Tools for treating model uncertainty in *ab initio* calculations of nuclei

Daniel Phillips
Ohio University

Research supported by the DOE Office of Science, the NSF OAC, and a NSF FRHTP

Sources of uncertainty in ab initio calculations

- χEFT Hamiltonian and β-decay operator are only computed up to a finite order in the EFT: "truncation uncertainty" **model uncertainty**
- χEFT Hamiltonian (and operators) depend on LECs (parameters) that must be estimated from data of finite precision: "parametric uncertainty" **aleatoric uncertainty**
- Many-body methods used to compute finite nuclei using that Hamiltonian truncate the Hilbert space or otherwise approximate the problem: "many-body uncertainty" **model uncertainty**
- Emulators, matrix inversion, integrals, etc. only represent actual result of model at a finite precision. **model uncertainty**

I will not discuss this last source of uncertainty here

Abstraction

- Suppose I have a χEFT Hamiltonian and many-body method that predicts a set of outputs $y_j(\boldsymbol{\theta})$, j=1,…J, with $\boldsymbol{\theta}$ the LECs of the Hamiltonian
- We estimate ("fit") the LECs by using a finite number of outputs and comparing them to experimental data. Goal is $p(\boldsymbol{\theta}|y_1,..,y_N;M)$, $N < J$.
- But $y_{j,\text{exp}} = y_{j,\text{th}}(\boldsymbol{\theta}) + \delta y_{j,\text{th}} + \delta y_{j,\text{exp}}$ with $y_{j,\text{exp}}$ the central value measured by the experiment and each of $\delta y_{j,\text{th}}$ and $\delta y_{j,\text{exp}}$ a random variable distributed as, say, a Gaussian. (This statement, by the way, is how one works out the correct form of the likelihood $p(y_{1,\text{exp}},..,y_{N,\text{exp}}|\boldsymbol{\theta};M_{\text{theory}},M_{\text{stats}})$.)
- This makes $\boldsymbol{\theta}$ a random variable with its own probability distribution too: "Parametric uncertainty" and "model uncertainty" are both reflected in the posterior for $\boldsymbol{\theta}$.
- Then we wish to predict $y_{N+1,\text{exp}}$. But $y_{N+1,\text{exp}} = y_{N+1,\text{th}}(\boldsymbol{\theta}) + \delta y_{N+1,\text{th}}$. Need to propagate parametric uncertainty (sample $p(\boldsymbol{\theta}|y_1,..,y_N;M)$ and do forward evaluations) *and* account for correlations $\boldsymbol{\theta}$ has (or doesn't have) with this new observable's model uncertainty, $\delta y_{N+1,\text{th}}$.
- When this is all done we have the result from one "model": $p(y_{N+1}|M)$.
- Maybe we also want to assess other kinds of model uncertainty, not included in δy_{th}. Then we could combine probability distributions from several MB methods: $p(y_{N+1}) = \sum_k w_k p(y_{N+1}\,|\,M_k)$.

Structure of talk

- Introduction
- χEFT discrepancy modeling: estimating 3NF parameters
- Predictions with truncation errors in a toy model
- Building a discrepancy model for NN observables
  - Which potentials does the BUQEYE™ model actually work for?
  - Building physics into the length scale
- Model mixing
  - Combining models using weights
  - How to assess model correlations
- A few words about experimental design & a summary

Structure of talk

- Introduction
- χEFT discrepancy modeling: estimating 3NF parameters

  More accurate parameter posteriors, AND you can infer Q
- Predictions with truncation errors in a toy model
- Building a discrepancy model for NN observables
  - Which potentials does the BUQEYE™ model actually work for?
  - Building physics into the length scale
- Model mixing
  - Combining models using weights
  - How to assess model correlations
- A few words about experimental design & a summary

Structure of talk

- Introduction
- χEFT discrepancy modeling: estimating 3NF parameters

  More accurate parameter posteriors, AND you can infer Q
- Predictions with truncation errors in a toy model

  But it's easy to overdo it: correlations matter!
- Building a discrepancy model for NN observables
  - Which potentials does the BUQEYE™ model actually work for?
  - Building physics into the length scale
- Model mixing
  - Combining models using weights
  - How to assess model correlations
- A few words about experimental design & a summary

Structure of talk

- Introduction
- χEFT discrepancy modeling: estimating 3NF parameters

  More accurate parameter posteriors, AND you can infer Q
- Predictions with truncation errors in a toy model

  But it's easy to overdo it: correlations matter!
- Building a discrepancy model for NN observables
  - Which potentials does the BUQEYE™ model actually work for?
  - Building physics into the length scale

  Truncation errors for NN observables can be subtle
- Model mixing
  - Combining models using weights
  - How to assess model correlations
- A few words about experimental design & a summary

Structure of talk

- Introduction
- χEFT discrepancy modeling: estimating 3NF parameters

More accurate parameter posteriors, AND you can infer Q

- Predictions with truncation errors in a toy model

But it's easy to overdo it: correlations matter!

