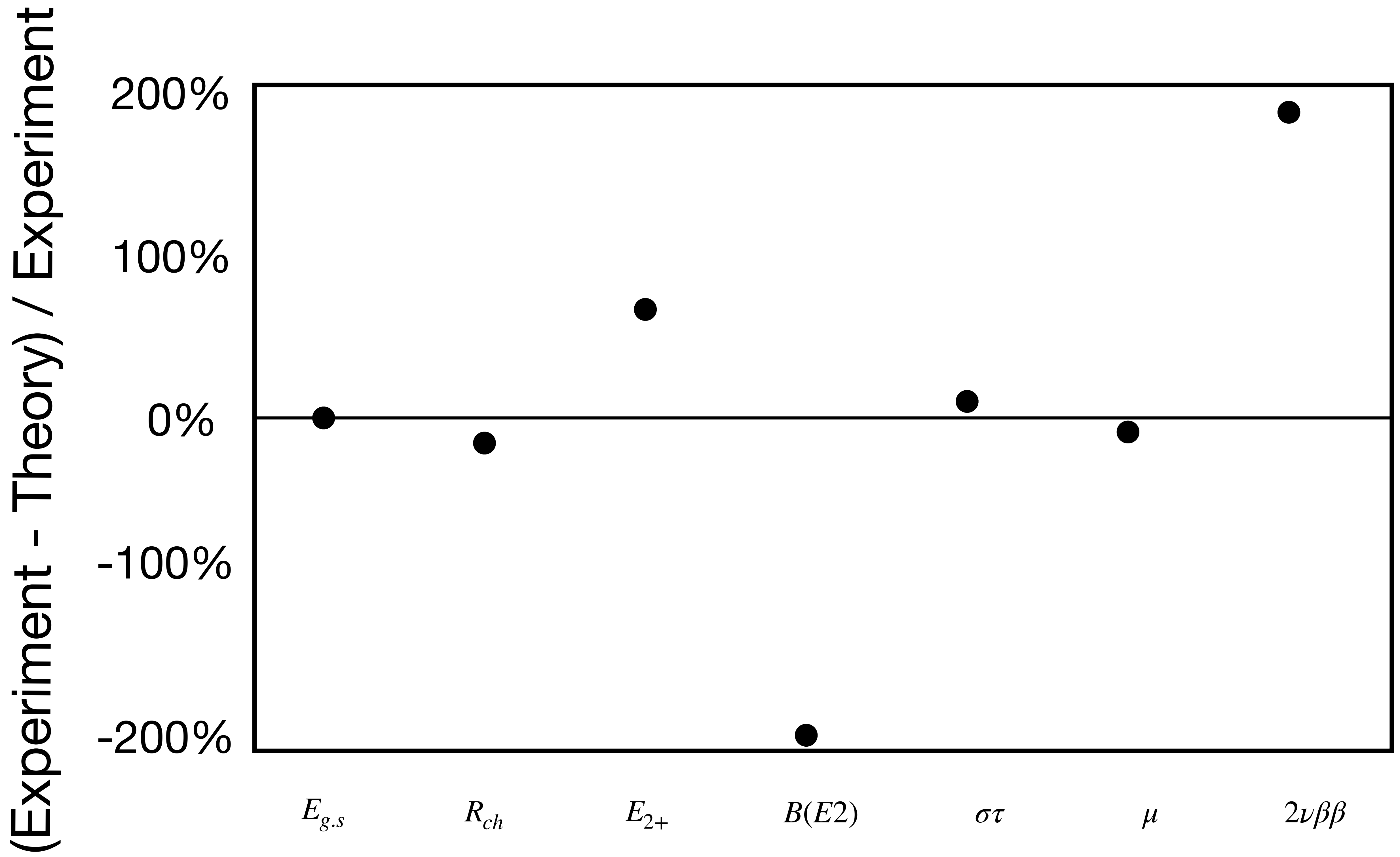


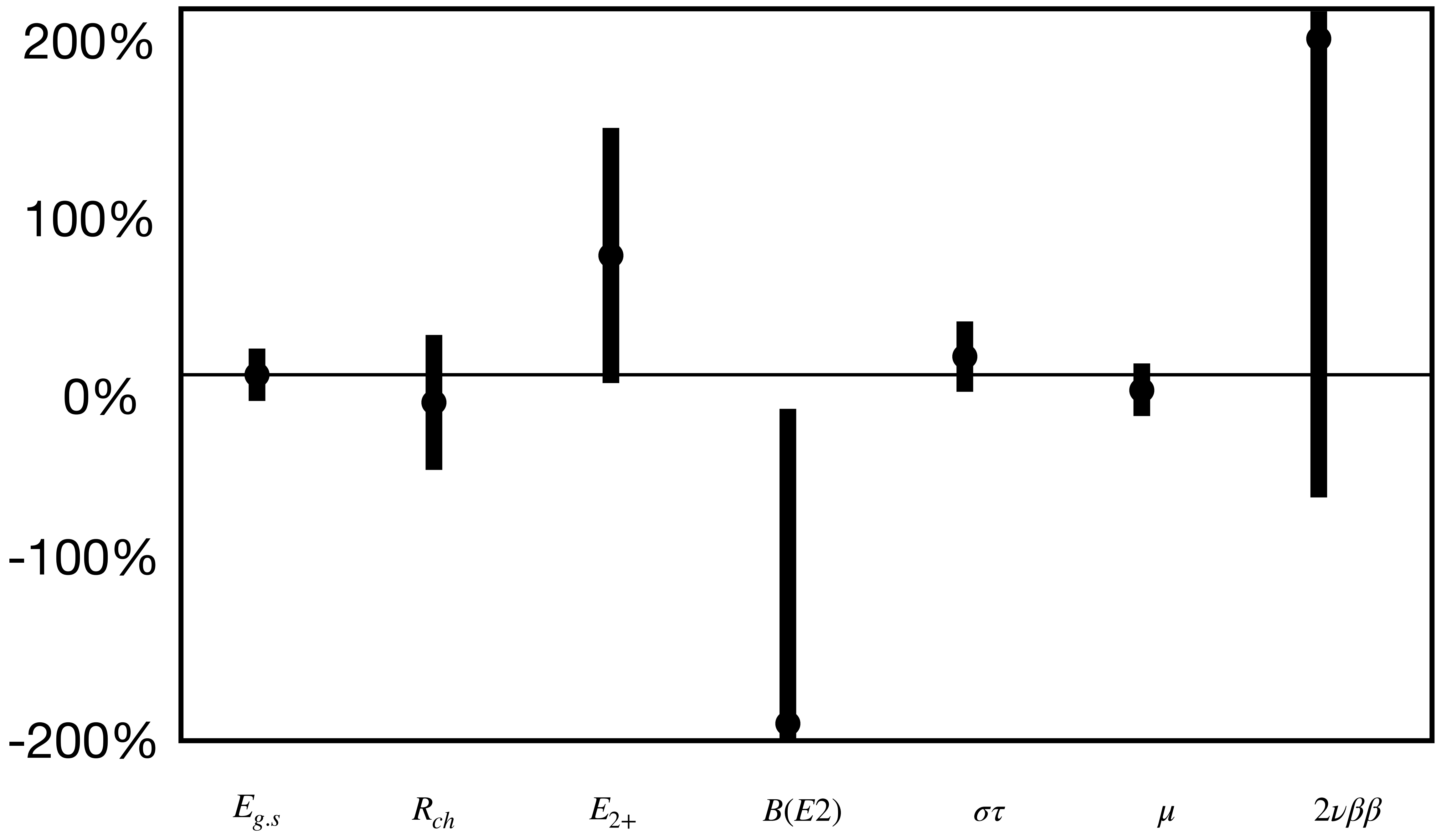
Using other observables



Should I trust this model for $0\nu\beta\beta$?

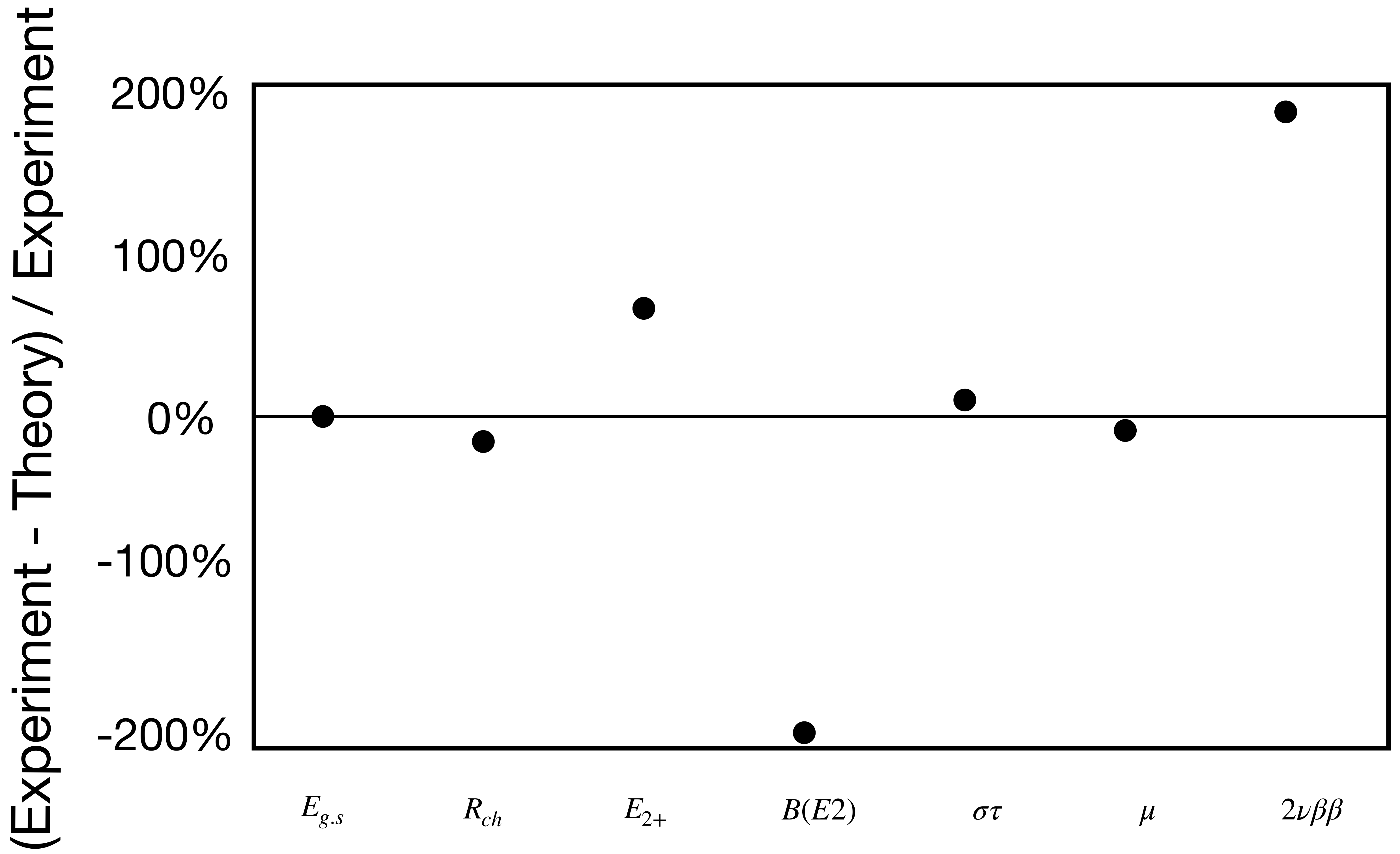
Using other observables

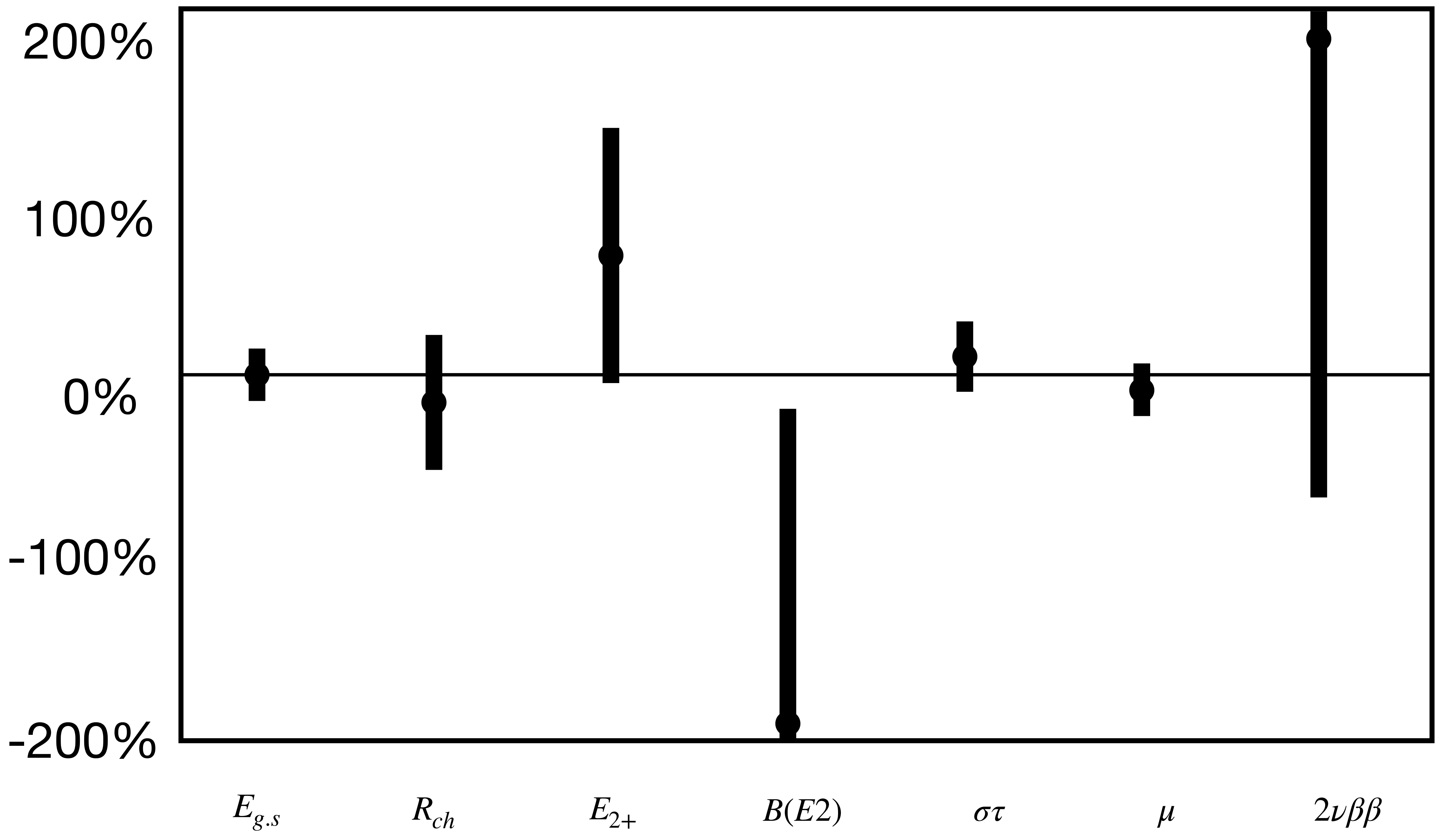




Should I trust this model for $0\nu\beta\beta$?

Using other observables



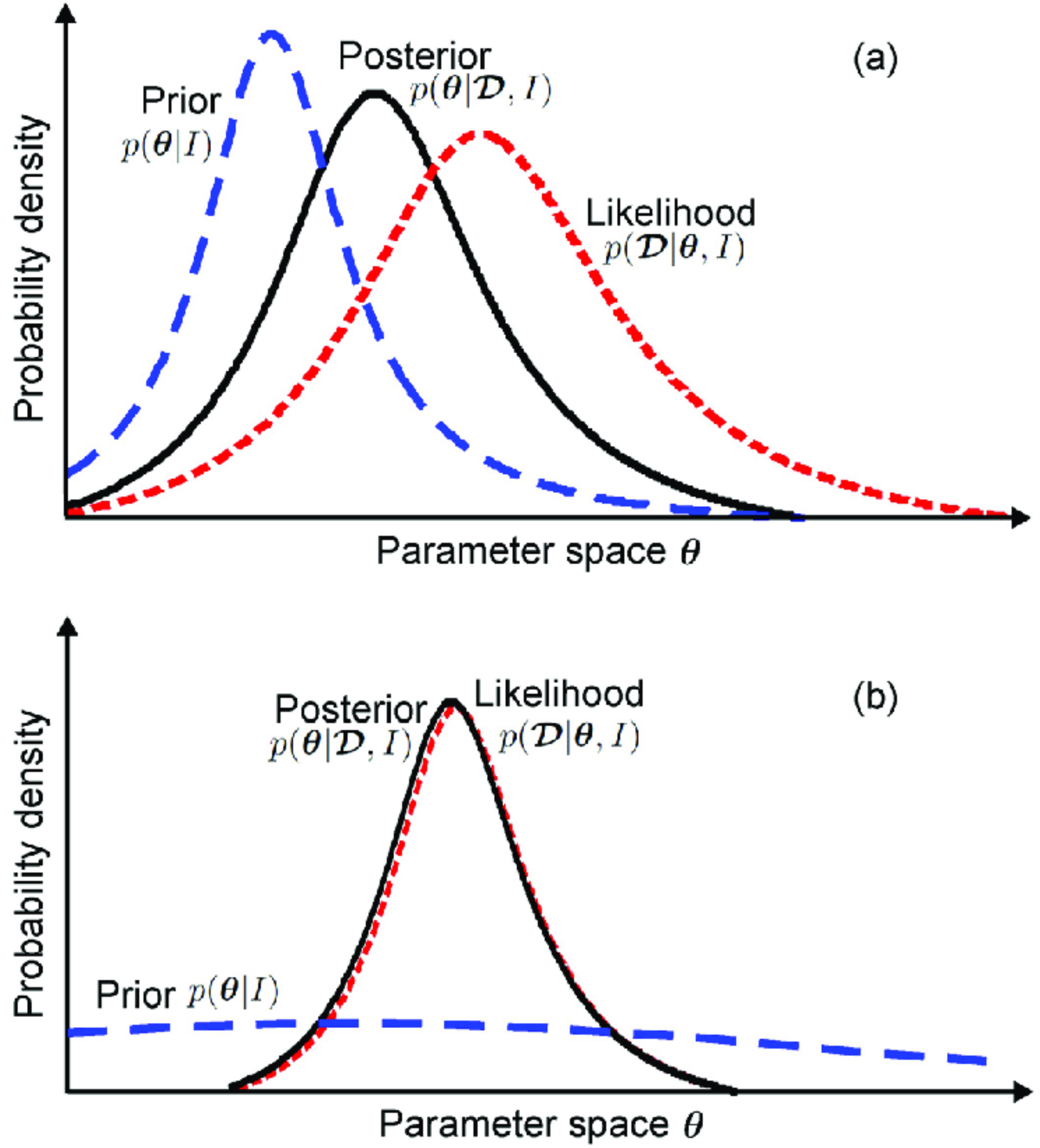


Should I trust this model for $0\nu\beta\beta$?

Should I trust this prediction and uncertainties for $0\nu\beta\beta$?

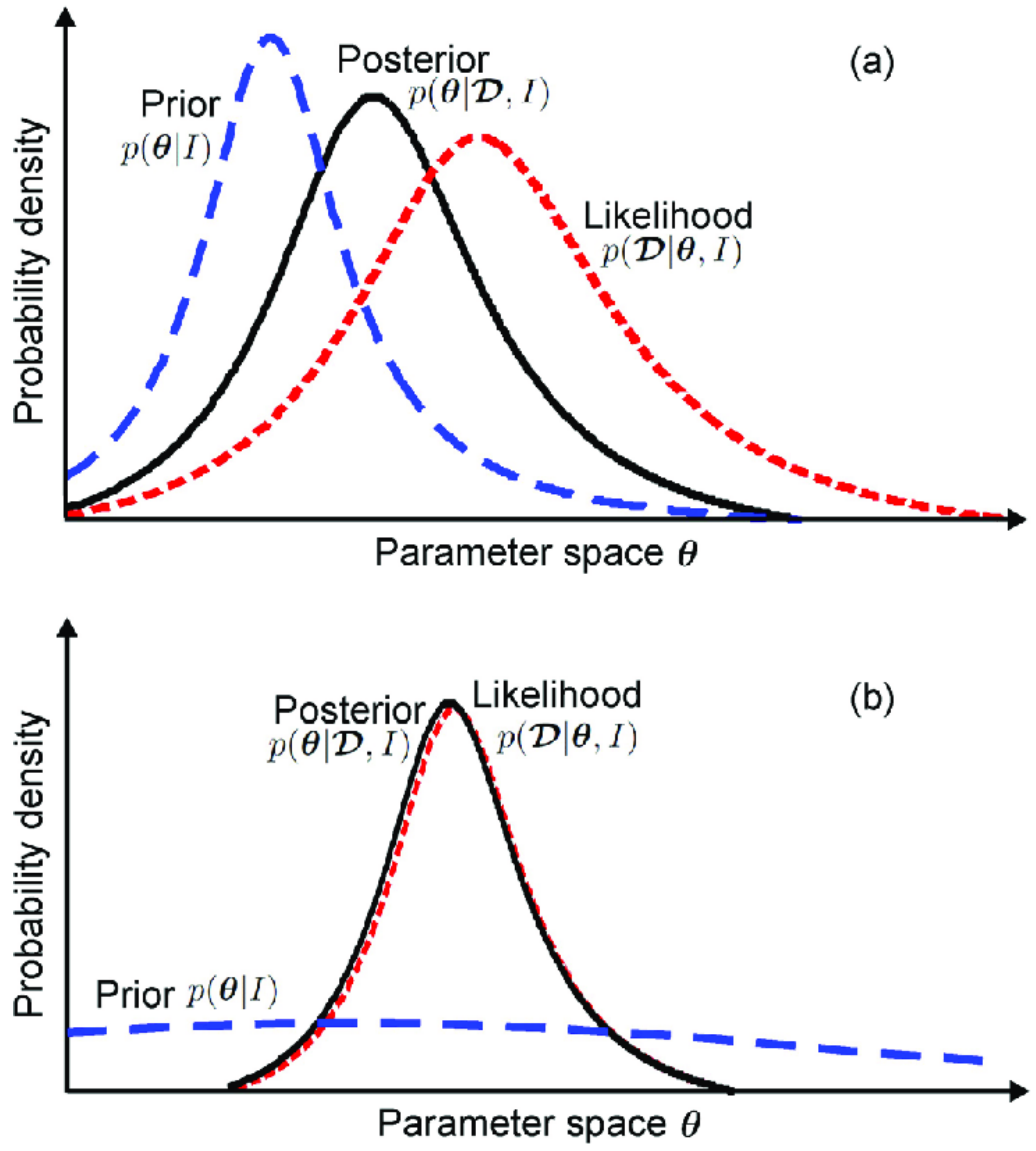
Uncertainty quantification

Procedure for UQ in the Bayesian Approach



Ning Xiang - https://doi.org/10.1121/10.0001731

Procedure for UQ in the Bayesian Approach



Ning Xiang - https://doi.org/10.1121/10.0001731

The catch

Need many samples.

Due to the large cost of many-body methods, for 1 isotope:

- Take ~**1 year** to compute all samples on HPC cluster.
- Cost > **$2 million!**
- Huge environmental impact (**220 tree-years** calculated using green-algorithms.org v3.0)

Procedure for UQ in the Bayesian Approach

The catch

Need many samples.

Due to the large cost of many-body methods, for 1 isotope:

- Take ~**1 year** to compute all samples on HPC cluster.
- Cost > **$2 million!**
- Huge environmental impact (**220 tree-years** calculated using green-algorithms.org v3.0)

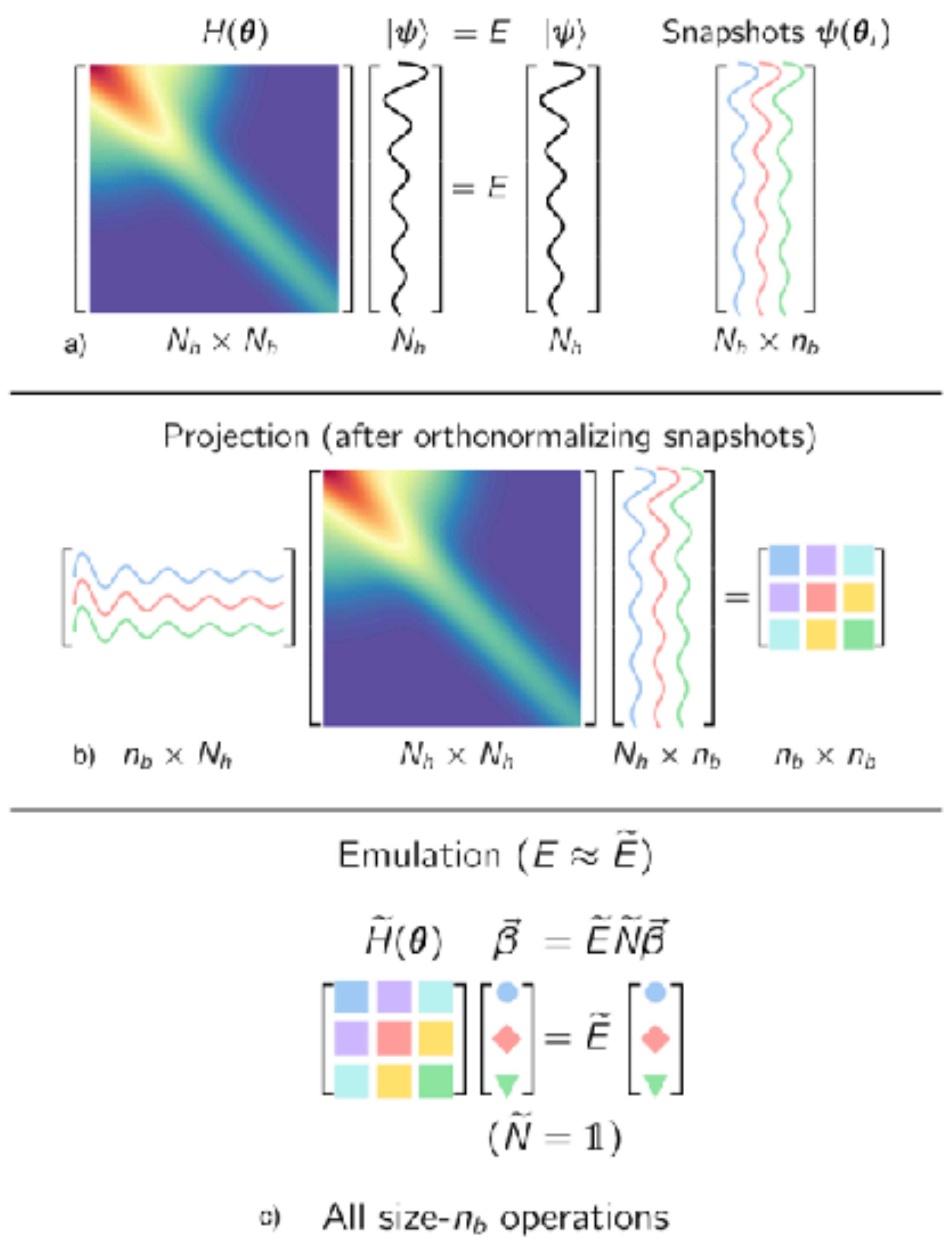
Emulators for Many-Body Methods

There are two ways to build an emulator for nuclear physics:

Emulators for Many-Body Methods

There are two ways to build an emulator for nuclear physics:

1. **Physics driven**

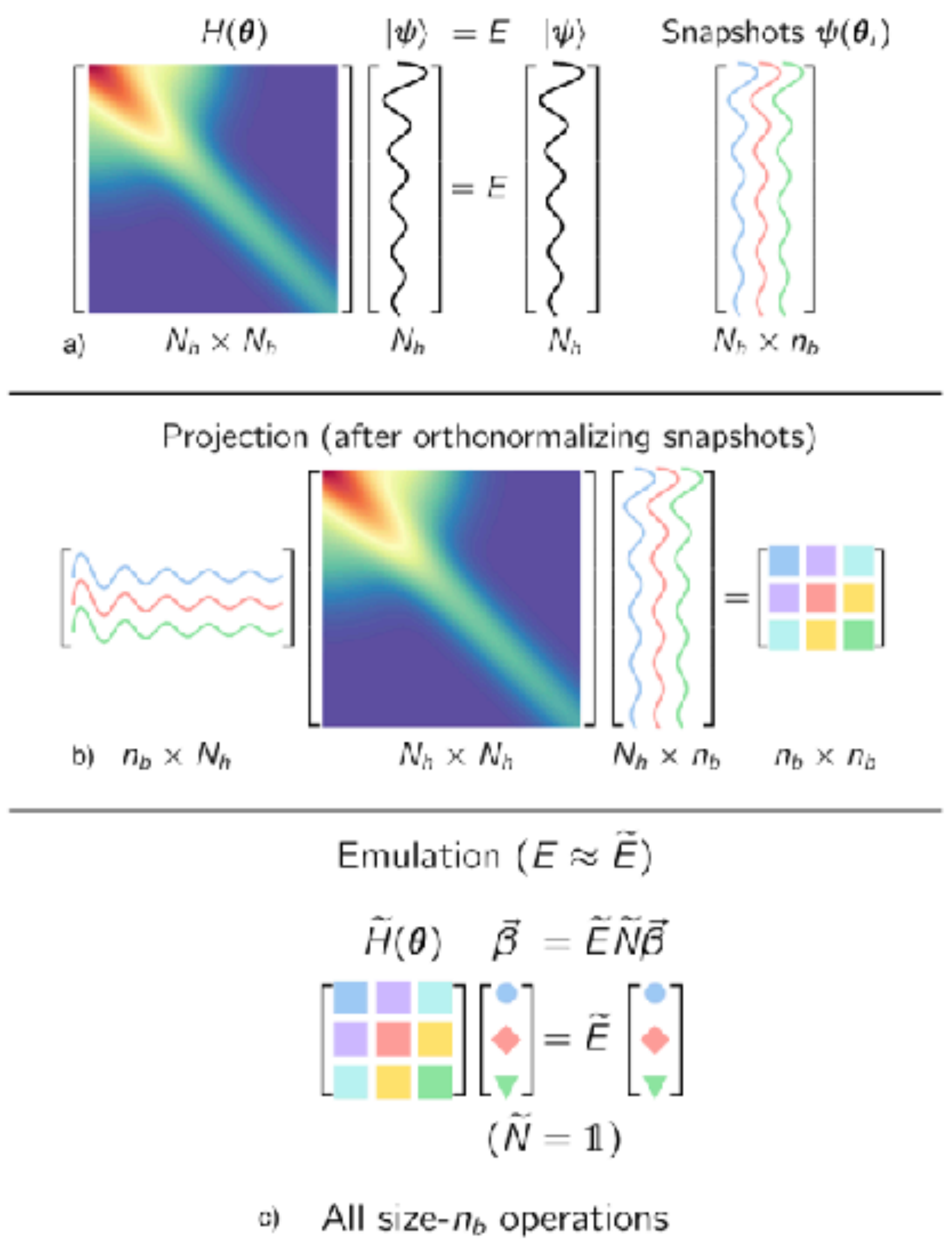


Duguet, et al., Rev. Mod. Phys. **96**, 031002 (2024)

Emulators for Many-Body Methods

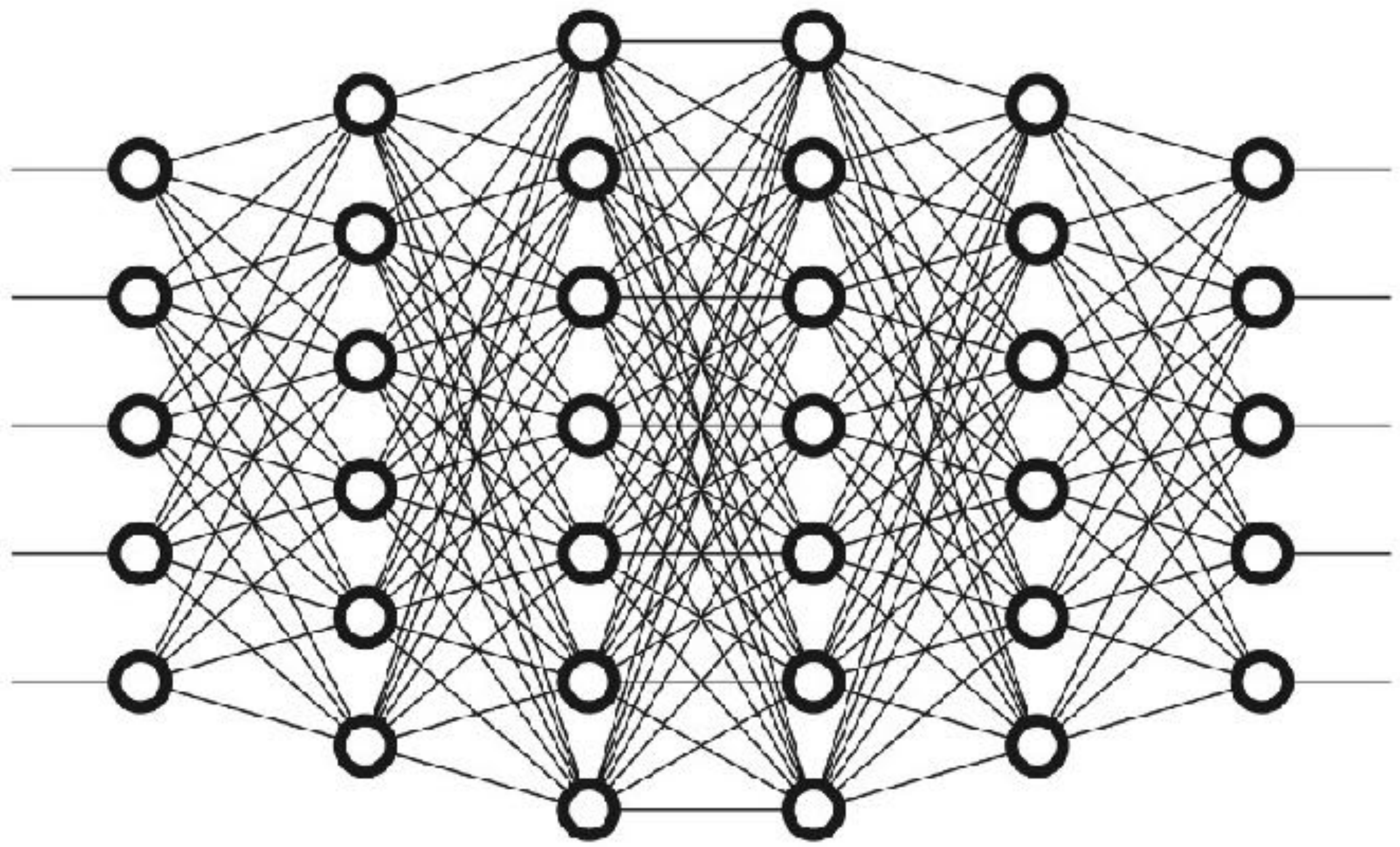
There are two ways to build an emulator for nuclear physics:

1. Physics driven

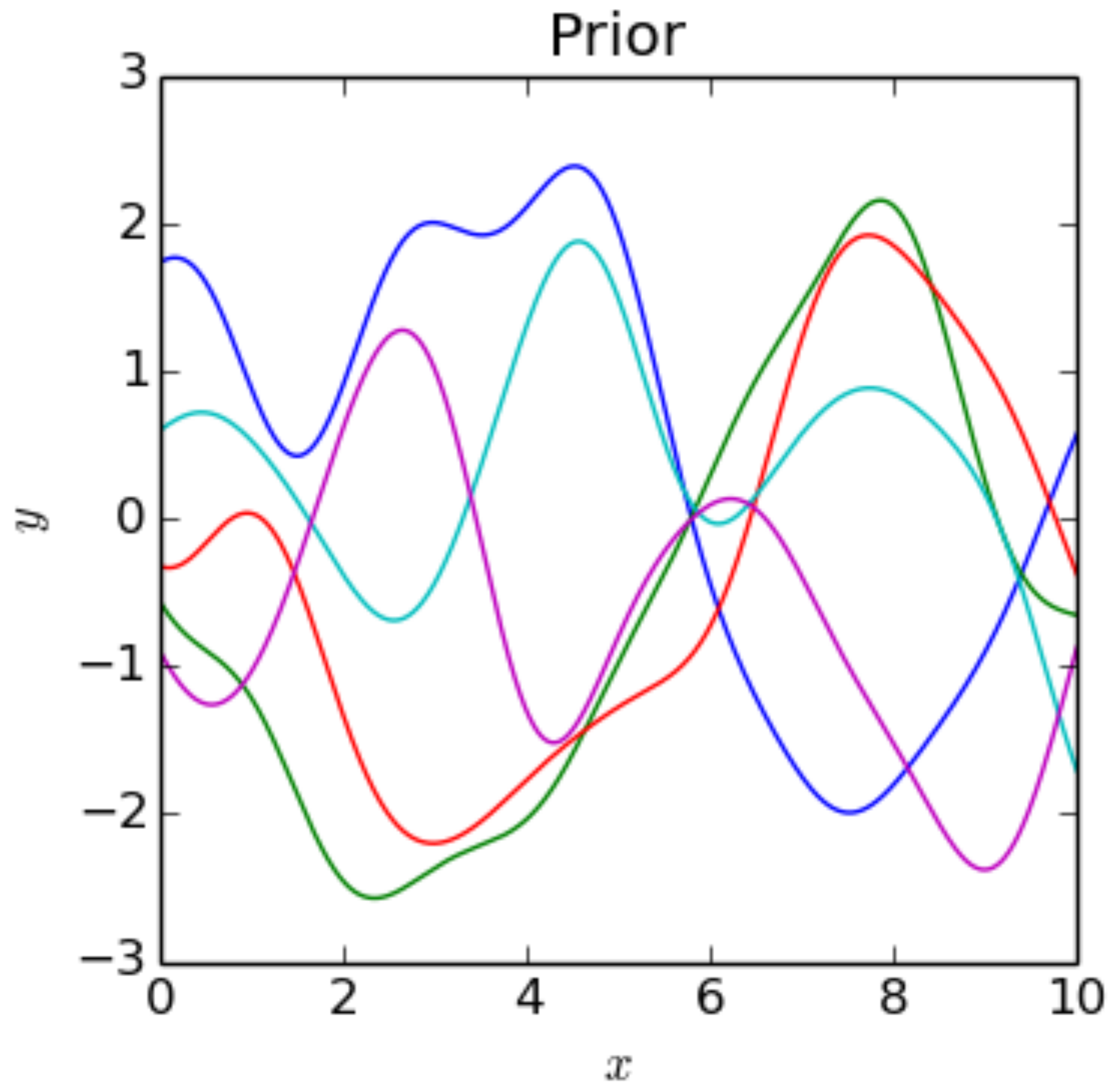


Duguet, et al., Rev. Mod. Phys. **96**, 031002 (2024)

2. Data driven

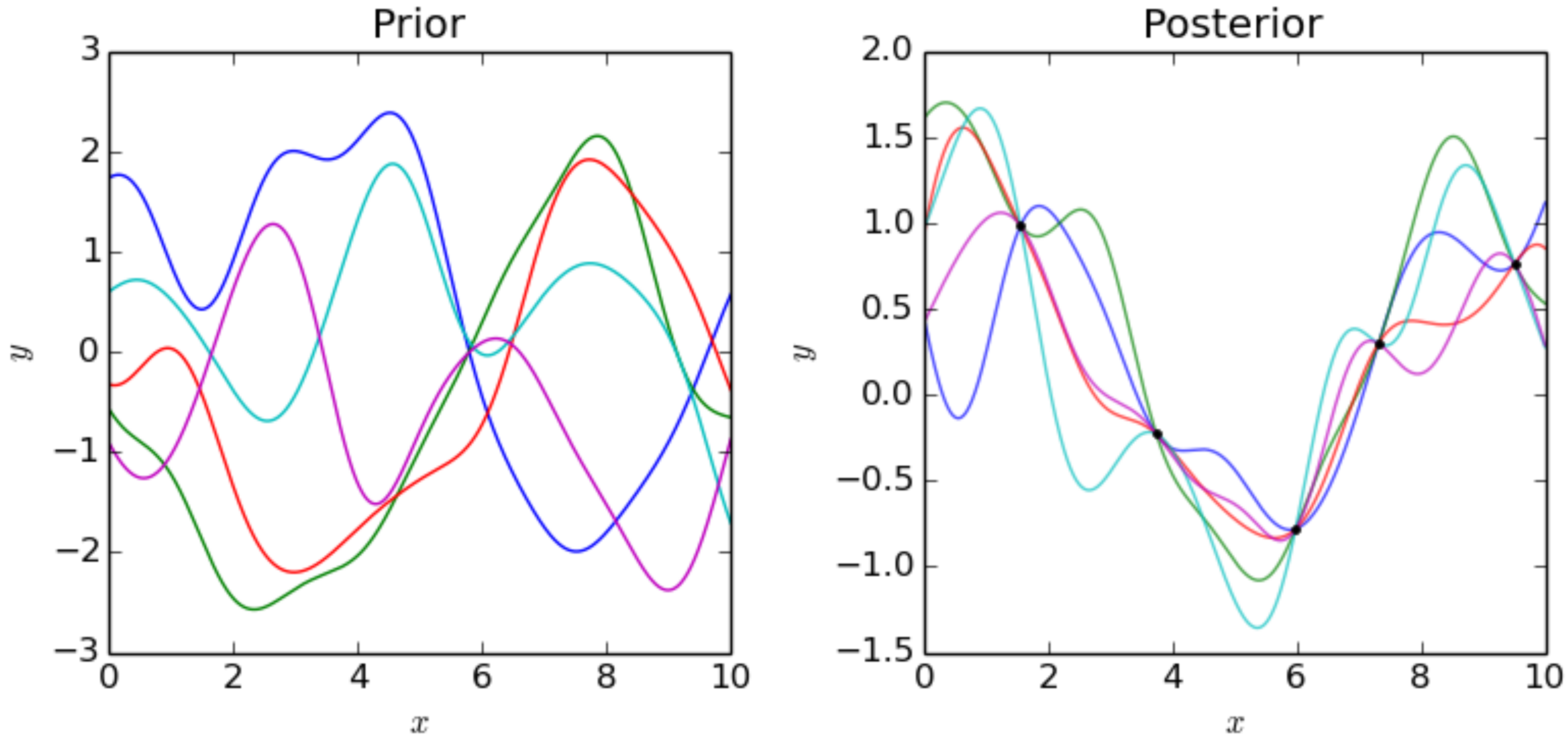


Using Gaussian Process as an Emulator



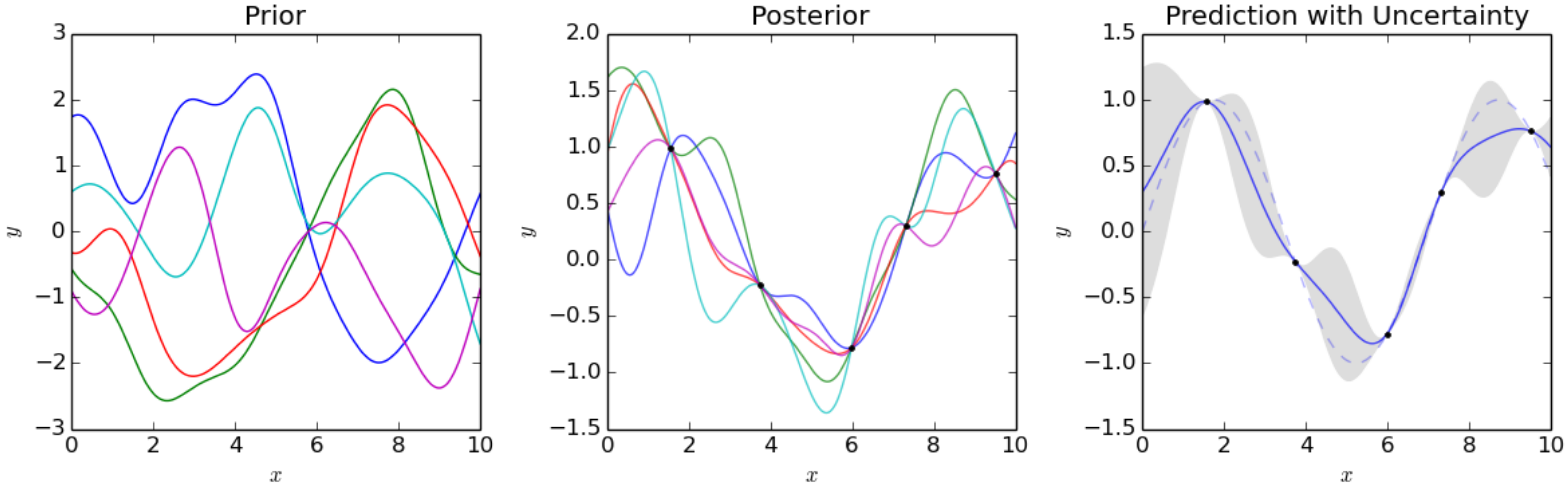
$$f(\mathbf{x}) = \mathcal{N}(\mu, K(\mathbf{x}, \mathbf{x}))$$

Using Gaussian Process as an Emulator



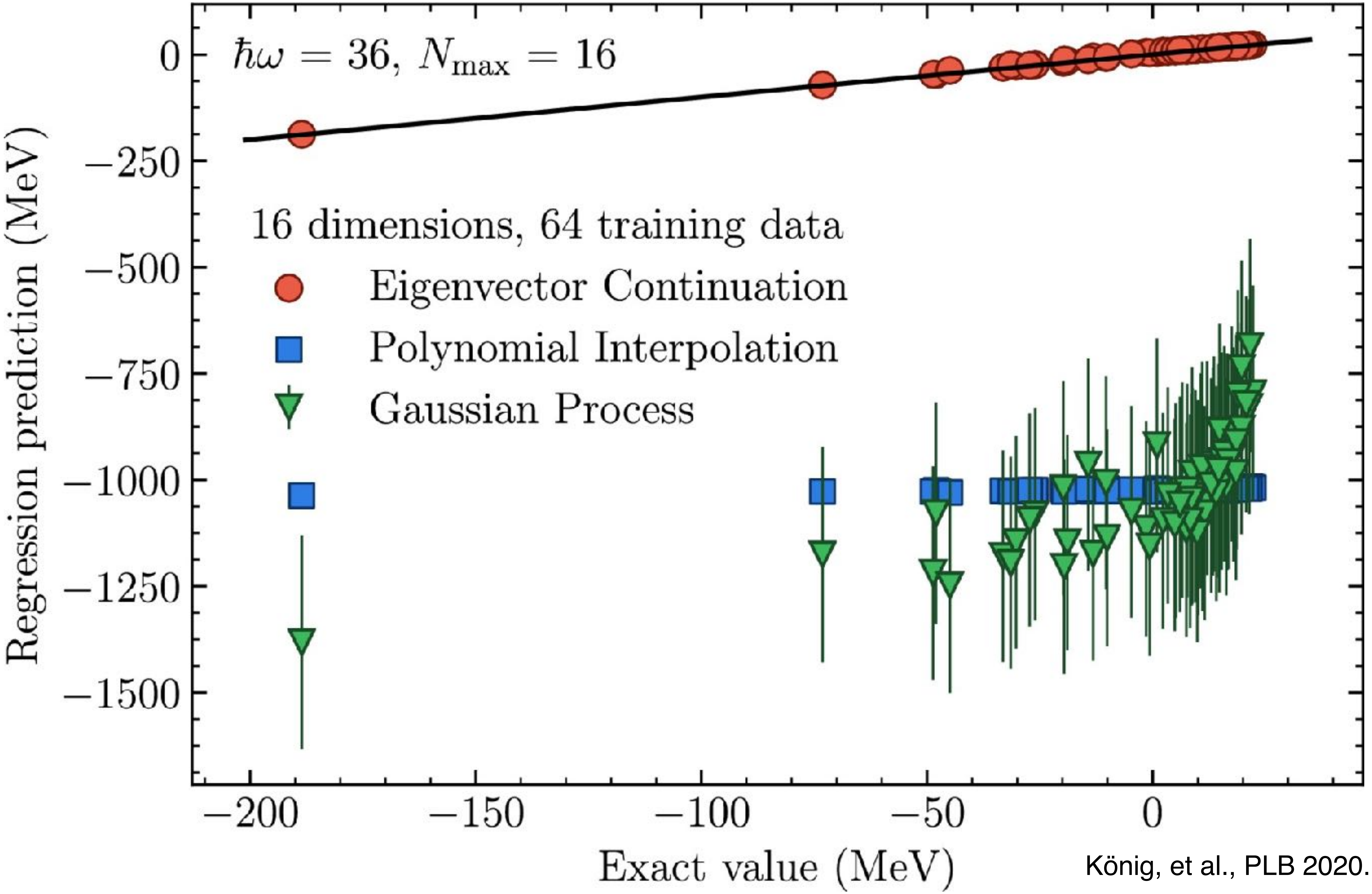
$$f(\mathbf{x}) = \mathcal{N}(\mu, K(\mathbf{x}, \mathbf{x})) \qquad P_{Y^*|Y} \sim \mathcal{N}\left(\Sigma_{X^*X}\Sigma_{XX}^{-1}Y, \Sigma_{X^*X^*} - \Sigma_{X^*X}\Sigma_{XX}^{-1}\Sigma_{XX^*}\right)$$

Using Gaussian Process as an Emulator



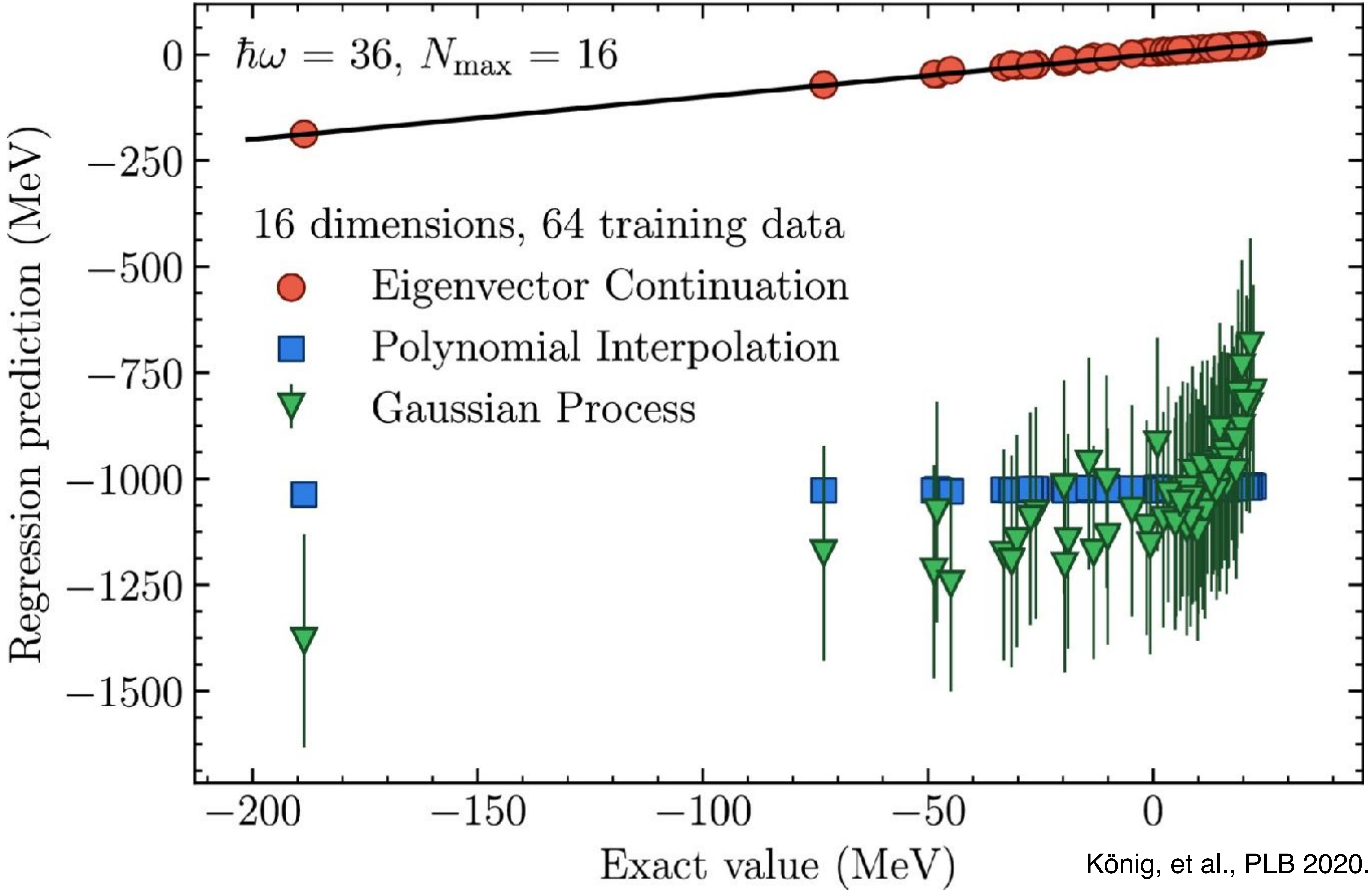
$$f(\mathbf{x}) = \mathcal{N}(\mu, K(\mathbf{x}, \mathbf{x})) \qquad P_{Y^*|Y} \sim \mathcal{N}\left(\Sigma_{X^*X}\Sigma_{XX}^{-1}Y, \Sigma_{X^*X^*} - \Sigma_{X^*X}\Sigma_{XX}^{-1}\Sigma_{XX^*}\right)$$

Using Gaussian Process as an Emulator



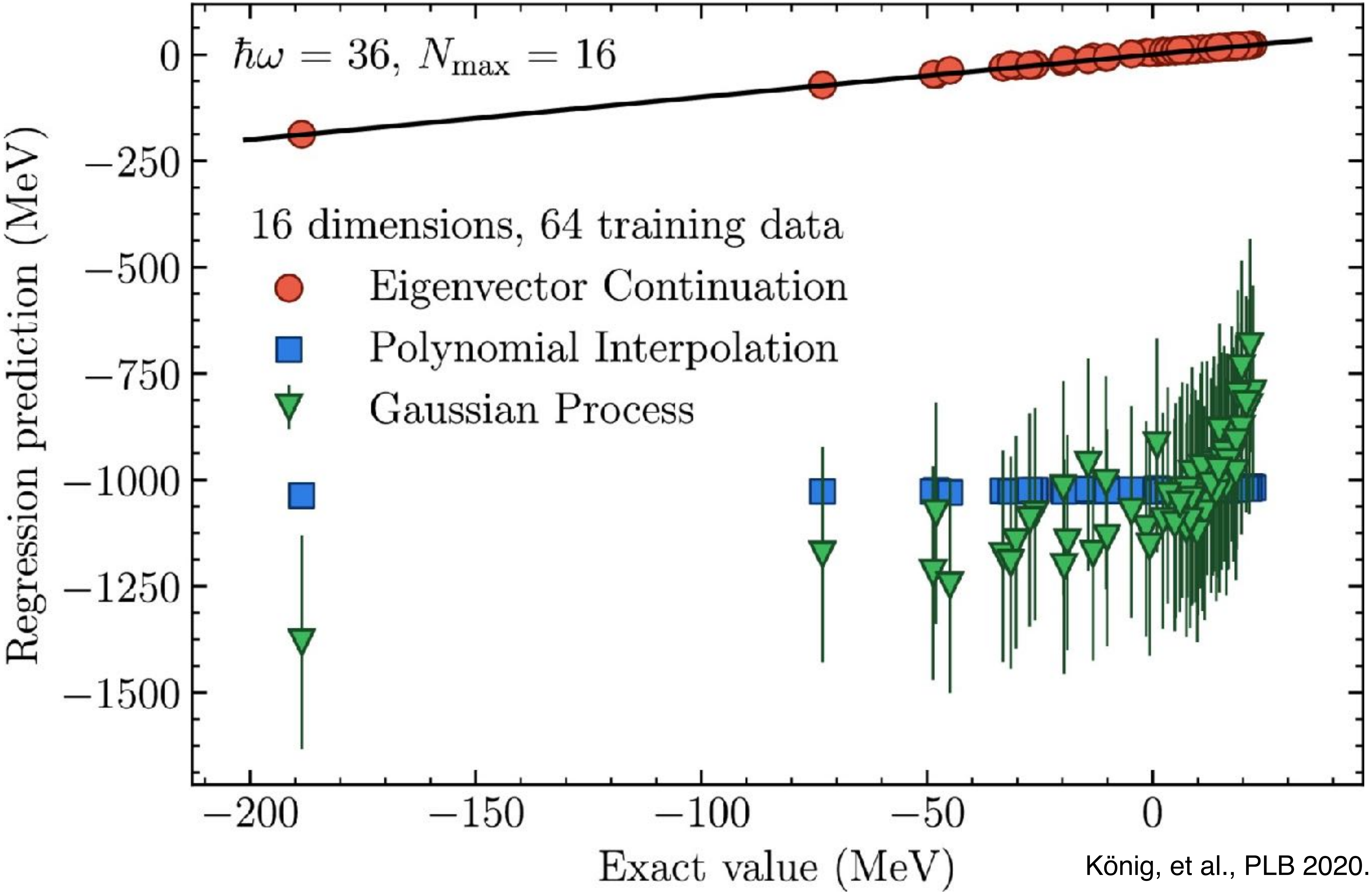
König, et al., PLB 2020.135814

Using Gaussian Process as an Emulator



König, et al., PLB 2020.135814

Using Gaussian Process as an Emulator



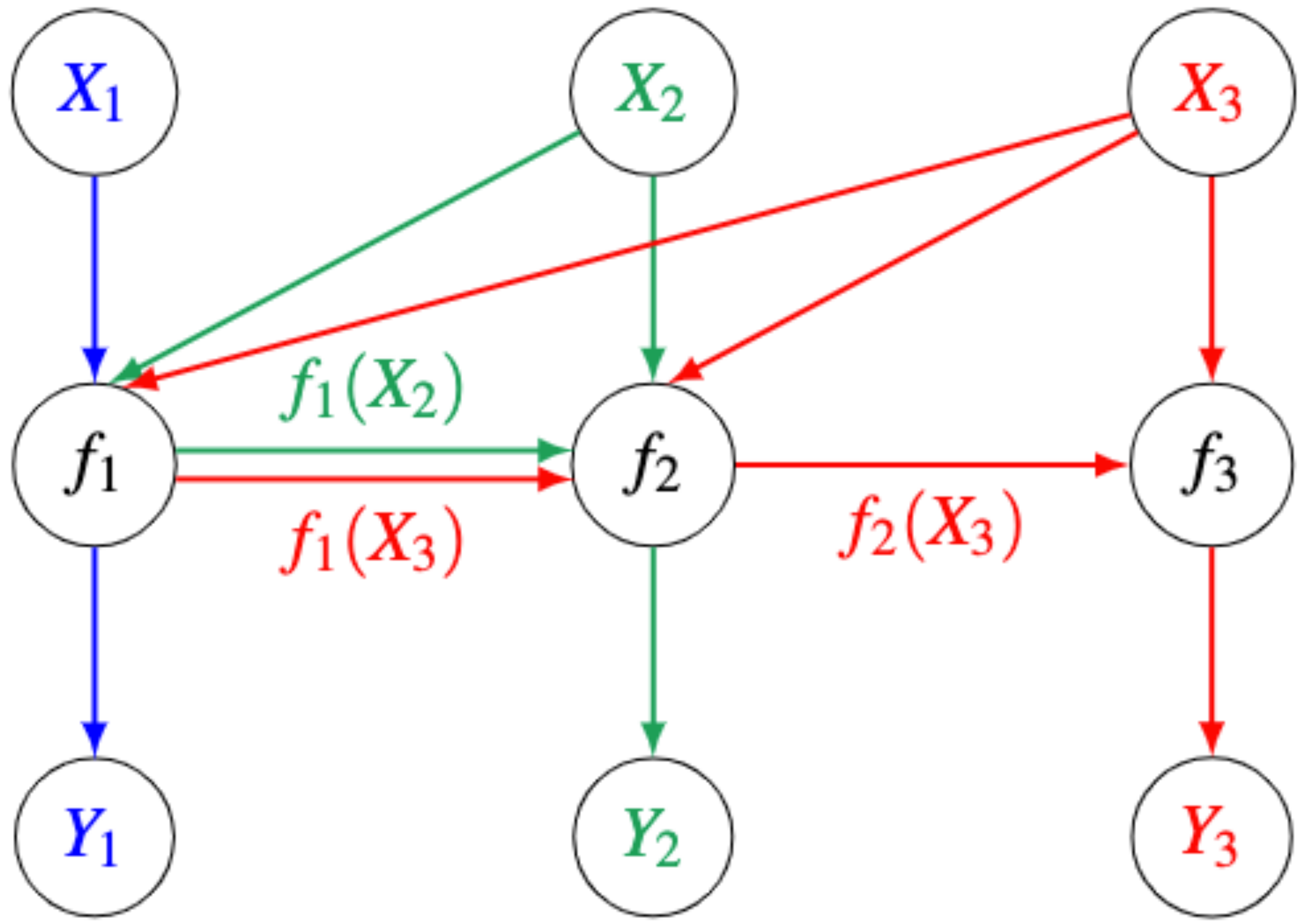
König, et al., PLB 2020.135814

The MM-DGP Algorithm

- **Deep Gaussian Processes [1]**: Stack multiple GPs in a neural network-like architecture for improved hierarchical learning.

- **Multi-Fidelity Modelling**: Model low-to-high fidelity differences by passing outputs from one fidelity as inputs to the next.

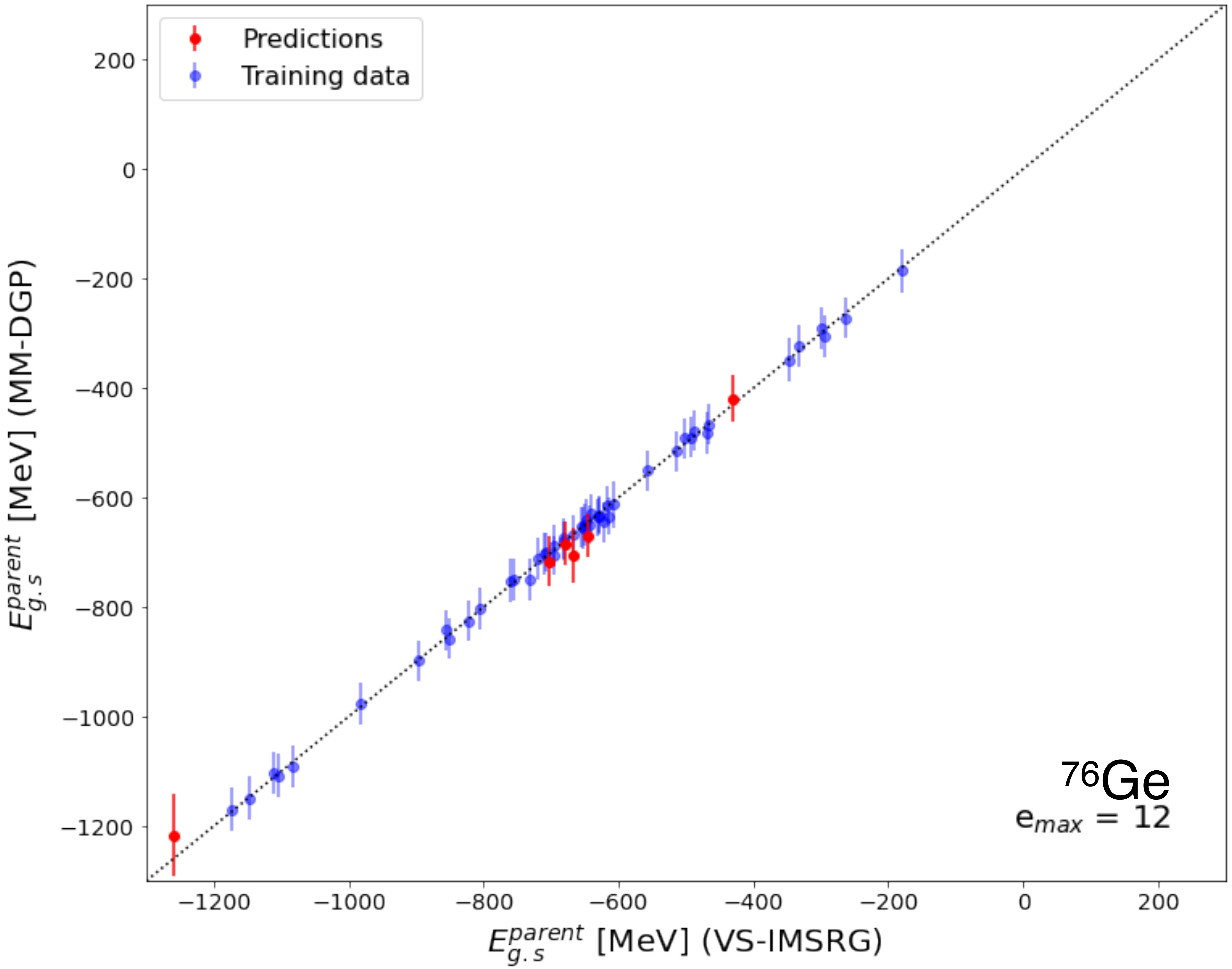
- **MM-DGP Extension**: Adapted to handle multiple outputs across fidelity levels, creating the Multi-output Multi-fidelity Deep Gaussian Process (MM-DGP).



[1] Kurt Cutajar, Mark Pullin, Andreas Damianou, Neil Lawrence, Javier González arXiv:1903.07320 (2021).

The MM-DGP Algorithm: Energies

e_{max}	# of training points
4	250
6	100
8	90
10	50
12	50



Root Mean Square Error = 11 MeV

Getting Posterior Distributions

Getting Posterior Distributions

Prior: 8188 “non-implausible” samples (Phys. Rev. C **109**, 064314)

Getting Posterior Distributions

Prior: 8188 "non-implausible" samples (Phys. Rev. C **109**, 064314)

"Solve" many-body with emulator for all samples

Getting Posterior Distributions

Prior: 8188 “non-implausible” samples (Phys. Rev. C **109**, 064314)

“Solve” many-body with emulator for all samples

Weight the samples by a likelihood for some calibration observables
(correlation with target observable)

Getting Posterior Distributions

Prior: 8188 “non-implausible” samples (Phys. Rev. C **109**, 064314)

“Solve” many-body with emulator for all samples

Weight the samples by a likelihood for some calibration observables
(correlation with target observable)

Final results is given by

$$y = y_{LECs} + \epsilon_{emulator} + \epsilon_{EFT} + \epsilon_{many-body} + \epsilon_{operator}$$

where y_{LECs} is sampled from the weighted interactions

Getting Posterior Distributions

Prior: 8188 “non-implausible” samples (Phys. Rev. C **109**, 064314)

“Solve” many-body with emulator for all samples

Weight the samples by a likelihood for some calibration observables (correlation with target observable)

Final results is given by

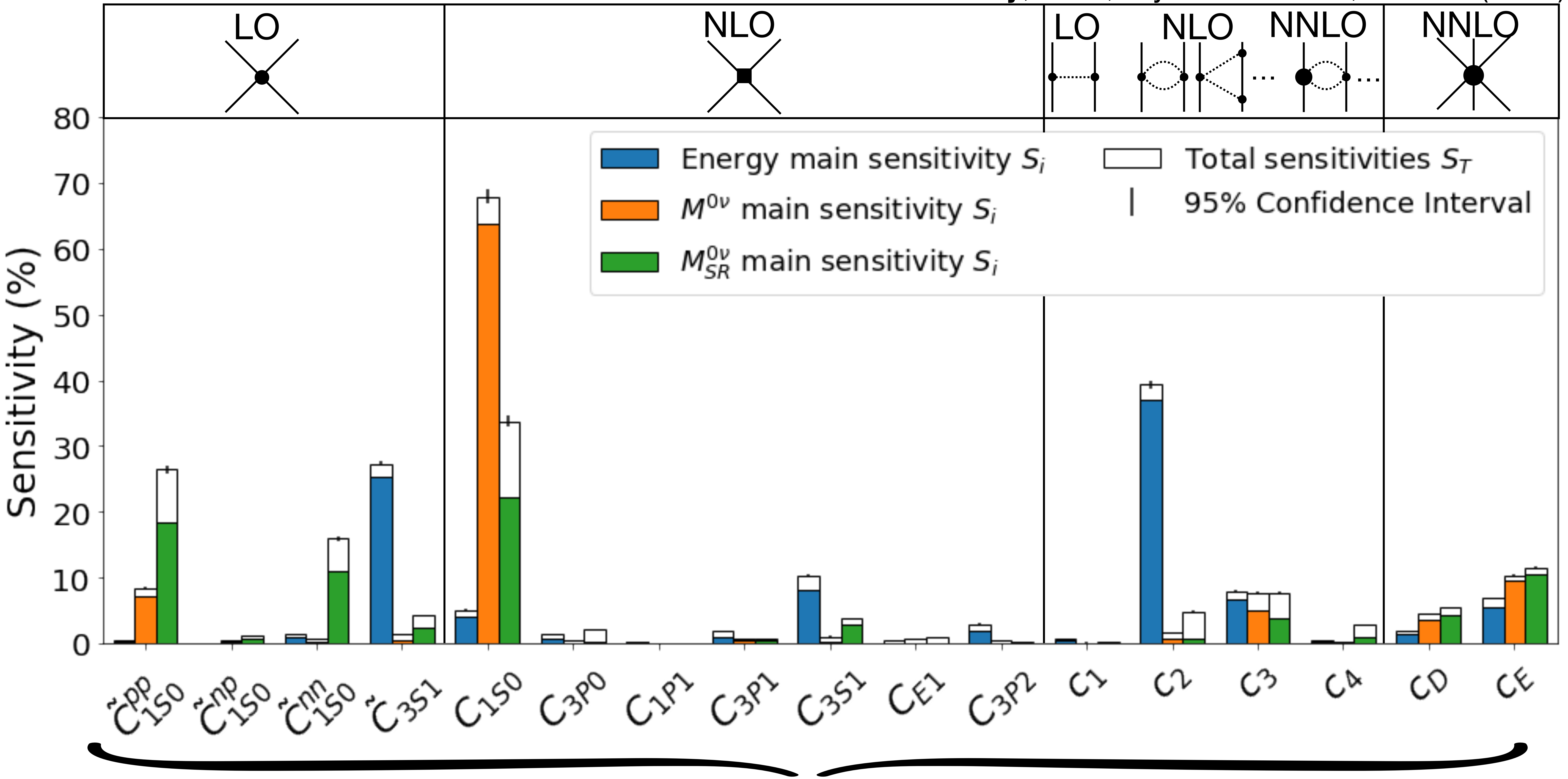
$$y = y_{LECs} + \epsilon_{emulator} + \epsilon_{EFT} + \epsilon_{many-body} + \epsilon_{operator}$$

where y_{LECs} is sampled from the weighted interactions

The different sources of errors ϵ are assumed to be normally distributed and independent

The MM-DGP Algorithm: GSA

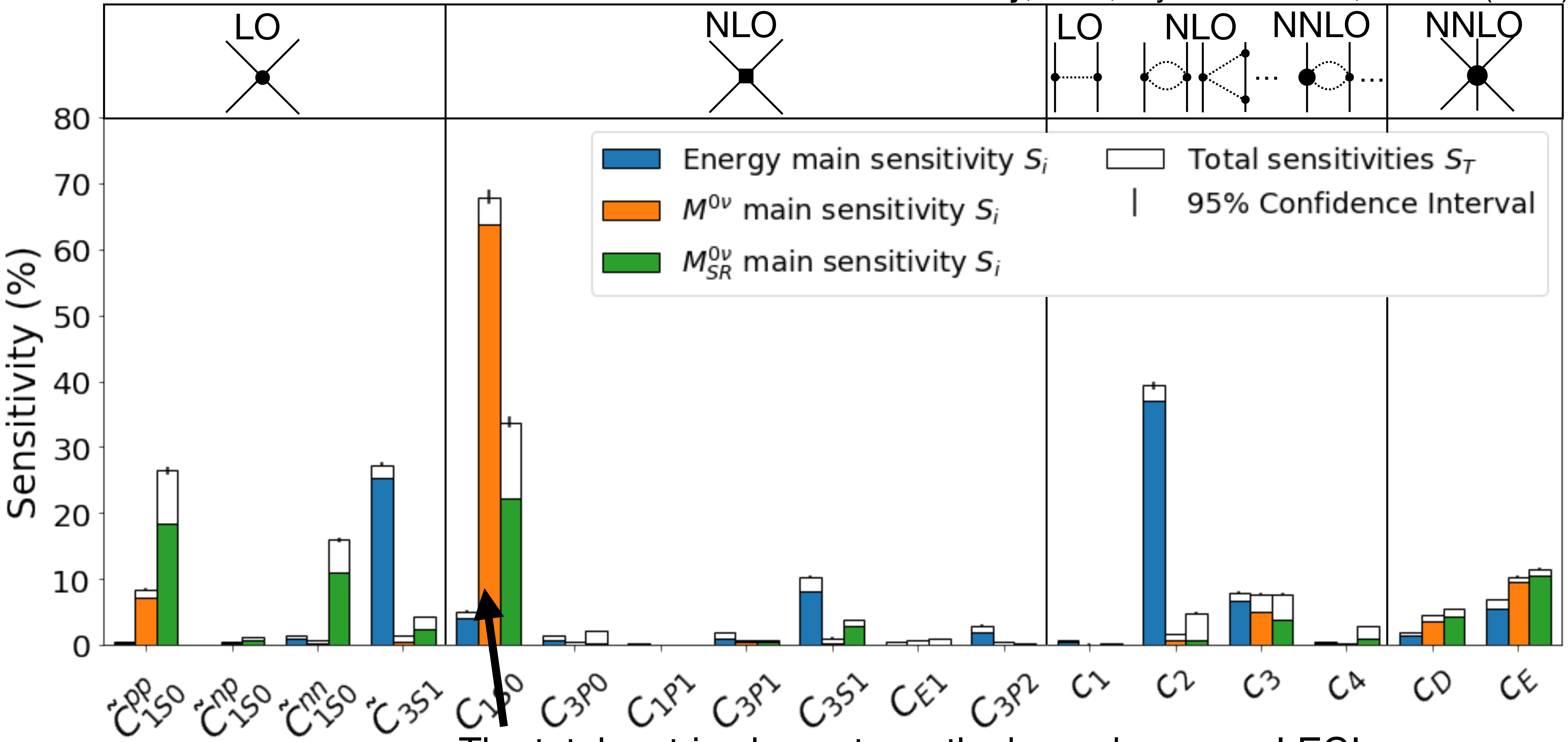
Belley, et al.,Phys. Rev. C **113**, 014319 (2026)



Low-energy constants, i.e. parameters of the nuclear

The MM-DGP Algorithm: GSA

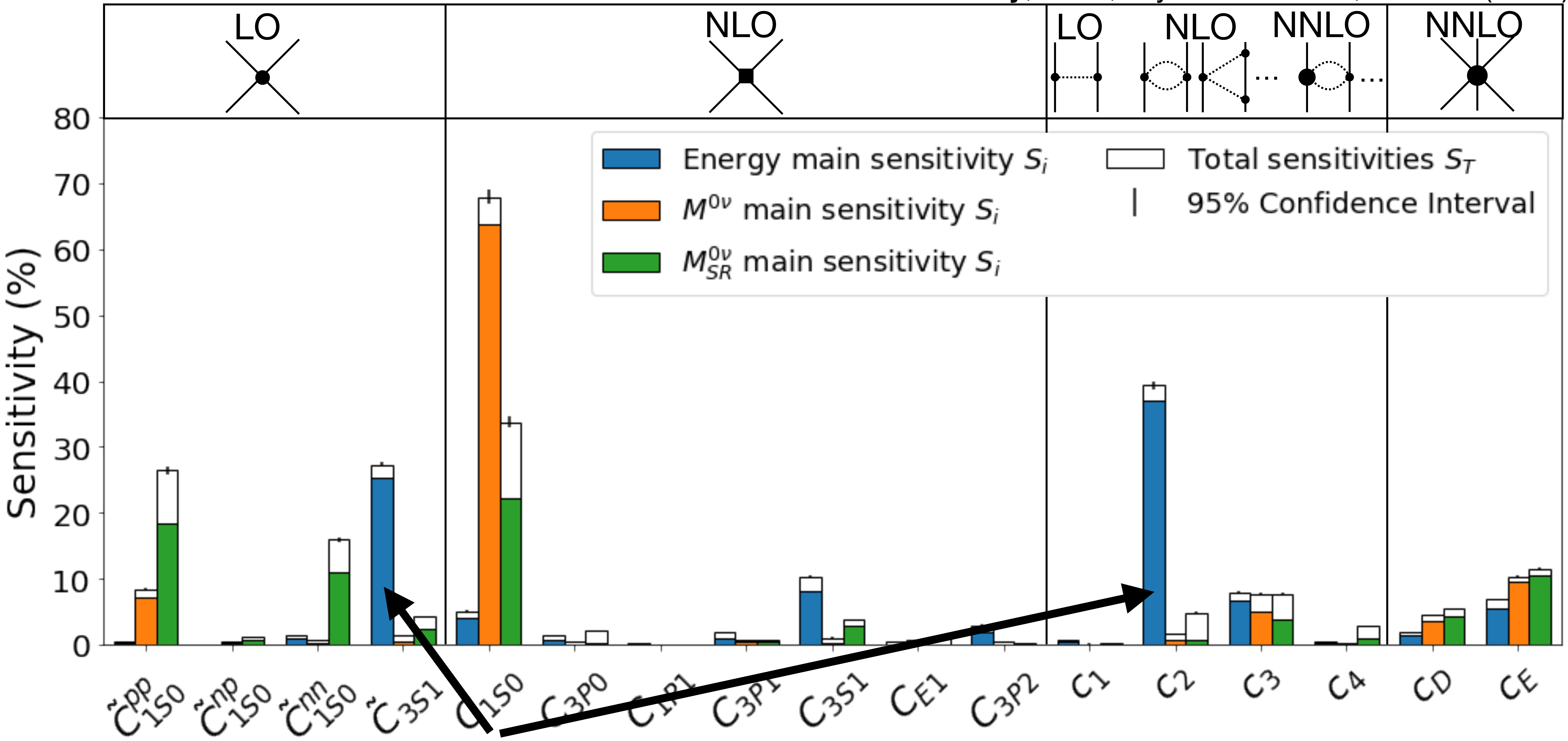
Belley, et al.,Phys. Rev. C **113**, 014319 (2026)



The total matrix element mostly depends on one LEC!

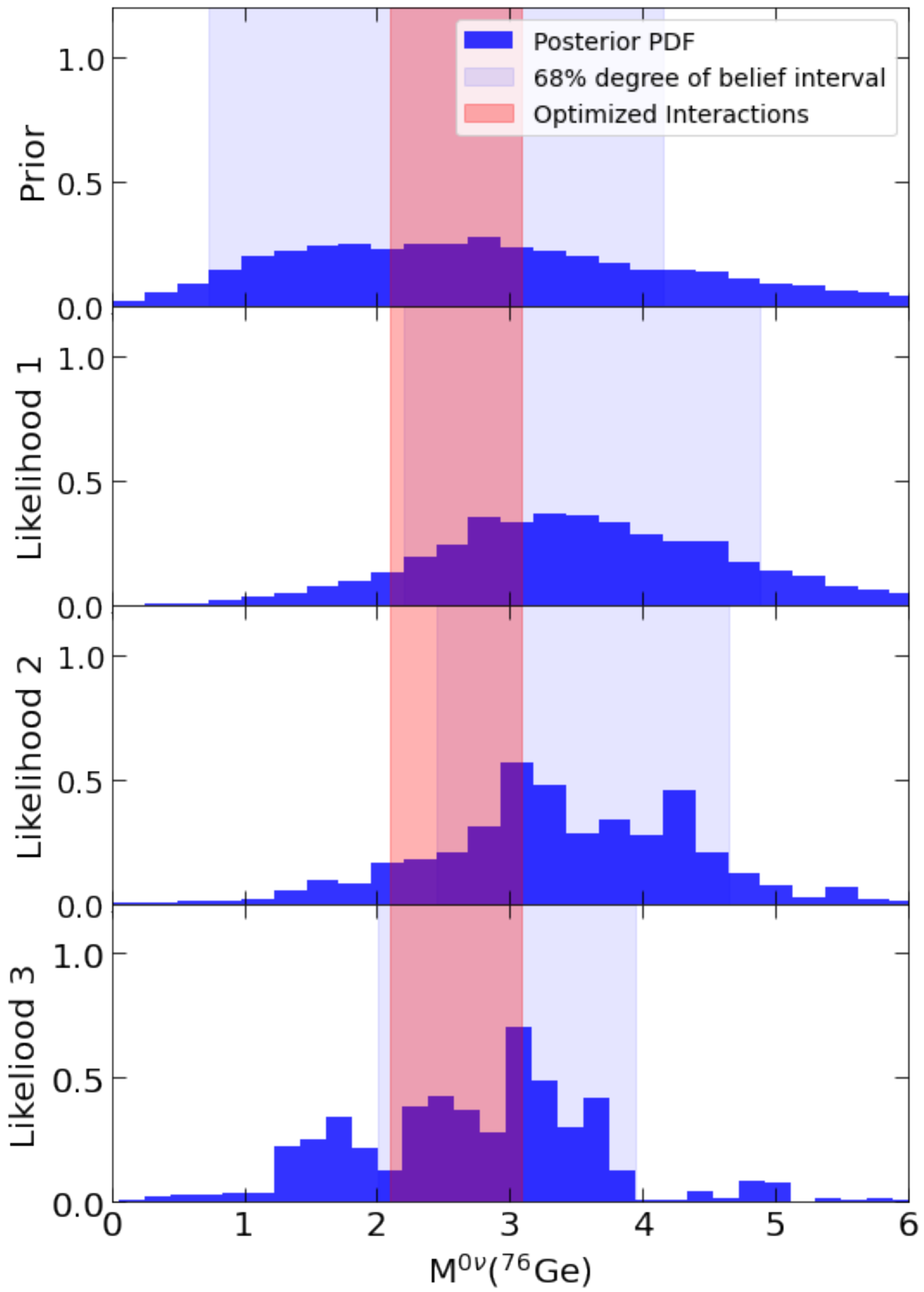
The MM-DGP Algorithm: GSA

Belley, et al.,Phys. Rev. C **113**, 014319 (2026)



Results for energies are consistent with results of physics-based emulators of the coupled cluster method.

Choosing a Likelihood: Example $0\nu\beta\beta$



Likelihood 1: Only contains $^1S_0$ neutron-proton phase shifts at 50 MeV.

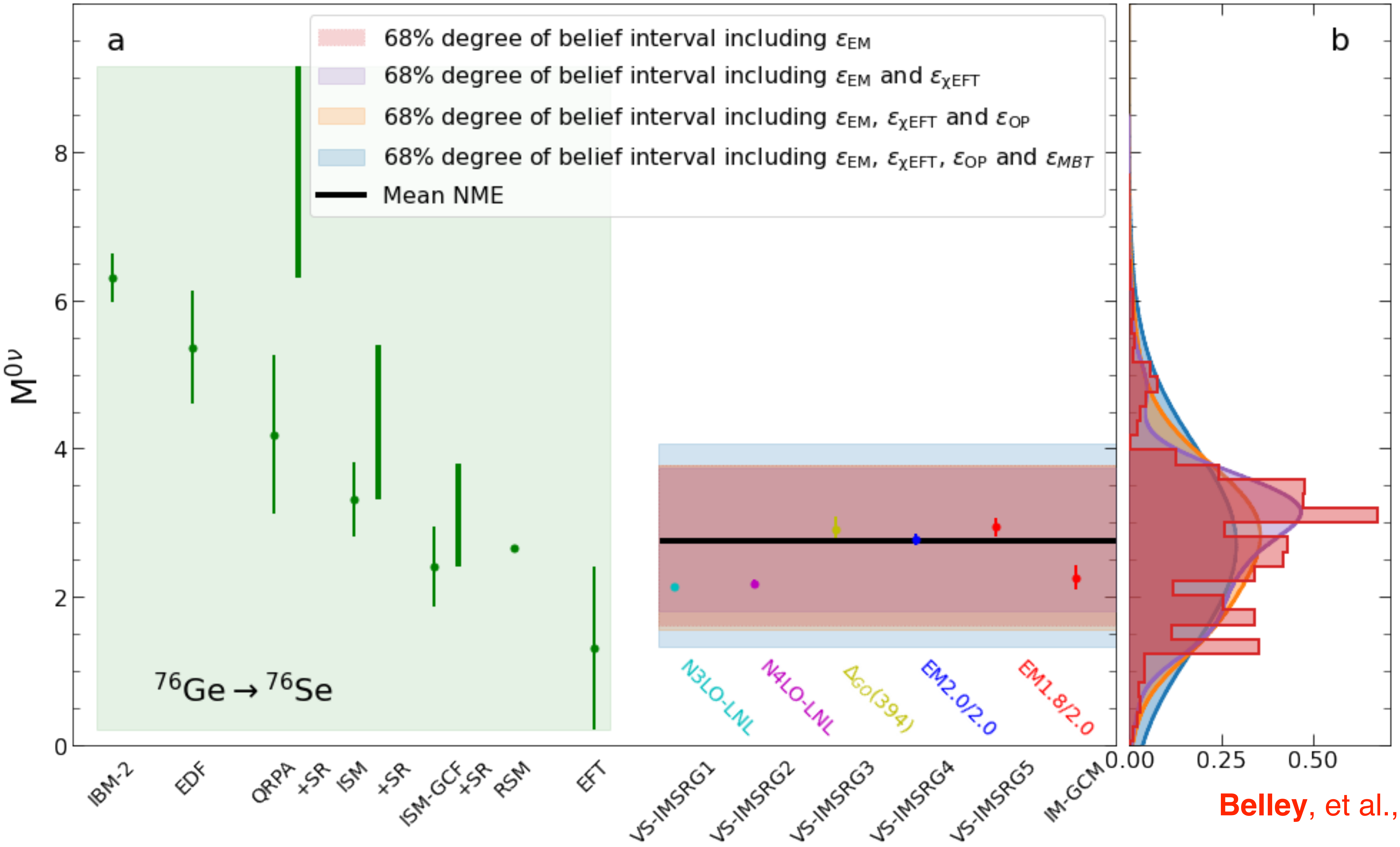
Likelihood 2: Contains $^1S_0$ neutron-proton phase shifts at 50 MeV and observables for A=2-4.

Likelihood 3: Contains $^1S_0$ neutron-proton phase shifts at 50 MeV and observables for A=2-4,16.

A2-4: $E(^2H)$, $r_p(^2H)$, $Q(^2H)$, $E(^3H)$, $E(^4He)$, $r_p(^4He)$

A16: $E(^{16}O)$, $r_p(^{16}O)$

Combining All Sources of Uncertainty



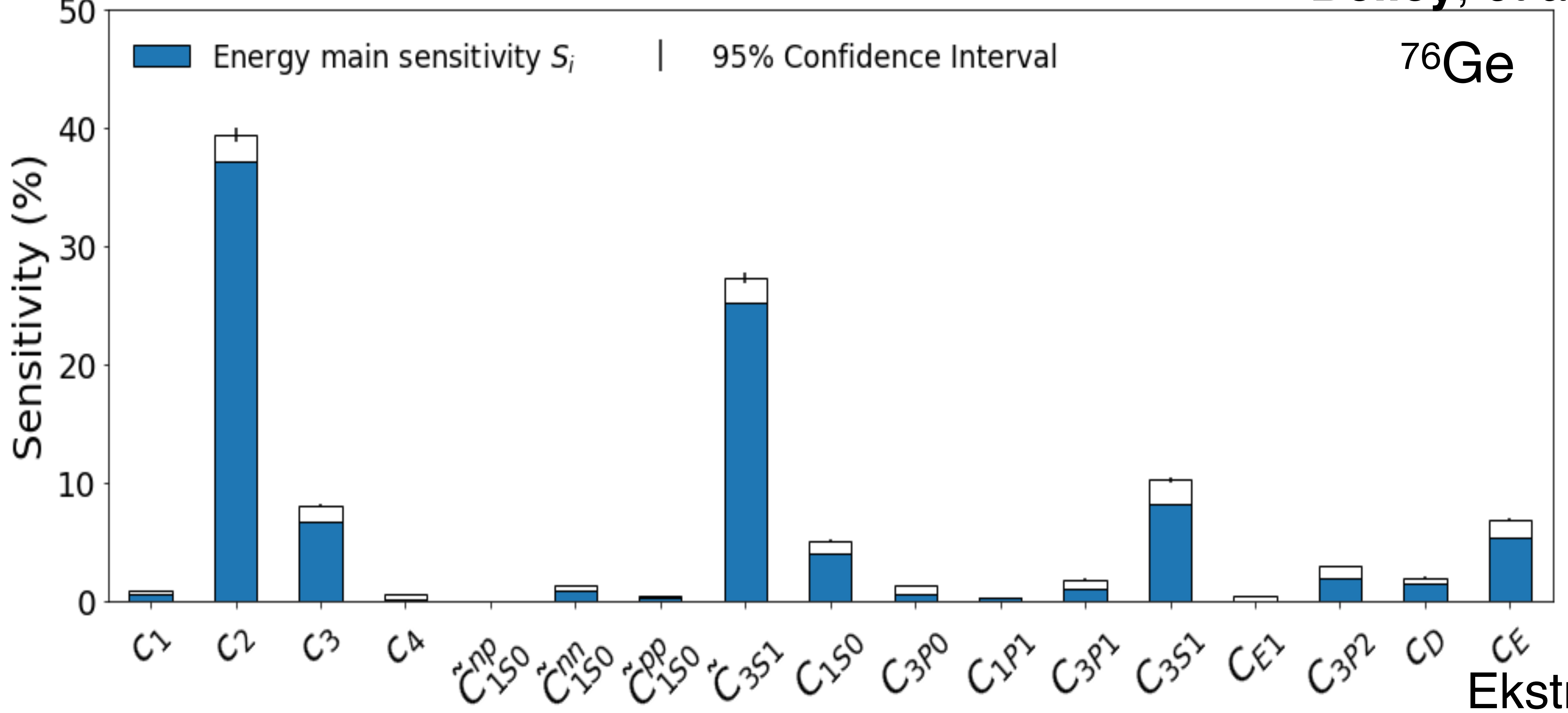
$$M^{0\nu\beta\beta} = 2.60^{+1.28}_{-1.36}$$

Belley, et al., Phys. Rev. Lett. 132, 182502

The MM-DGP Algorithm: GSA

Belley, et al.,Phys. Rev. C **113**, 014319 (2026)

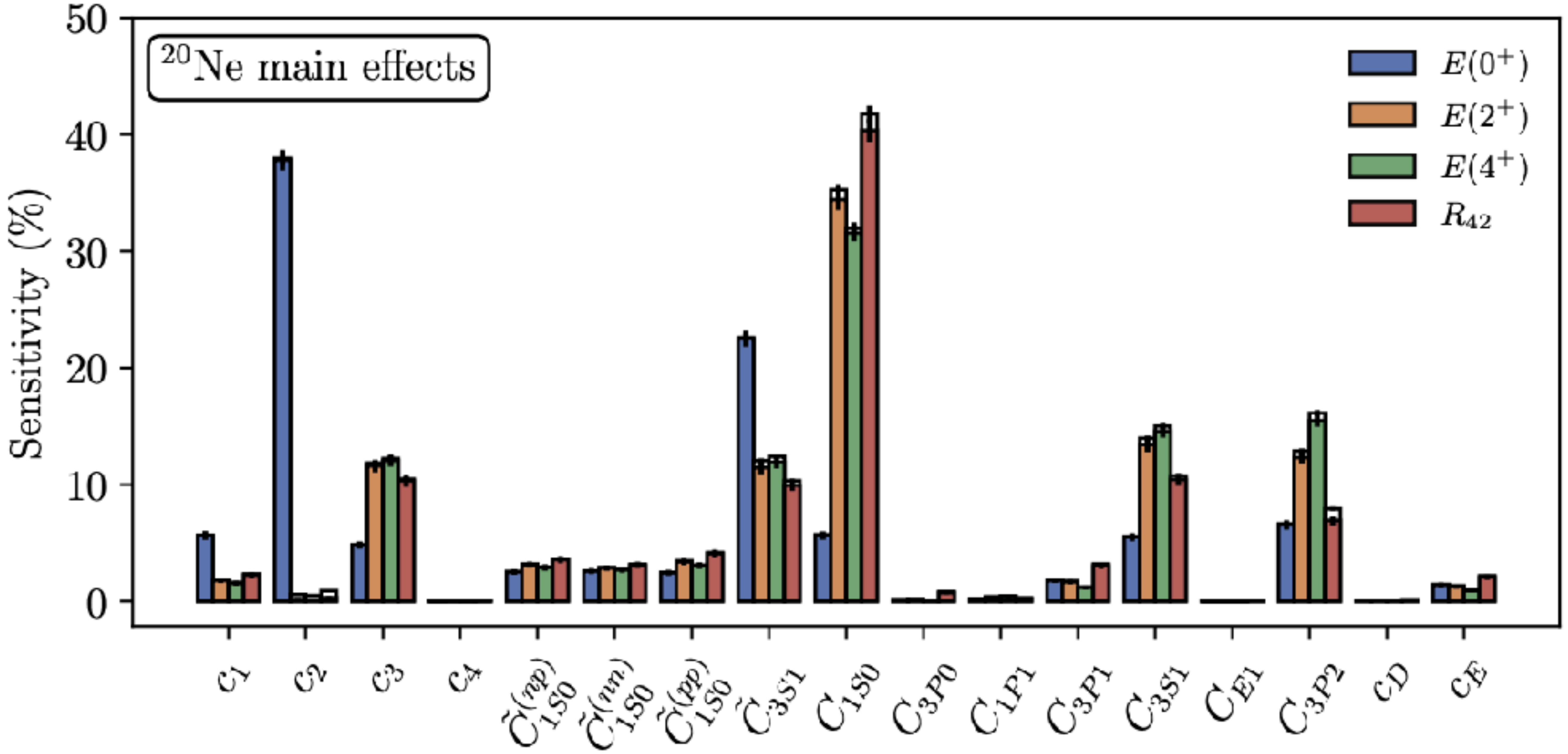
Data driven MM-DGP



Ekström, et al., arXiv:2305.06955 (2023)

vs

Physics Driven

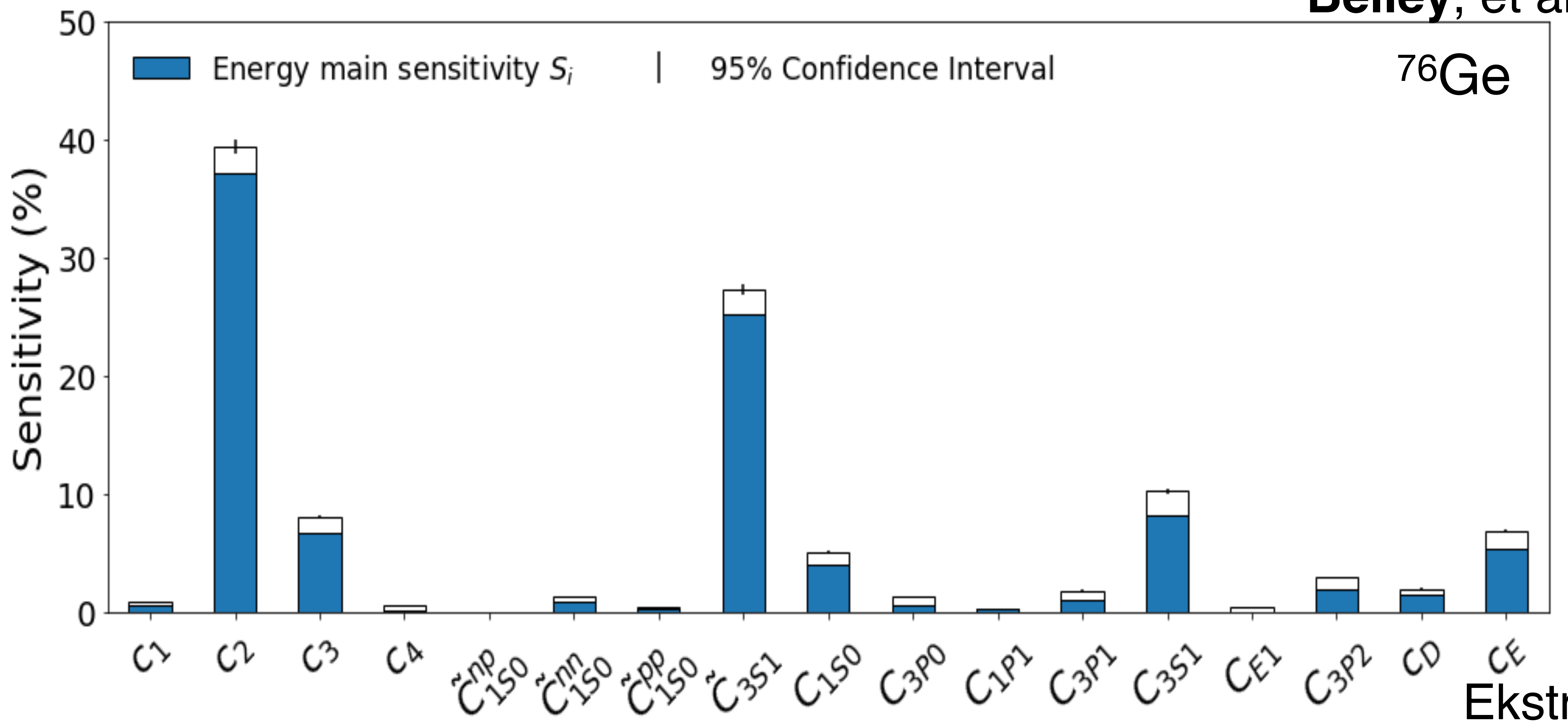


Global emulation

The MM-DGP Algorithm: GSA

Belley, et al.,Phys. Rev. C **113**, 014319 (2026)