- Building a discrepancy model for NN observables
  - Which potentials does the BUQEYE™ model actually work for?
  - Building physics into the length scale

Truncation errors for NN observables can be subtle

- Model mixing
  - Combining models using weights
  - How to assess model correlations

Model mixing is a frontier

- A few words about experimental design & a summary

Structure of talk

- Introduction
- χEFT discrepancy modeling: estimating 3NF parameters

More accurate parameter posteriors, AND you can infer Q

- Predictions with truncation errors in a toy model

But it's easy to overdo it: correlations matter!

- Building a discrepancy model for NN observables
  - Which potentials does the BUQEYE™ model actually work for?
  - Building physics into the length scale

Truncation errors for NN observables can be subtle

- Model mixing
  - Combining models using weights
  - How to assess model correlations

Model mixing is a frontier

Correlations matter!

- A few words about experimental design & a summary

χEFT to NNLO: error model and strategy

$$y_{\rm exp} = y_{\rm th}(\mathbf{a}_{3N}, \mathbf{a}_{NN}) + \delta y_{\rm th} + \delta y_{\rm exp}$$

$$Q = \frac{p, m_\pi}{\Lambda_b}$$

$\delta y_{\rm exp}$: normally distributed, uncorrelated errors (?)

$$y_{\rm th}(p) = y_{\rm ref}(p) \sum_{i=0}^{3} c_i(\{a_j\}) Q^i$$

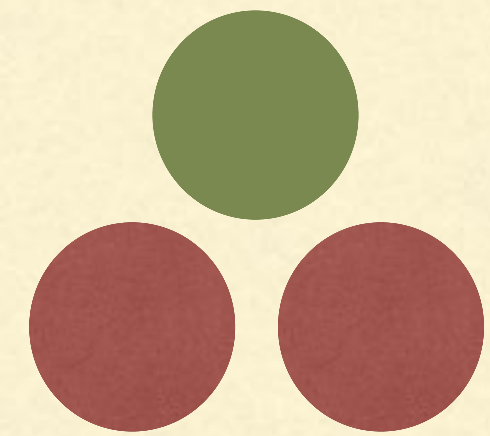
$$\delta y_{\rm th} = y_{\rm ref}(p)[c_{k+1} Q^{k+1} + c_{k+2} Q^{k+2} + \ldots]$$

- Q is not obvious for bound state observables: make it a parameter & sample
- Also sample $\bar{c}^2$, the mean-square value of the higher-order coefficients
- $\bar{c}^2$ and Q are inferred from the NNLO-NLO shift, but that result is modified because they form the (dominant piece of) the theory error in the likelihood
- NN force LECs $\mathbf{a}_{NN}$ refit (πN LECs from Roy-Steiner analysis). Use as prior on higher-body observables, so $\mathbf{a}_{NN}$ will get updated if they are sensitive to it

Example: 3N bound-state observables

Wesolowski, Svennson, Ekström, Forssén, Furnstahl, Melendez, DP, Phys. Rev. C (2022)

- Binding energy of three-nucleon nuclei: ^{3}H
- Binding energy of ^{4}He
- Charge radius of ^{4}He
- Beta-decay half-life of ^{3}H, aka "GT matrix element"



Solve Schrödinger equation for ^{3}He and ^{4}He and compute radii, GT matrix element

Done at $O(Q^0)$, $O(Q^2)$, $O(Q^3)$

Emulation via Eigenvector Continuation make fast evaluation possible

Posterior and priors

$$\text{pr}(\mathbf{a}, \bar{c}^2, Q \,|\, D, I) \propto \exp\left(-\frac{1}{2}\mathbf{r}^T(\mathbf{\Sigma}_{\text{exp}} + \mathbf{\Sigma}_{\text{th}})^{-1}\mathbf{r}\right) \exp\left(-\frac{\mathbf{a}^2}{2\bar{a}^2}\right) \text{pr}(\bar{c}^2 \,|\, Q, \bar{a}, I)\text{pr}(Q \,|\, \mathbf{a}, I)$$