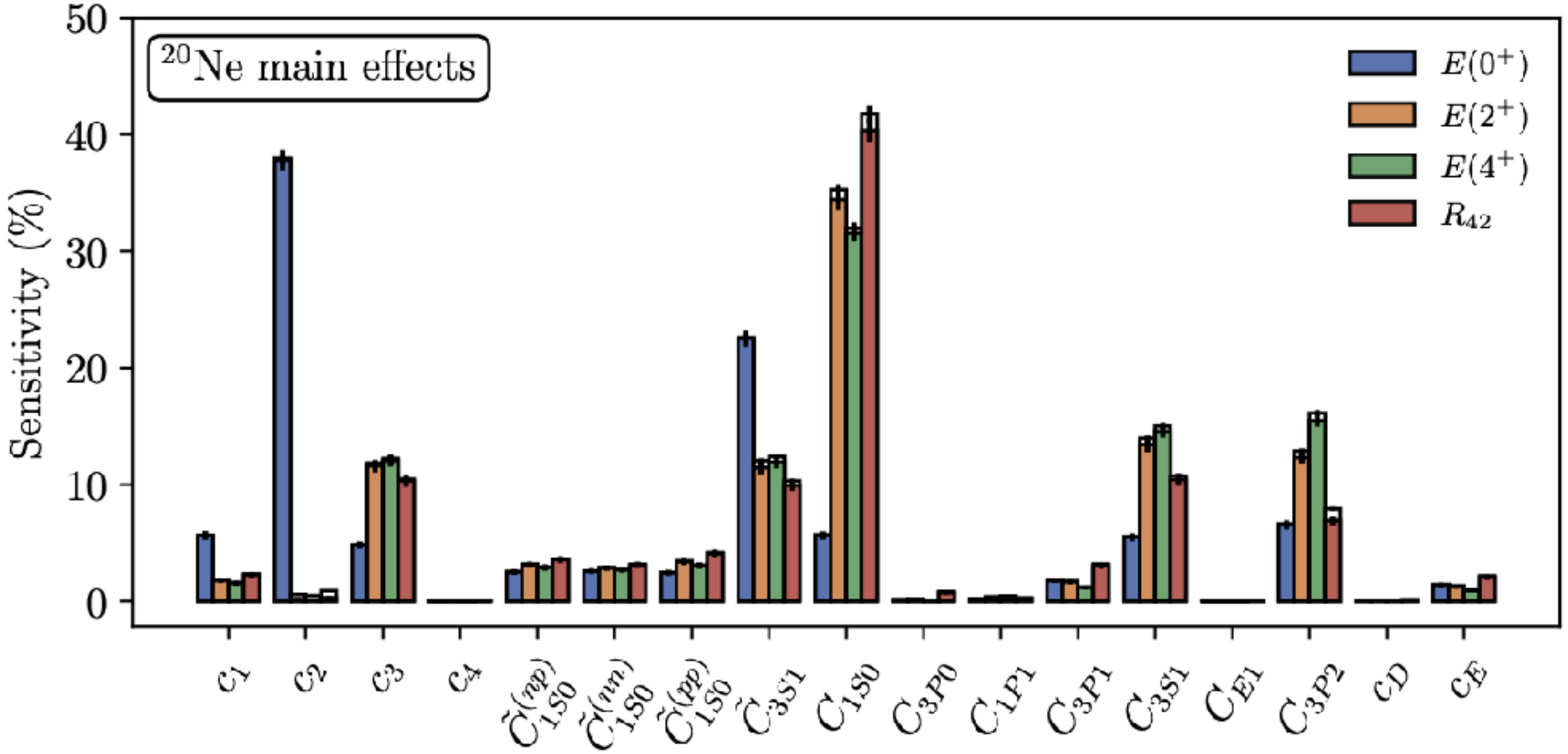
Data driven MM-DGP



Ekström, et al., arXiv:2305.06955 (2023)

vs

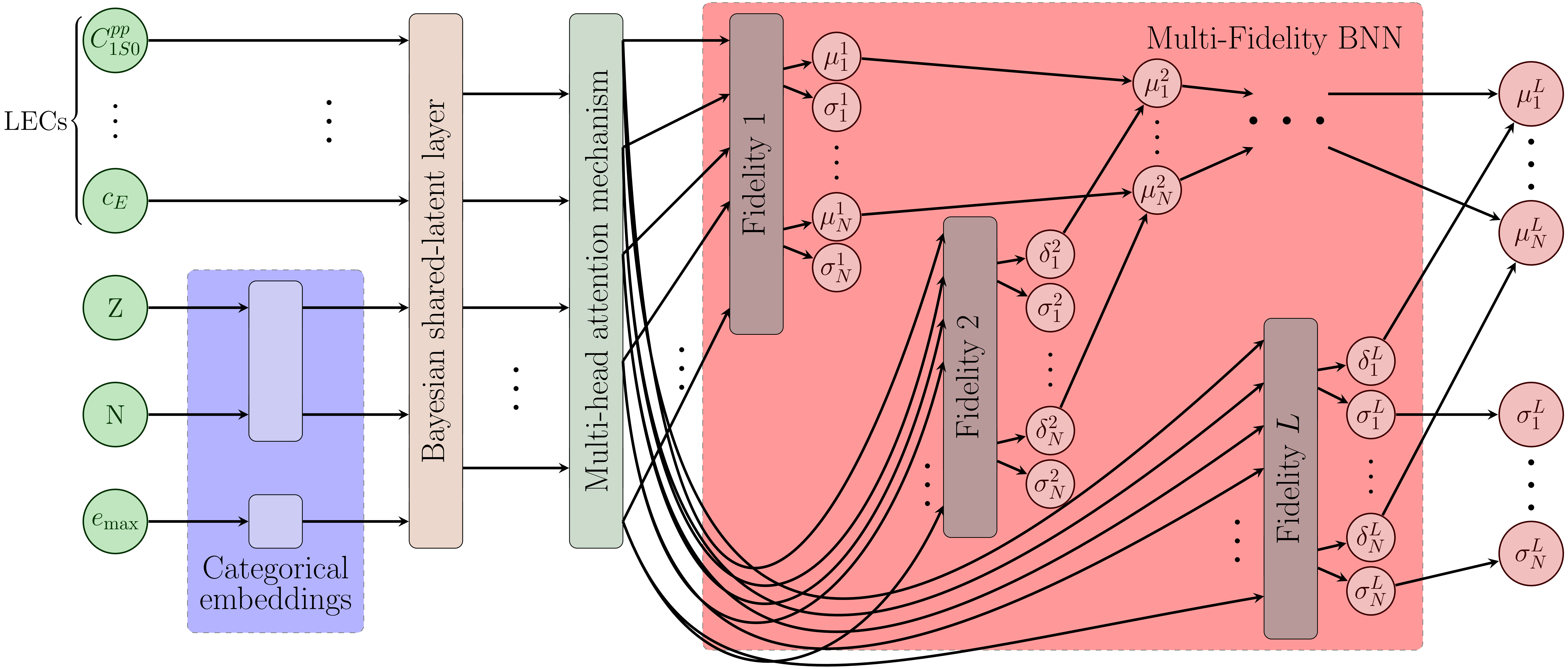
Physics Driven



Jose Miguel Muñoz Arias

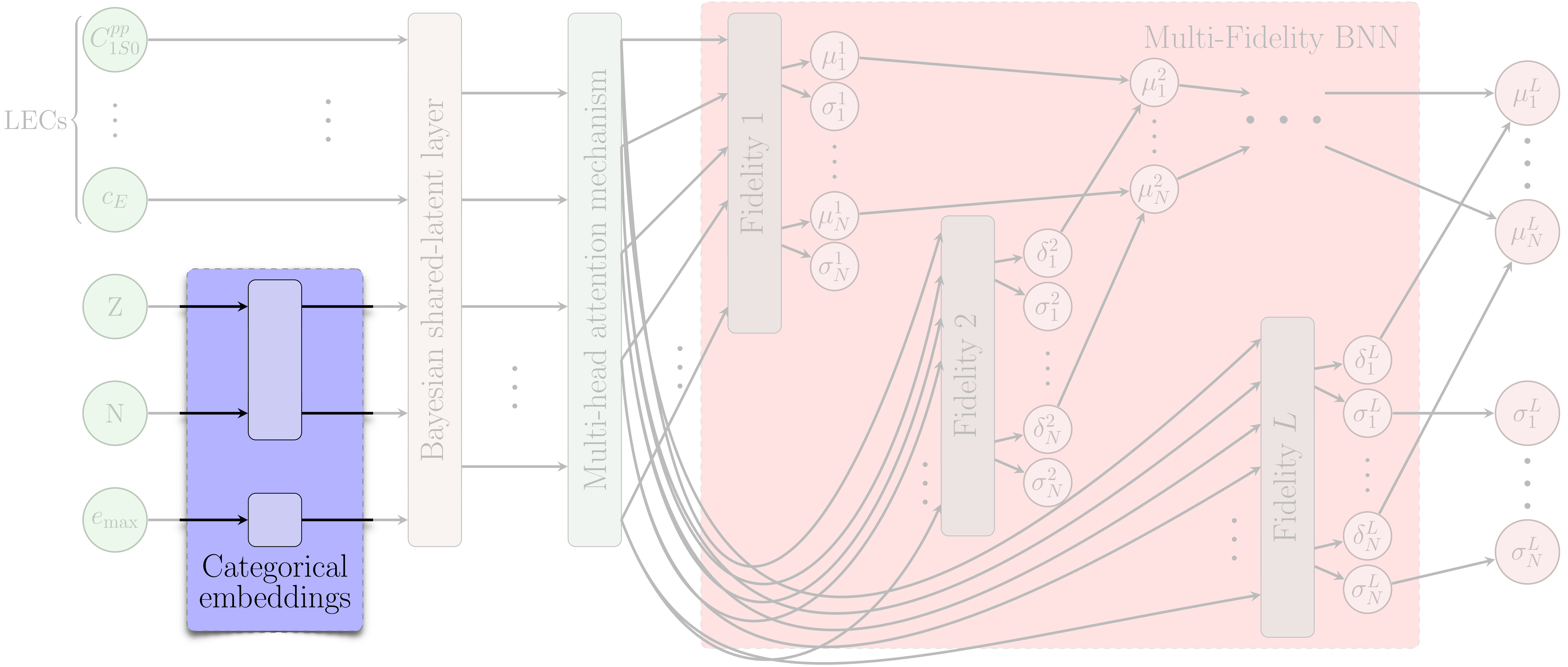
BANNANE

BAyesian Neural Network for Atomic Nuclei Emulation

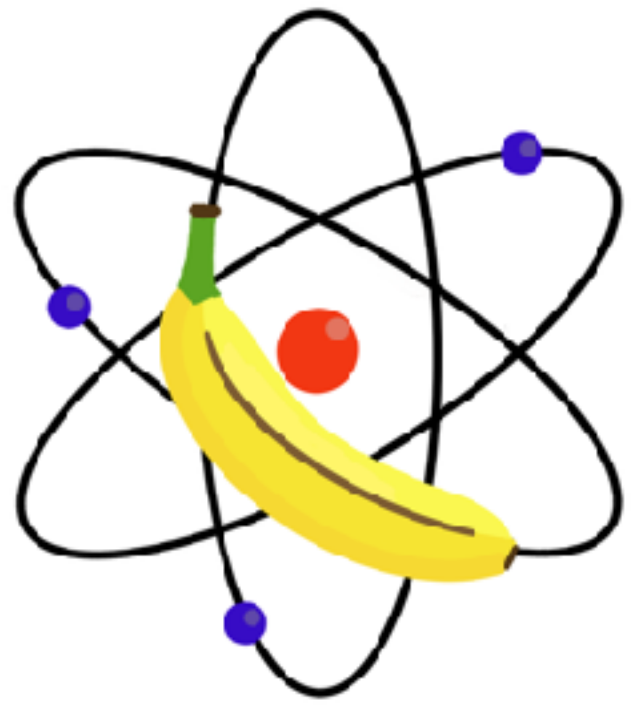


BANNANE

BAyesian Neural Network for Atomic Nuclei Emulation



Embeddings



Embeddings

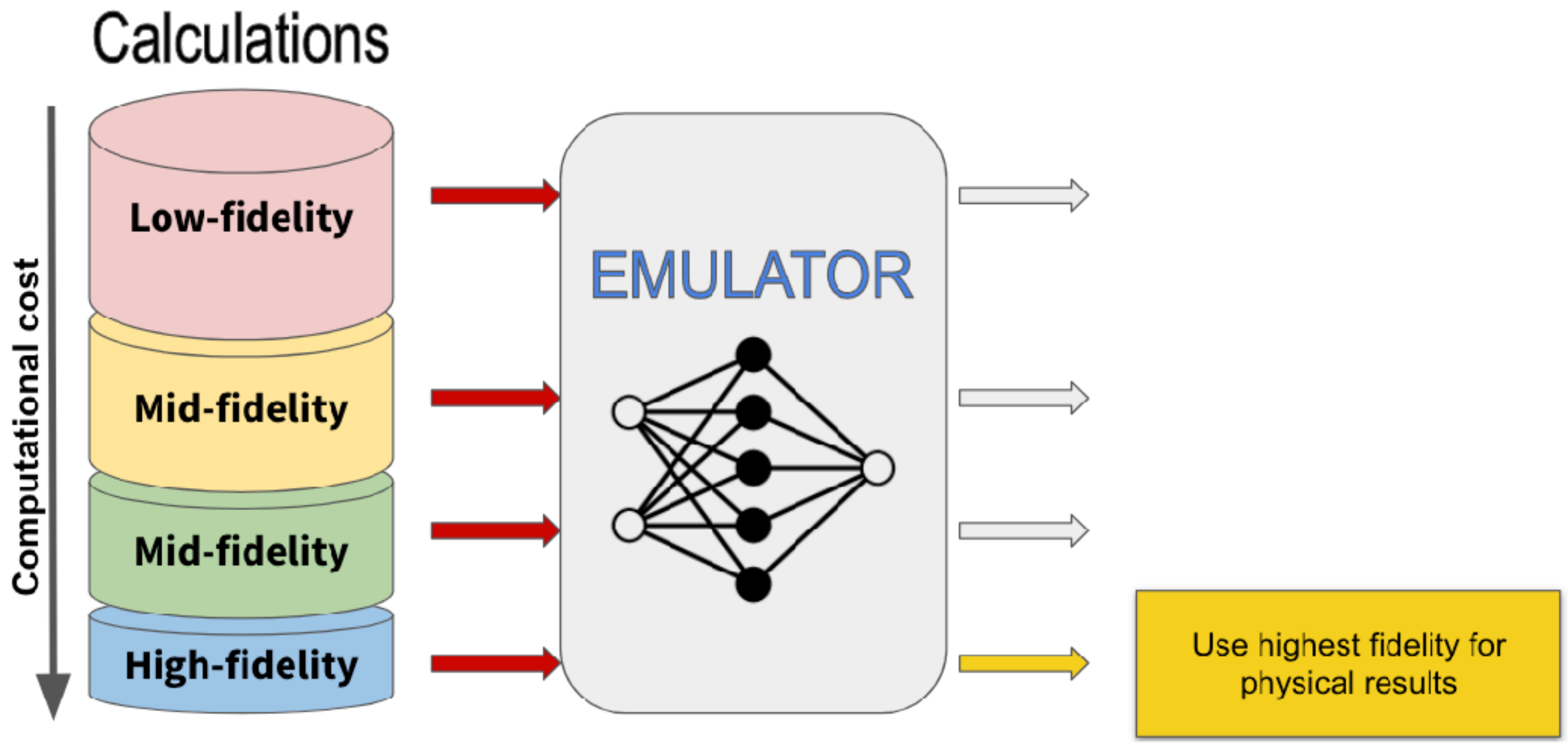
Encodes discrete data into a vector space, while learning their relative positions, e.g.

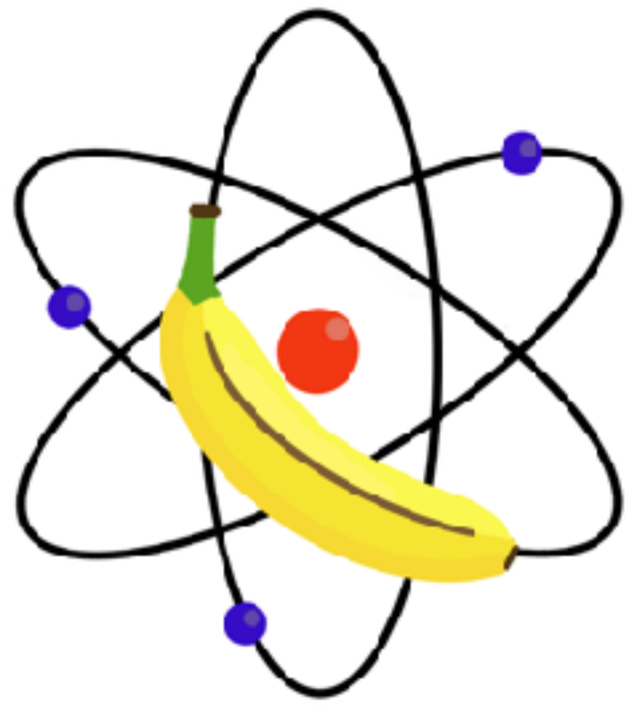
- $e_{max} = 4 \rightarrow 6 \rightarrow 8 \rightarrow 10$

Embeddings

Encodes discrete data into a vector space, while learning their relative positions, e.g.

- $e_{max} = 4 \rightarrow 6 \rightarrow 8 \rightarrow 10$





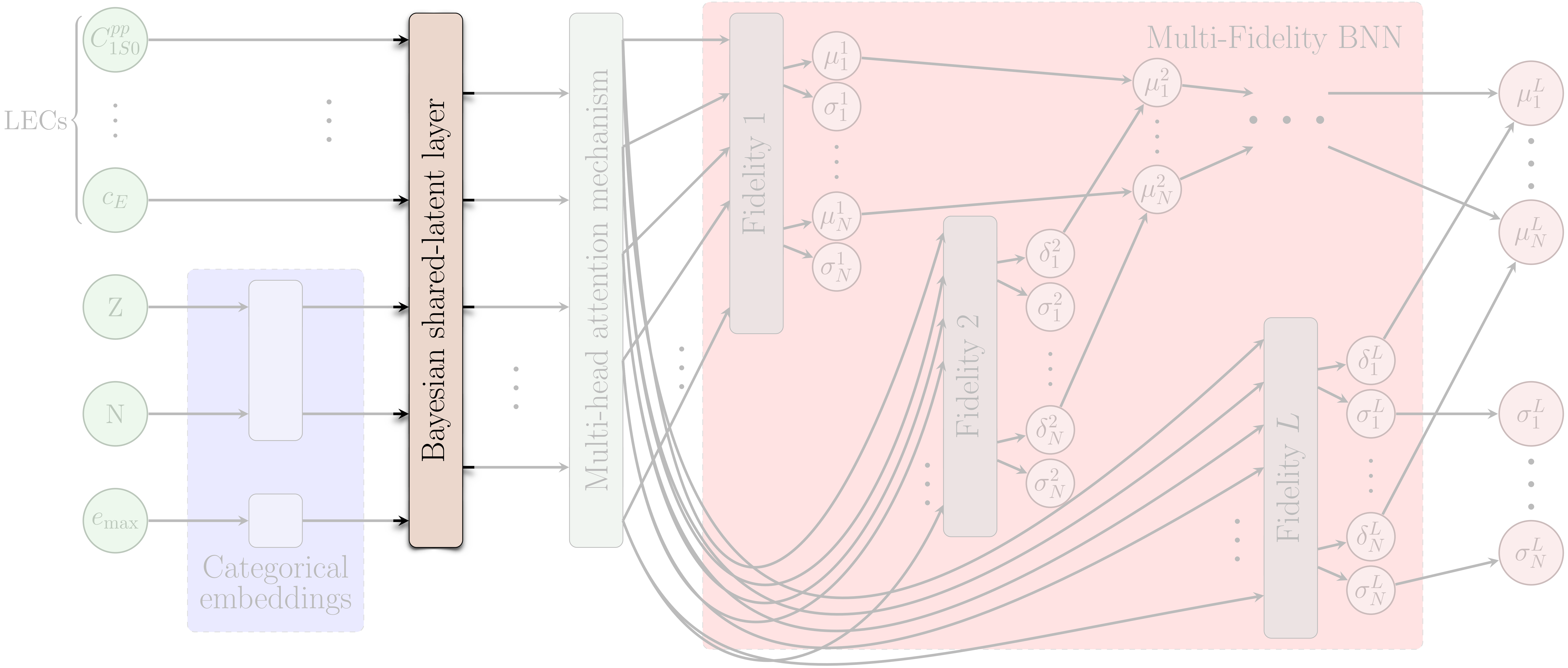
Embeddings

Encodes discrete data into a vector space, while learning their relative positions, e.g.

- $e_{max} = 4 \rightarrow 6 \rightarrow 8 \rightarrow 10$
- **Positions in the nuclear chart: ordering of N and Z**

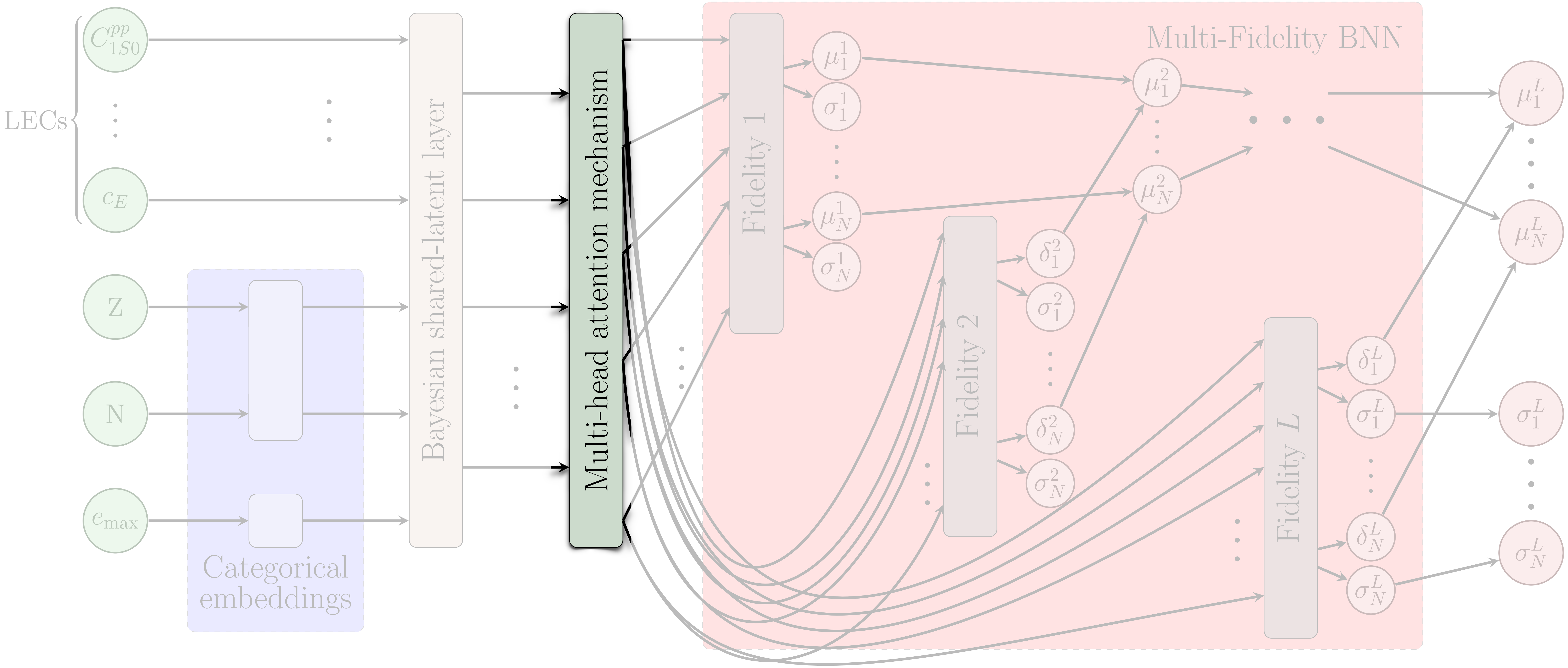
BANNANE

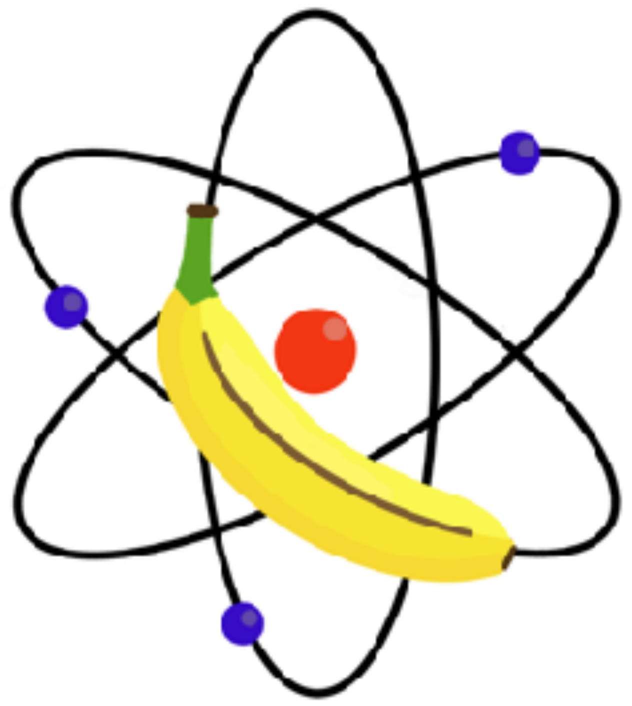
BAyesian Neural Network for Atomic Nuclei Emulation



BANNANE

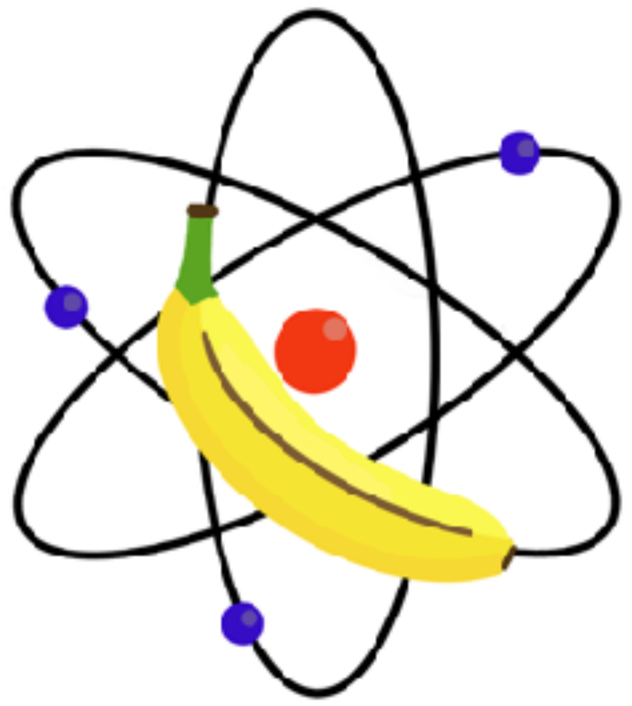
BAyesian Neural Network for Atomic Nuclei Emulation





Attention!

- **Attention Mechanisms learns how the embeddings need to be adapted due to other inputs**
- **Responsible to for the improvements of large language models in recent years!**



Attention!

- **Attention Mechanisms learns how the embeddings need to be adapted due to other inputs**
- **Responsible to for the improvements of large language models in recent years!**

Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com

Noam Shazeer* Google Brain noam@google.com

Niki Parmar* Google Research nikip@google.com

Jakob Uszkoreit* Google Research usz@google.com

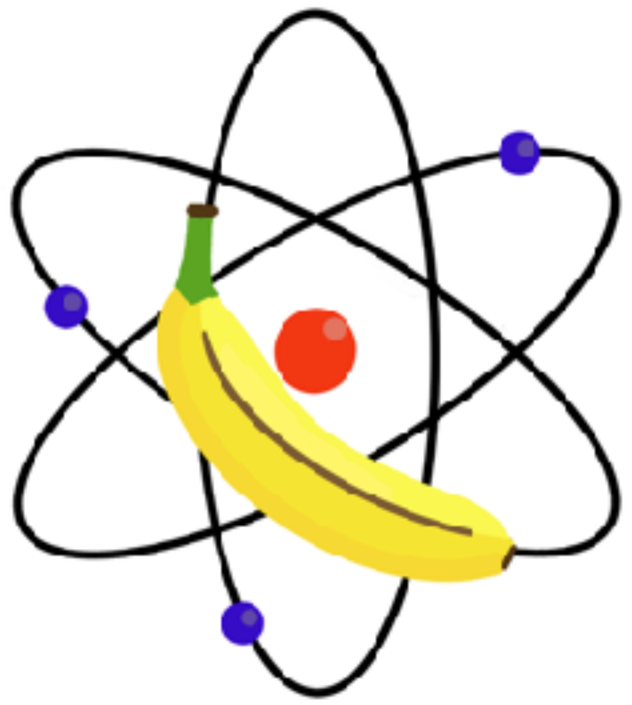
Llion Jones* Google Research llion@google.com

Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu

Łukasz Kaiser* Google Brain lukaszkaiser@google.com

Illia Polosukhin* ‡ illia.polosukhin@gmail.com

244, 184 citations



Attention!

- **Attention Mechanisms learns how the embeddings need to be adapted due to other inputs**
- **Responsible to for the improvements of large language models in recent years!**

Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com

Noam Shazeer* Google Brain noam@google.com

Niki Parmar* Google Research nikip@google.com

Jakob Uszkoreit* Google Research usz@google.com

Llion Jones* Google Research llion@google.com

Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu

Łukasz Kaiser* Google Brain lukaszkaiser@google.com

Illia Polosukhin* ‡ illia.polosukhin@gmail.com

244, 184 citations

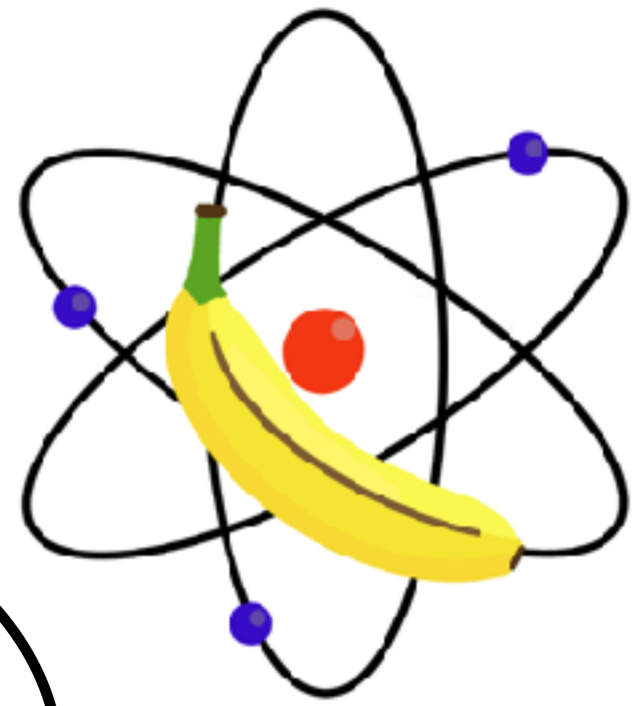
Highly accurate protein structure prediction with AlphaFold

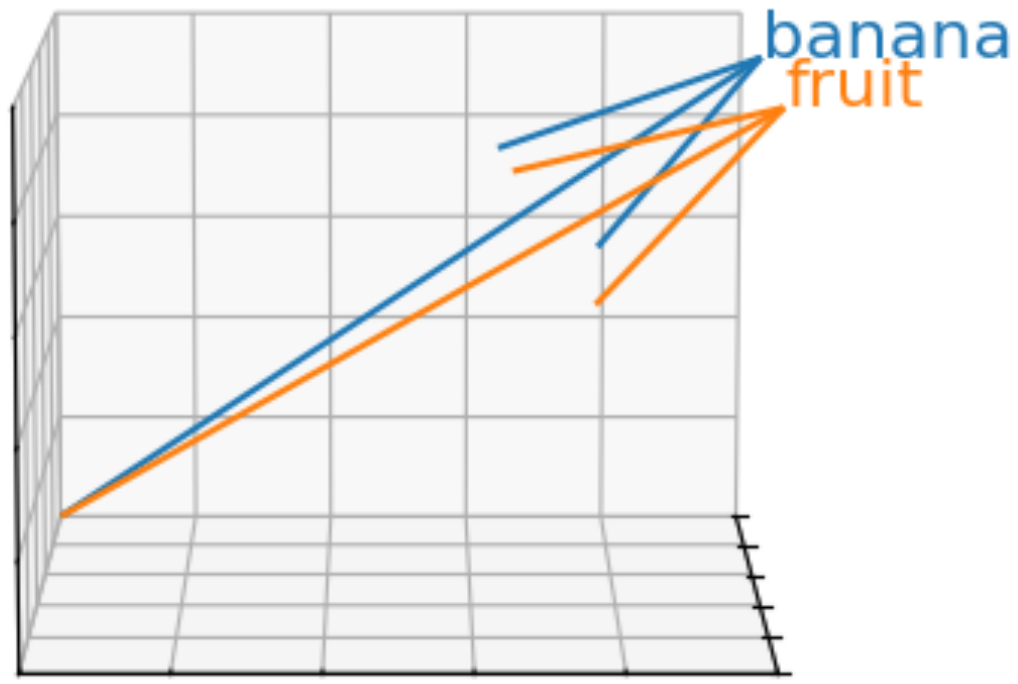
John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, ... Demis Hassabis + Show authors

48, 990 citations

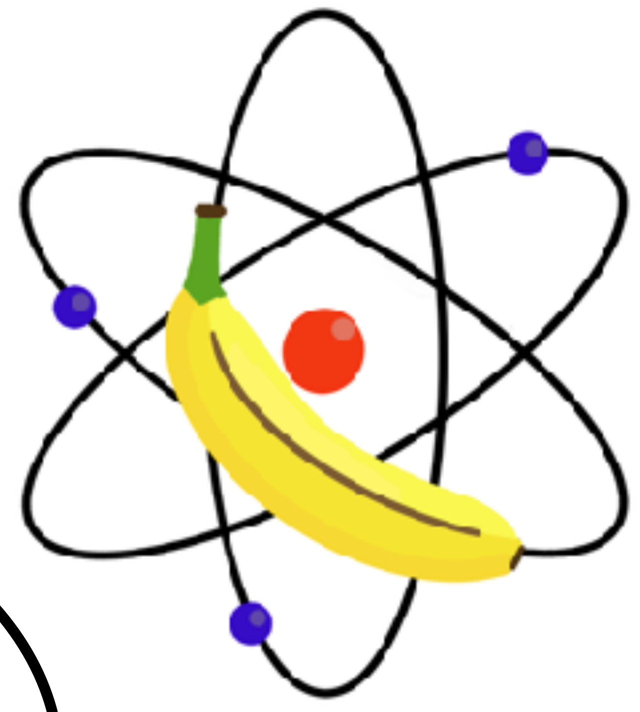
Attention!

Attention!

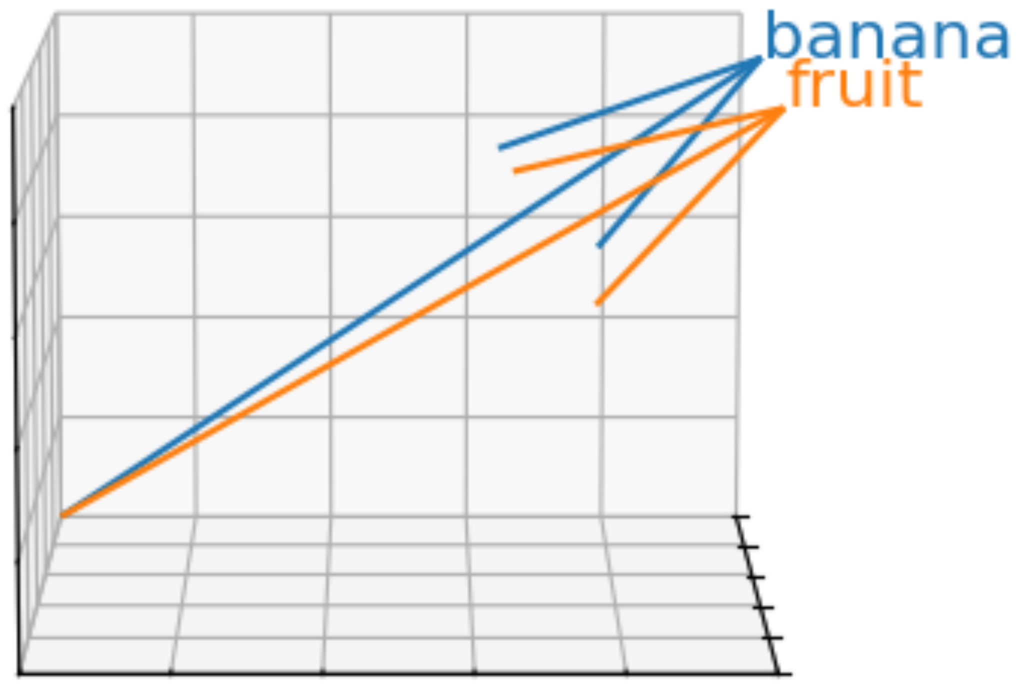




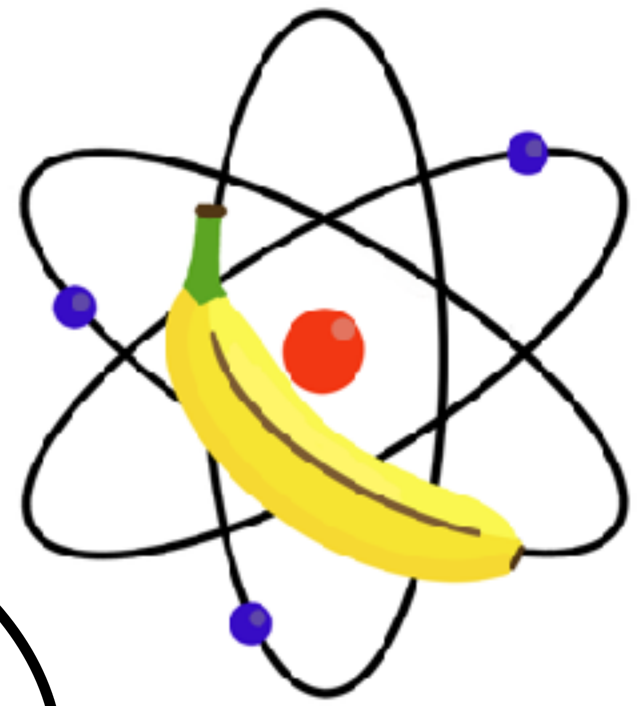
Attention!

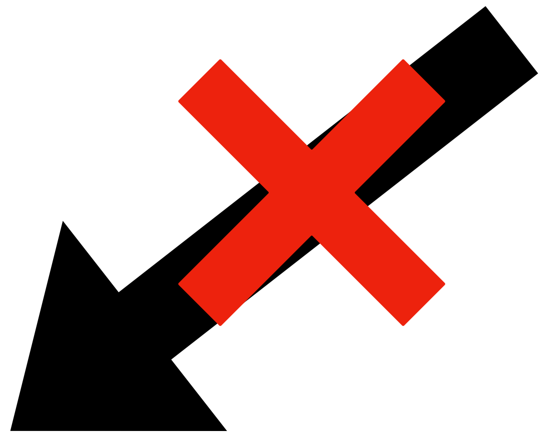


ChatGPT:

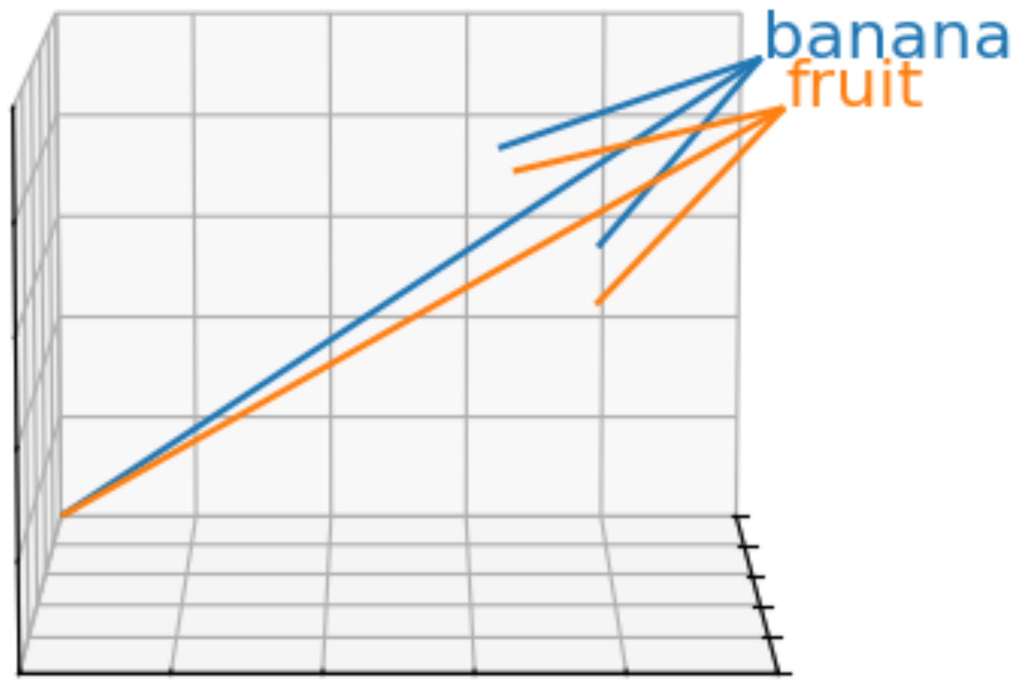


Attention!

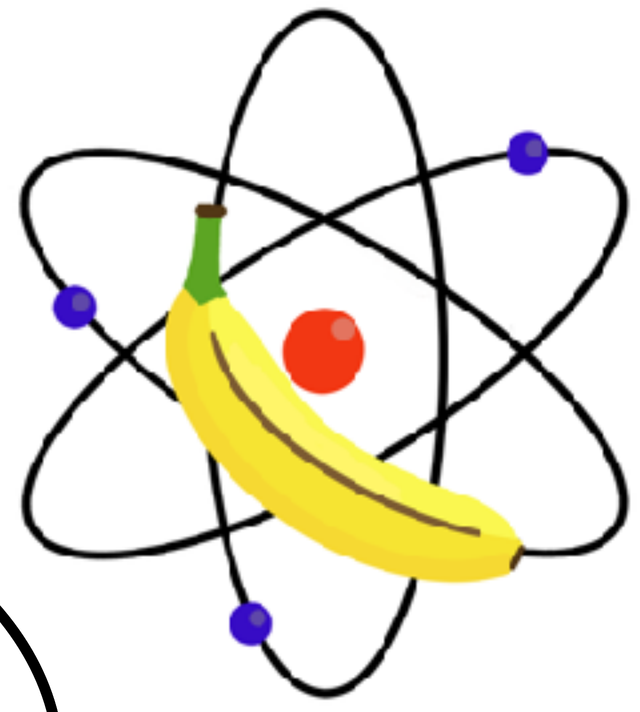




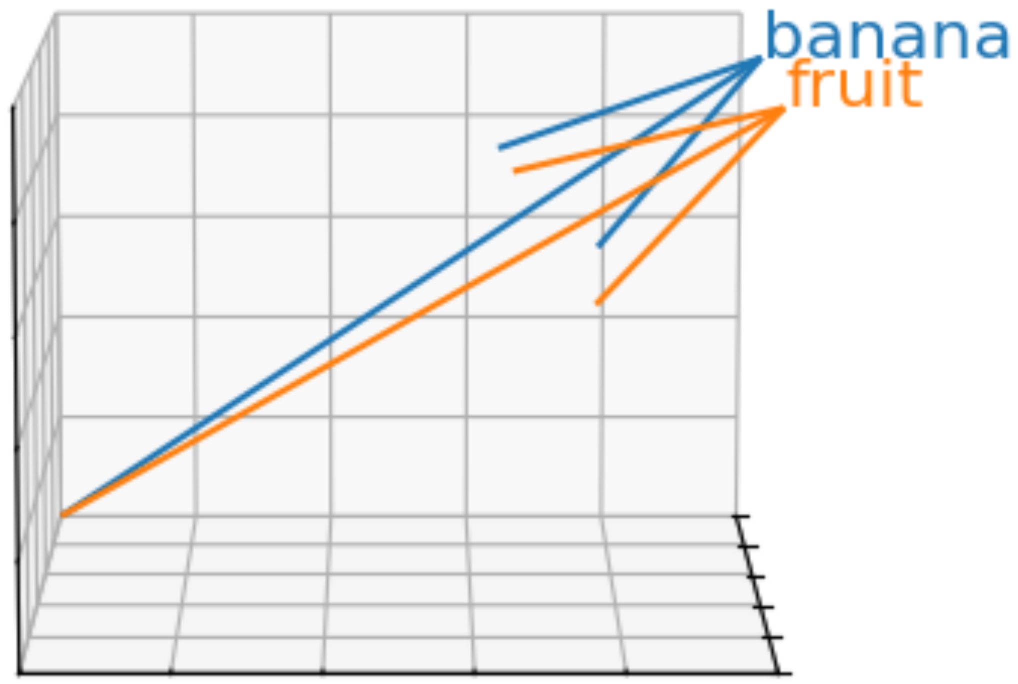
ChatGPT:

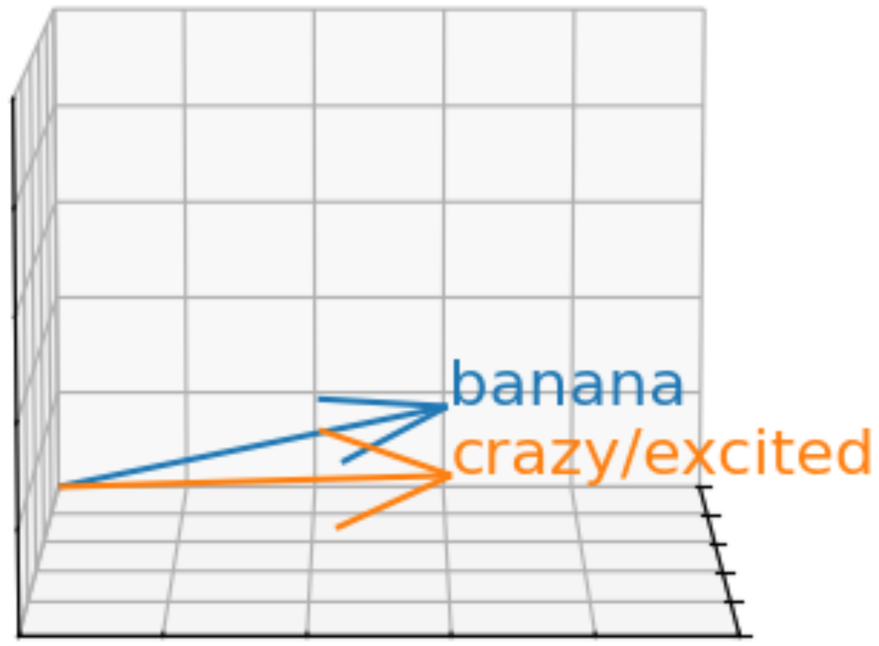


Attention!