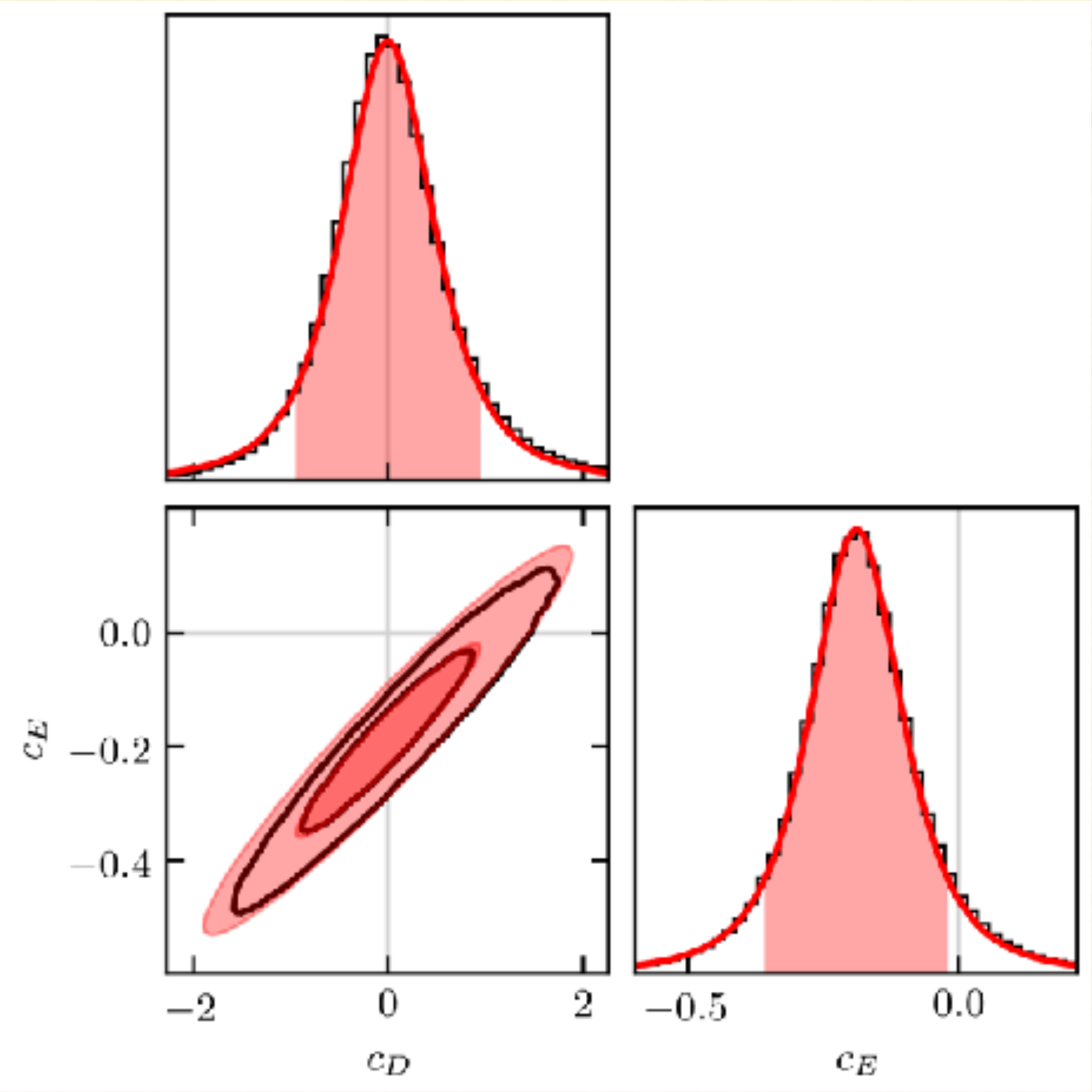
Truncation errors

$$\mathbf{r} = \mathbf{y}_{\text{exp}} - \mathbf{y}_{\text{th}}$$

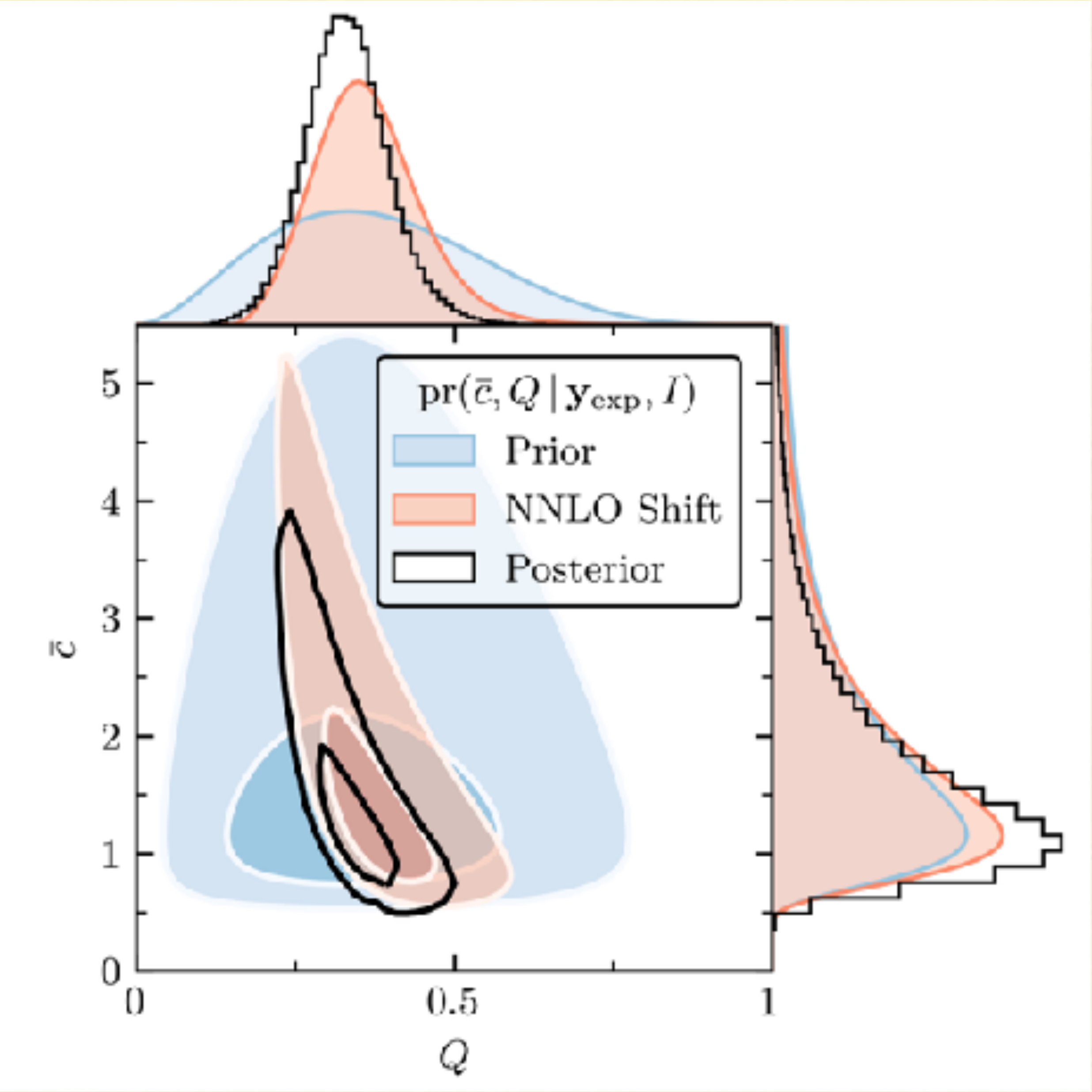
Naturalness

- We take uncorrelated theory errors: $(\mathbf{\Sigma}_{\text{th,uncorr}})_{ij} = (\mathbf{y}_{\text{ref}})^2 \bar{c}^2 \delta_{ij} \sum_{n=k+1}^{\infty} Q^{2n}$
- Experimental errors are negligible in comparison
- $\text{pr}(\bar{c}^2 \,|\, Q, \vec{a}, I)$ is taken to be an inverse-χ^2 distribution. Information on the order-to-order shifts included there
- $\text{pr}(Q \,|\, \mathbf{a}, I)$ then also affected by that information. Starts as weakly informative Beta distribution before any updating from NLO-LO and NNLO-NLO shifts

Results for 3NF parameters, Q, $\bar{c}^2$

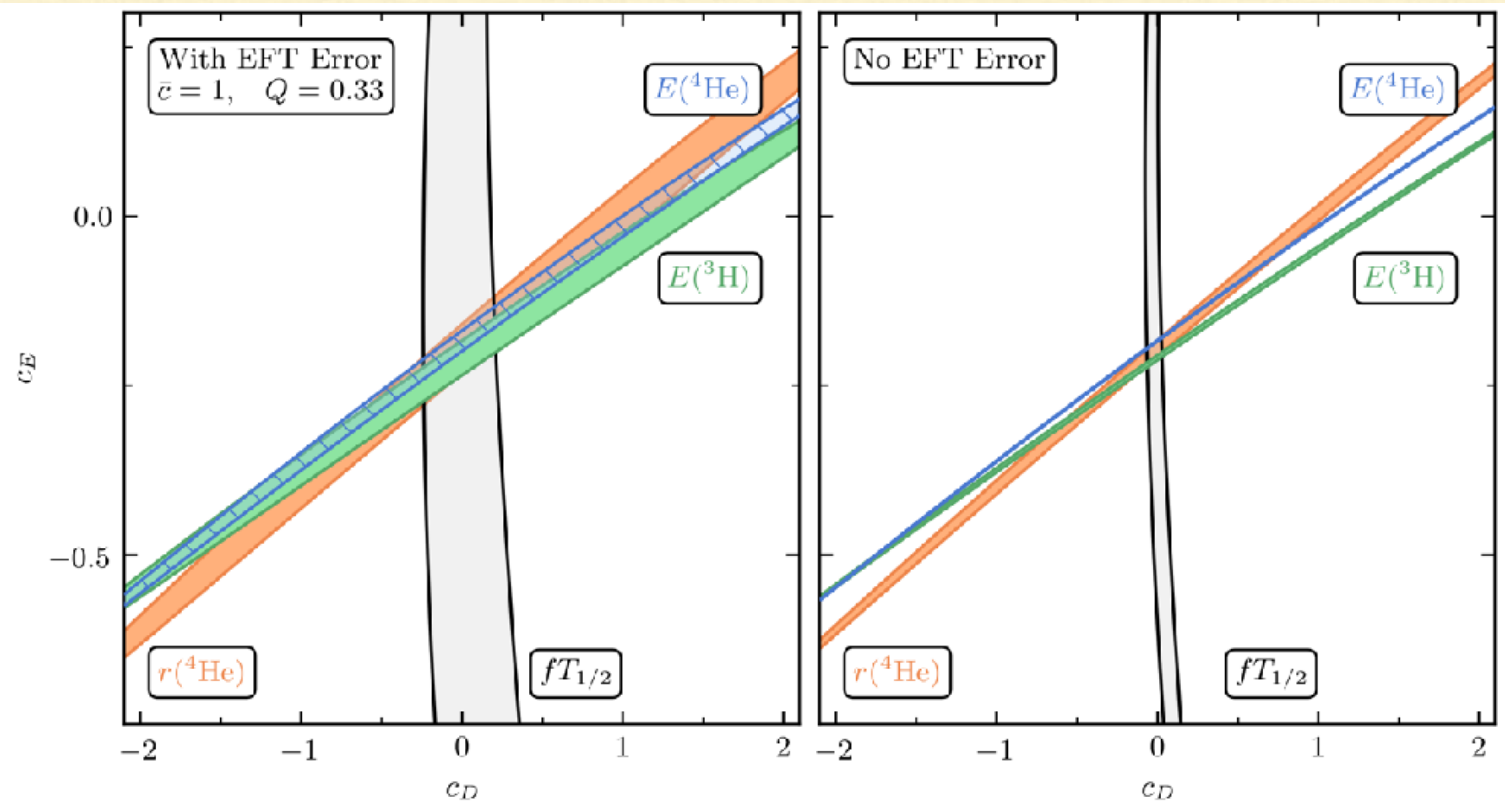


t distributions!



Q inferred from data, convergence pattern

The role of different constraints and of truncation errors



From LECs to prediction in a Toy Model

Schindler, DRP, Ann. Phys., 2009

Wesolowski, Klco, Furnstahl, Phillips, Thapaliya, JPG, 2016

- Given data D={(d_k,σ_k):k=1,...,N} taken at points x_k and a fit function f(x;**a**) that depends on LECs **a**={a_0,…,a_{kmax}}, determine the first k+1