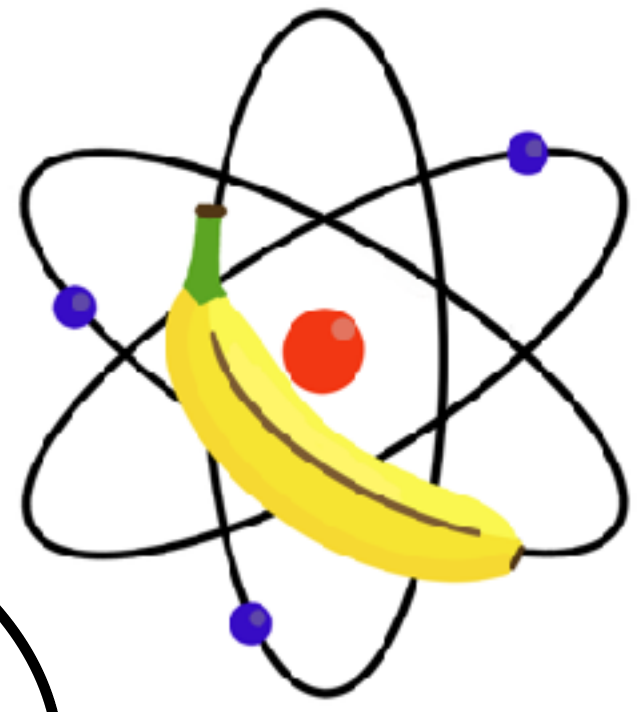


ChatGPT:

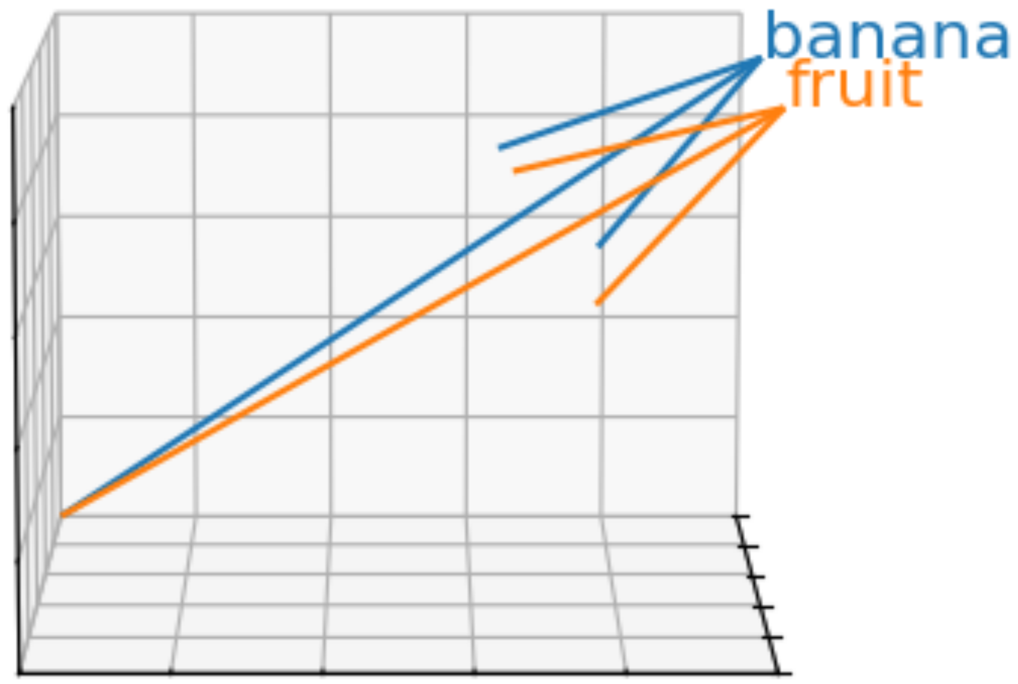


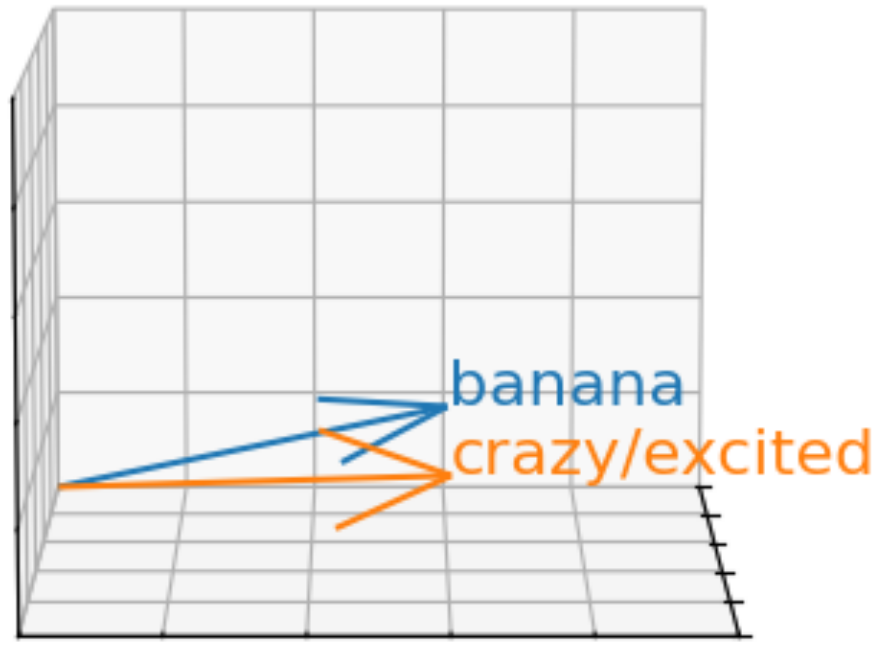


Attention!

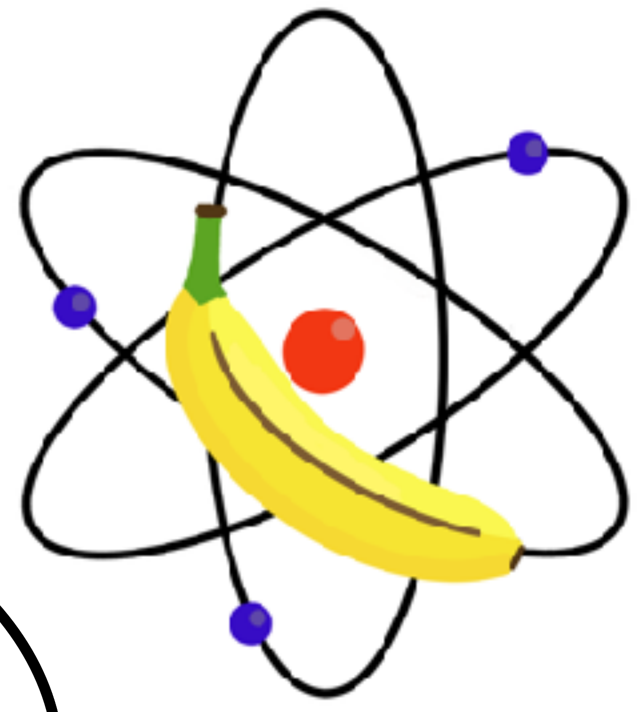


ChatGPT:

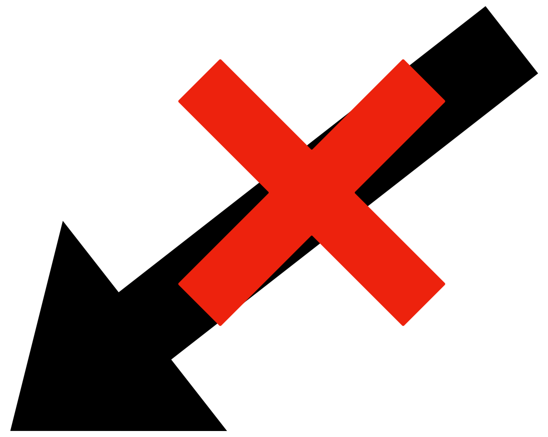




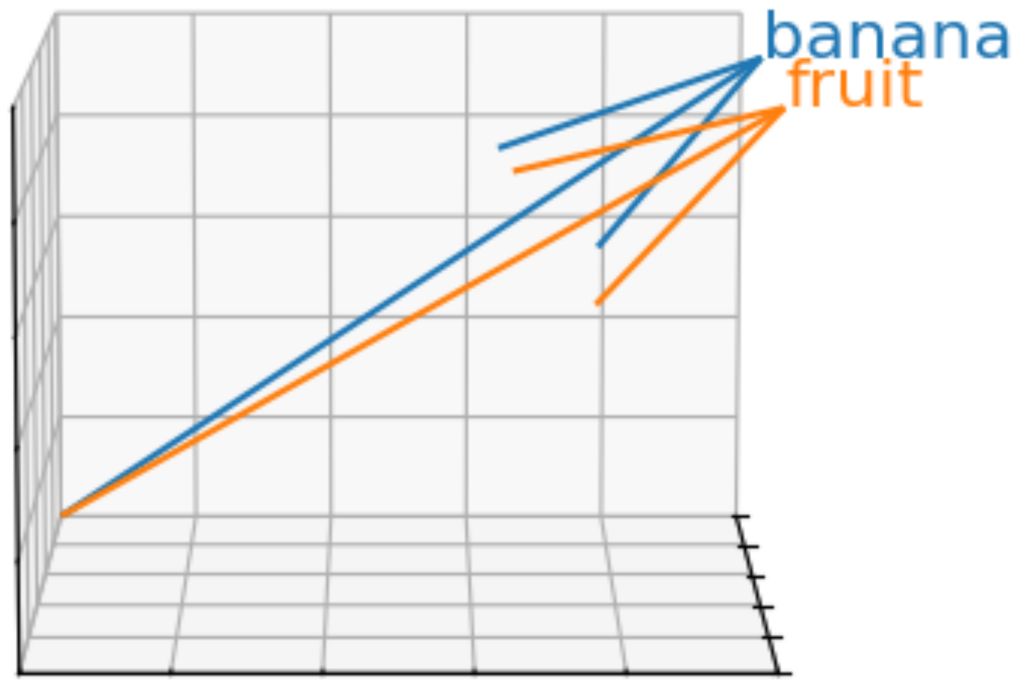
Attention!

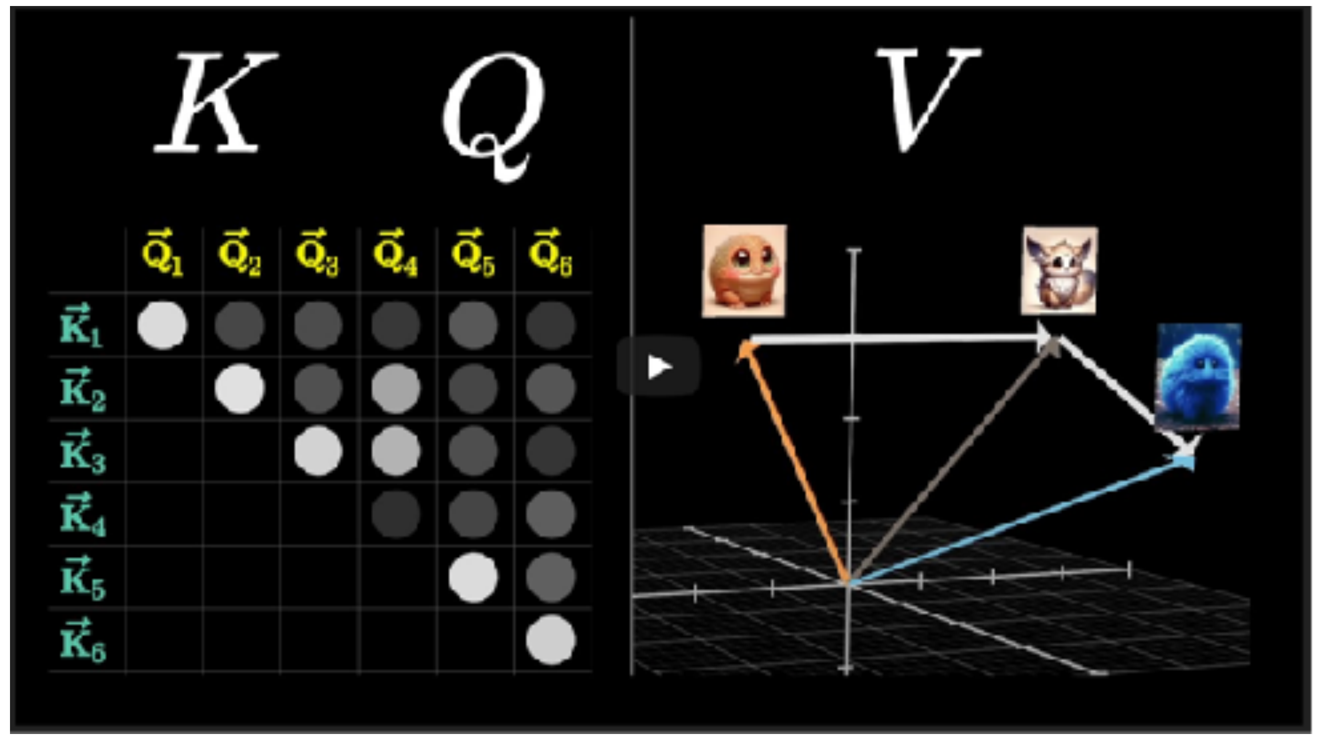


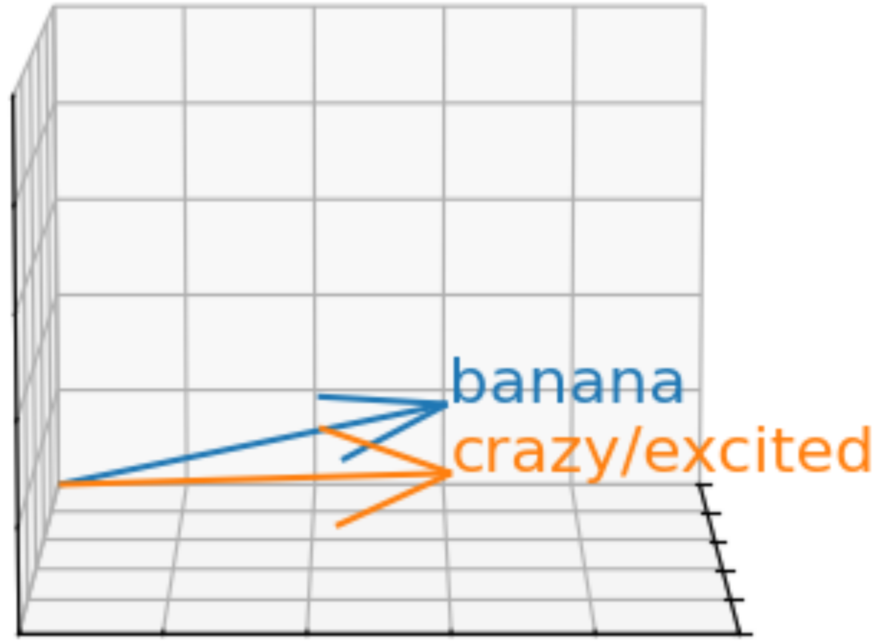
ChatGPT:



3Blue1Brown:

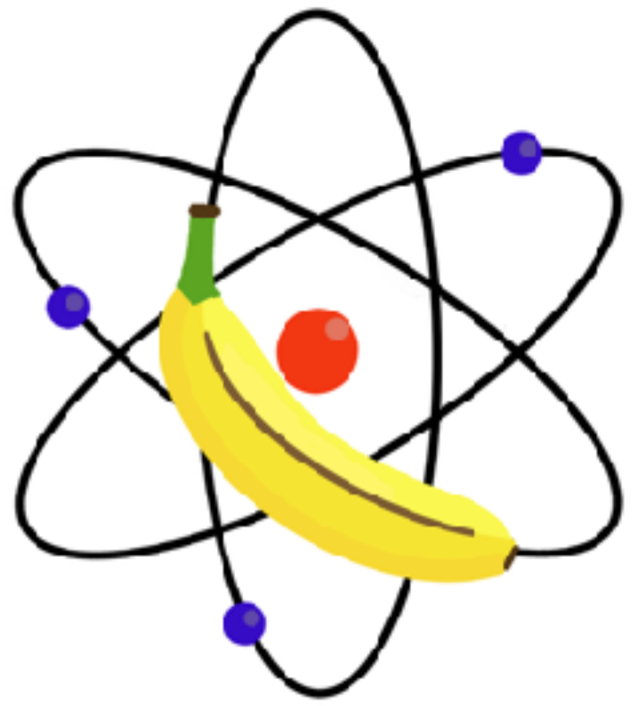




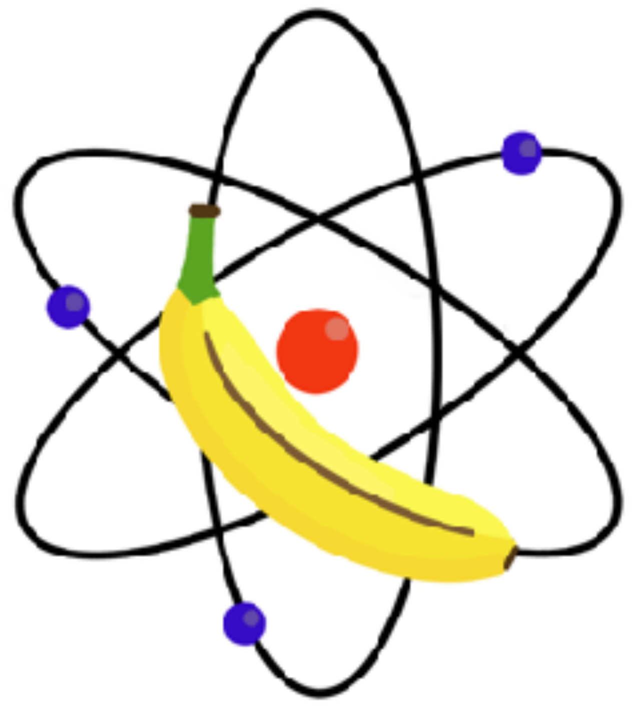


Attention

Embeddings: $\vec{x}_i$'s



Embeddings: $\vec{x}_i$'s $\longleftarrow$ N input of dimension M   i.e. X -> NxM



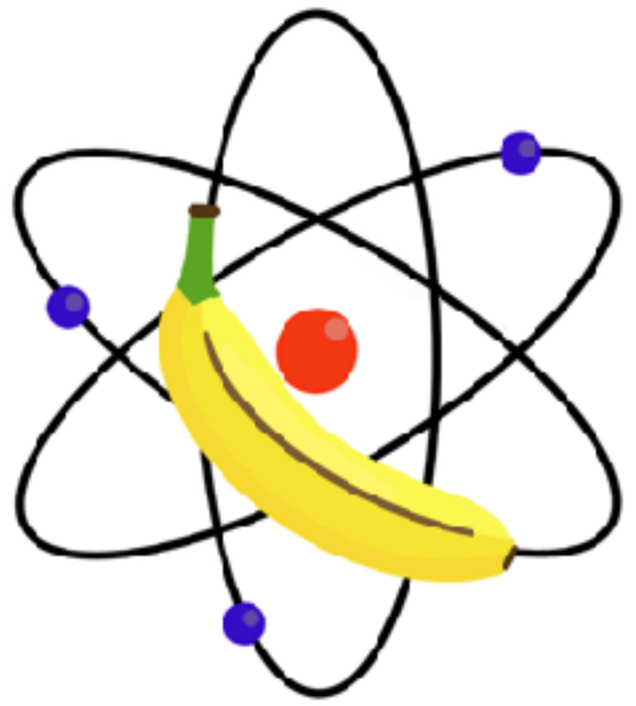
Attention

Embeddings: $\vec{x}_i$'s ⟵ N input of dimension M   i.e. X -> NxM

Query matrix: $W_Q \rightarrow Q = W_Q X$

Key matrix: $W_K \rightarrow K = W_K X$

Value matrix: $W_V \rightarrow V = W_V X$



Attention

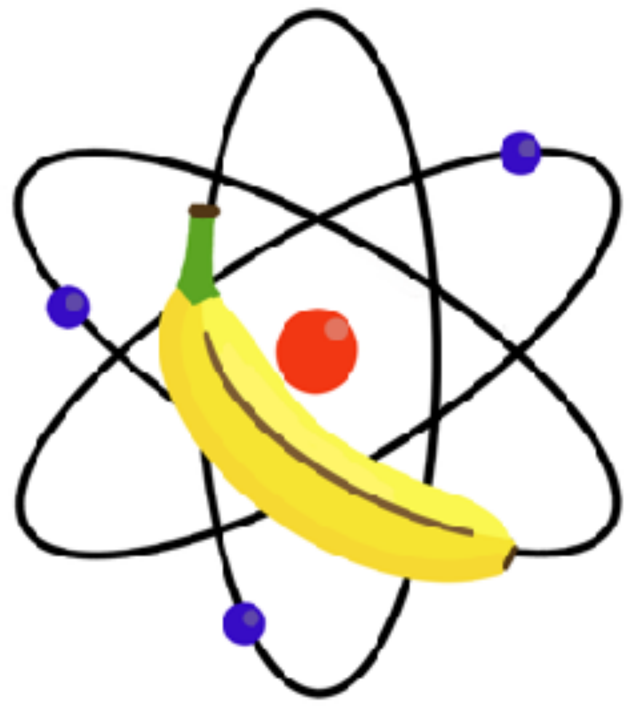
Embeddings: $\vec{x}_i$'s ⟵ N input of dimension M   i.e. X -> NxM

Query matrix: $W_Q \rightarrow Q = W_Q X$

Key matrix: $W_K \rightarrow K = W_K X$

Value matrix: $W_V \rightarrow V = W_V X$

All matrix elements of these matrices are tunable parameters.



Attention

Embeddings: $\vec{x}_i$'s ⟵ N input of dimension M   i.e. X -> NxM

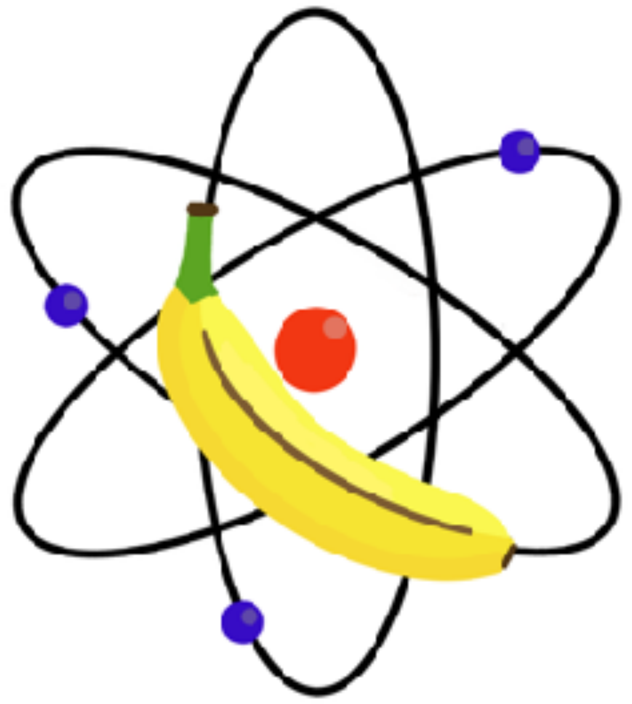
Query matrix: $W_Q \rightarrow Q = W_Q X$

Key matrix: $W_K \rightarrow K = W_K X$

Value matrix: $W_V \rightarrow V = W_V X$

All matrix elements of these matrices are tunable parameters.

$$\boxed{Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d}}\right)V}$$



Attention

Embeddings: $\vec{x}_i$'s ⟵ N input of dimension M   i.e. X -> NxM

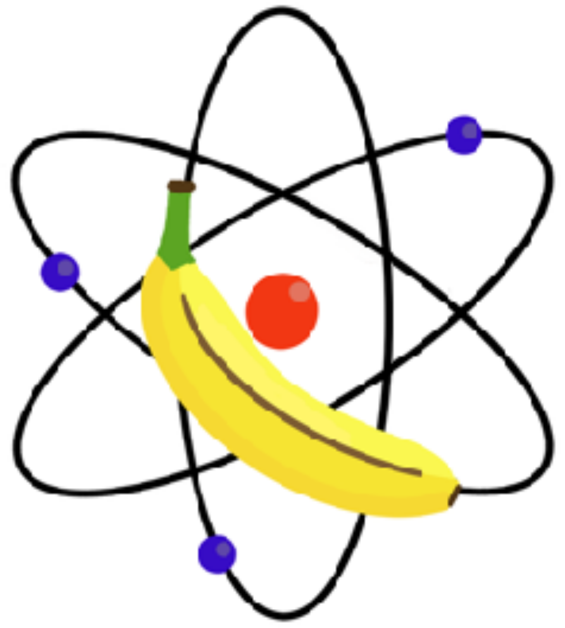
Query matrix: $W_Q \rightarrow Q = W_Q X$
Key matrix: $W_K \rightarrow K = W_K X$
Value matrix: $W_V \rightarrow V = W_V X$

All matrix elements of these matrices are tunable parameters.

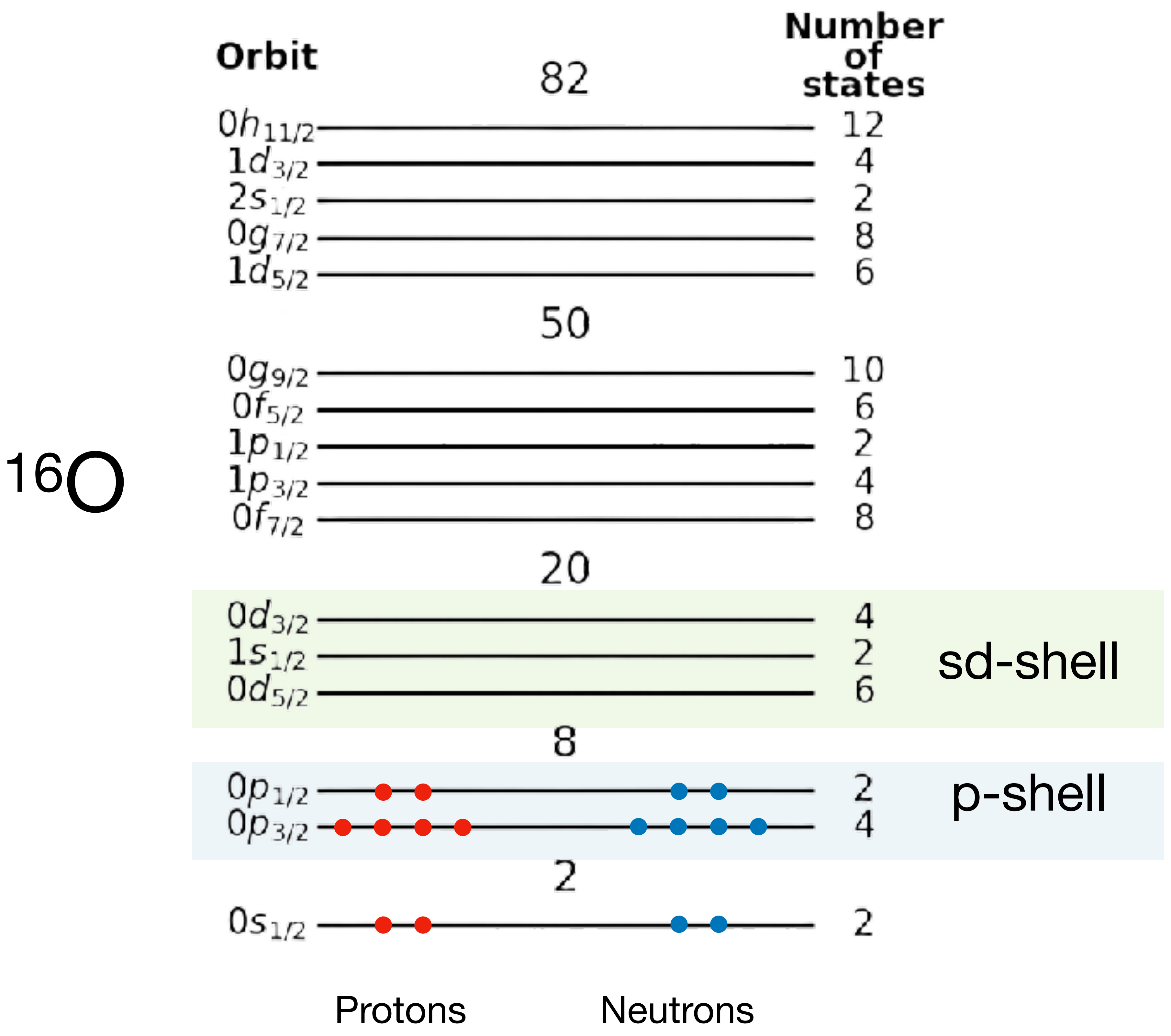
$$\boxed{Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d}}\right)V}$$

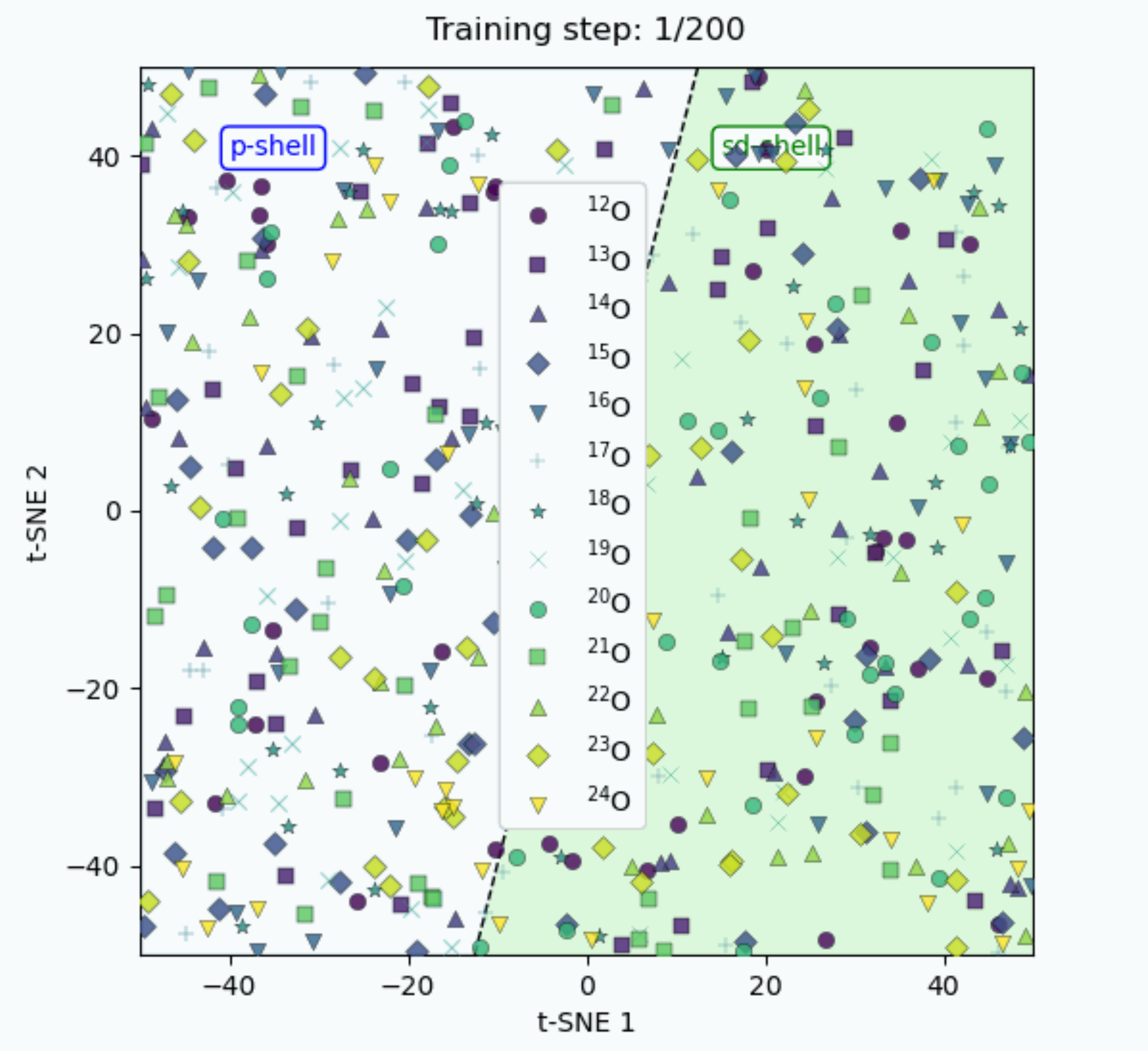
Finally $X \rightarrow X + Attention(Q, K, V)$

Visualizing the Embeddings



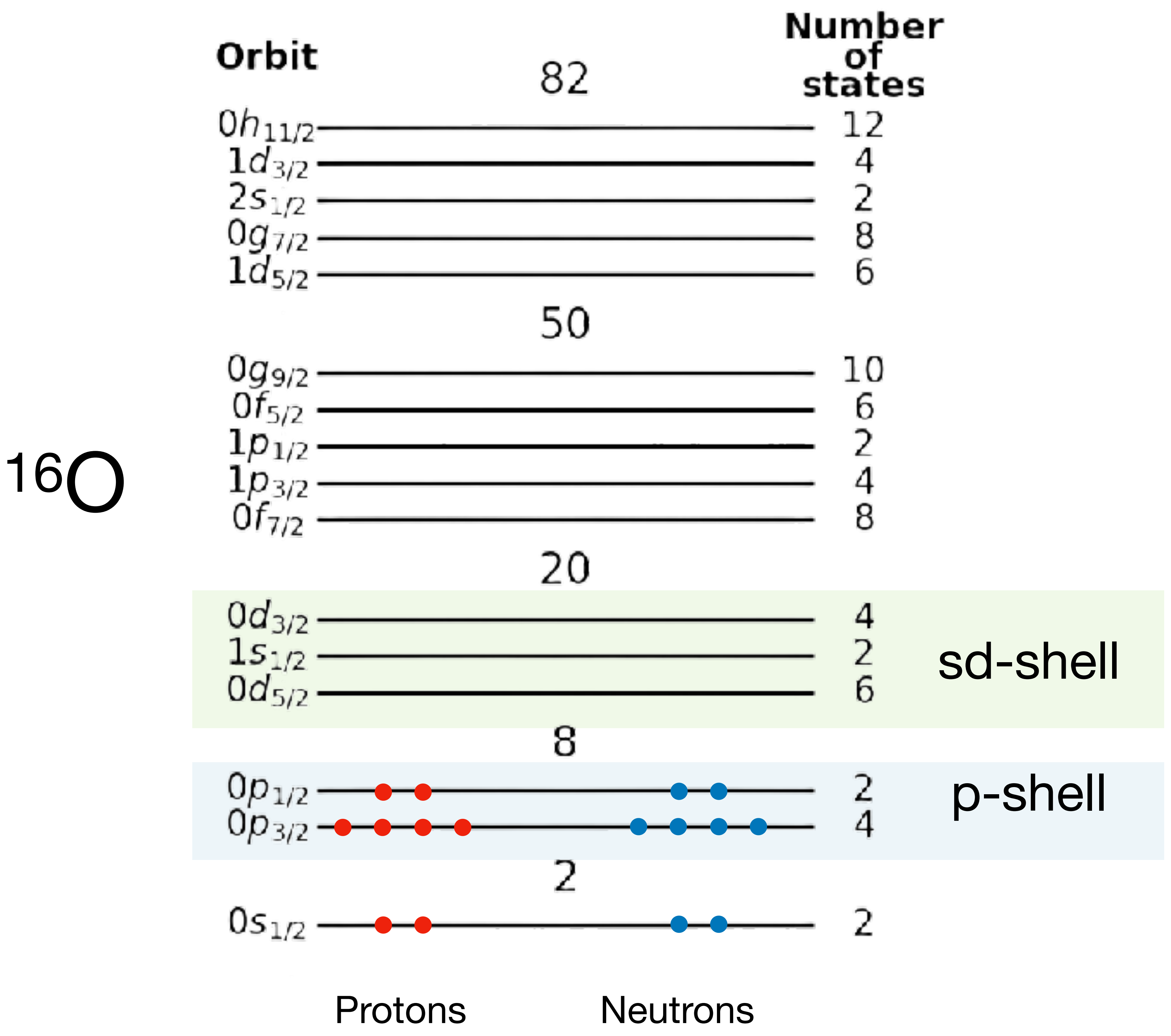
- Projection of embeddings from the attention mechanism.
- Model is learning nuclear shells!

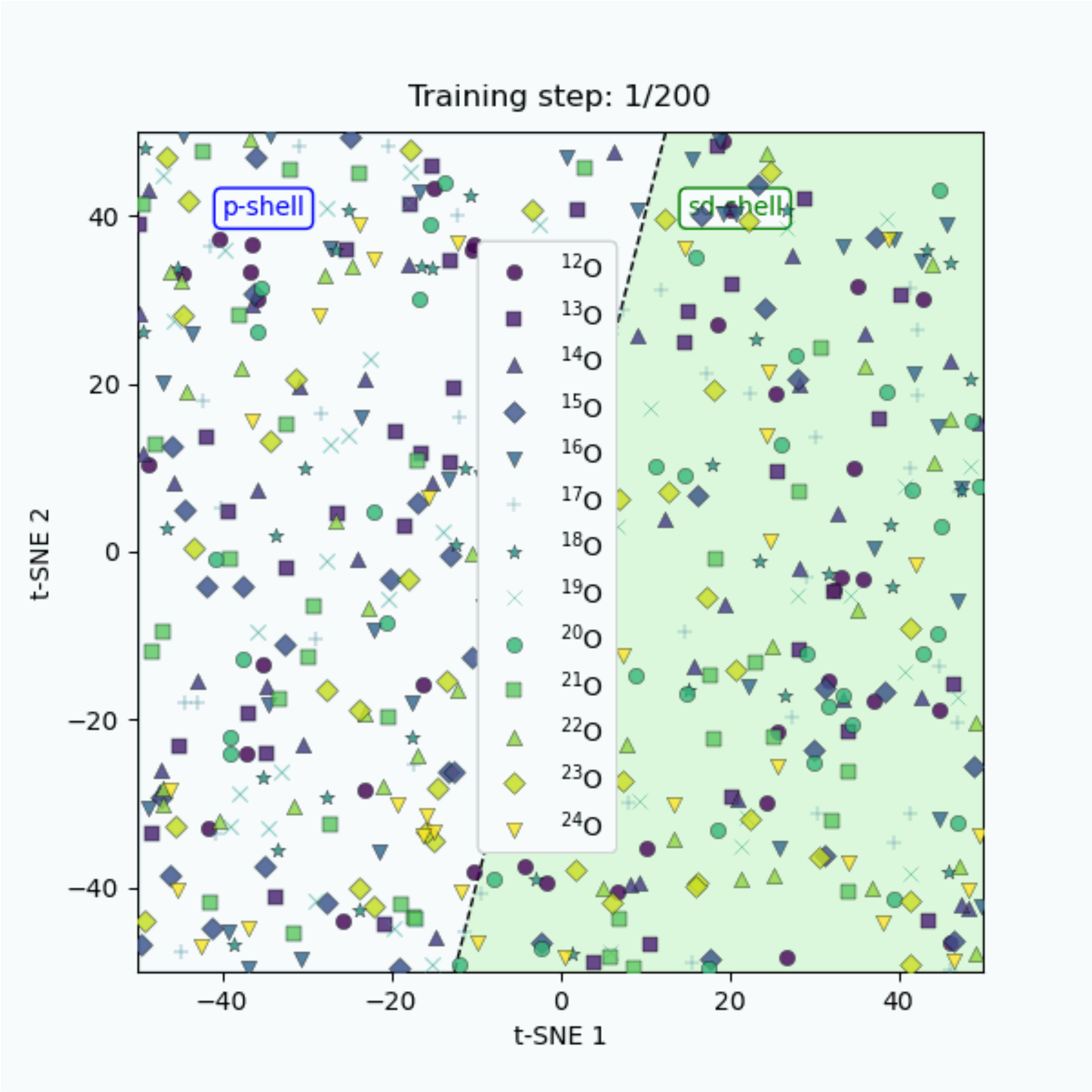




Visualizing the Embeddings

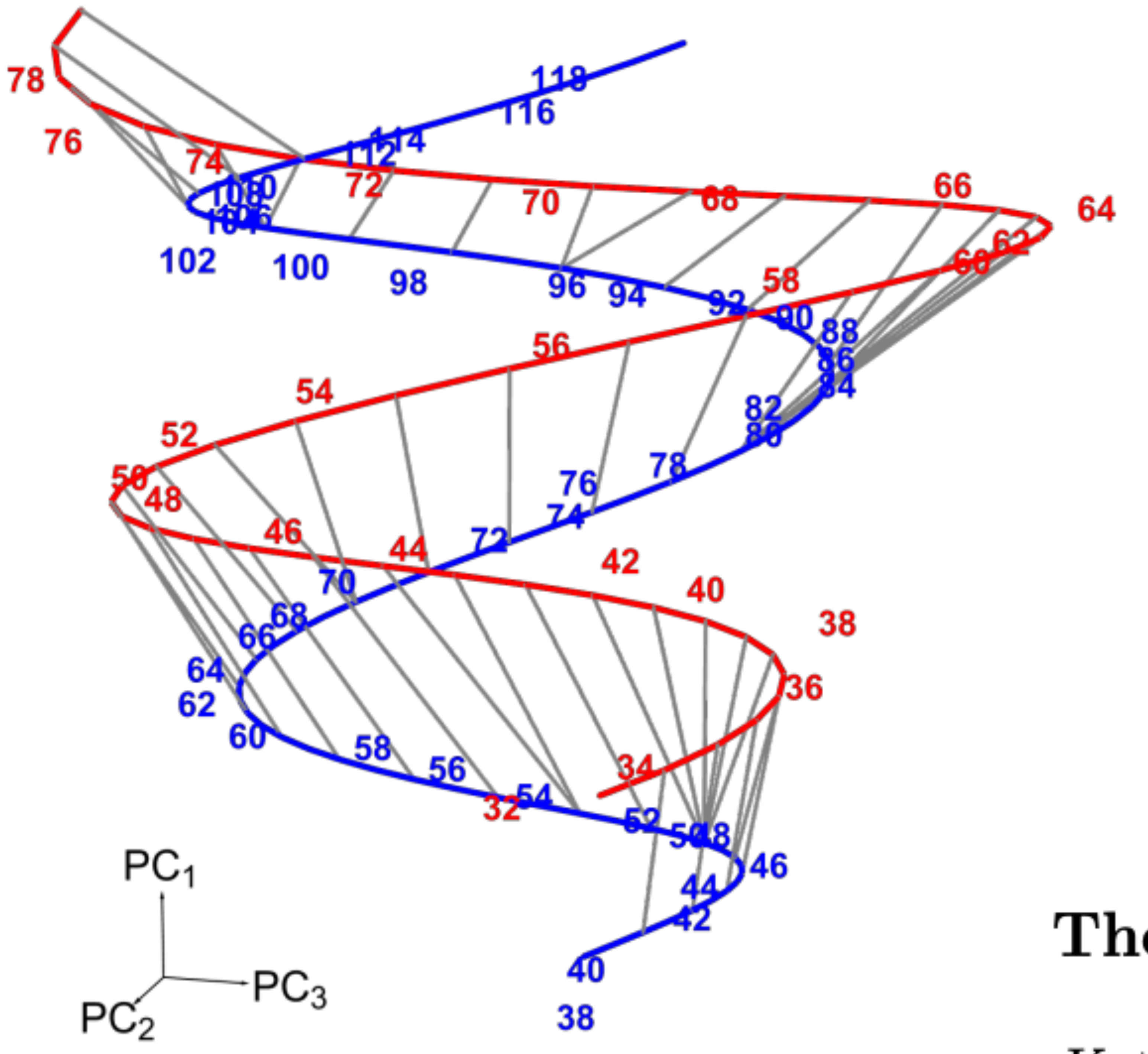
- Projection of embeddings from the attention mechanism.
- Model is learning nuclear shells!

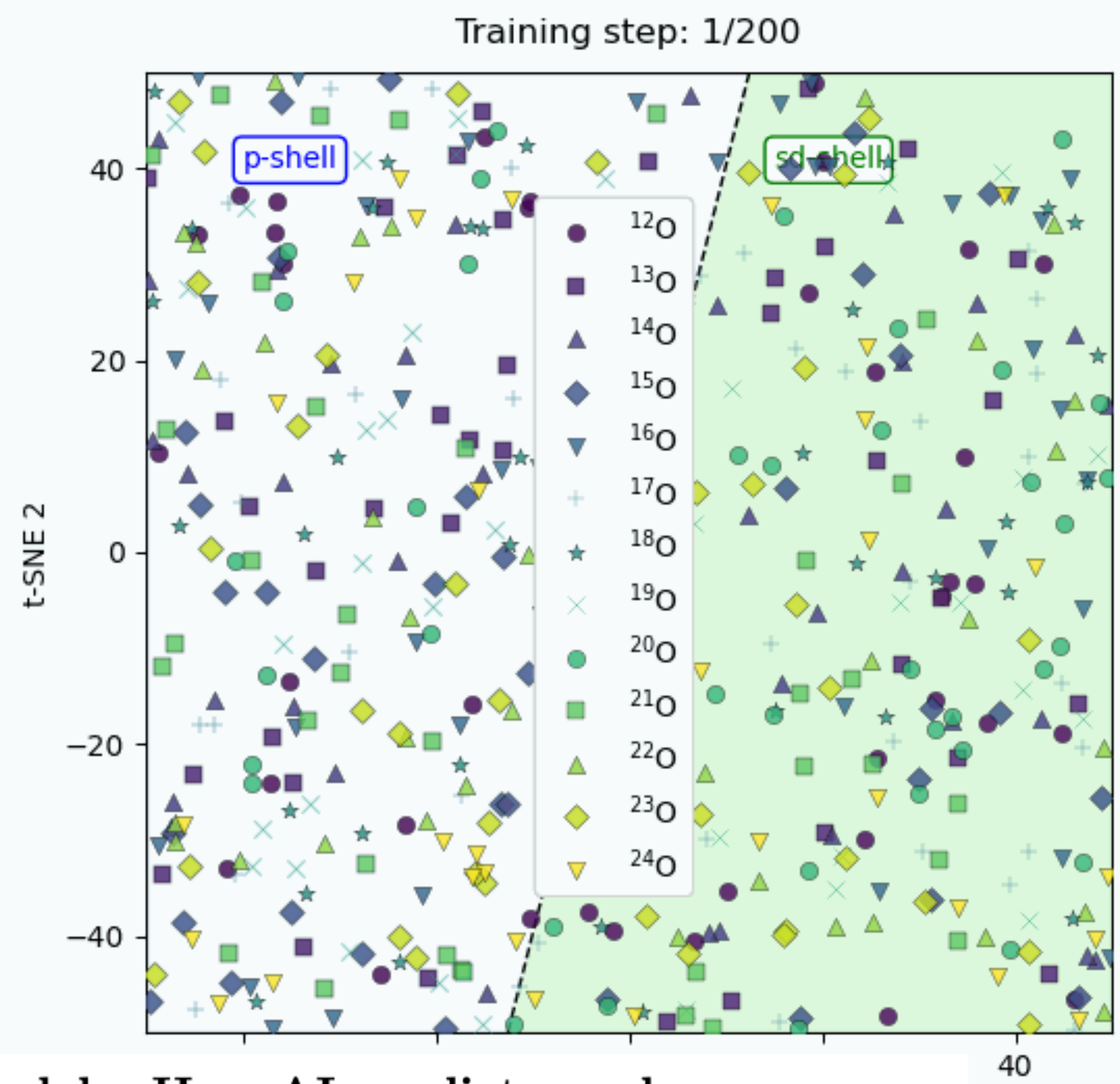




Visualizing the Embeddings

- Projection of embeddings from the attention mechanism.
- Model is learning nuclear shells!