- BUT, be careful! f only describes data in a limited domain

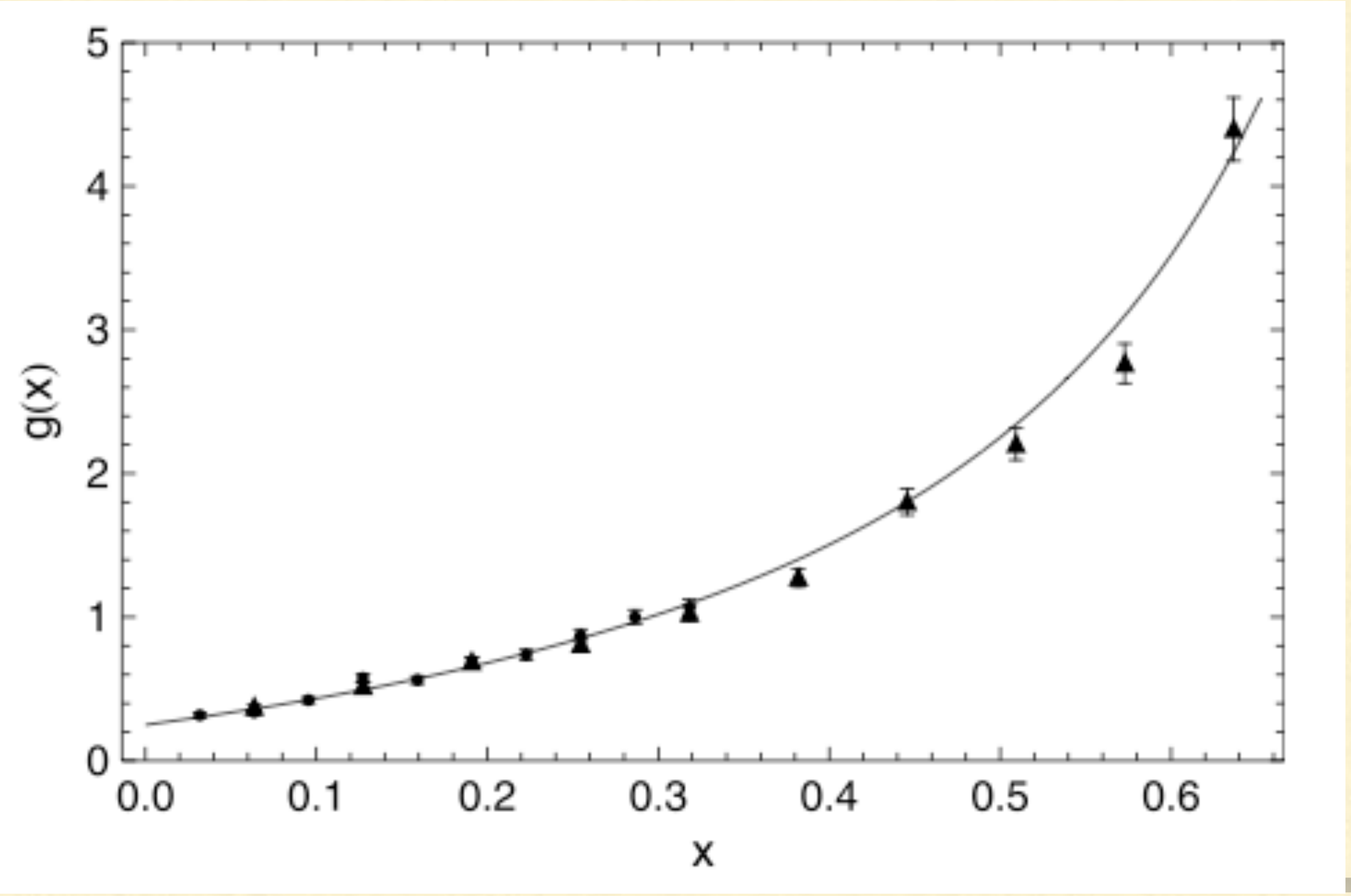
From LECs to prediction in a Toy Model

Schindler, DRP, Ann. Phys., 2009

Wesolowski, Klco, Furnstahl, Phillips, Thapaliya, JPG, 2016

- Given data D={(d_k,σ_k):k=1,...,N} taken at points x_k and a fit function f(x;**a**) that depends on LECs **a**={a_0,…,a_{kmax}}, determine the first k+1
- BUT, be careful! f only describes data in a limited domain

Fit function: $f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_k x^k$



Toy example: data from

$g(x) = (1/2 + \tan(\pi x/2))^2$

From LECs to prediction in a Toy Model

Schindler, DRP, Ann. Phys., 2009

Wesolowski, Klco, Furnstahl, Phillips, Thapaliya, JPG, 2016

- Given data D={(d_k,σ_k):k=1,...,N} taken at points x_k and a fit function f(x;**a**) that depends on LECs **a**={a_0,…,a_{kmax}}, determine the first k+1
- BUT, be careful! f only describes data in a limited domain

Fit function: $f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_k x^k$

Toy function with data D1 and first terms in expansion

Toy example: data from
$g(x) = (1/2 + \tan(\pi x/2))^2$

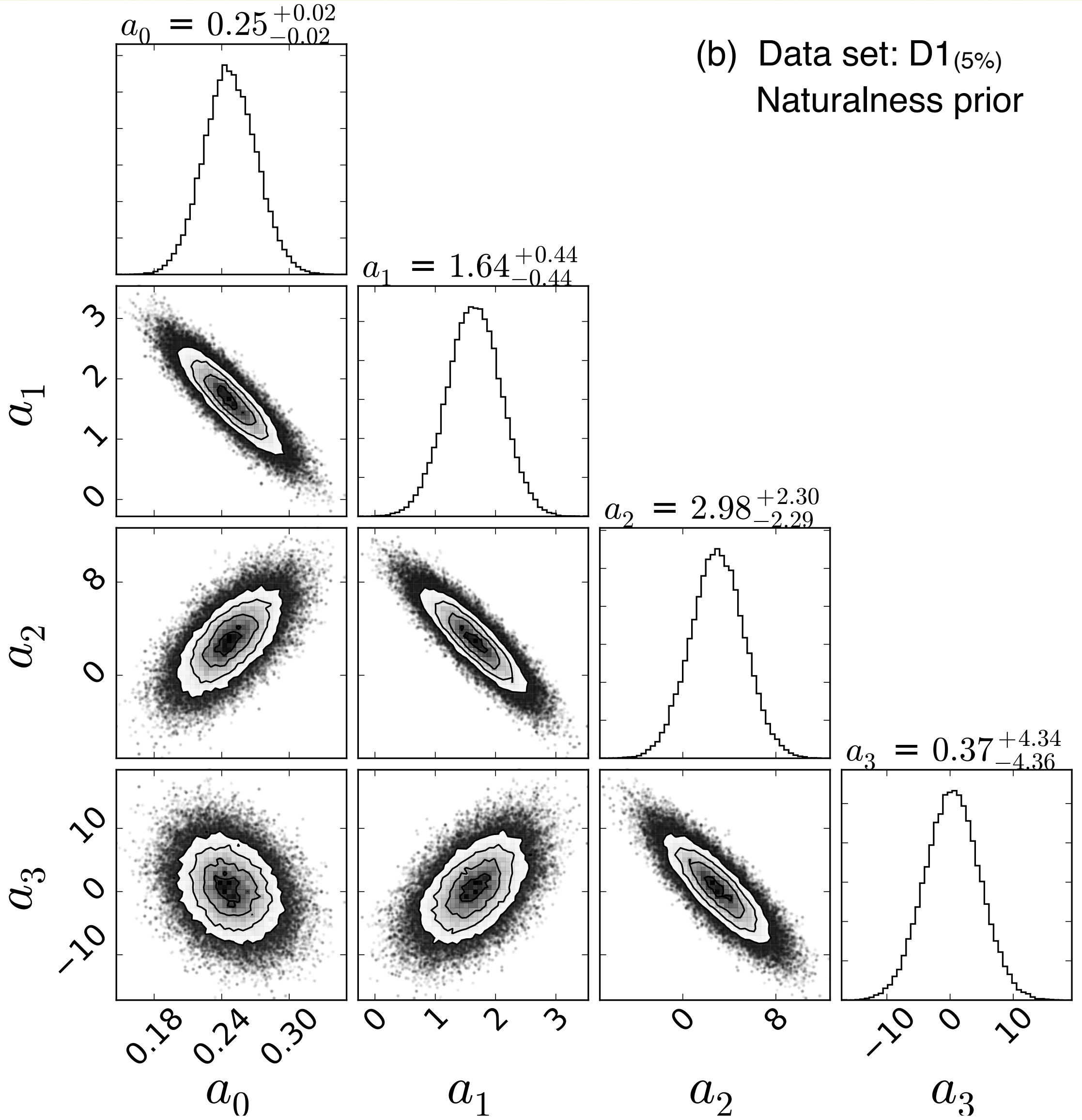
$g(x) = 0.25 + 1.57x + 2.47x^2 + \ldots$