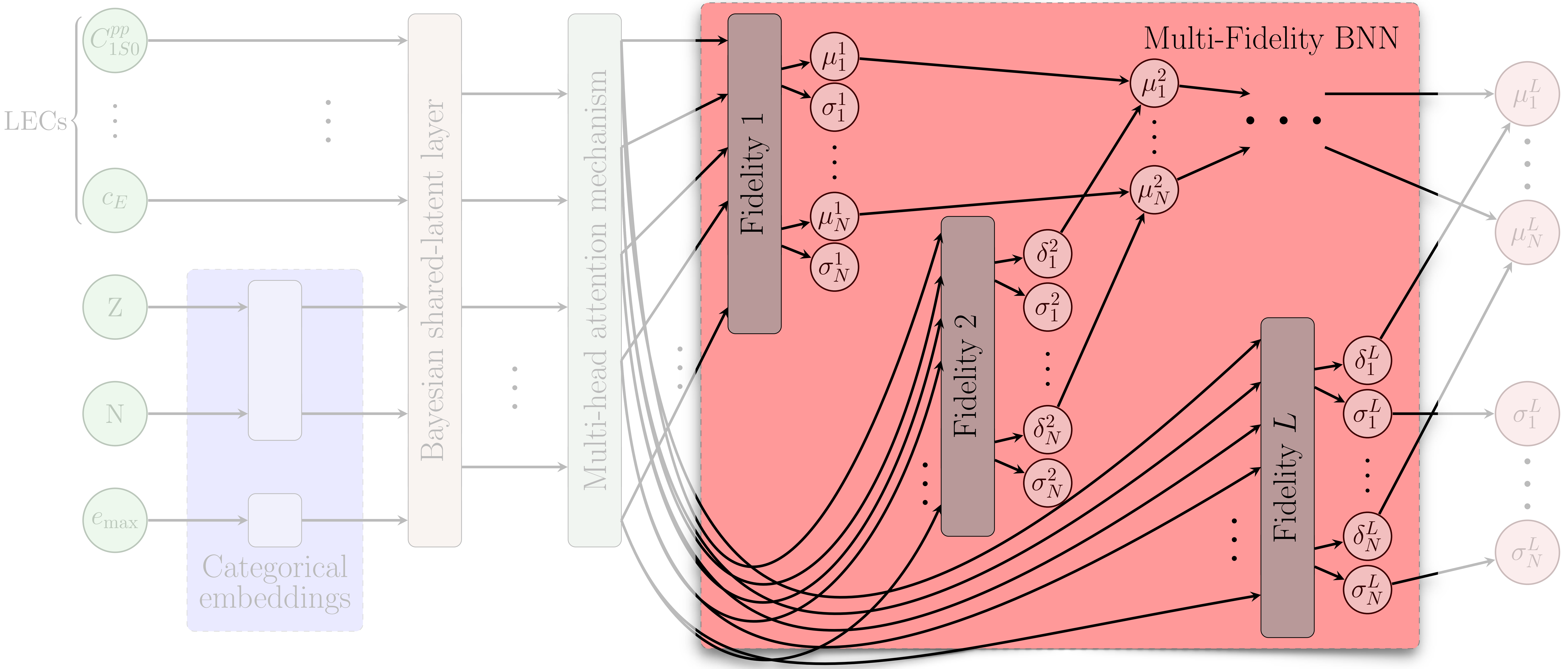




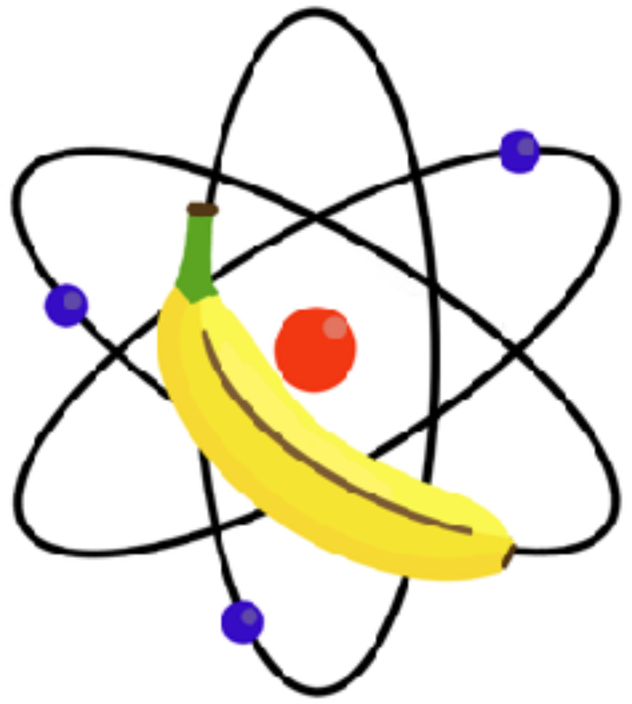
The DNA of nuclear models: How AI predicts nuclear masses

Kate A. Richardson,[1,2,*] Sokratis Trifinopoulos,[1,2,3,4,†] and Mike Williams[1,2,‡]

BAyesian Neural Network for Atomic Nuclei Emulation

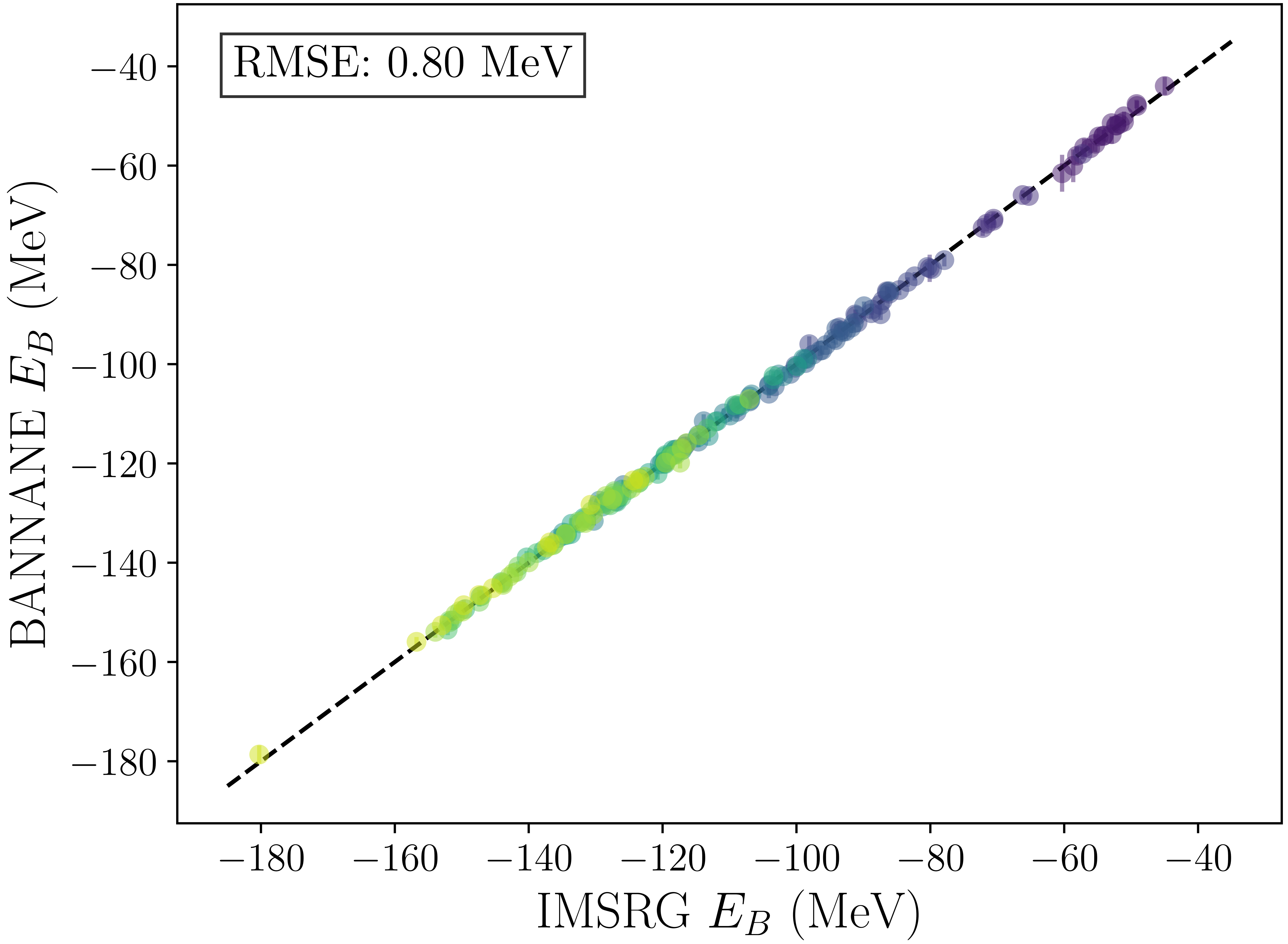


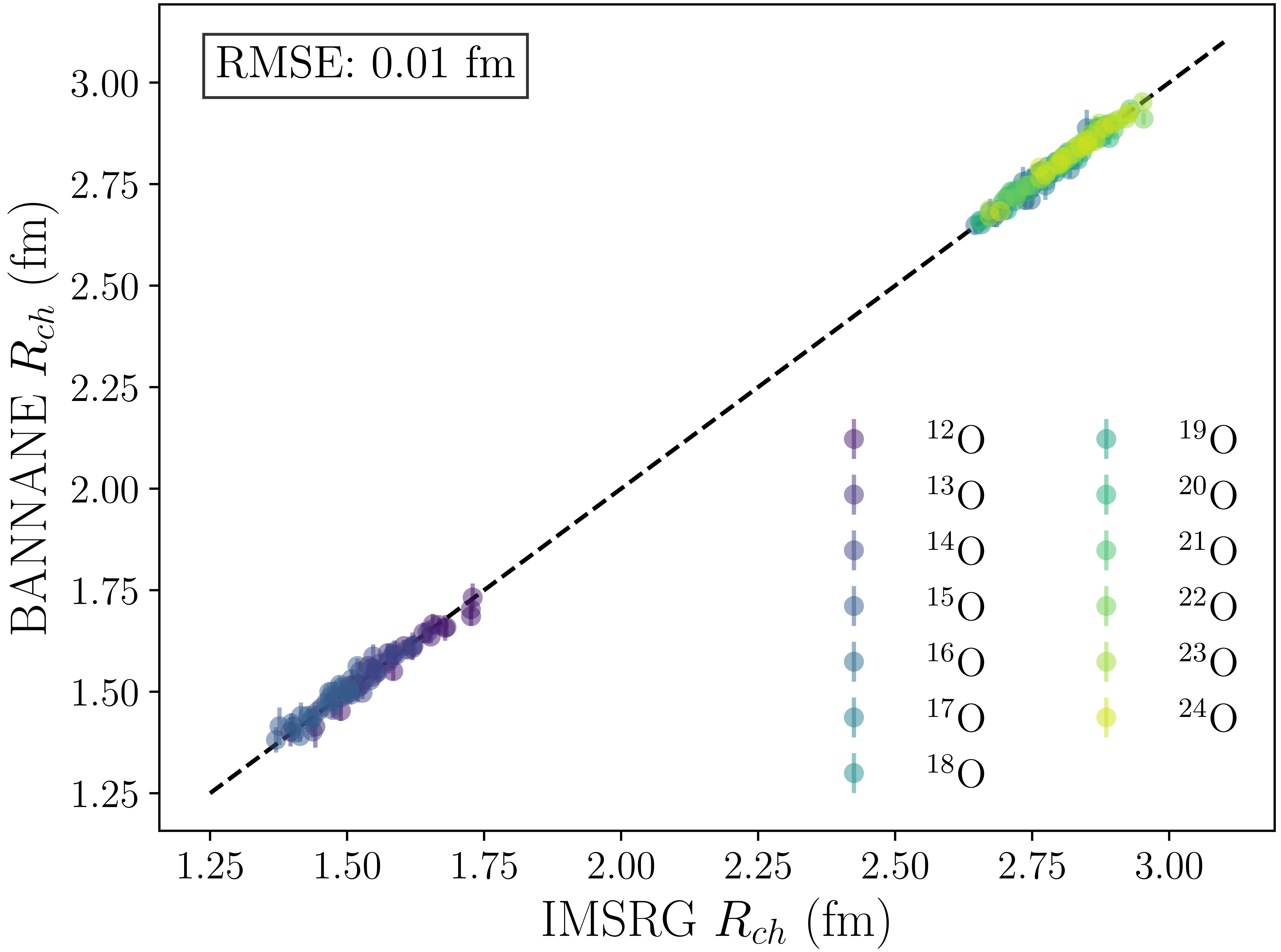
Results



Emulating Multiple Isotopes

Belley, et al.,Phys. Rev. Lett. **136**, 082501 (2026)

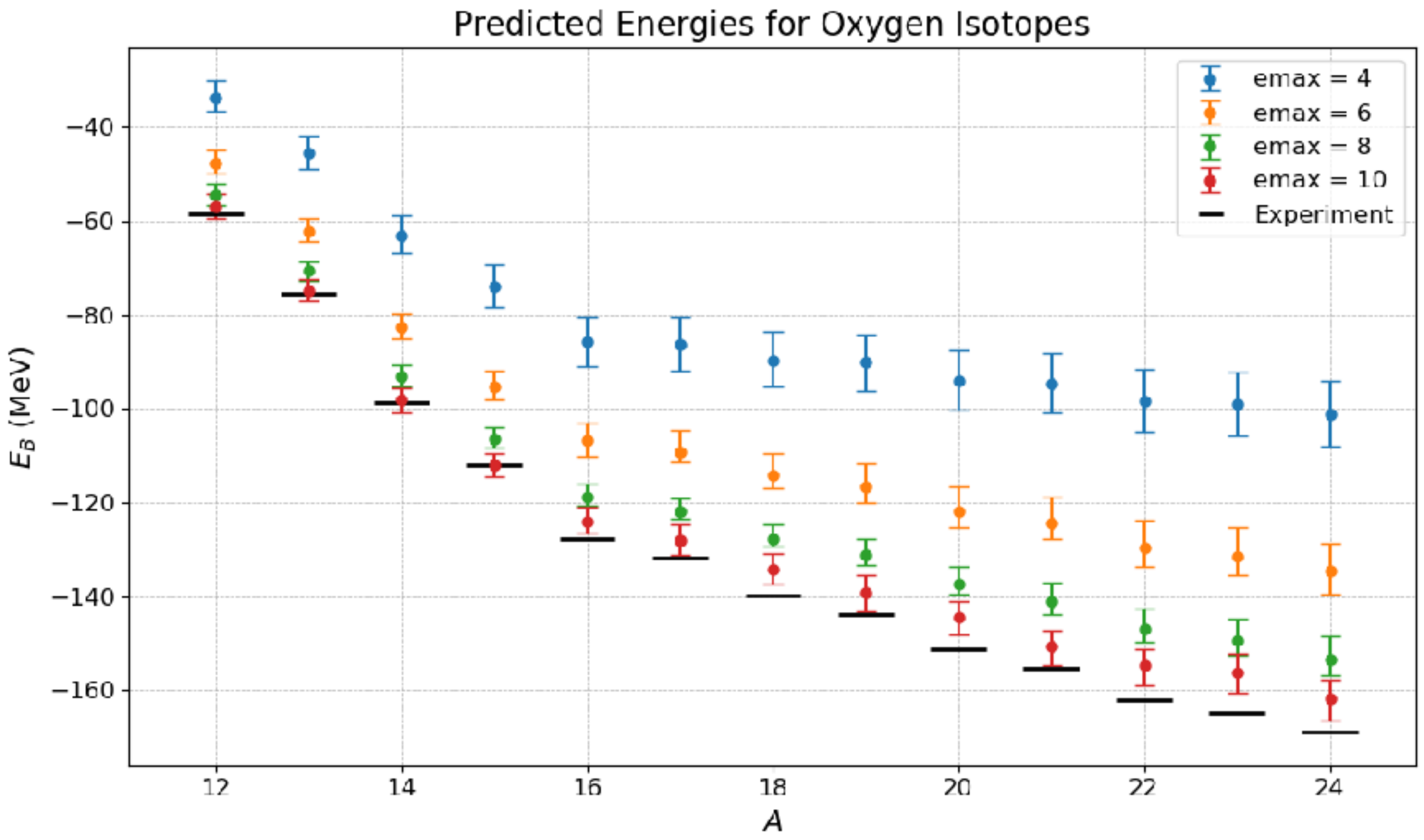


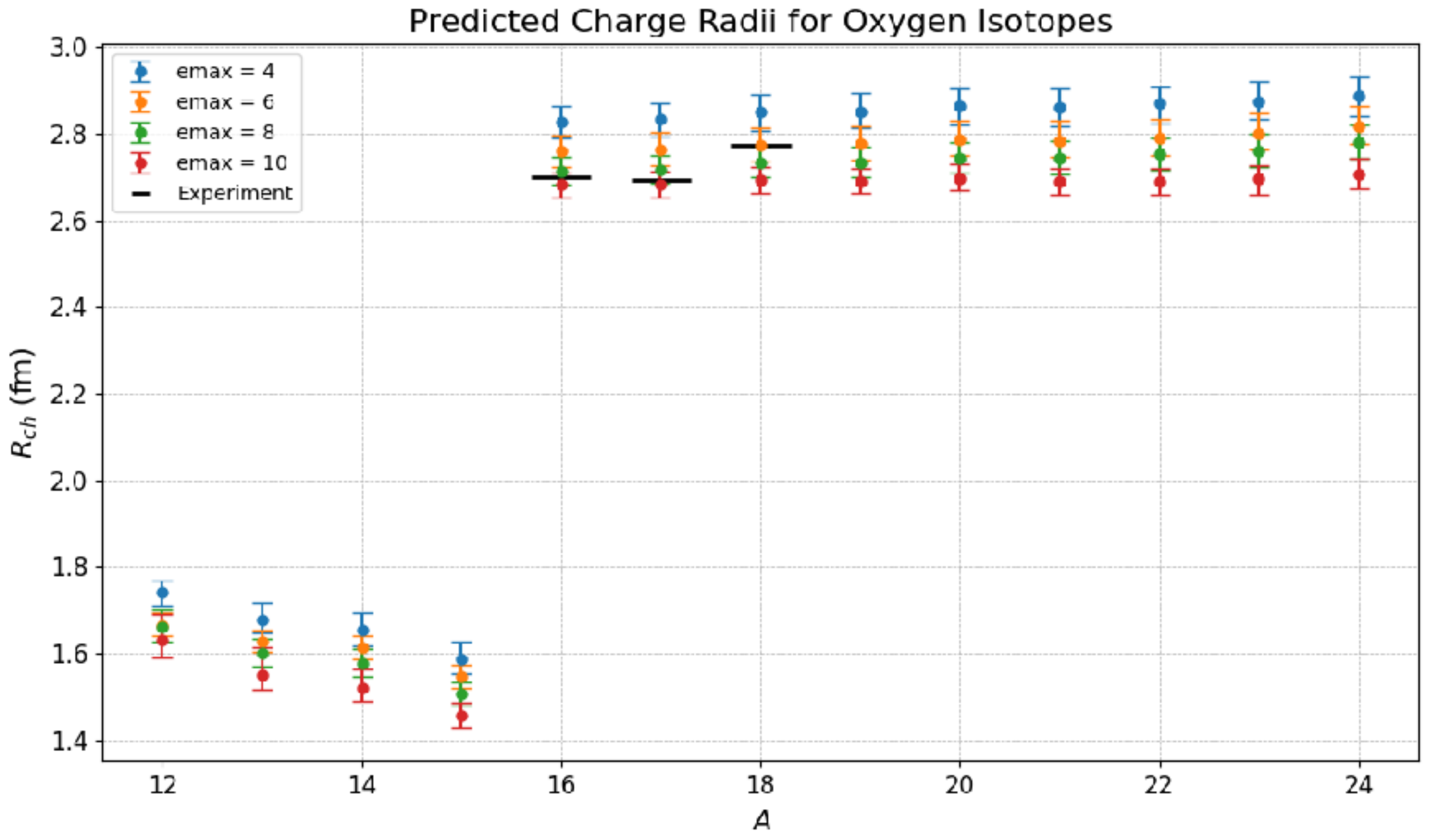


BANNANE achieves state-of-the-art emulation, with smaller errors than other emulators while emulating over a full isotopic chain.

Emulating Multiple Isotopes

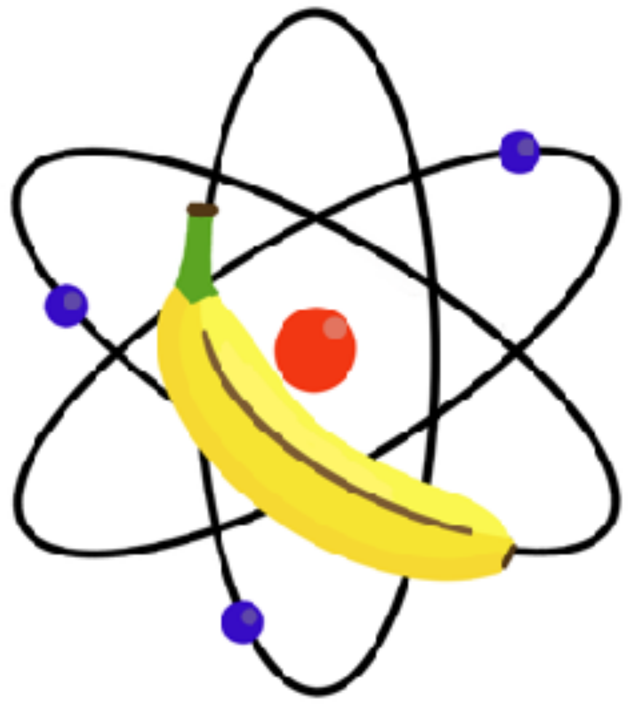
Belley, et al.,Phys. Rev. Lett. **136**, 082501 (2026)

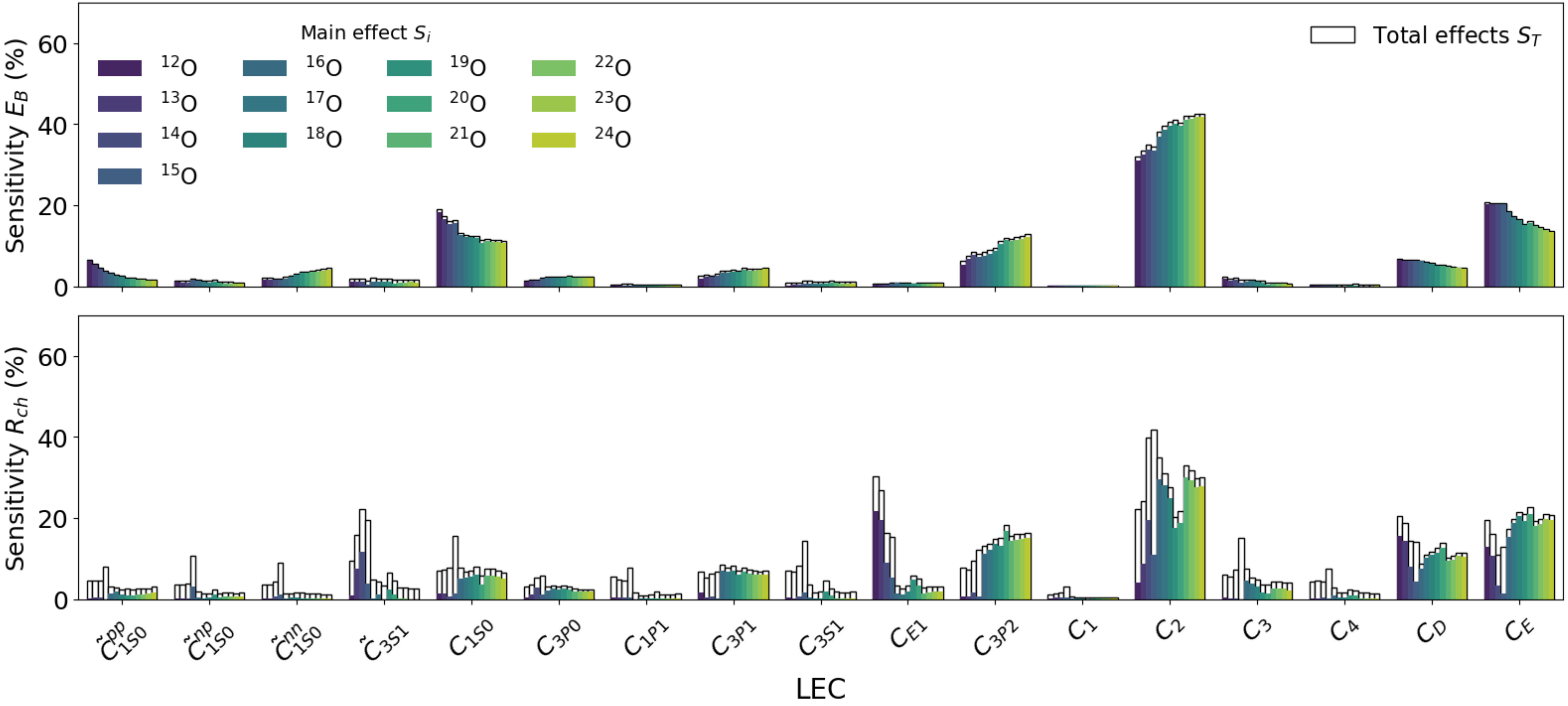




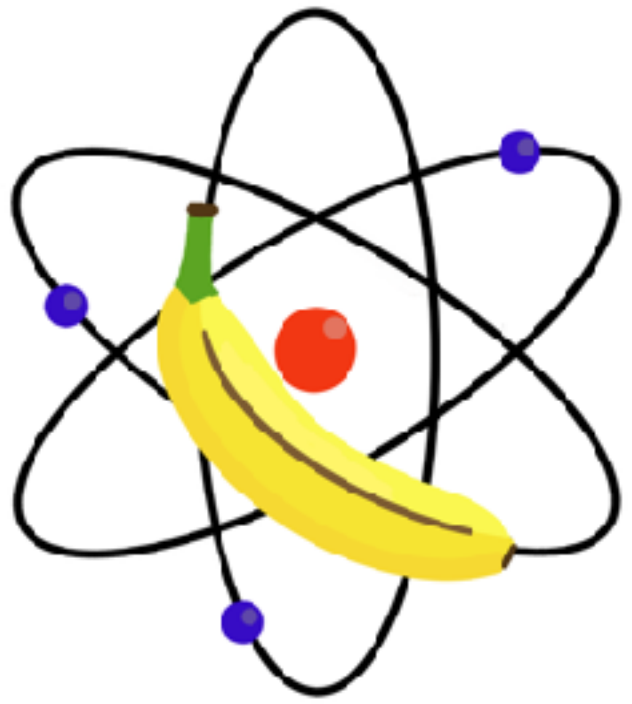
Combining this with UQ technique, we can predict observables with associated uncertainties over the full isotopic chains in a few minutes.

Emulating Multiple Isotopes

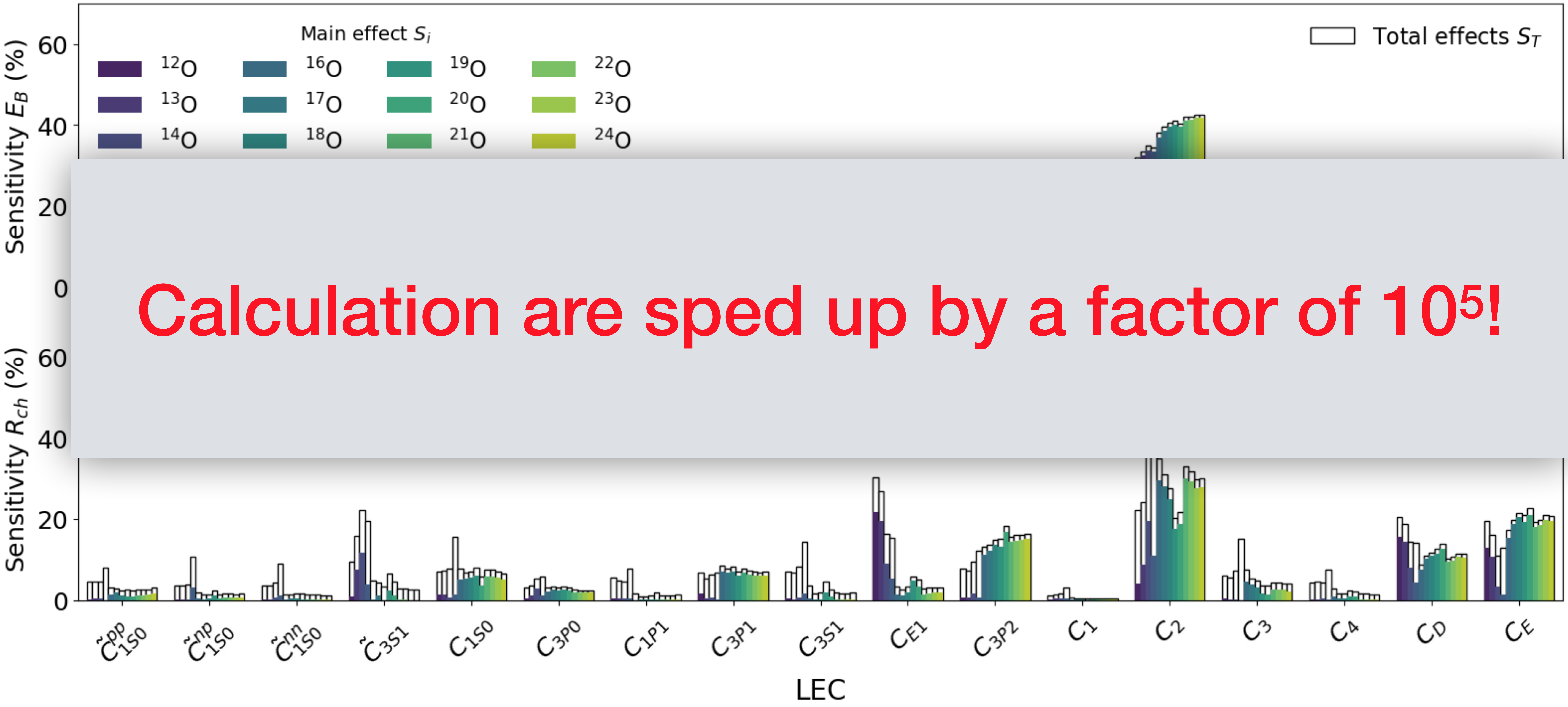




Global sensitivity analysis is consistent with other emulators!



Emulating Multiple Isotopes

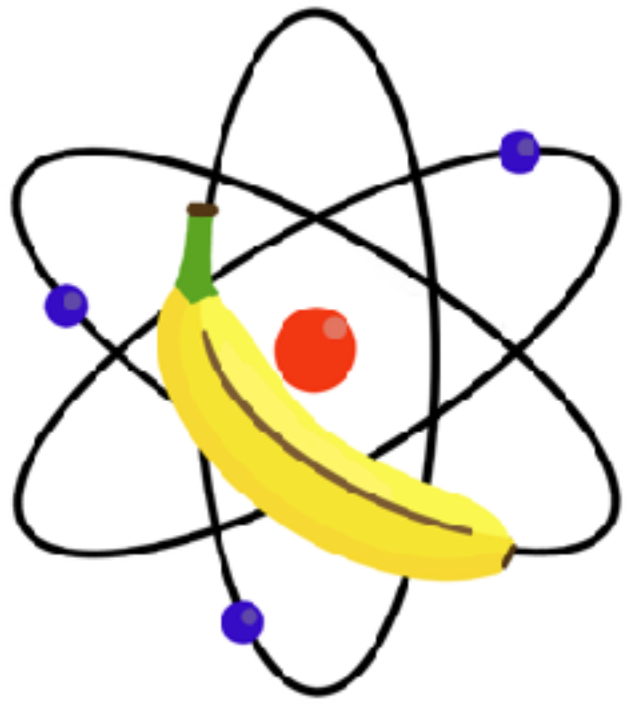


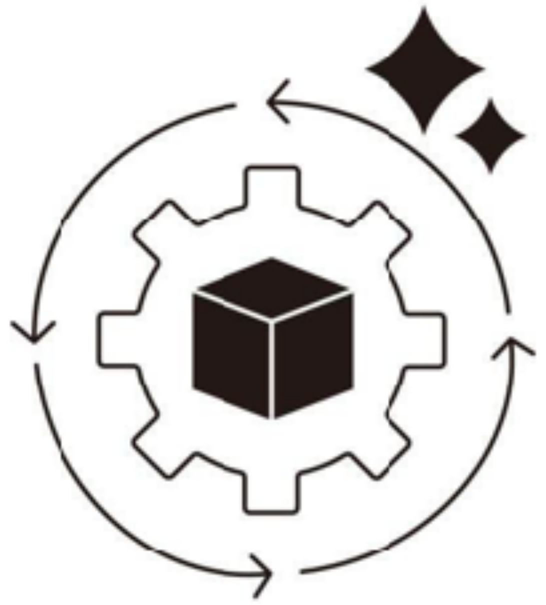
Calculation are sped up by a factor of 10^5!

Global sensitivity analysis is consistent with other emulators!

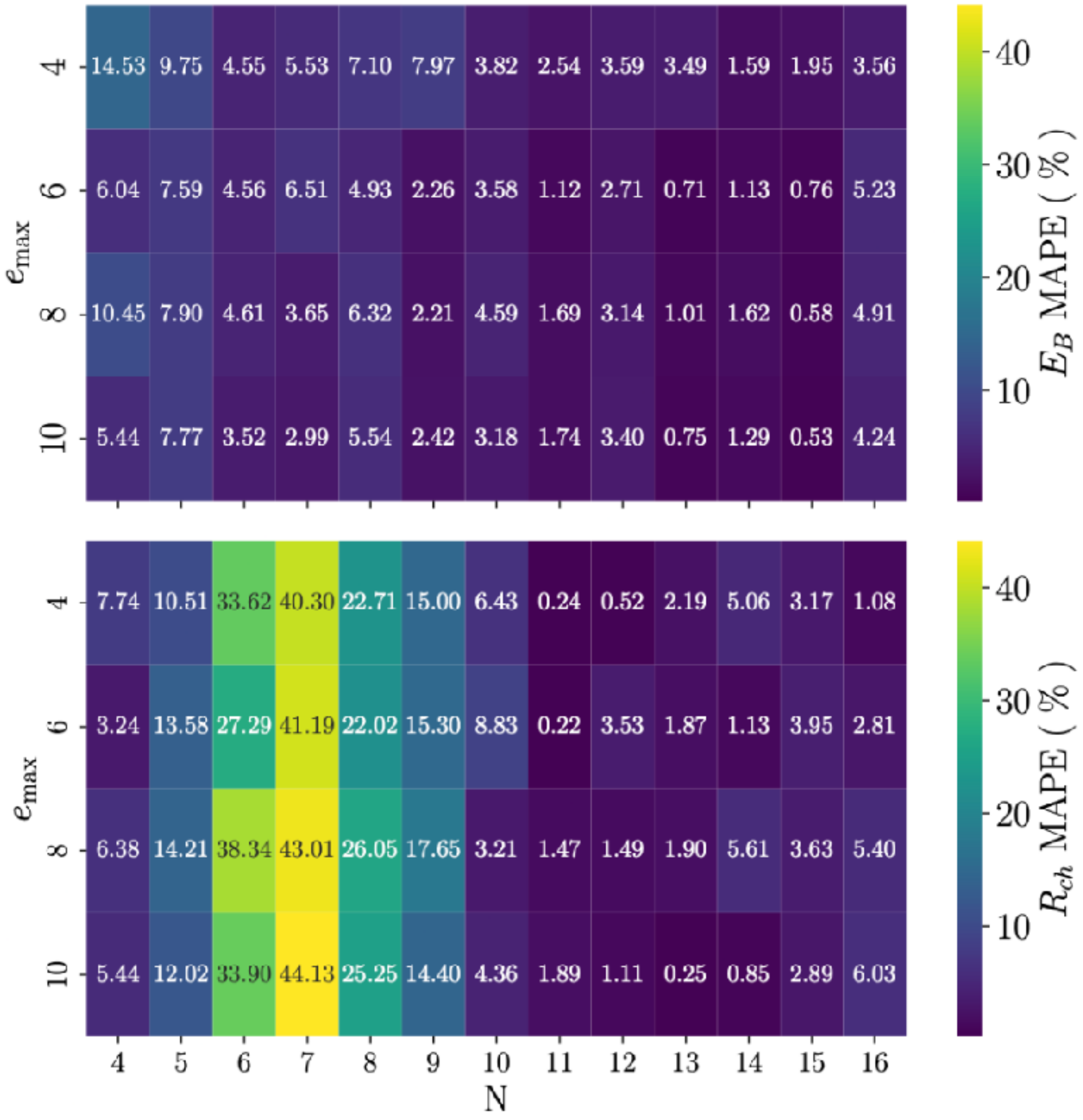
Challenges

Emulating To Unseen Data





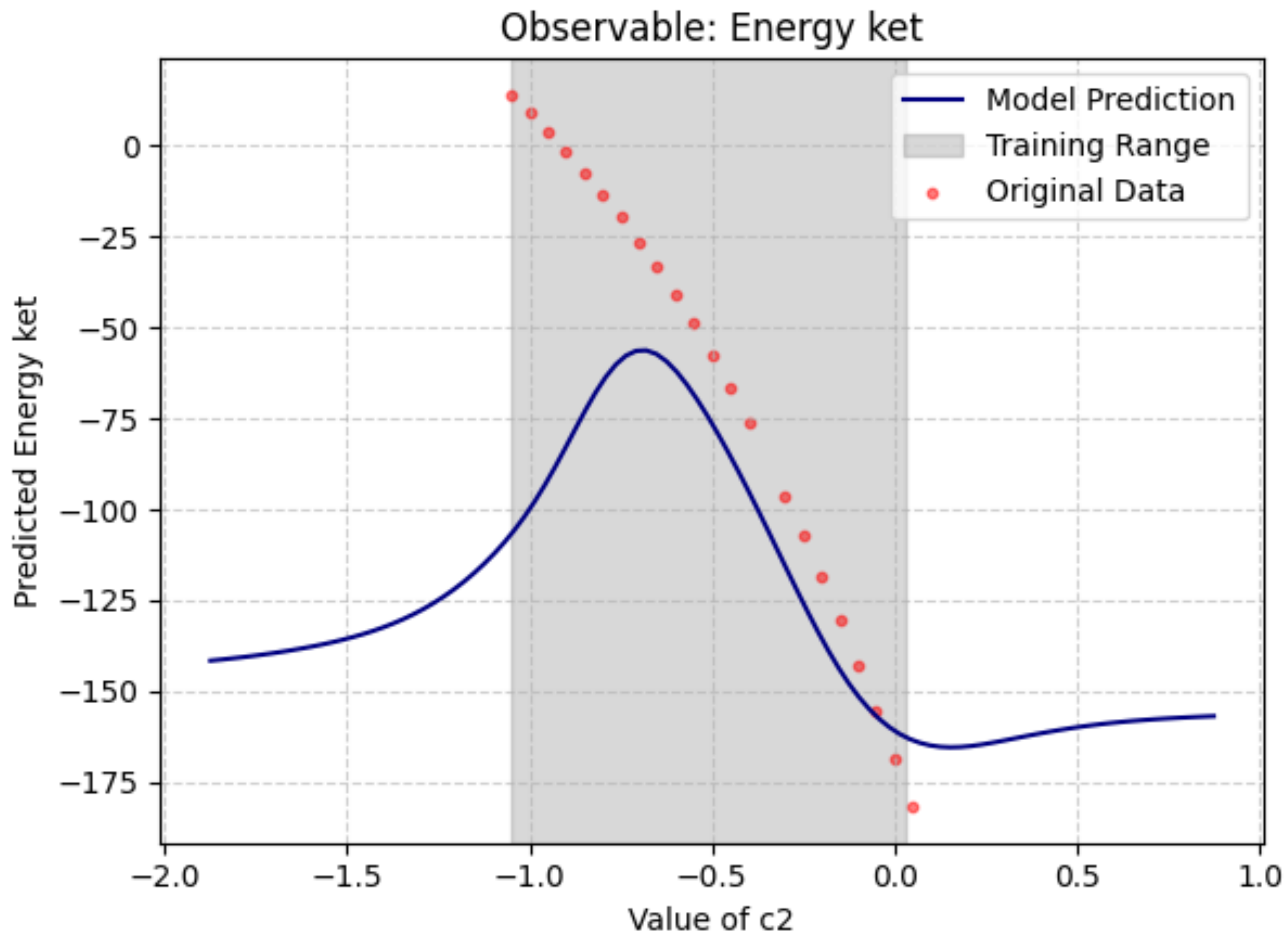
Zero-shot Learning

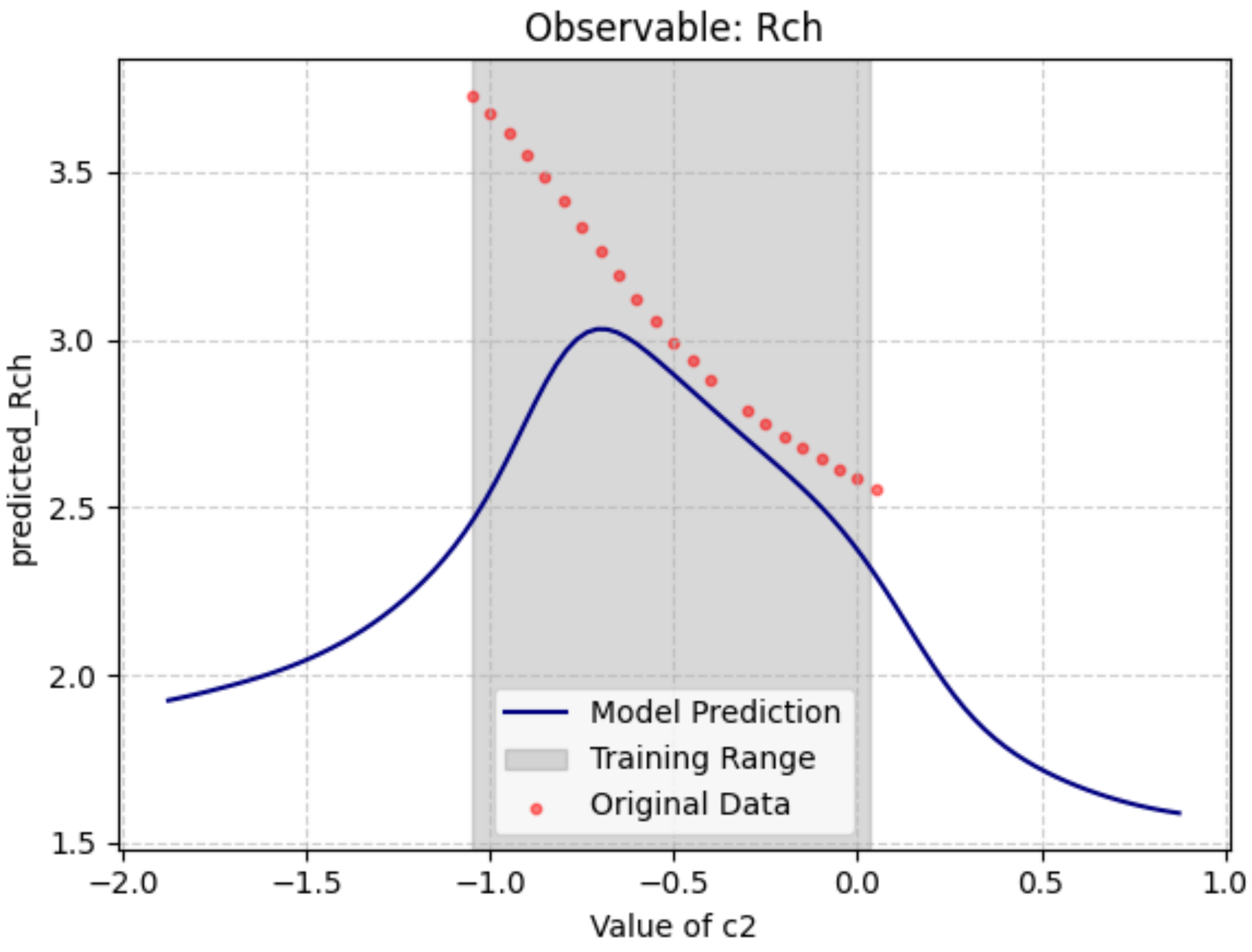


Difference with model and IMSRG when no data is used to train in a specific isotope.

Emulating Outside the Physical Range

Model Extrapolation vs. Low-Energy Constants (LECs)





Emulators for Many-Body Methods

There are two? ways to build an emulator for nuclear physics:

Physics driven **Data driven**

Emulators for Many-Body Methods

There are two? ways to build an emulator for nuclear physics:

Physics driven **Data driven**

Reduce Basis Methods

Emulators for Many-Body Methods

There are two? ways to build an emulator for nuclear physics:

Physics driven **Data driven**

Reduce Basis Methods NN, GPs

Emulators for Many-Body Methods

There are two? ways to build an emulator for nuclear physics:

Physics driven **Data driven**

Reduce Basis Methods PMMs NN, GPs

Emulators for Many-Body Methods

There are two? ways to build an emulator for nuclear physics:

Physics driven **Data driven**

Reduce Basis Methods PMMs PINNs NN, GPs

Emulators for Many-Body Methods

There are two? ways to build an emulator for nuclear physics:

Physics driven **Data driven**

Reduce Basis Methods

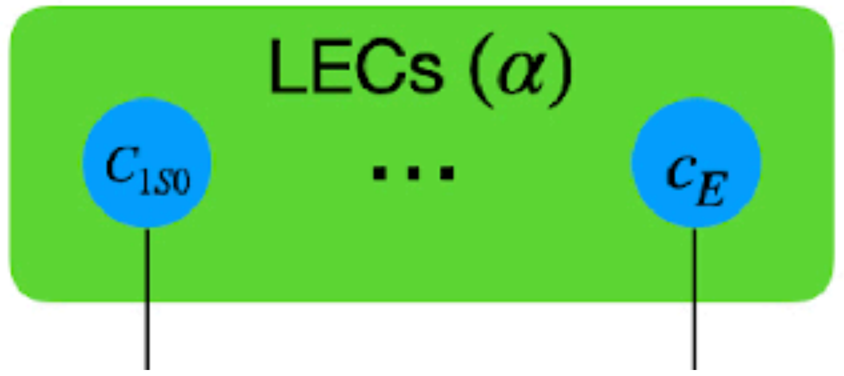
PMMs

PINNs

NN, GPs

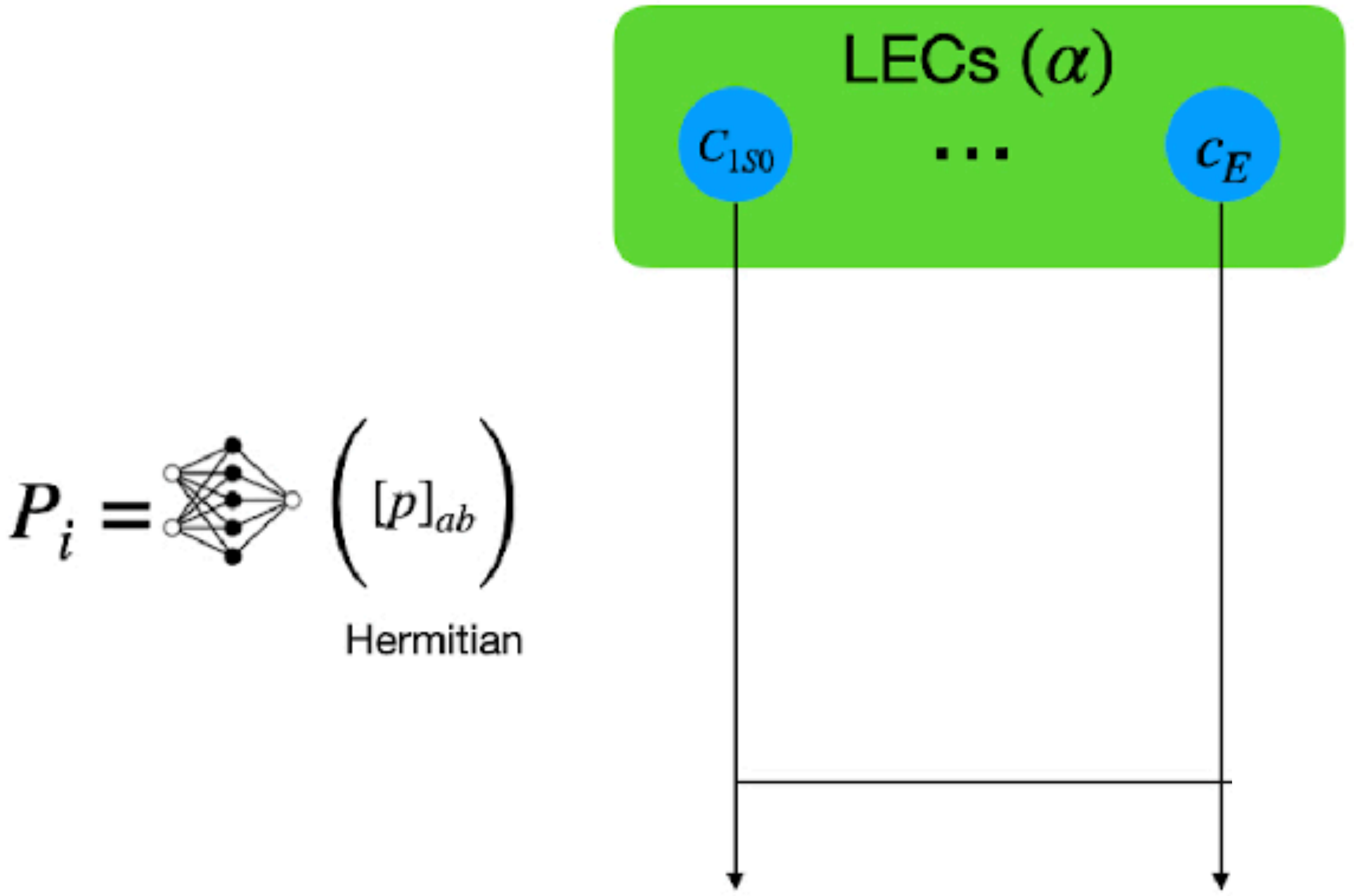
Physics Driven BANNANE?

Jose Miguel Muñoz Arias



Physics Driven BANNANE?