Parameter estimation

$$\mathrm{pr}(\mathbf{a}\,|\,D,k,k_{\max}) \propto \exp\left(-\frac{1}{2}\mathbf{r}^T(\boldsymbol{\Sigma}_{\exp})^{-1}\mathbf{r}\right)\exp\left(-\frac{\mathbf{a}^2}{2\bar{a}^2}\right) \qquad \mathbf{r} = \mathbf{y}_{\exp} - \mathbf{y}_{\mathrm{th}}$$

→equivalent to likelihood with Σ_{th} after marginalization over $\{a_{k+1}, a_{k+2}, \ldots, a_{k_{\max}}\}$

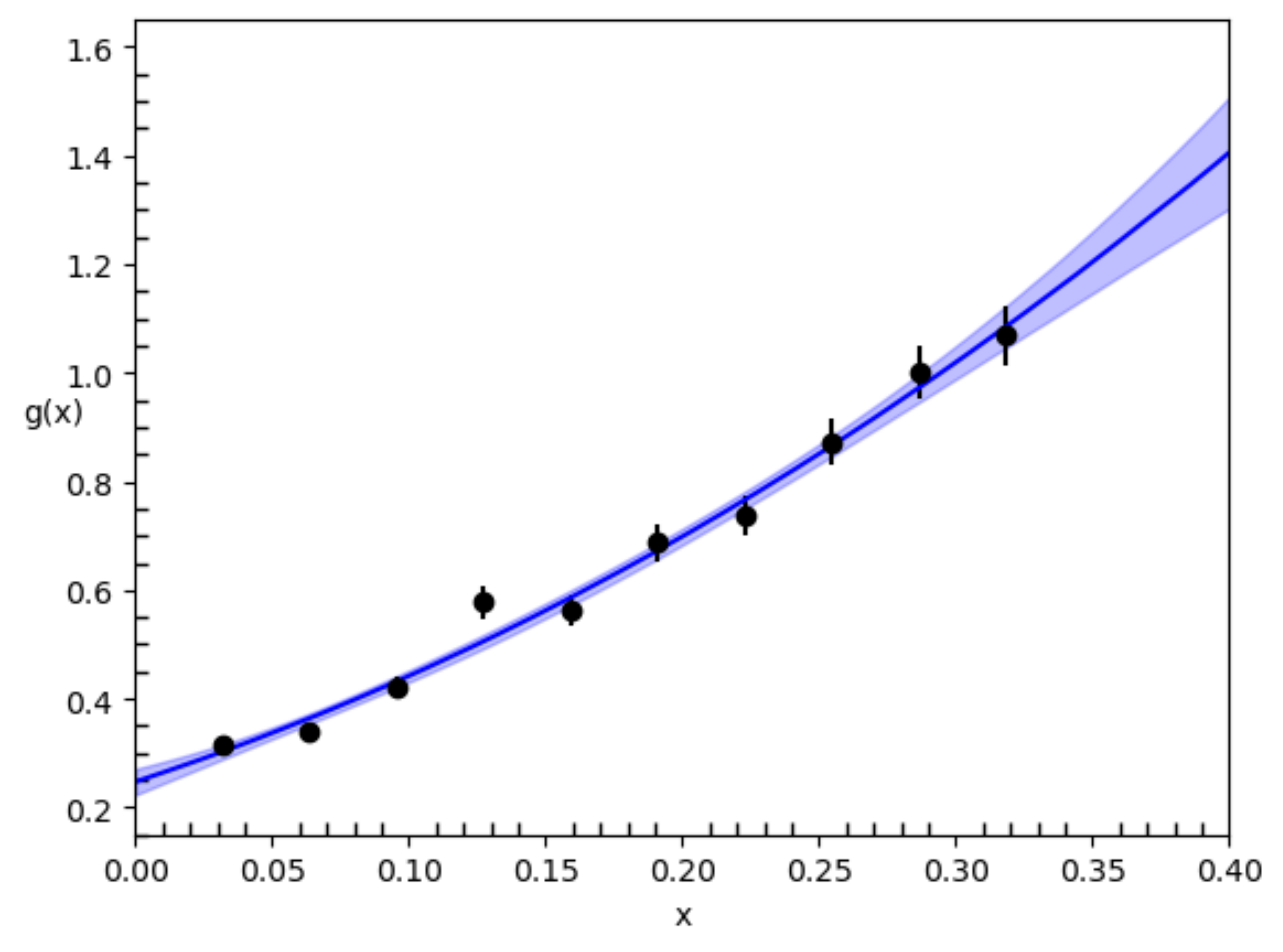
		Gaussian prior			
k	$k_{\max}$	Evidence	a_0	a_1	a_2
0	0	~ 0	0.48±0.01		
1	1	6.0×10^2	0.20±0.01	2.6±0.1	
2	2	3.3×10^3	0.25±0.02	1.6±0.4	3.1±1
2	3	2.9×10^3	0.25±0.02	1.7±0.5	3.0±2
2	4	2.8×10^3	0.25±0.02	1.7±0.5	3.0±2
2	5	2.8×10^3	0.25±0.02	1.7±0.5	3.0±2
2	6	2.8×10^3	0.25±0.02	1.7±0.5	3.0±2
True			0.25	1.57	2.47