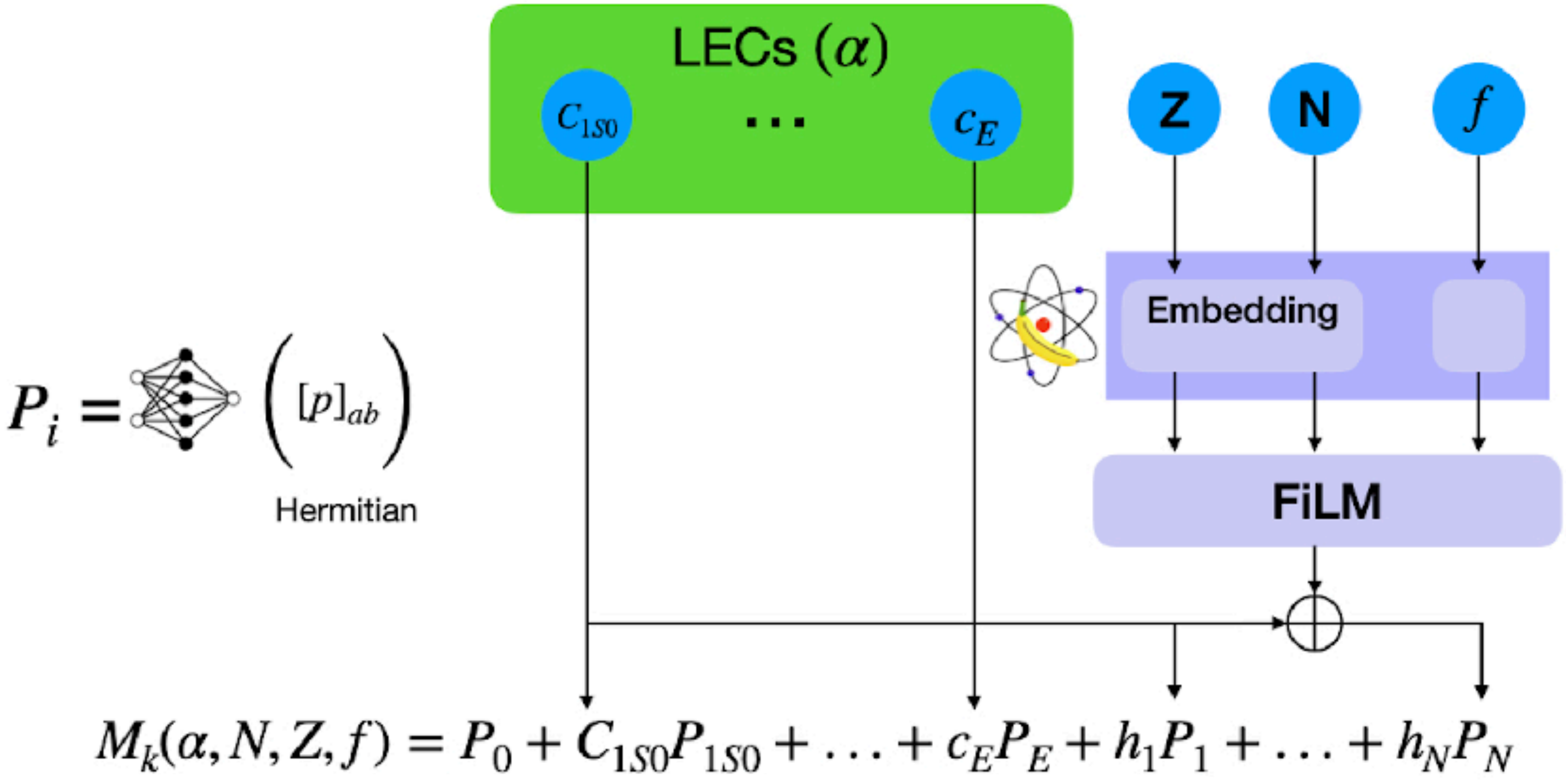
Jose Miguel Muñoz Arias



$$M_k(\alpha, N, Z, f) = P_0 + C_{1S0}P_{1S0} + \ldots + c_E P_E$$

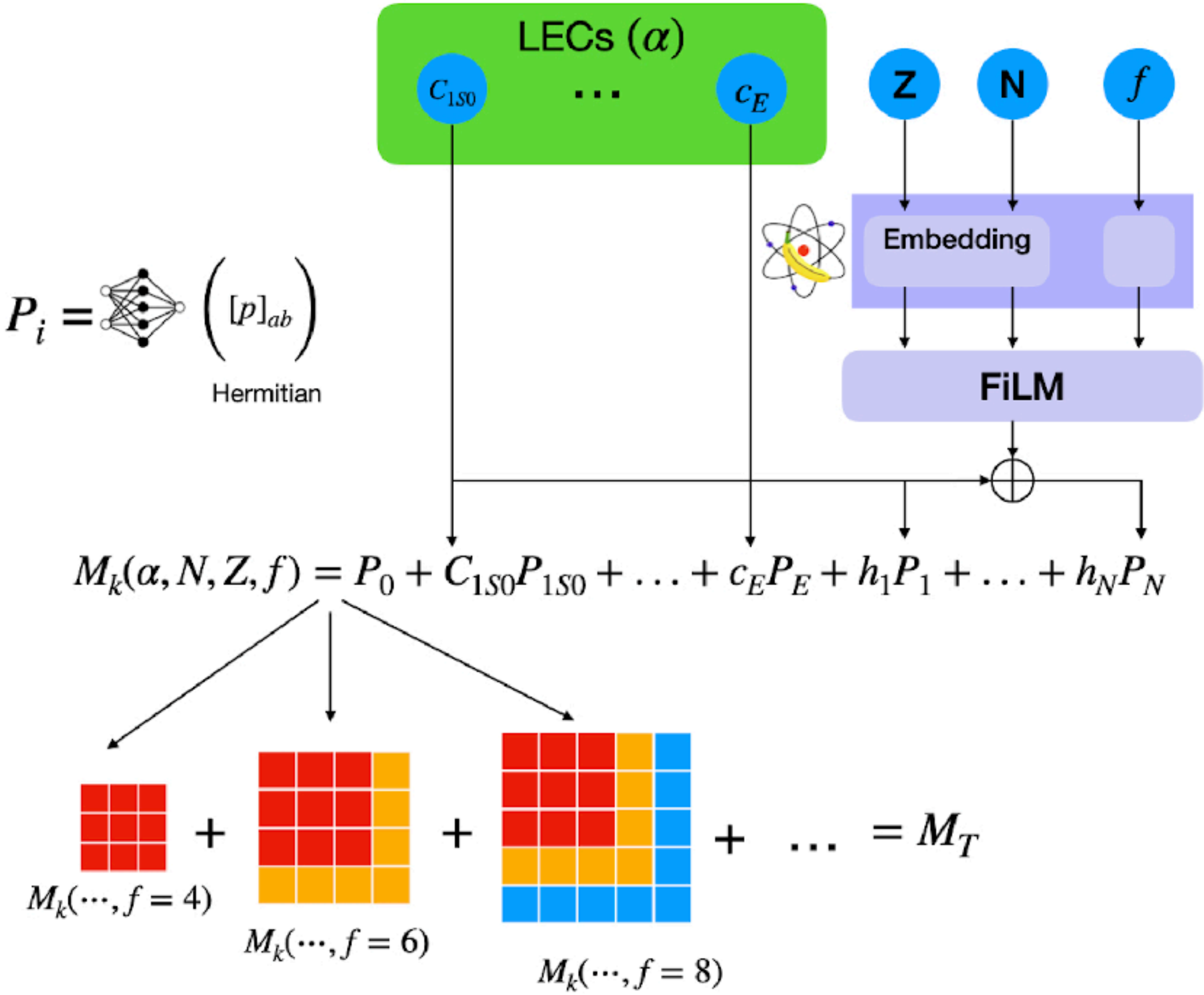
Physics Driven BANNANE?

Jose Miguel Muñoz Arias



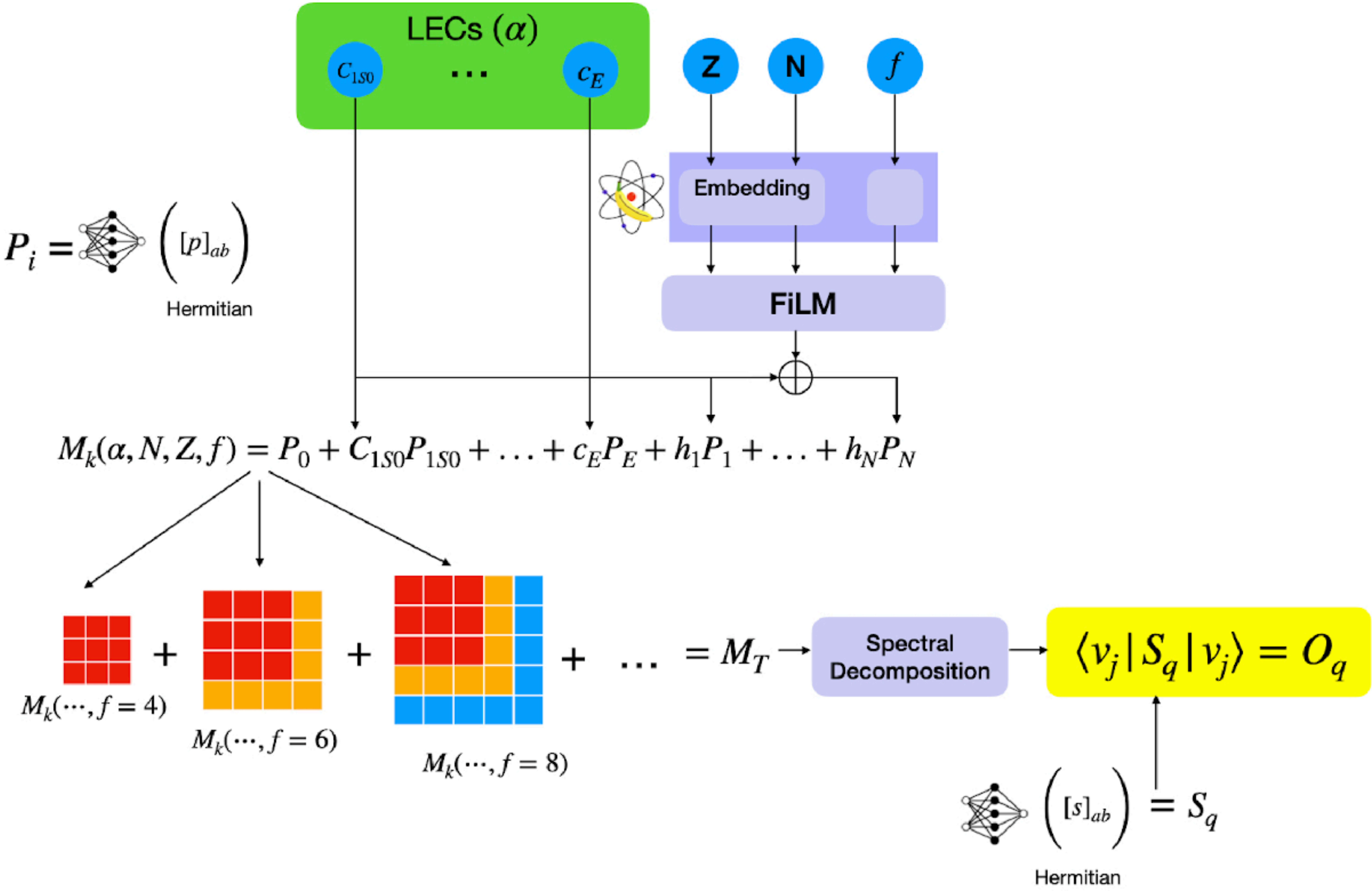
Physics Driven BANNANE?

Jose Miguel Muñoz Arias

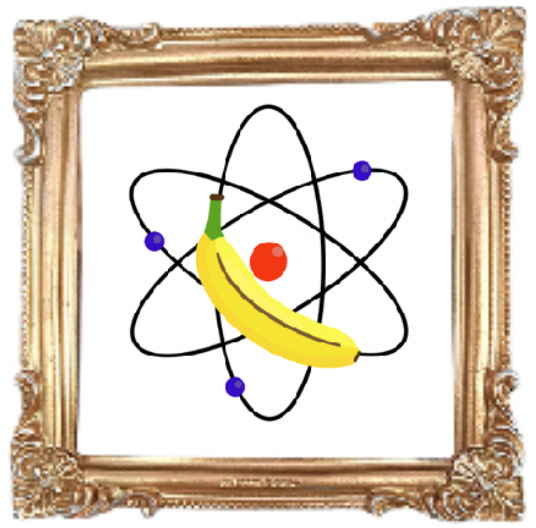


Physics Driven BANNANE?

Jose Miguel Muñoz Arias

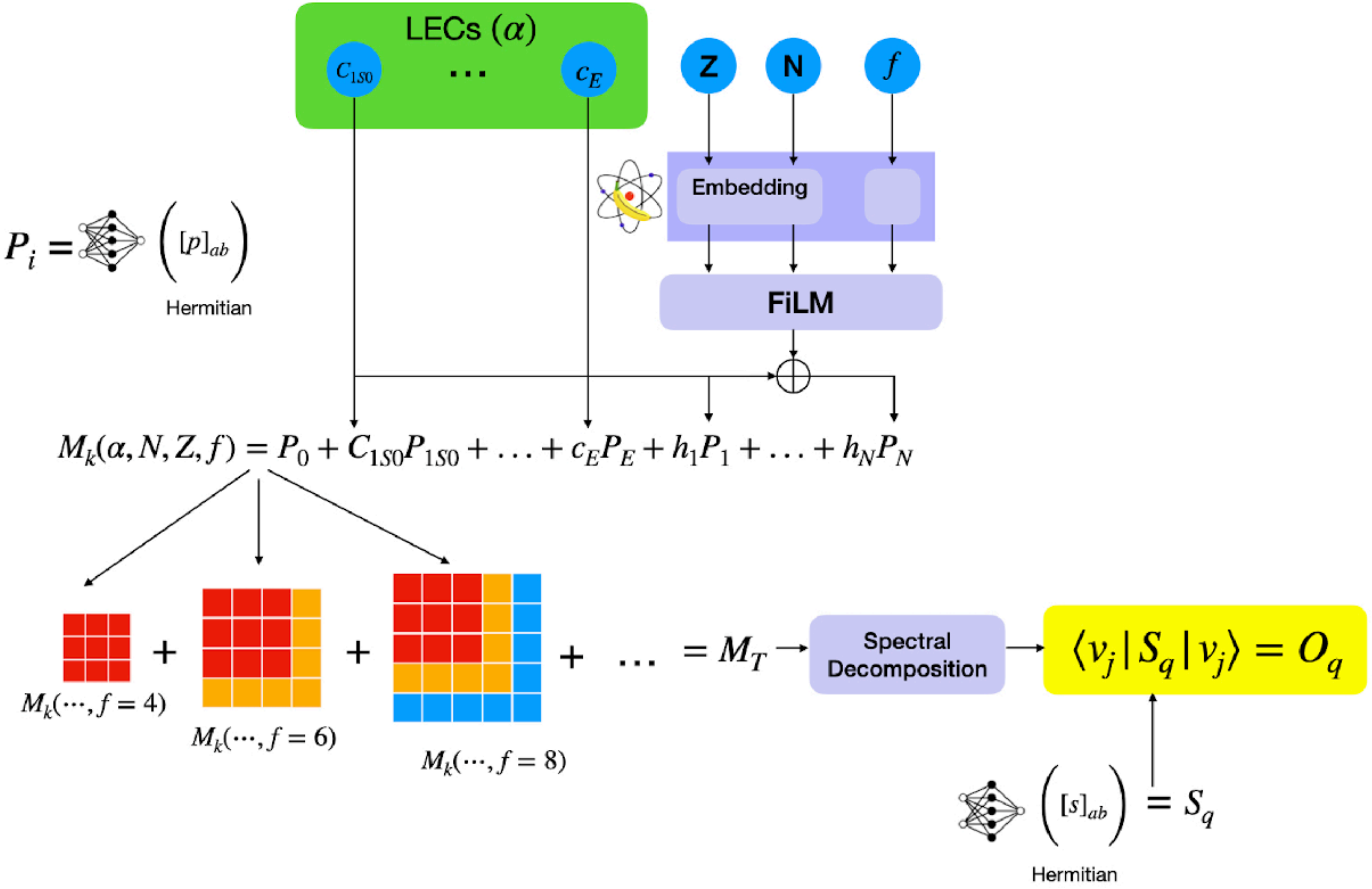


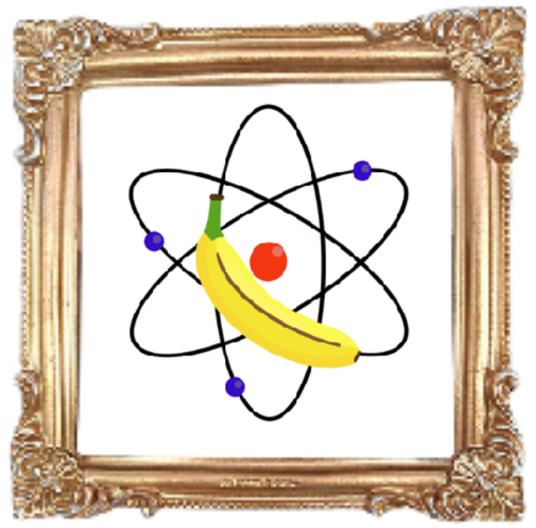
Jose Miguel Muñoz Arias



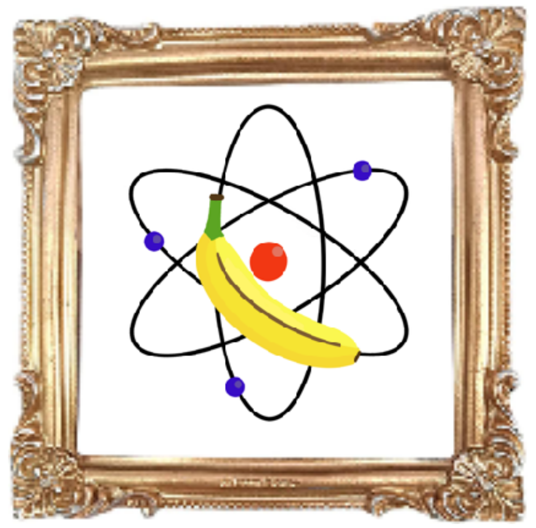
Physics Driven BANNANE?

Feature Resolved Affine Matrix Emulator (FRAME)



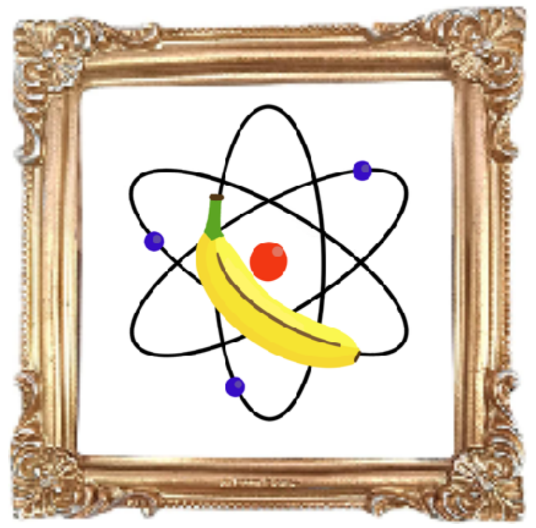


Anchor and Refinement



Anchor and Refinement

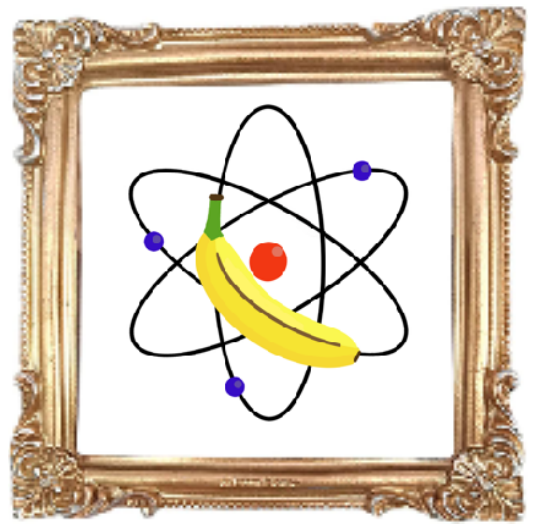
$$M(N,Z,f,\alpha) = [M_{f_{min}} + \sum_{p=1}^{f} \xi(f;N,Z,\alpha)\Delta M_p]$$



Anchor and Refinement

$$M(N, Z, f, \alpha) = [M_{f_{min}} + \sum_{p=1}^{f} \xi(f; N, Z, \alpha)\Delta M_p]$$

Anchor: Matrix at the lowest fidelity



Anchor and Refinement

$$M(N, Z, f, \alpha) = [M_{f_{min}} + \sum_{p=1}^{f} \xi(f; N, Z, \alpha)\Delta M_p]$$

Anchor: Matrix at the lowest fidelity

Refinement: Corrections from higher fidelities

Anchor and Refinement

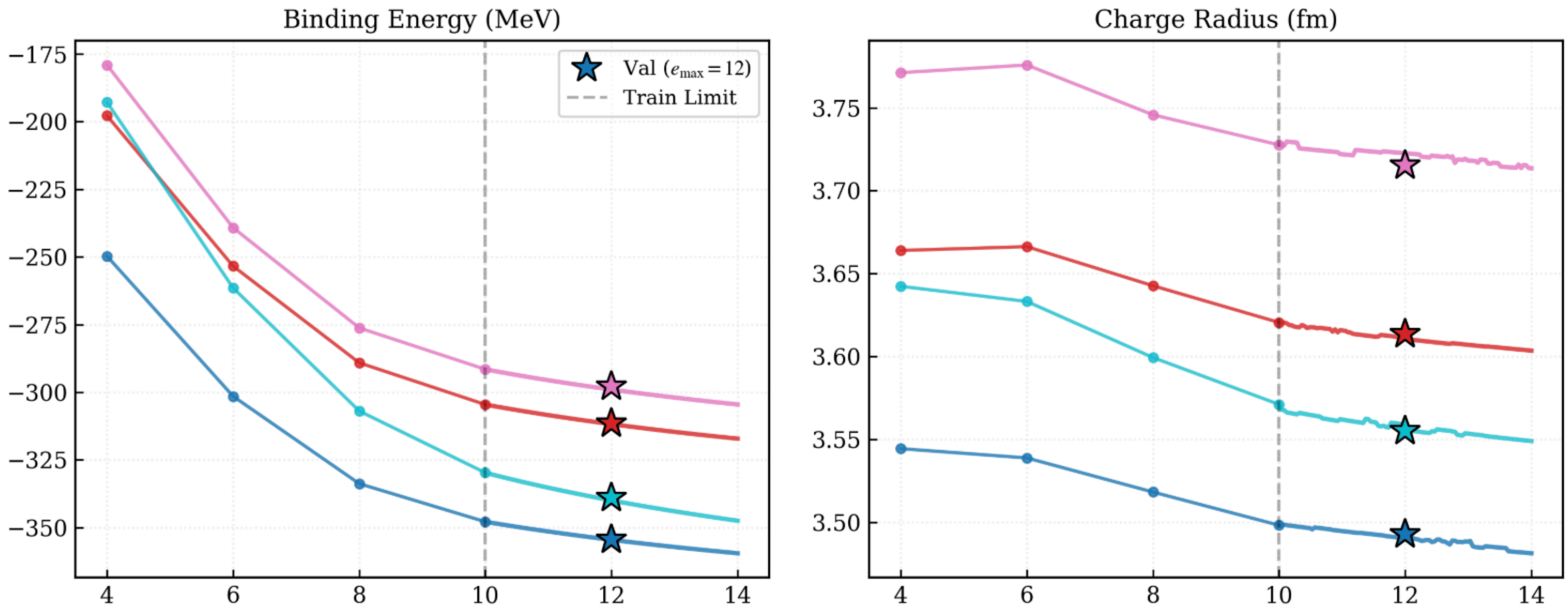
$$M(N, Z, f, \alpha) = [M_{f_{min}} + \sum_{p=1}^{f} \xi(f; N, Z, \alpha)\Delta M_p]$$

Anchor: Matrix at the lowest fidelity

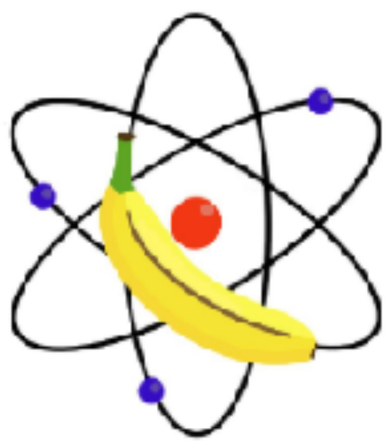
Refinement: Corrections from higher fidelities

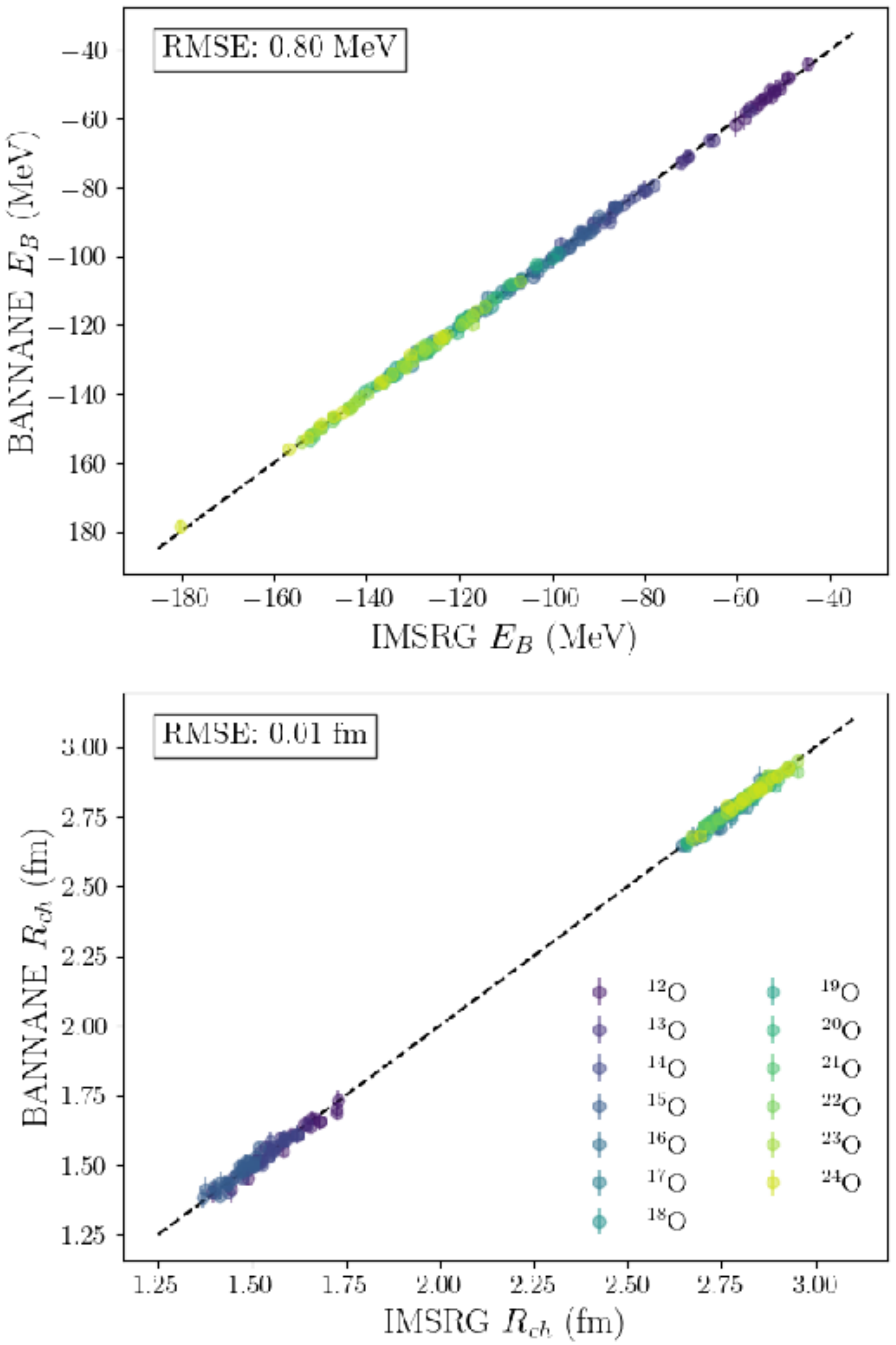
https://frame-architecture-explorer-553167399436.us-west1.run.app/

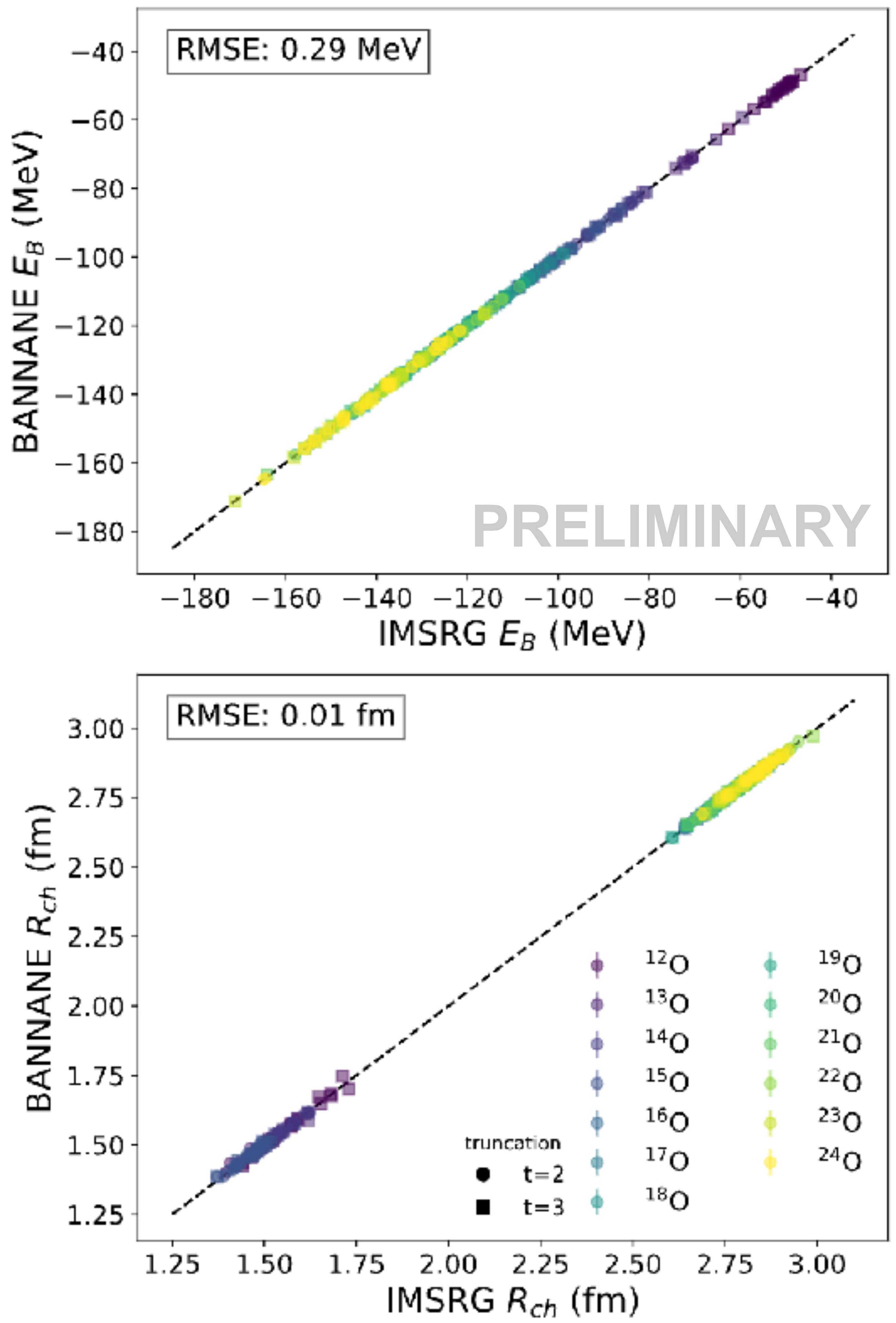
Extrapolating in e_{max}

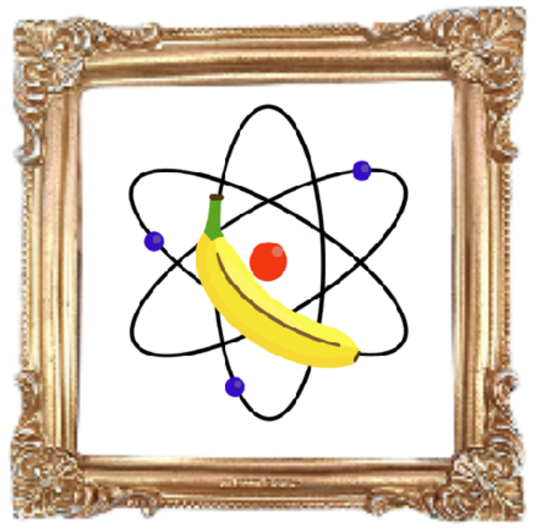


Physics Driven BANNANE!



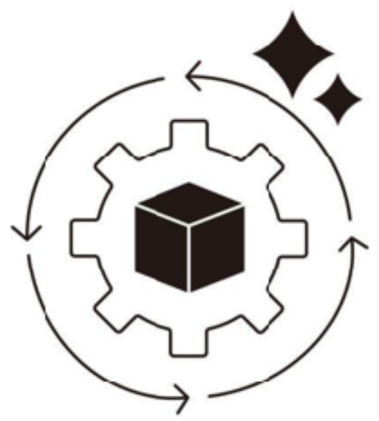






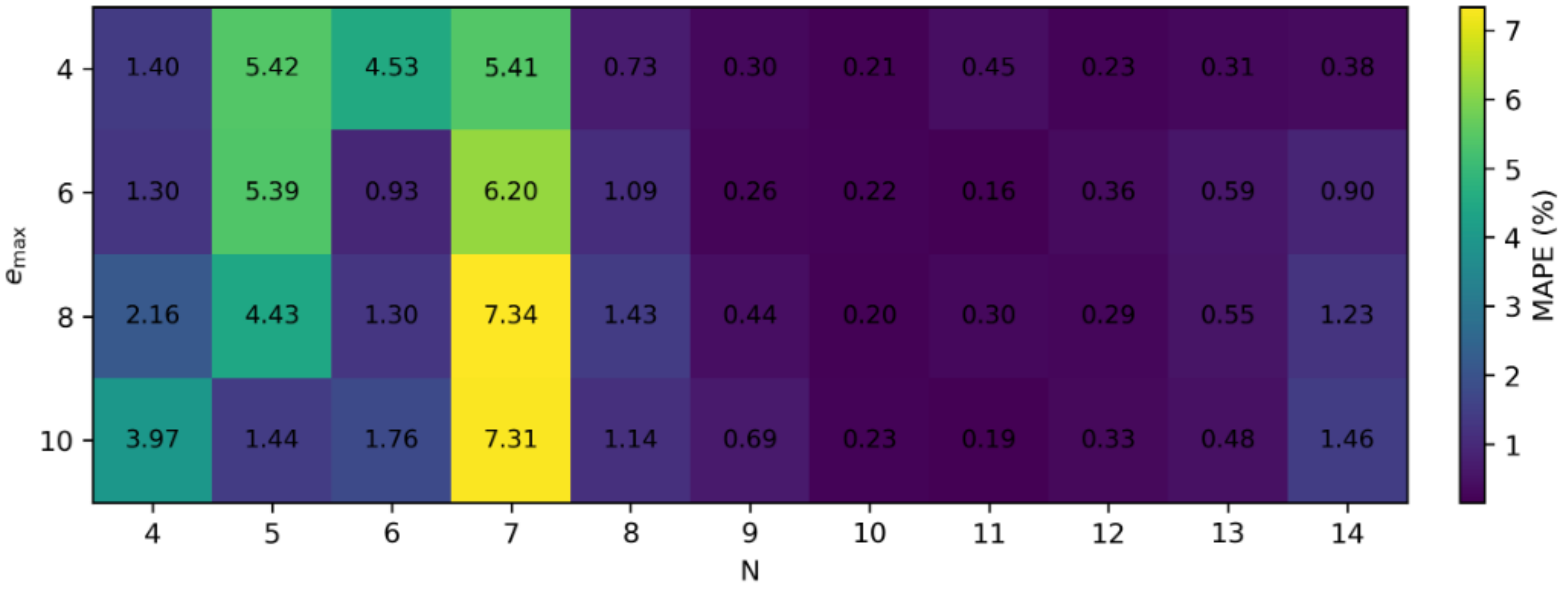
Training cost was reduced by a factor of 5!

Physics Driven BANNANE!



Zero-shot Learning

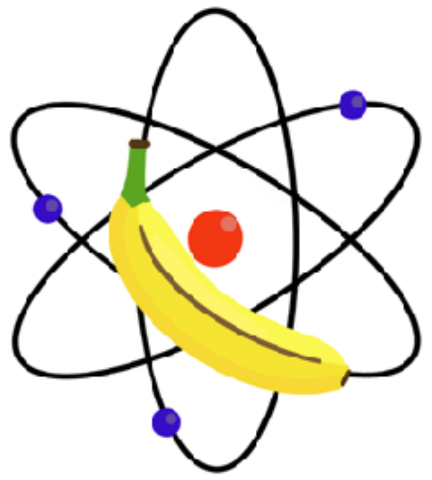
R_{ch}

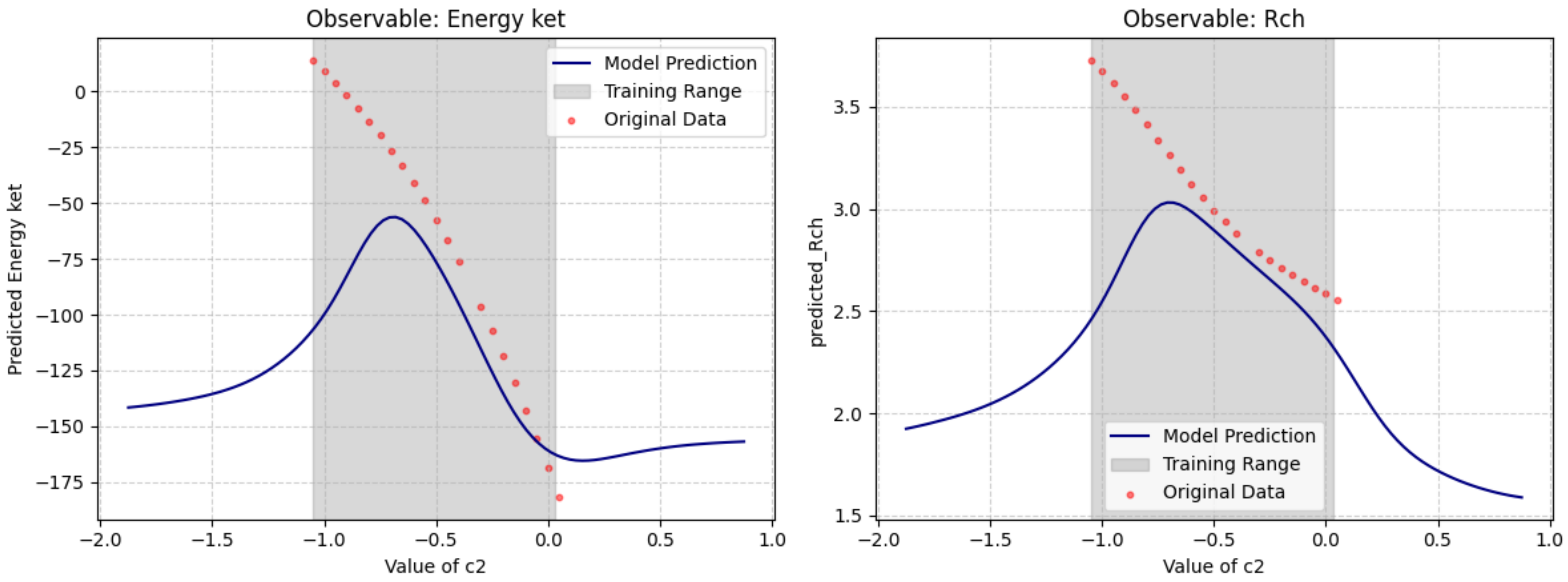


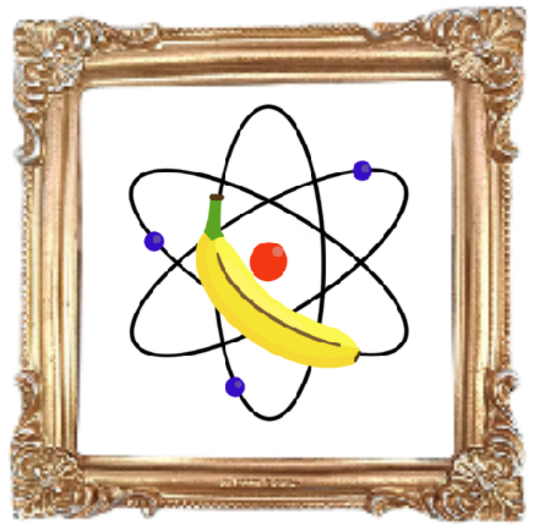
Physics driven version reduced worst zero-shot learning case from 44% error to 7% error!

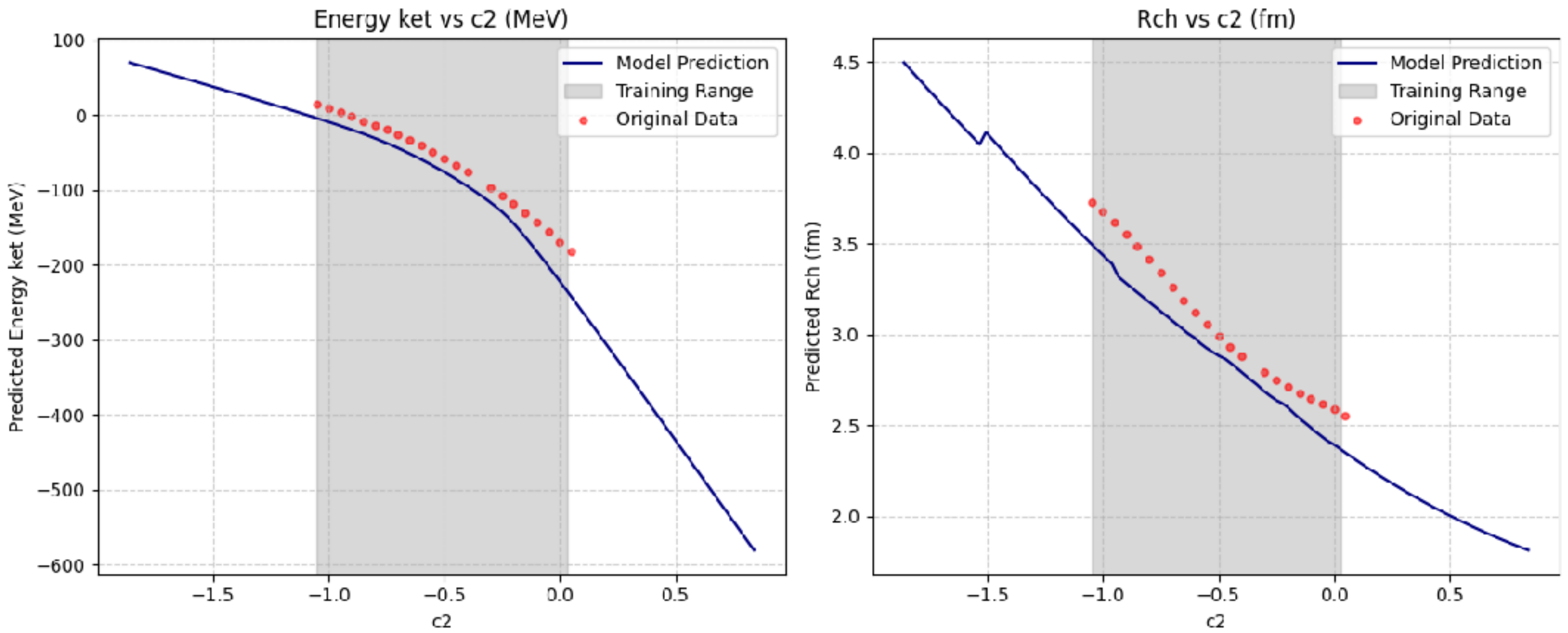
Physics Driven BANNANE!

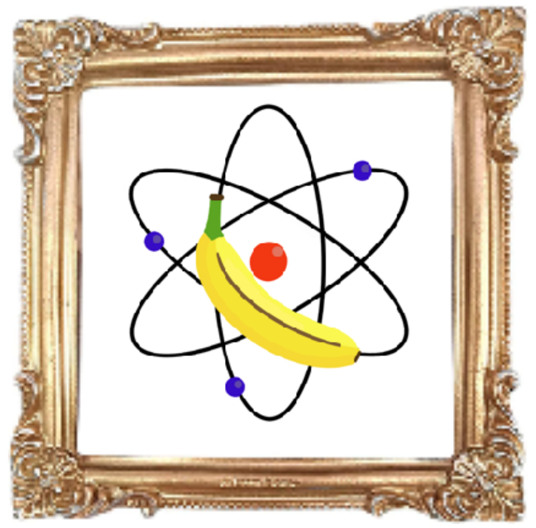
Model Extrapolation vs. Low-Energy Constants (LECs)



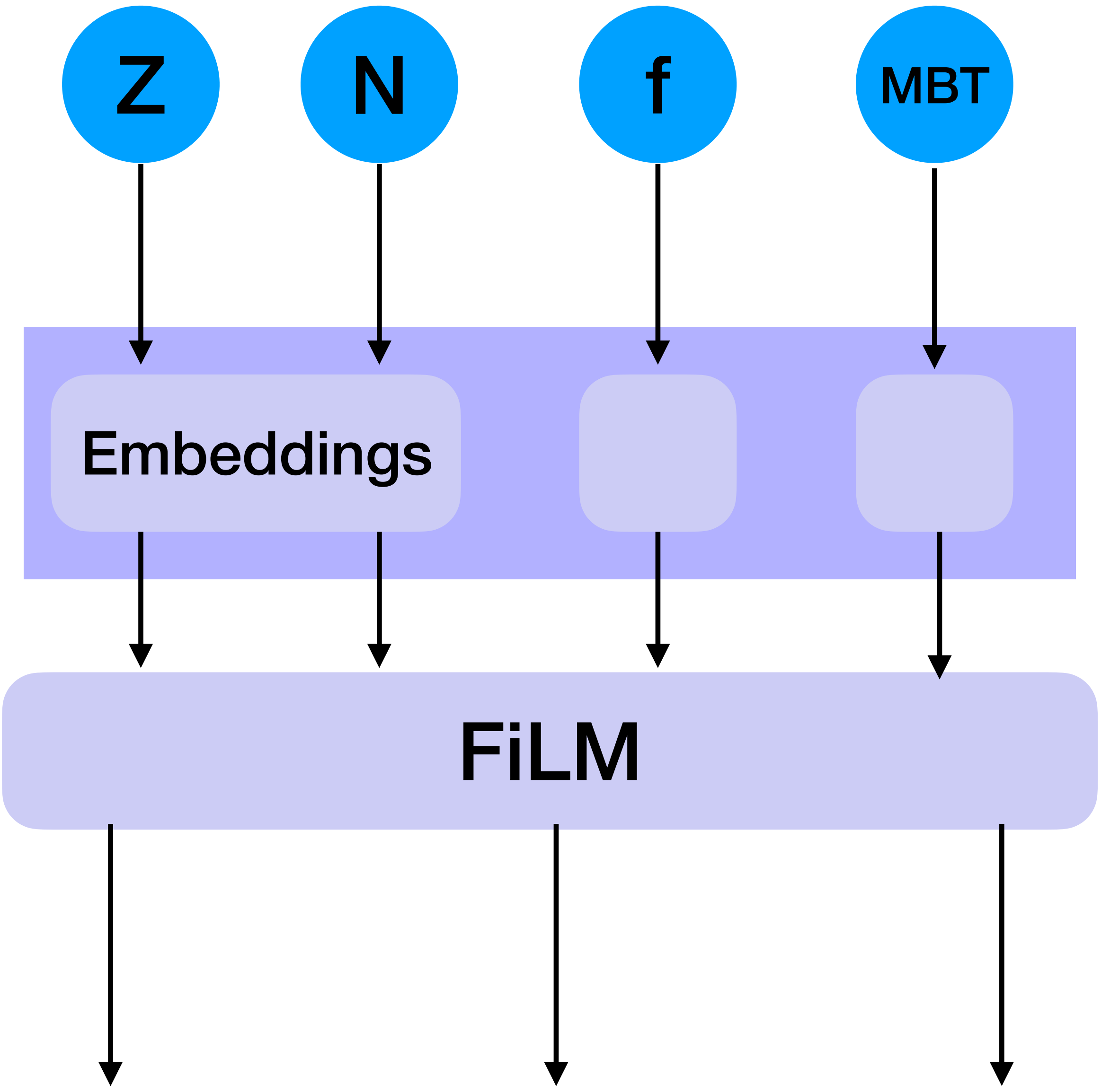


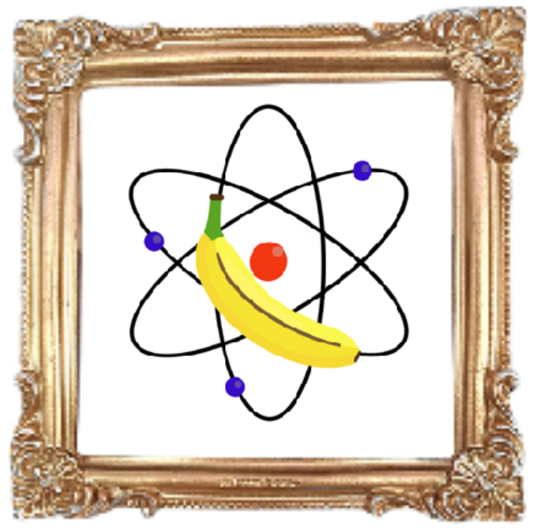




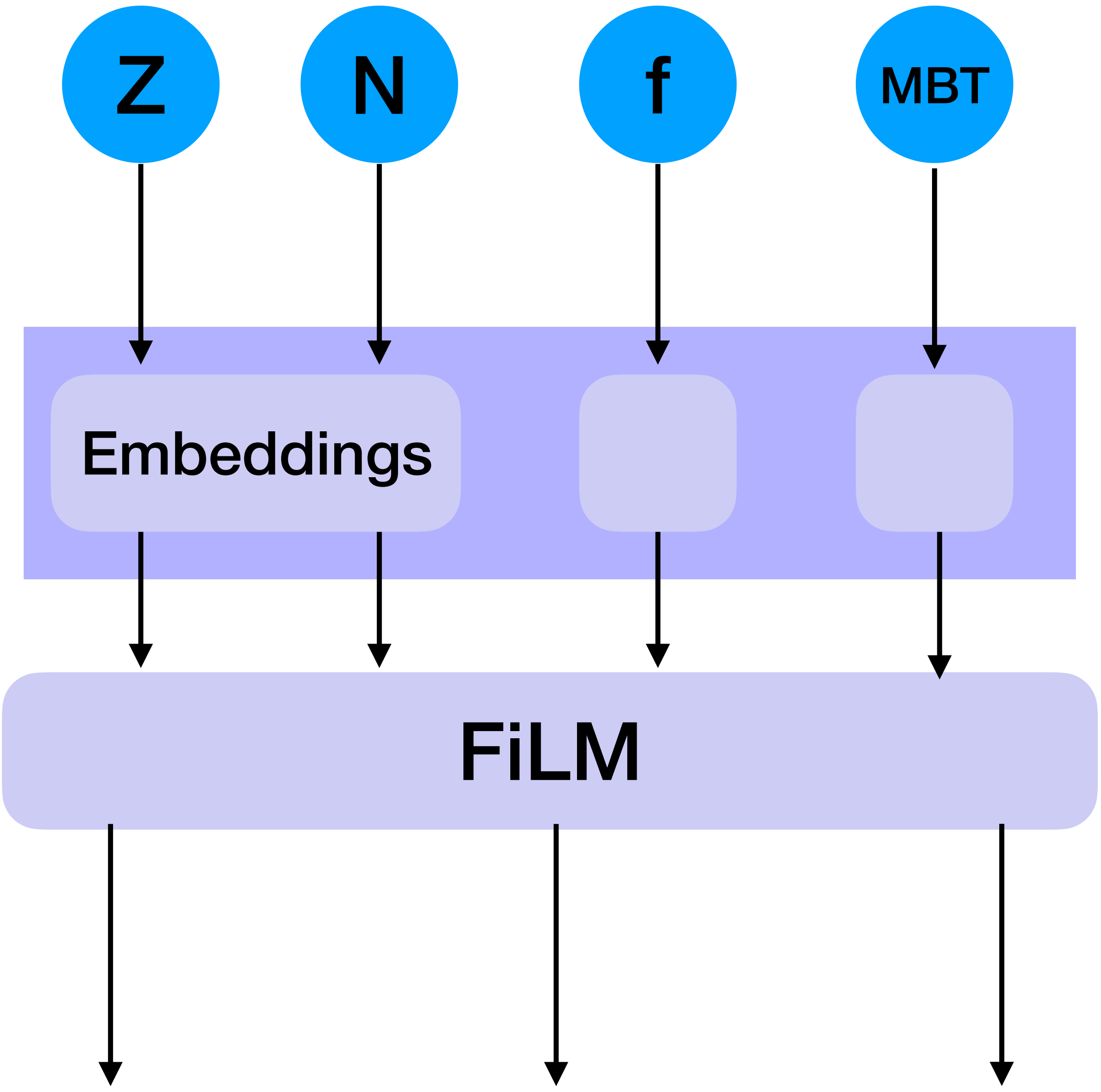


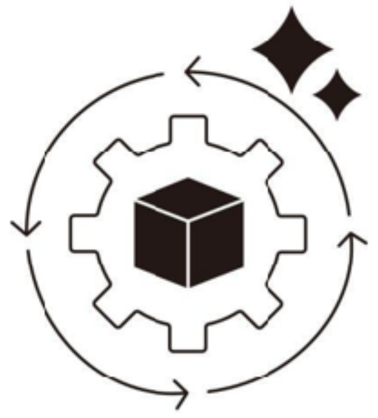
Extrapolating to IMSRG(3)





Extrapolating to IMSRG(3)

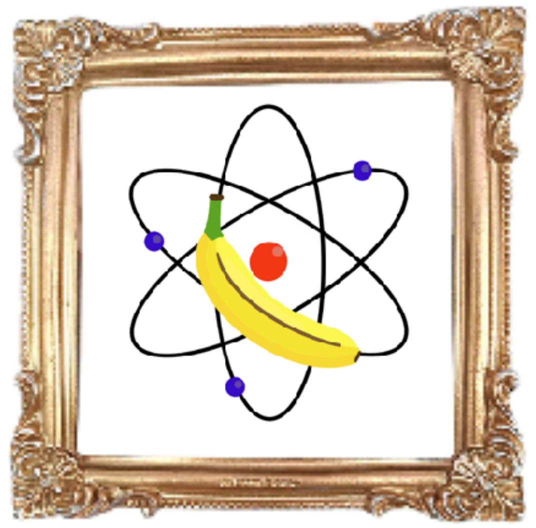
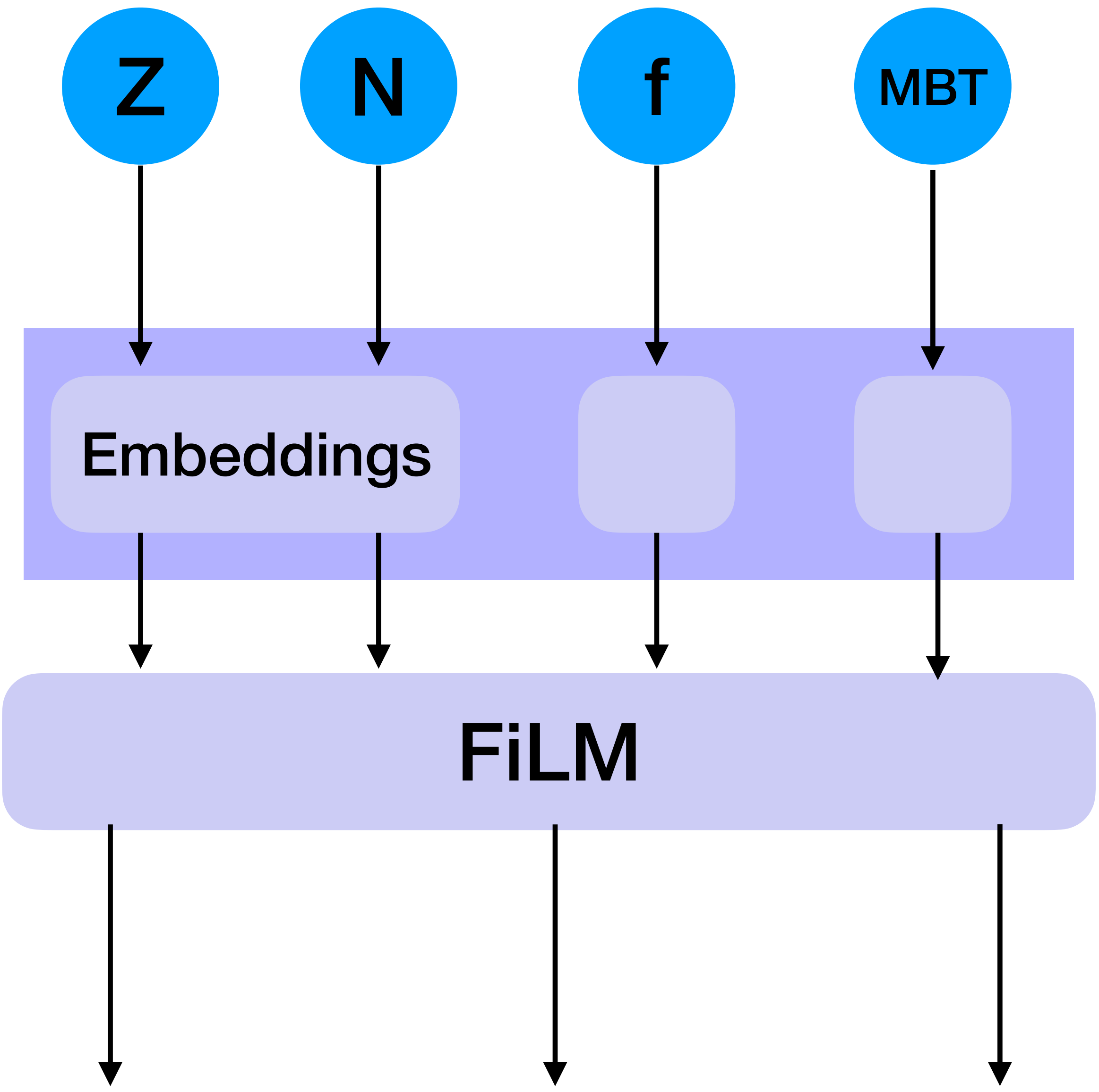


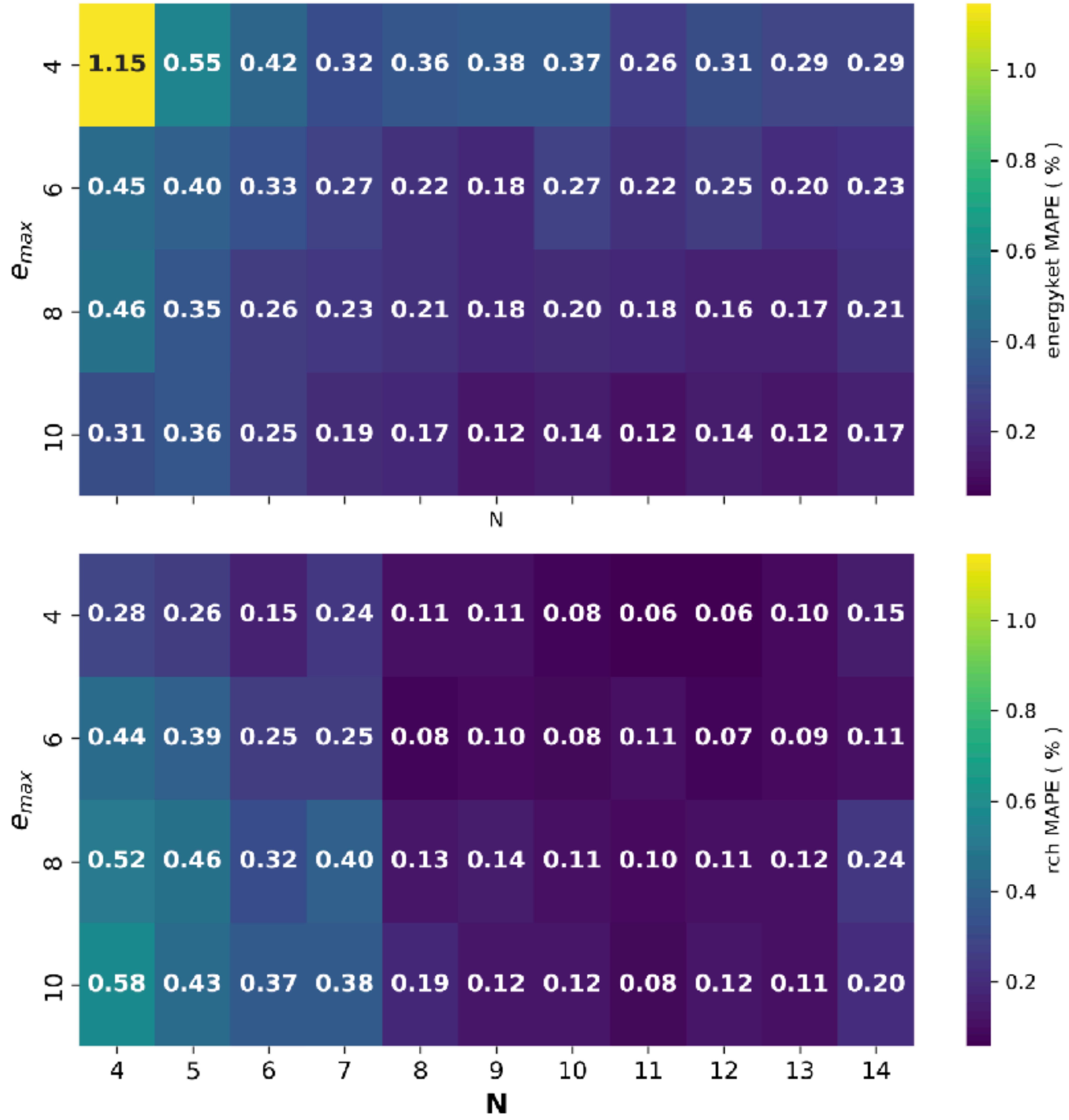


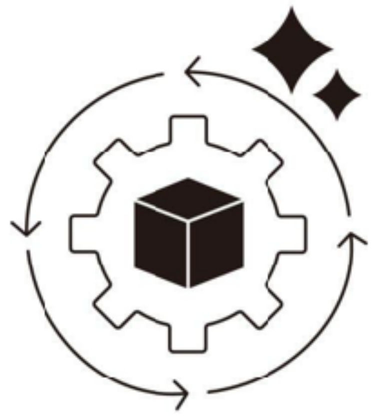
Zero-shot Learning

Difference with model and calculation of IMSRG(3f2) when no IMSRG(3f2) is used to train in a specific isotope.

Extrapolating to IMSRG(3)



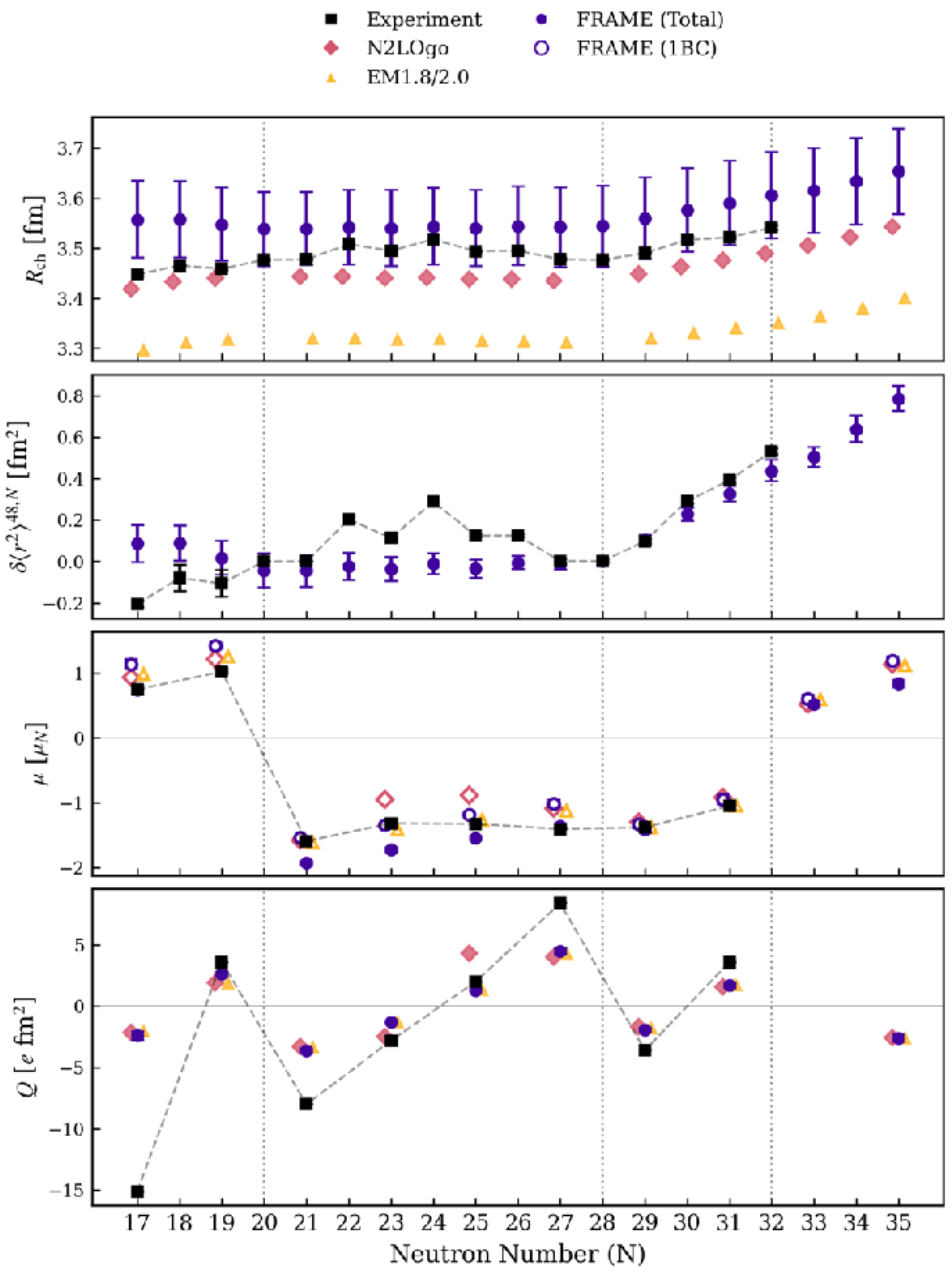


Zero-shot Learning

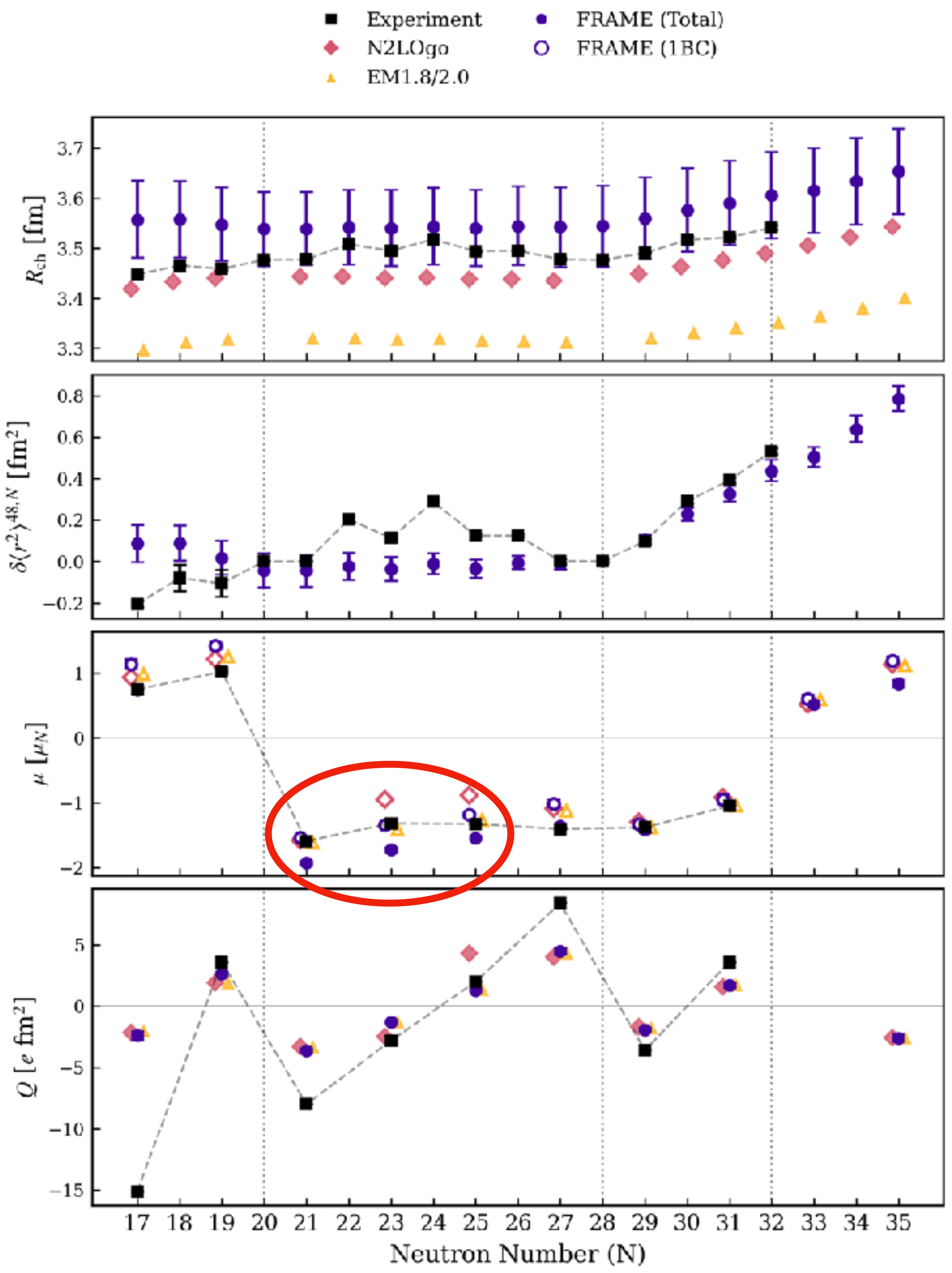
Difference with model and calculation of IMSRG(3f2) when no IMSRG(3f2) is used to train in a specific isotope.

Case study: EM Moments of Ca isotopes

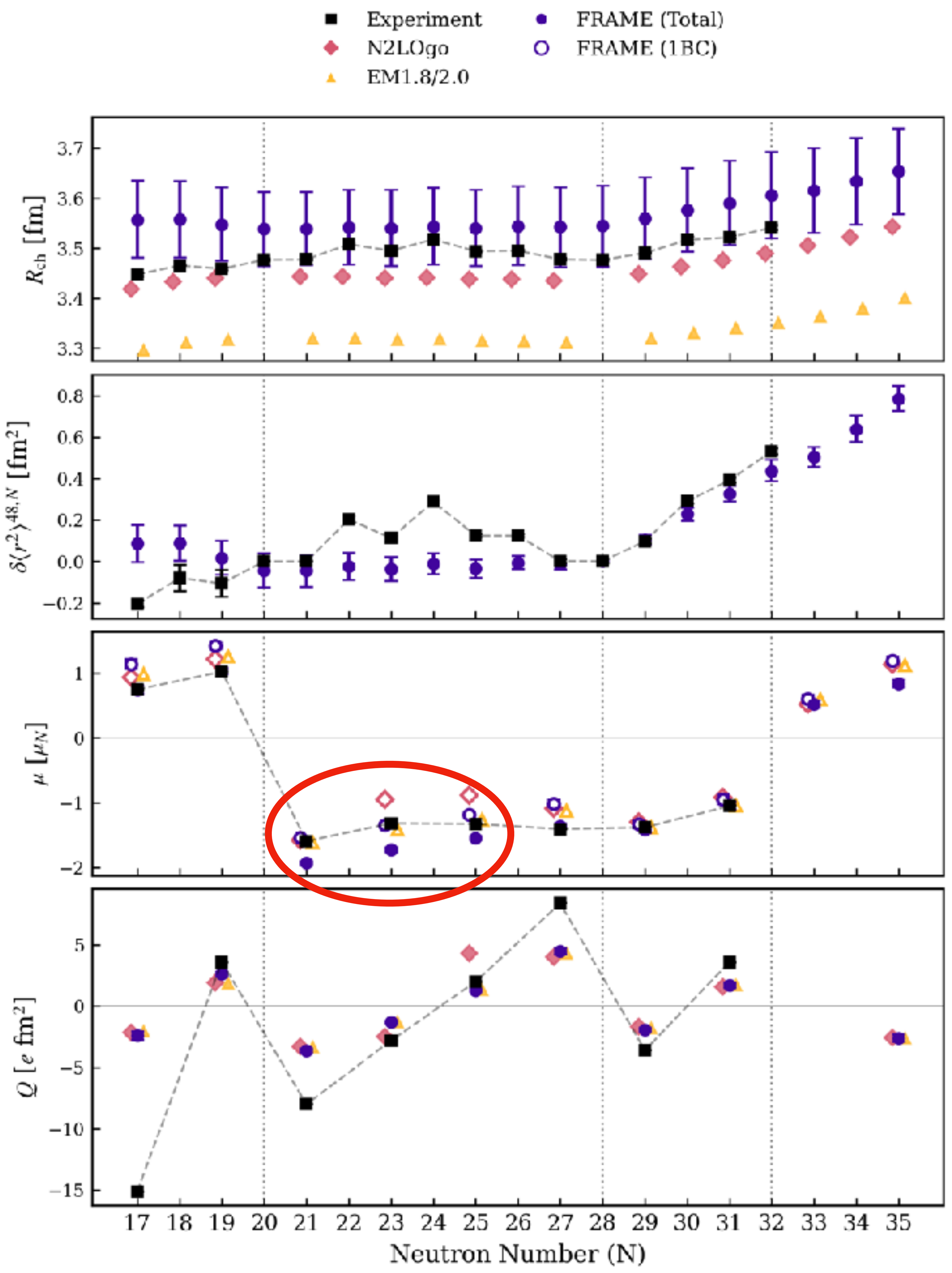
EM Moments of Ca Isotopes



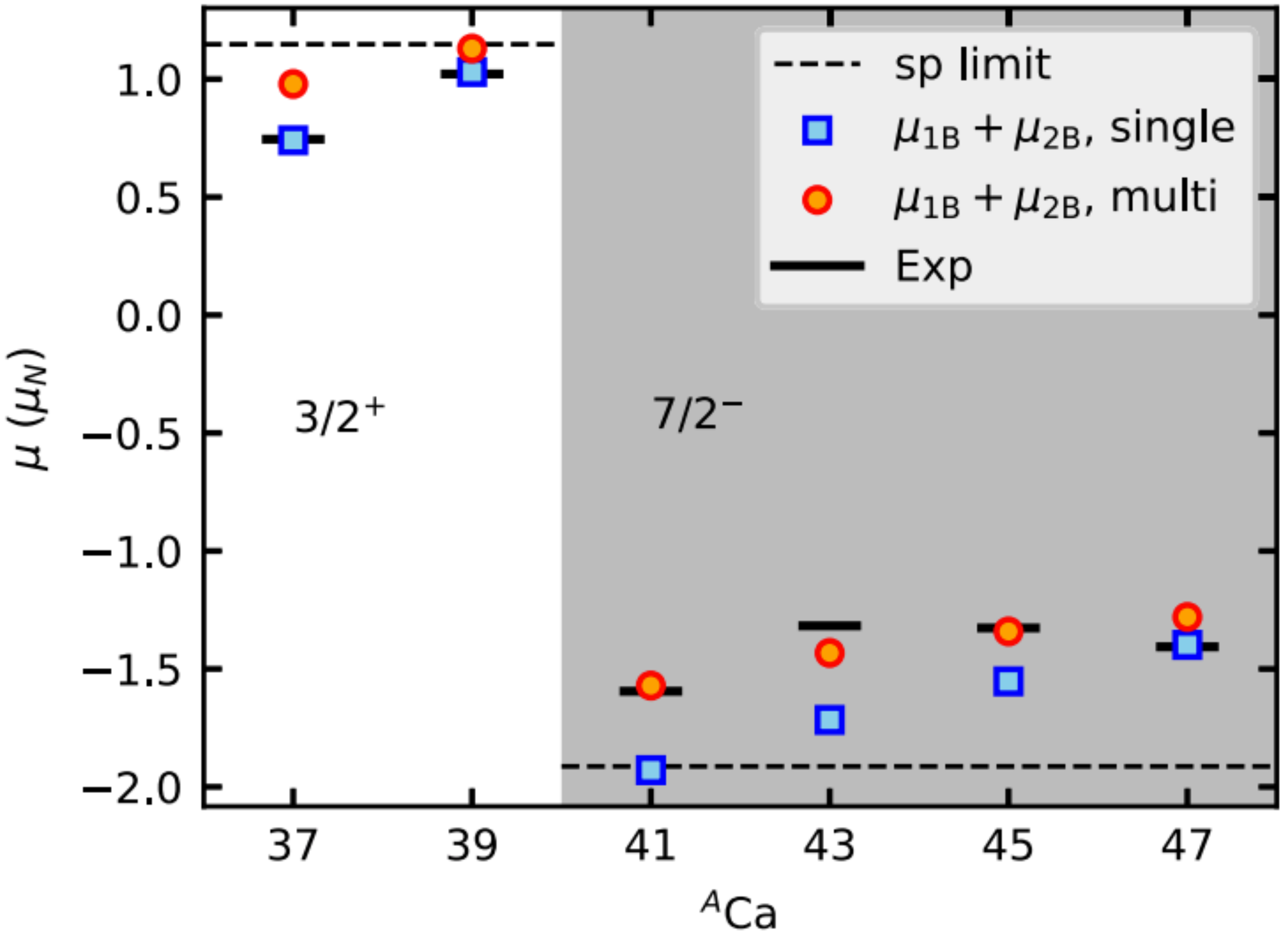
EM Moments of Ca Isotopes



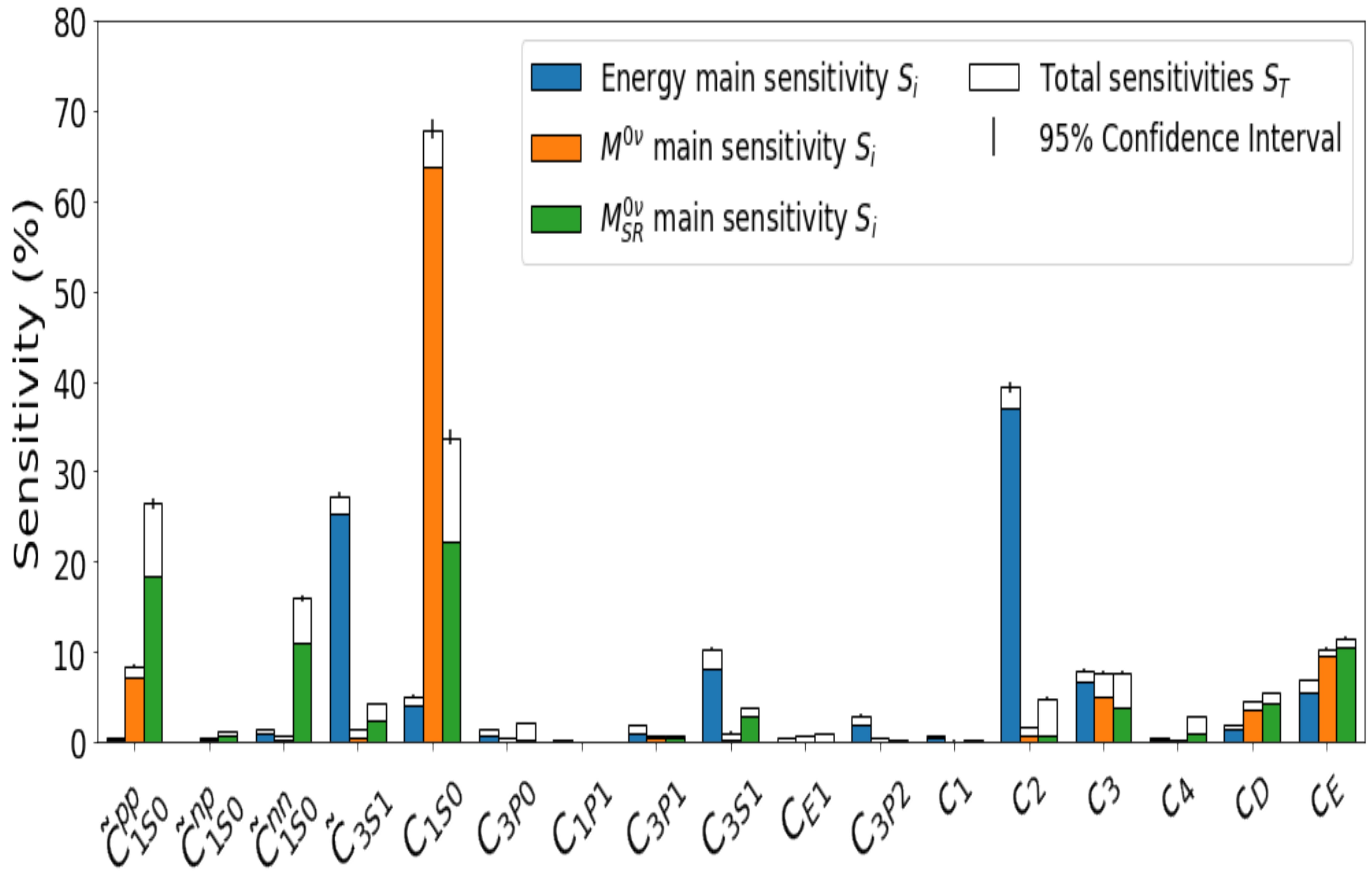
EM Moments of Ca Isotopes



Miyagi et al., Phys. Rev. Lett. **132**, 232503

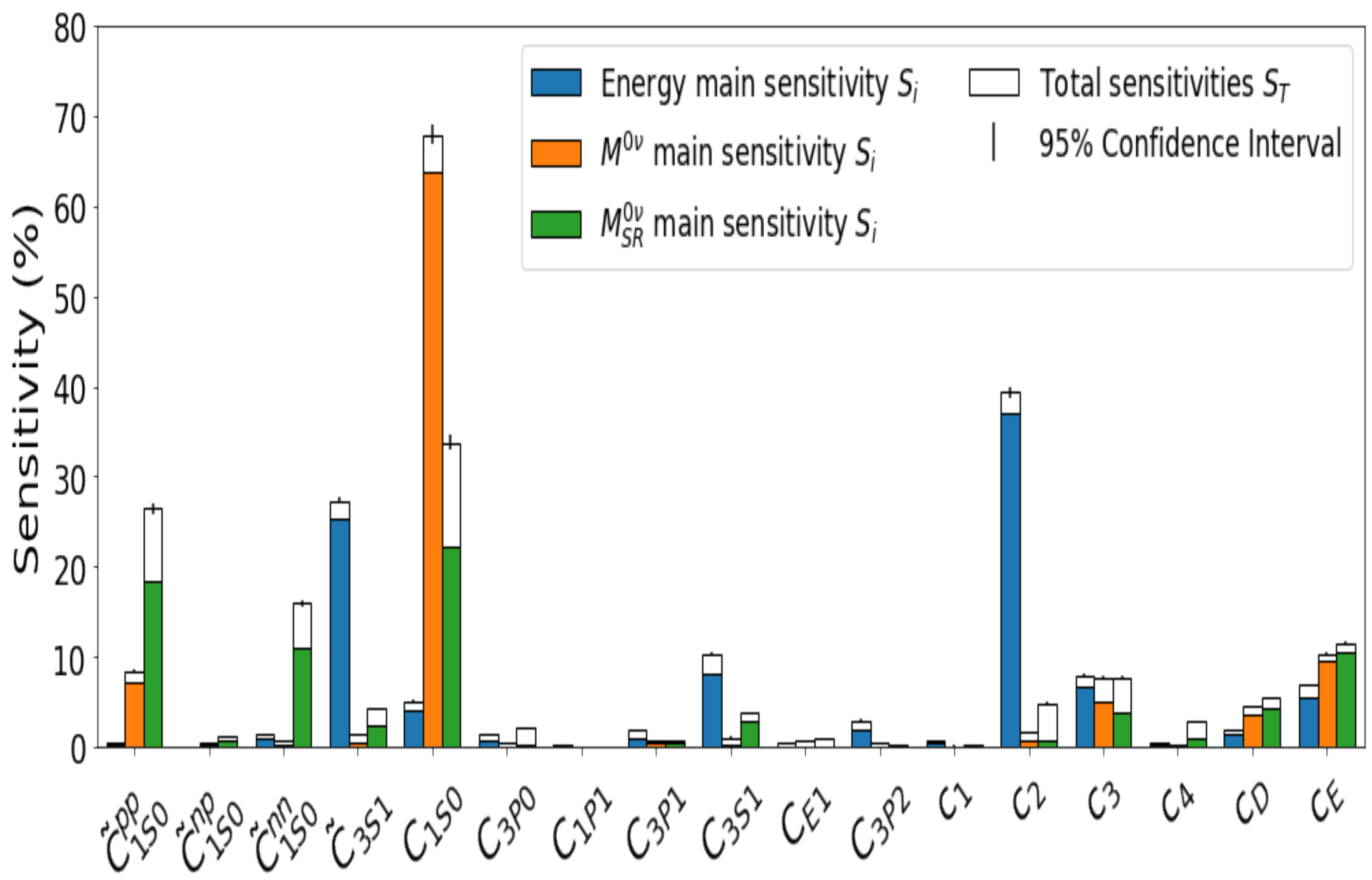


Issues with GSA



Issues with GSA

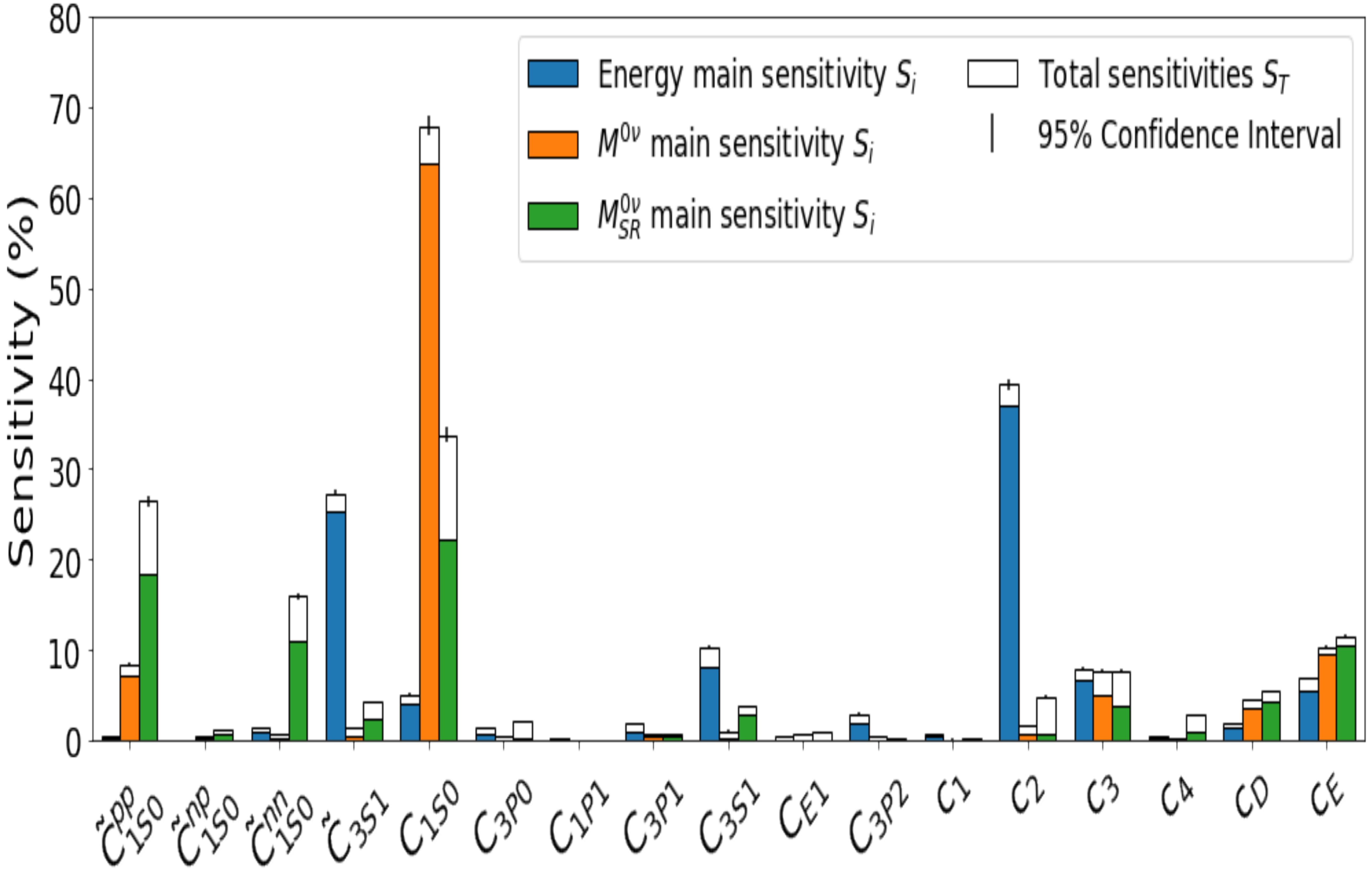
- Very costly



Issues with GSA

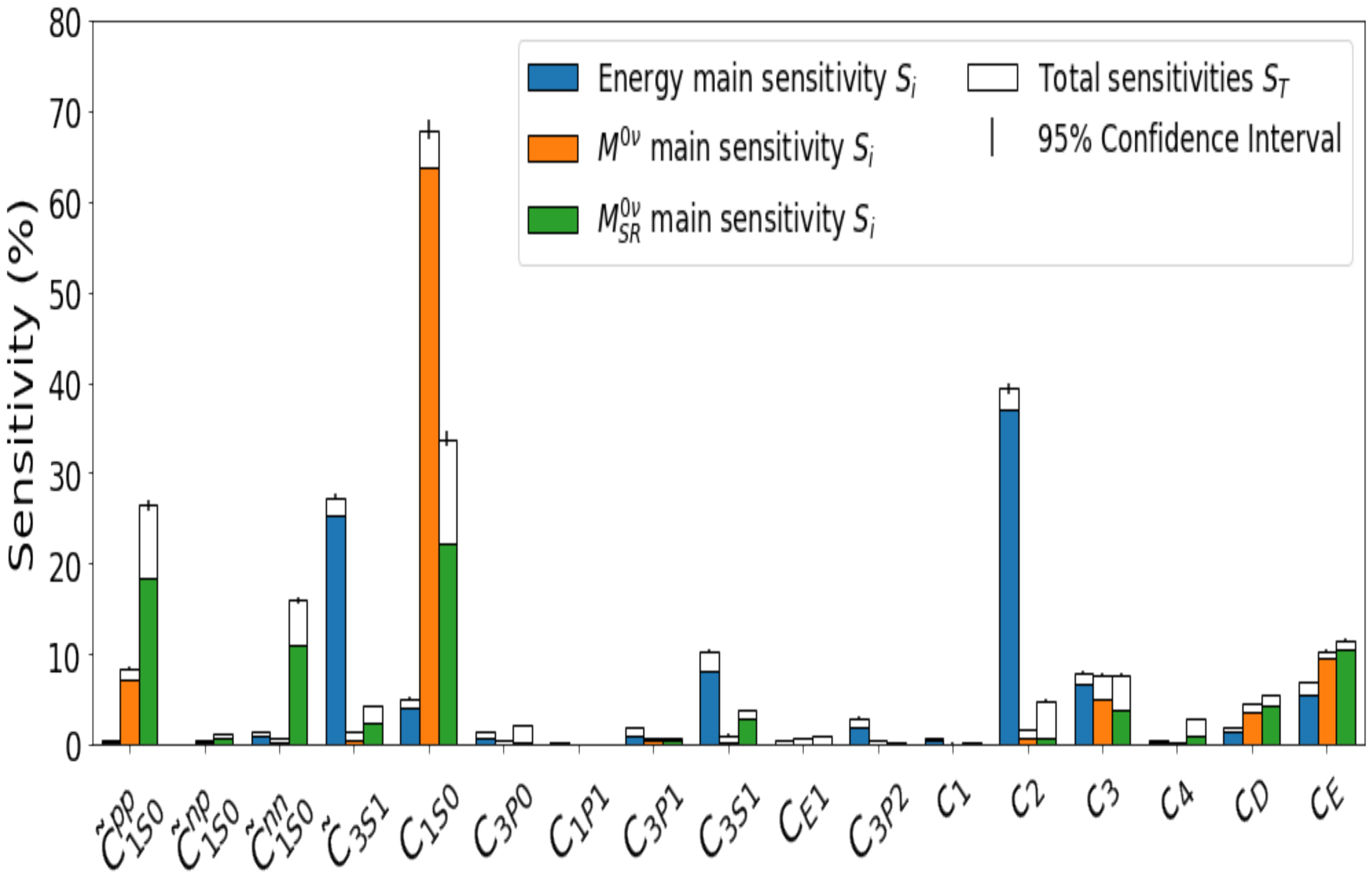
- Very costly

- No indication of how the parameters influence the results



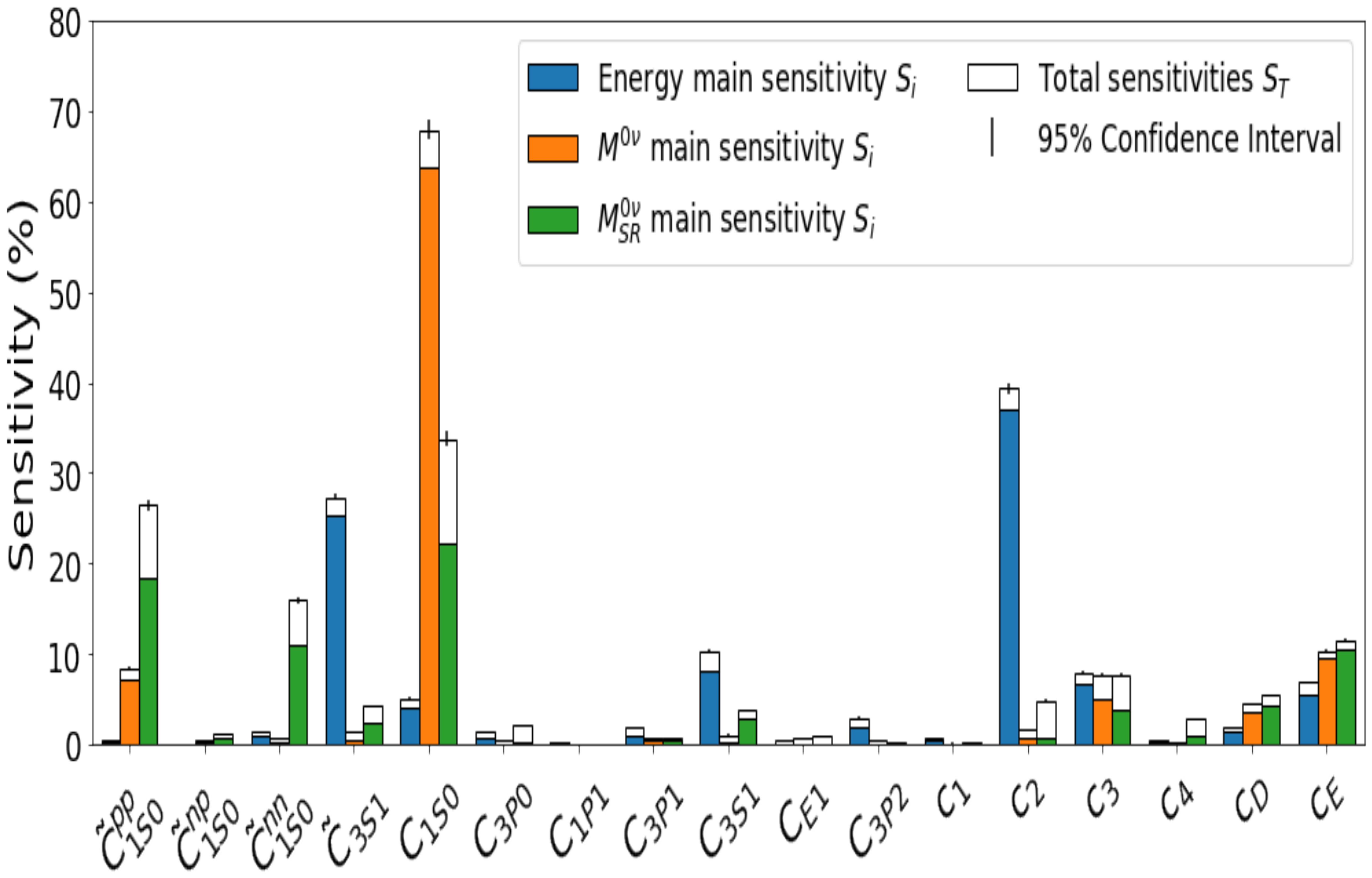
Issues with GSA

- Very costly
- No indication of how the parameters influence the results
- Initial range in the prior influence the final sensitivity



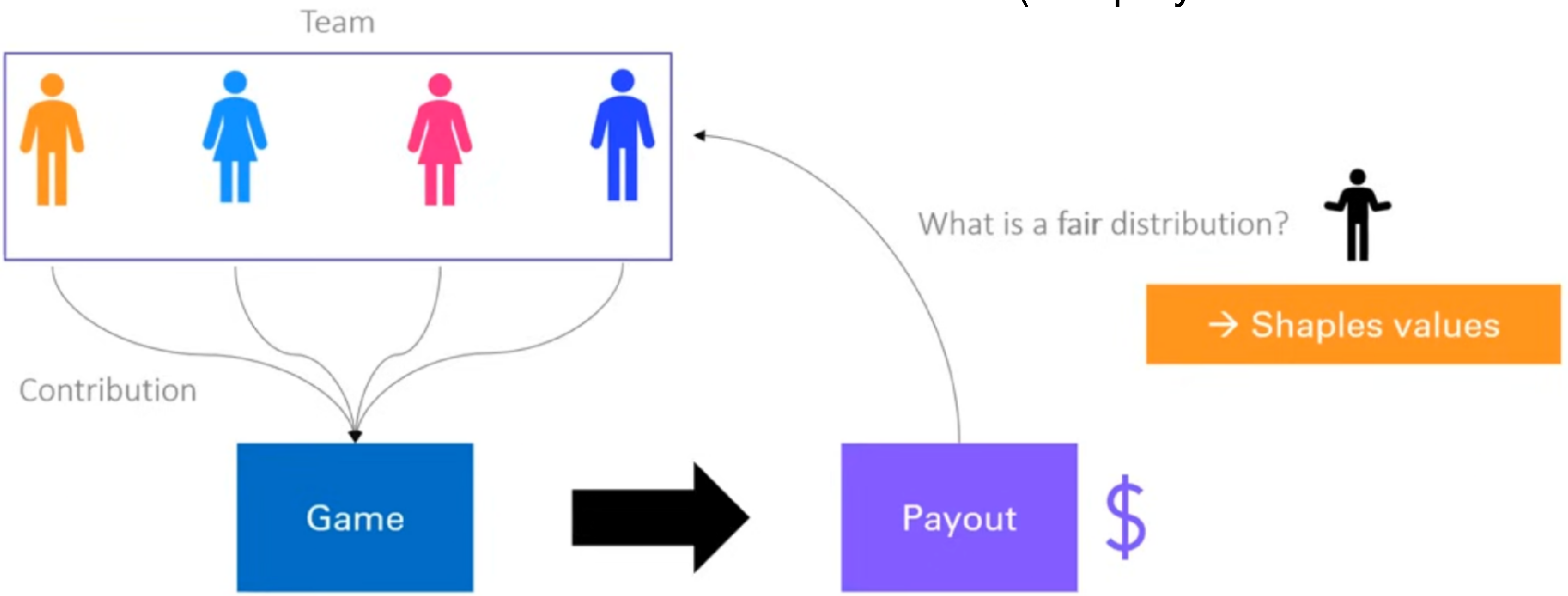
Issues with GSA

- Very costly
- No indication of how the parameters influence the results
- Initial range in the prior influence the final sensitivity
- “Bad” samples impact sensitivity as much as “good” samples



SHAP Values

(SHapely Additive exPlanation)

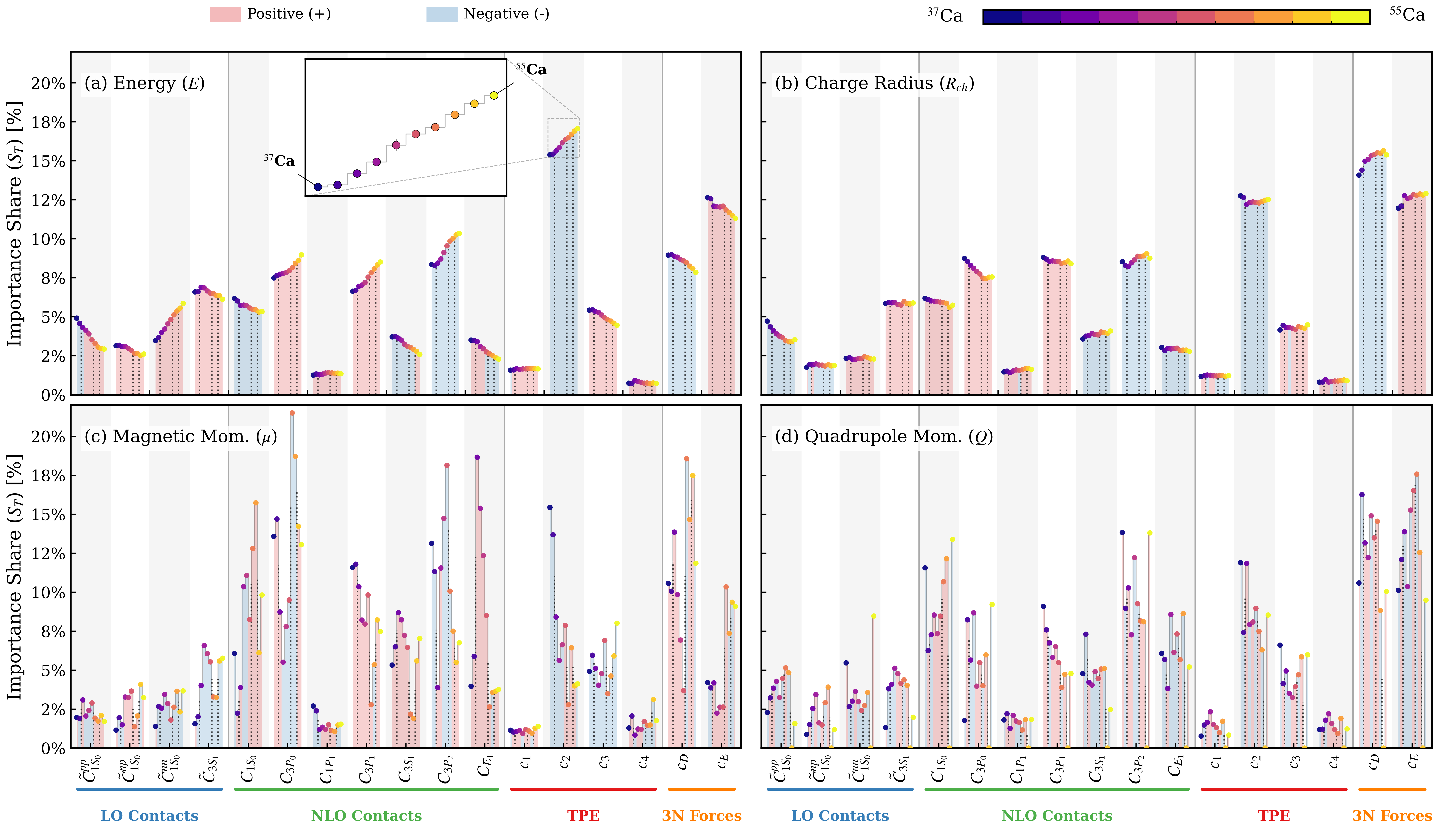


Sobol vs SHAP

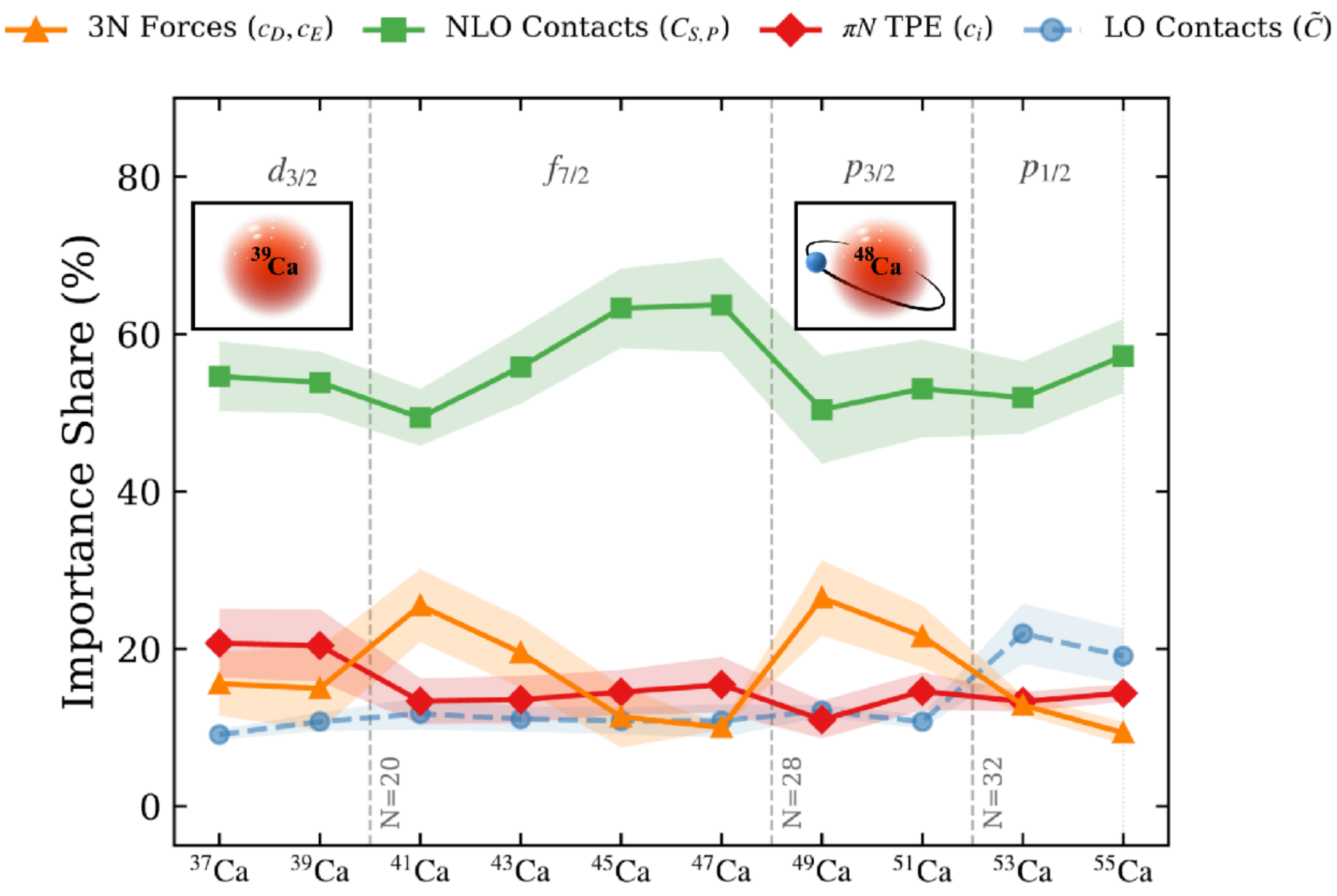
Aspect	Sobol′	SHAP (standard ML usage)
Primary question	Which inputs drive **variance** globally?	Which inputs drive a **specific prediction level** relative to baseline?
Scope	**Global** by design	**Local** per x; global only after aggregation
Units	Fraction of **variance** $[0, 1]$	In **output units** (additive contributions)
Interactions	Explicit $S_{ij}, S_{ijk}, \ldots; S_{T_i} - S_i$ = "all interactions involving i"	Interaction SHAP exists (ϕ_{ij}), but usually you view main ϕ_i or sum of interactions
Input dependence	Assumes **independent** inputs (classic Sobol′)	Can use **interventional** (assumes independence) or **conditional** SHAP (uses empirical dependence); each has trade-offs
Baseline	None (centered by variance)	Explicit baseline $\mathbb{E}[f(X)]$ (or dataset mean)
Cost	$(2 + d)N$ model evals for S_i, S_{T_i} with Saltelli/Jansen	Depends on explainer; TreeSHAP is fast & exact for trees; KernelSHAP is expensive (many coalitions)
Interpretability	"% of uncertainty due to X_i"	"How much X_i pushed this prediction up/down"

SHAP values in Ca Isotopes

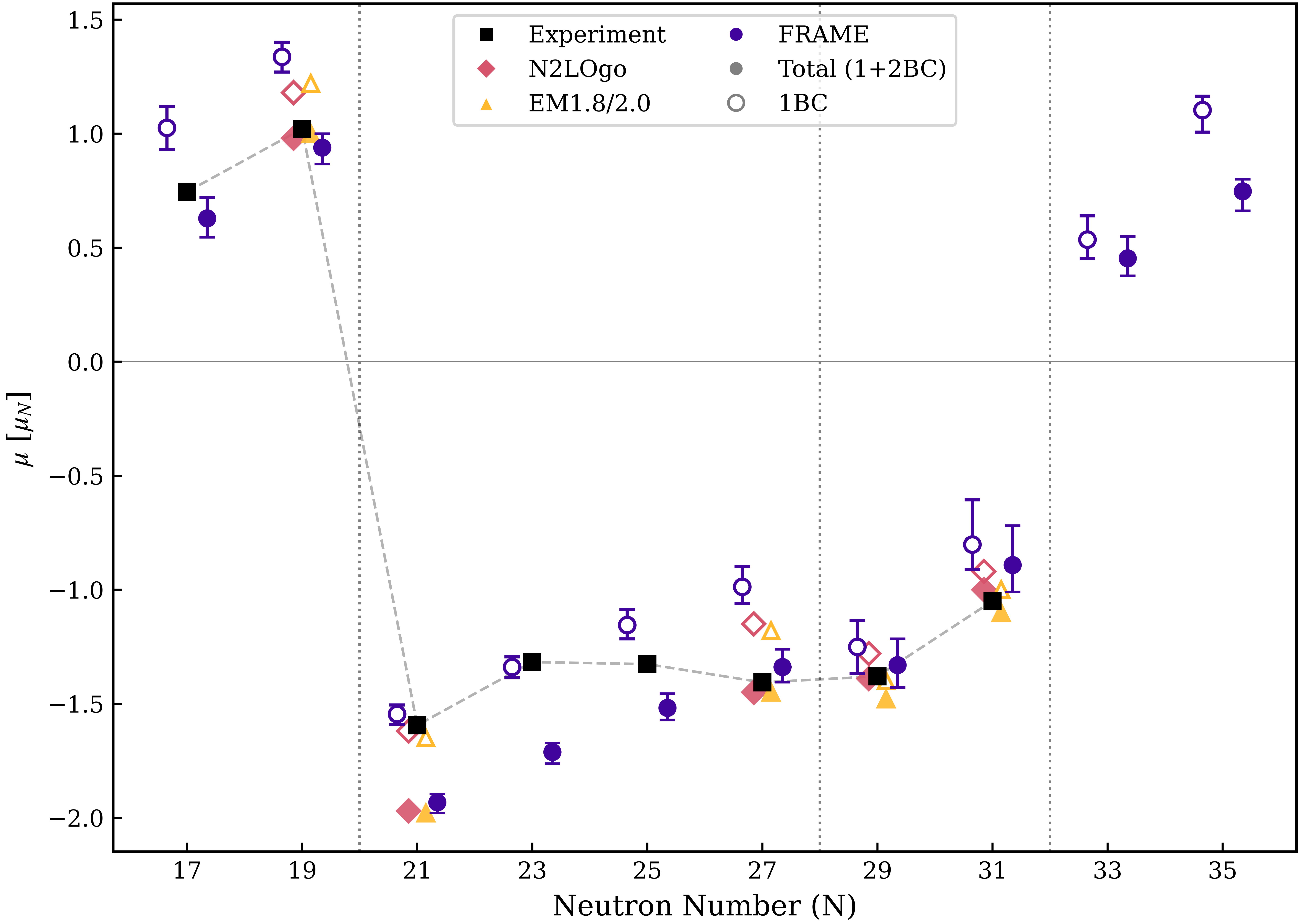
Munoz, **Belley**, et al., arXiv: 2603.26905 (2026)



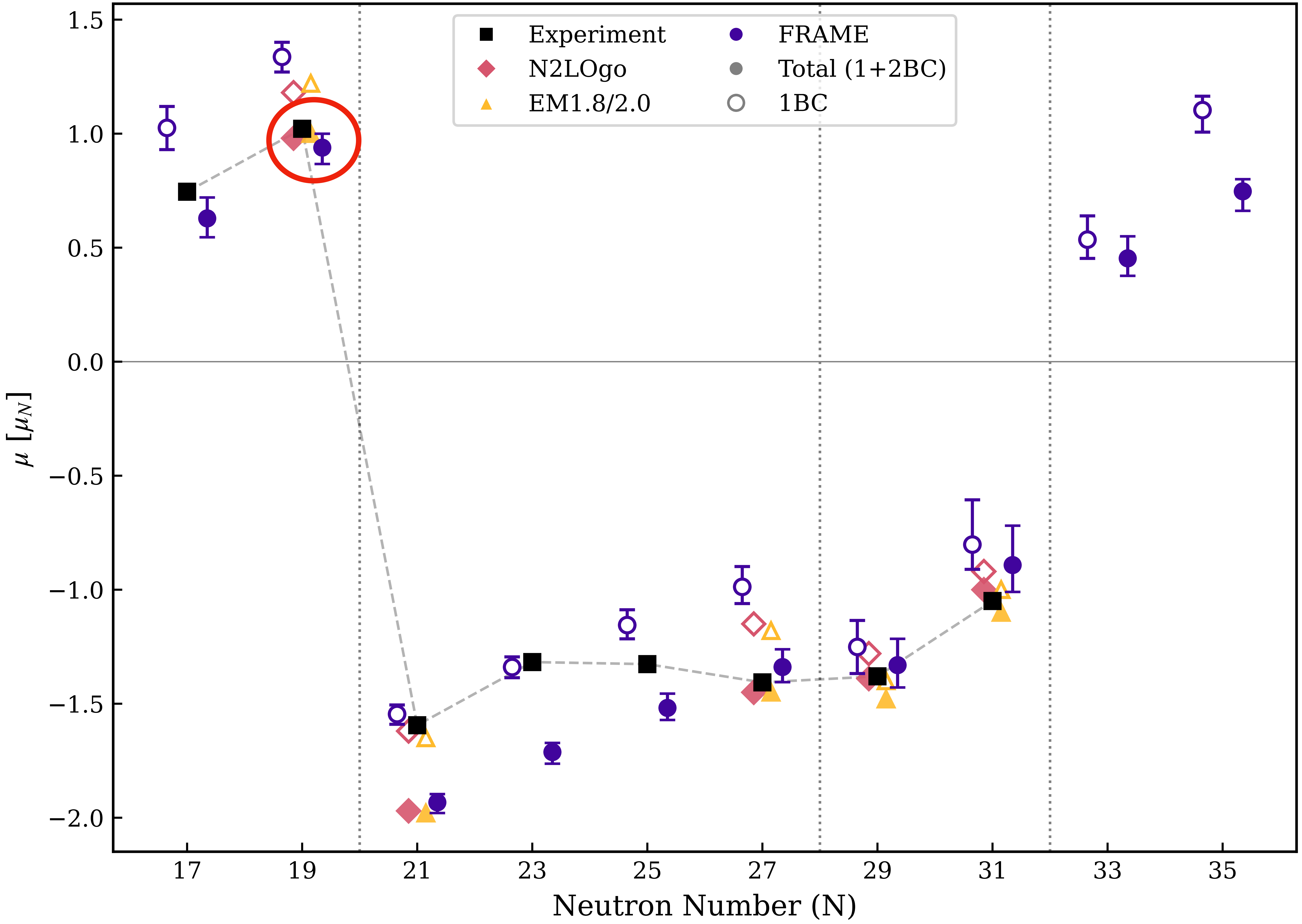
Forces Through the Shells



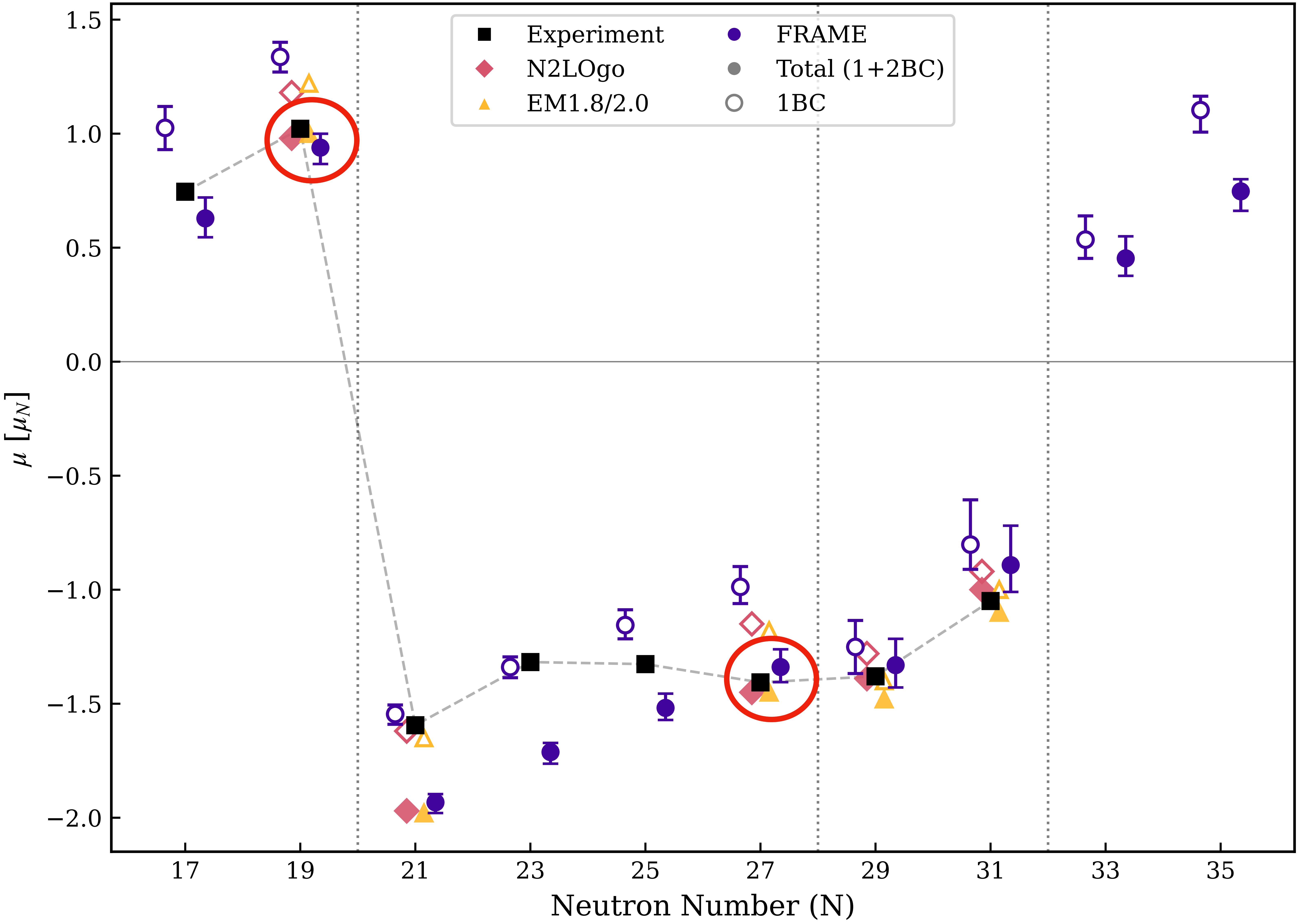
Using Emulators to Constrain Forces



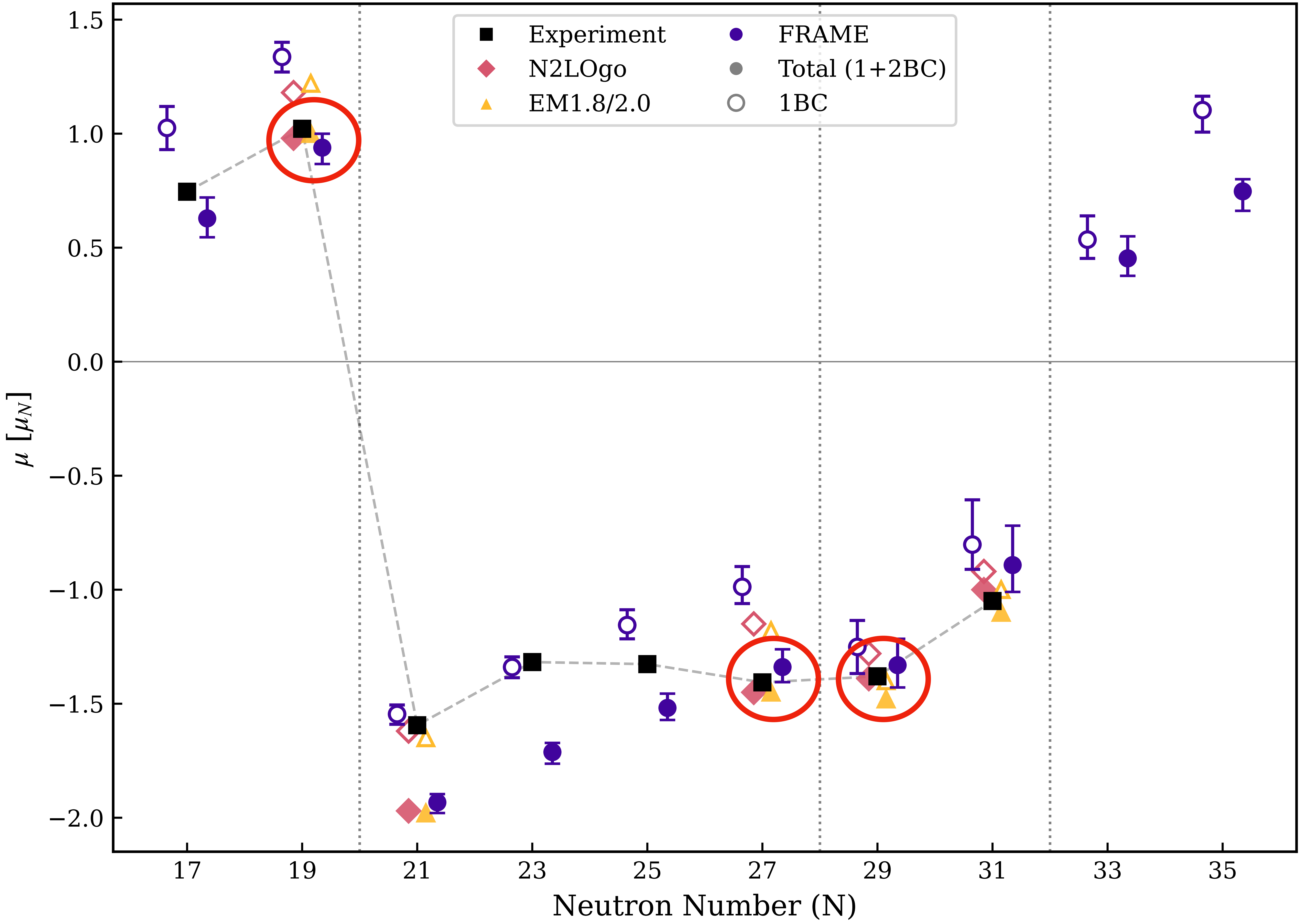
Using Emulators to Constrain Forces



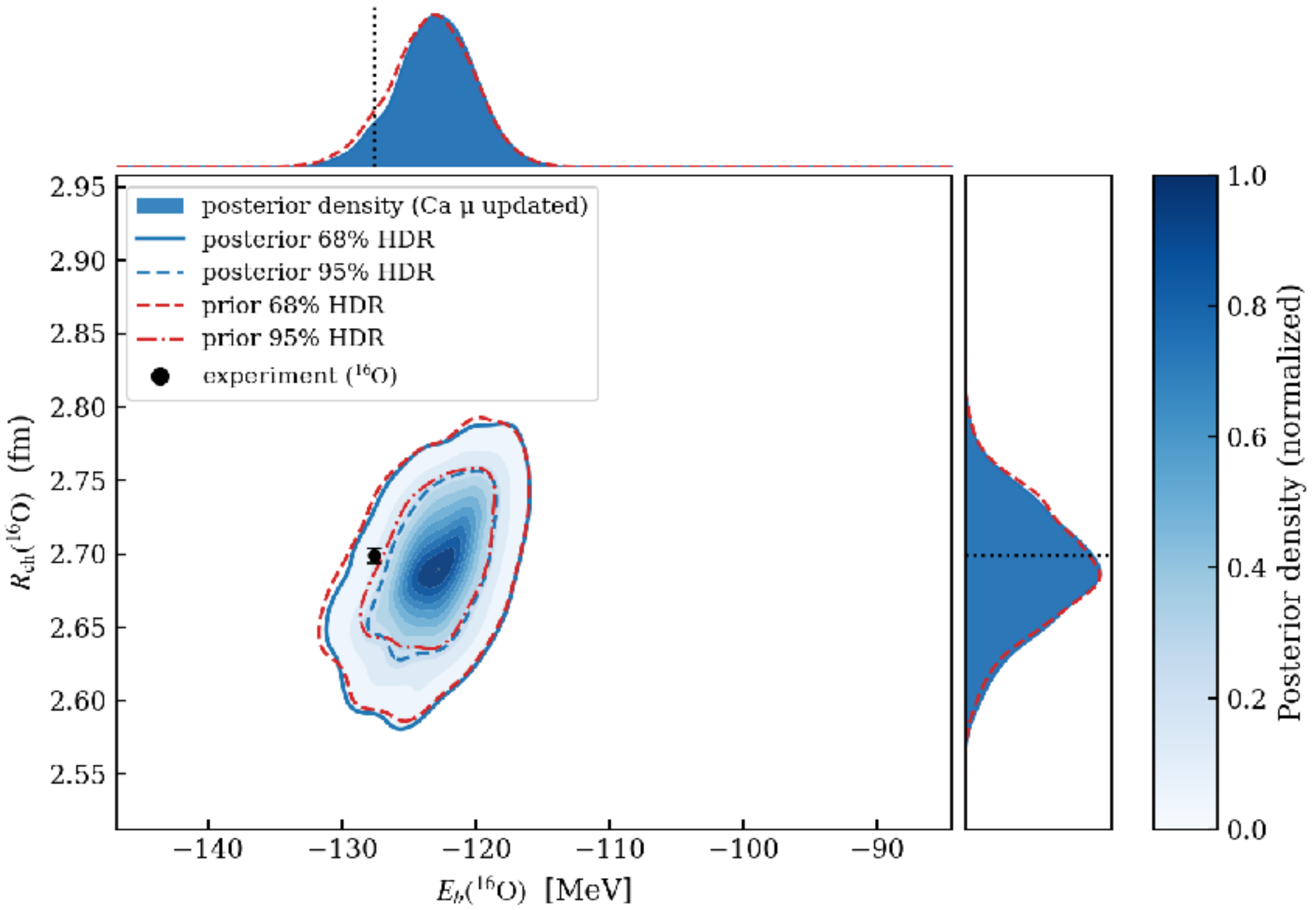
Using Emulators to Constrain Forces

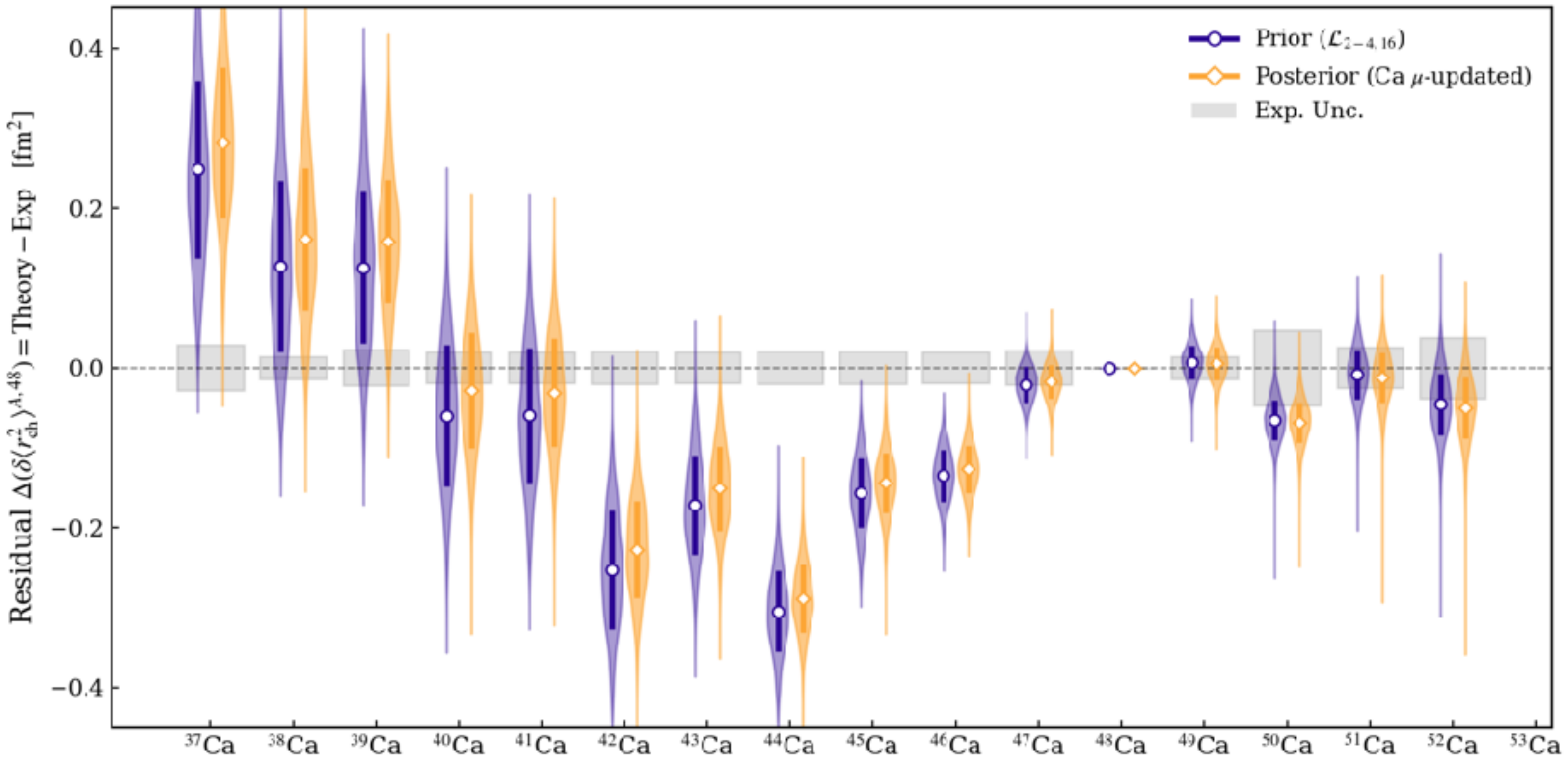


Using Emulators to Constrain Forces

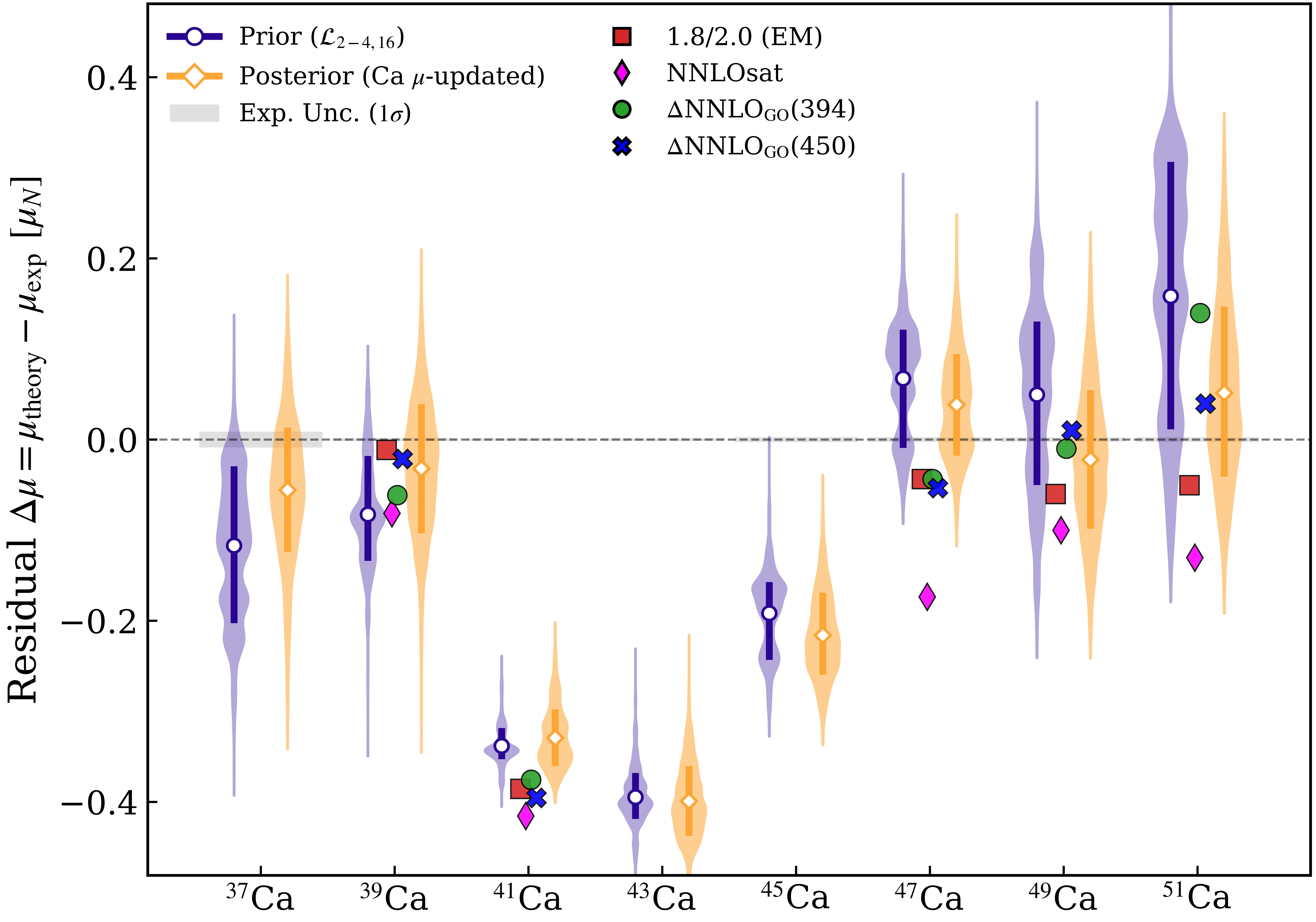


Changes to Likelihood

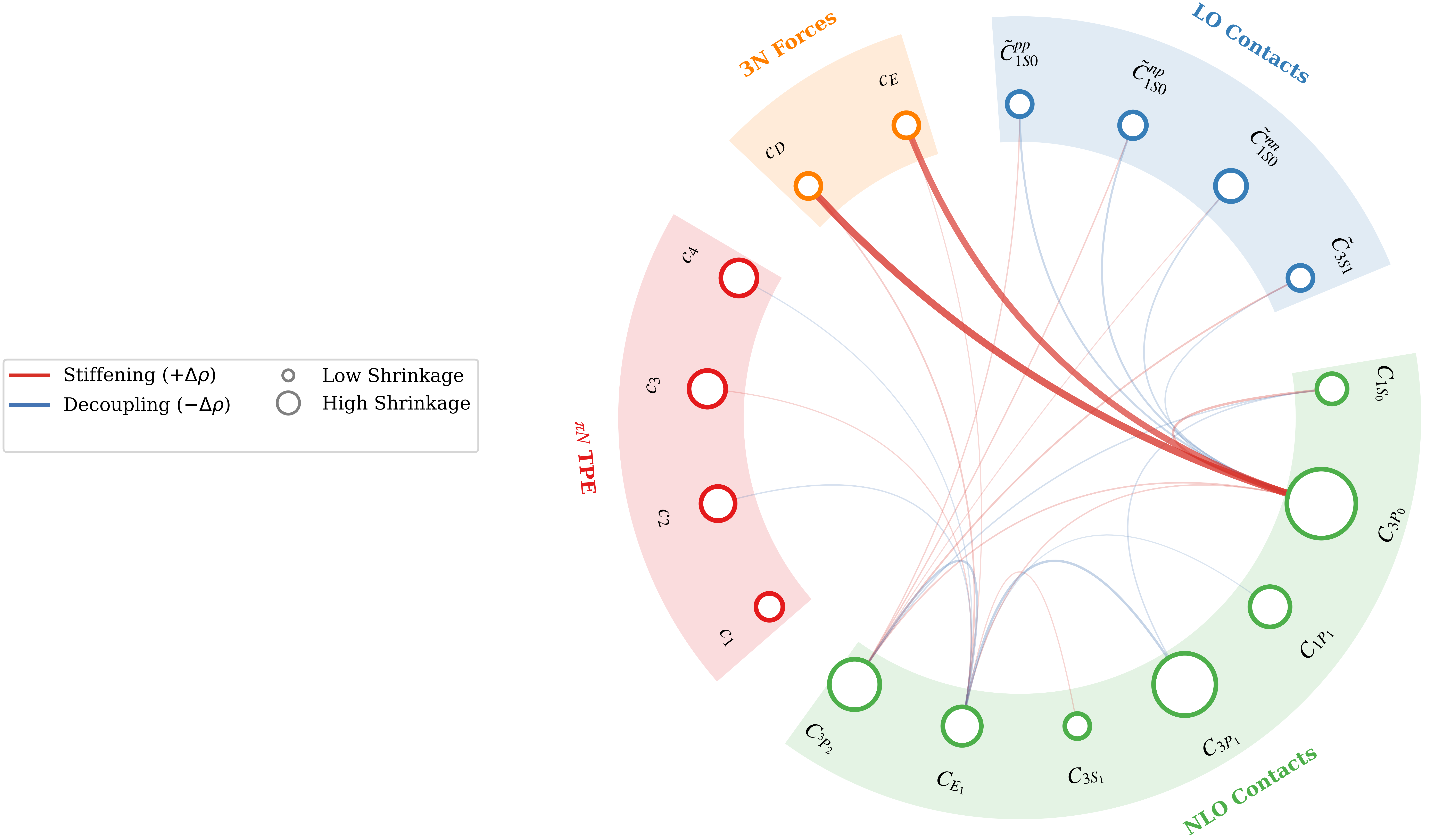




Changes to Likelihood



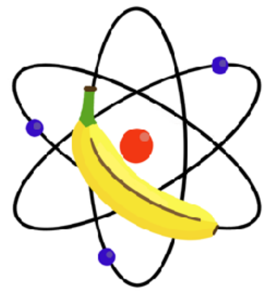
Changes to Likelihood



Summary …

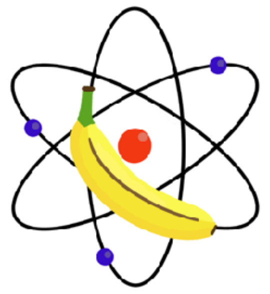
Thank you!

Summary …

- Created new ML emulator: speed up calculations by a factor of ~10^5. 

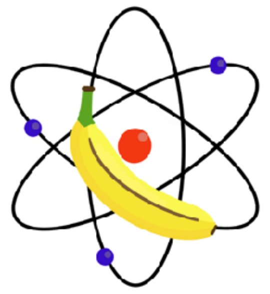
Thank you!

Summary …

- Created new ML emulator: speed up calculations by a factor of $\sim 10^5$. 
- Improved emulator using physics driven technology.

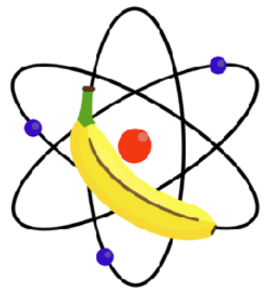
Thank you!

Summary …

- Created new ML emulator: speed up calculations by a factor of ~10^5. 
- Improved emulator using physics driven technology.
- Can predict IMSRG(3f2) from IMSRG(2) with sub-percent accuracy.

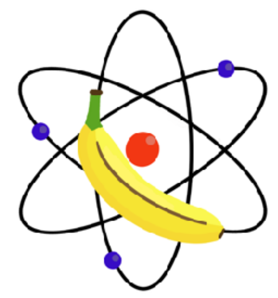
Thank you!

Summary …

- Created new ML emulator: speed up calculations by a factor of $\sim 10^5$. 
- Improved emulator using physics driven technology.
- Can predict IMSRG(3f2) from IMSRG(2) with sub-percent accuracy.
- Applied emulator in the Ca isotopic chain to find new constraints on interaction parameters.

Thank you!

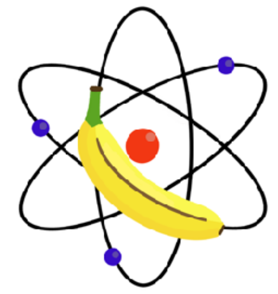
Summary …

- Created new ML emulator: speed up calculations by a factor of $\sim 10^5$. 
- Improved emulator using physics driven technology.
- Can predict IMSRG(3f2) from IMSRG(2) with sub-percent accuracy.
- Applied emulator in the Ca isotopic chain to find new constraints on interaction parameters.

… and Outlook

Thank you!

Summary …

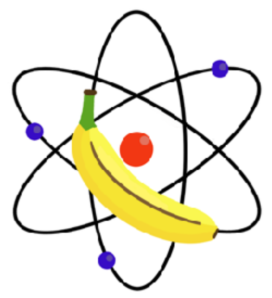
- Created new ML emulator: speed up calculations by a factor of ~10^5. 
- Improved emulator using physics driven technology.
- Can predict IMSRG(3f2) from IMSRG(2) with sub-percent accuracy.
- Applied emulator in the Ca isotopic chain to find new constraints on interaction parameters.

… and Outlook

- Use new emulator to quantify uncertainty of many-observables.

Thank you!

Summary …

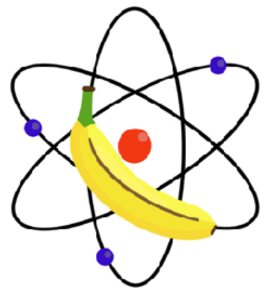
- Created new ML emulator: speed up calculations by a factor of $\sim 10^5$. 
- Improved emulator using physics driven technology.
- Can predict IMSRG(3f2) from IMSRG(2) with sub-percent accuracy.
- Applied emulator in the Ca isotopic chain to find new constraints on interaction parameters.

… and Outlook

- Use new emulator to quantify uncertainty of many-observables.
- Develop a unified embedding for the nuclear interactions so that they can all be included together.

Thank you!

Summary …

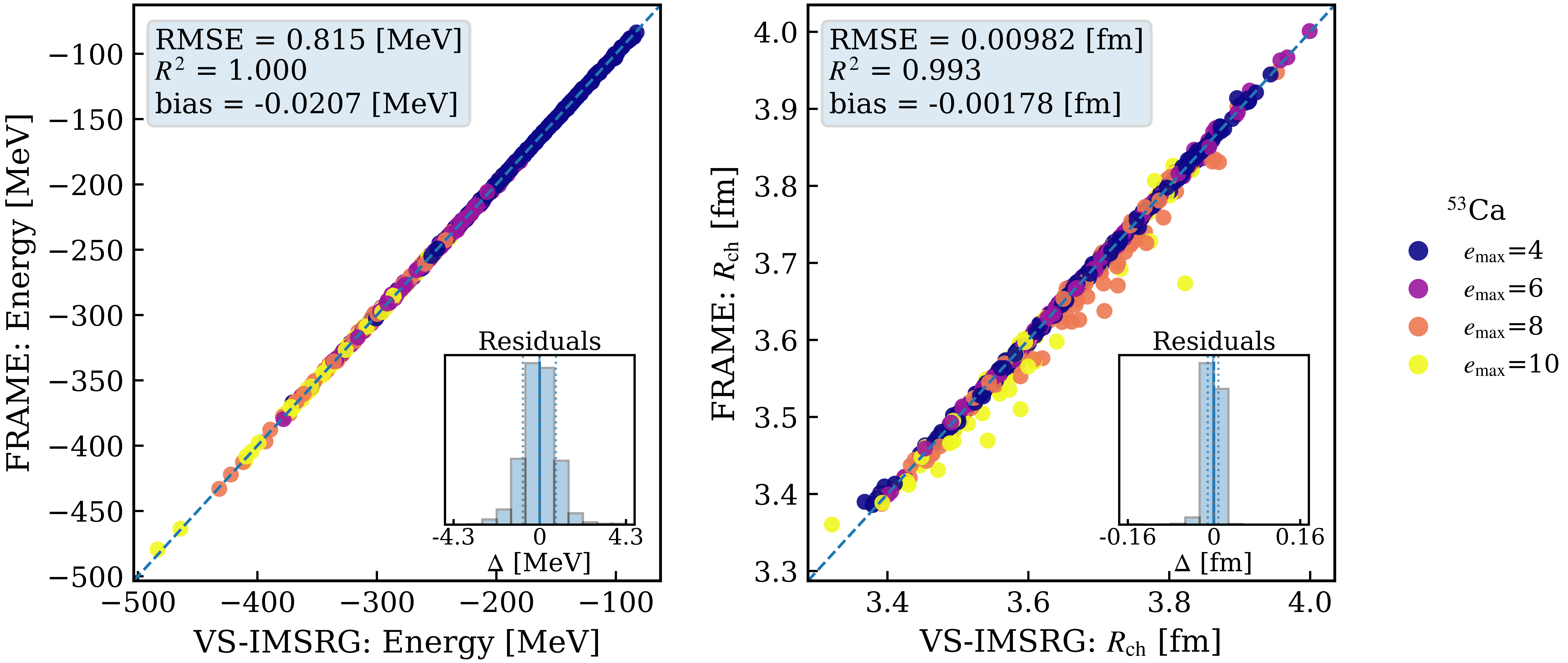
- Created new ML emulator: speed up calculations by a factor of $\sim 10^5$. 
- Improved emulator using physics driven technology.
- Can predict IMSRG(3f2) from IMSRG(2) with sub-percent accuracy.
- Applied emulator in the Ca isotopic chain to find new constraints on interaction parameters.

… and Outlook

- Use new emulator to quantify uncertainty of many-observables.
- Develop a unified embedding for the nuclear interactions so that they can all be included together.
- Use emulator to get a better understanding of the correlations between different nuclear observables in different nuclei and how they constrain the nuclear forces.

Thank you!

Extrapolation



Correlation between observables