Prediction

$$g(x) = f_k(x;\boldsymbol{\theta}) + \delta y(x)$$

- Use the posterior at third order for **θ**, fit simultaneously with a fourth-order "model discrepancy", a_4x^4, to obtain the posterior for $f_k(x;\boldsymbol{\theta}) + \delta y(x)$

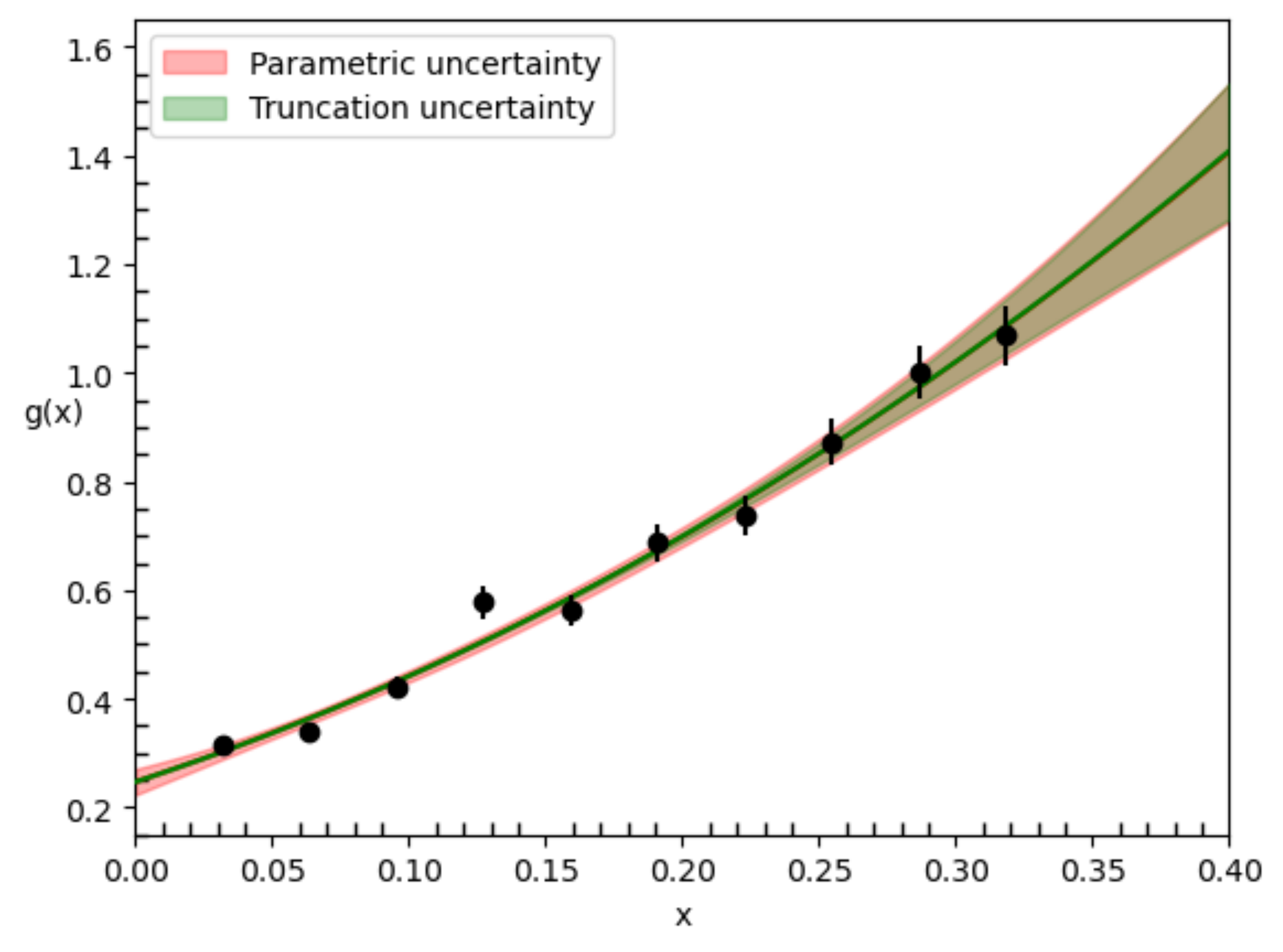


Prediction

$$g(x) = f_k(x;\boldsymbol{\theta}) + \delta y(x)$$

- Use the posterior at third order for **θ**, fit simultaneously with a fourth-order "model discrepancy", a_4x^4, to obtain the posterior for $f_k(x;\boldsymbol{\theta}) + \delta y(x)$

- But the parametric & model errors are each as large, or larger, than the total uncertainty



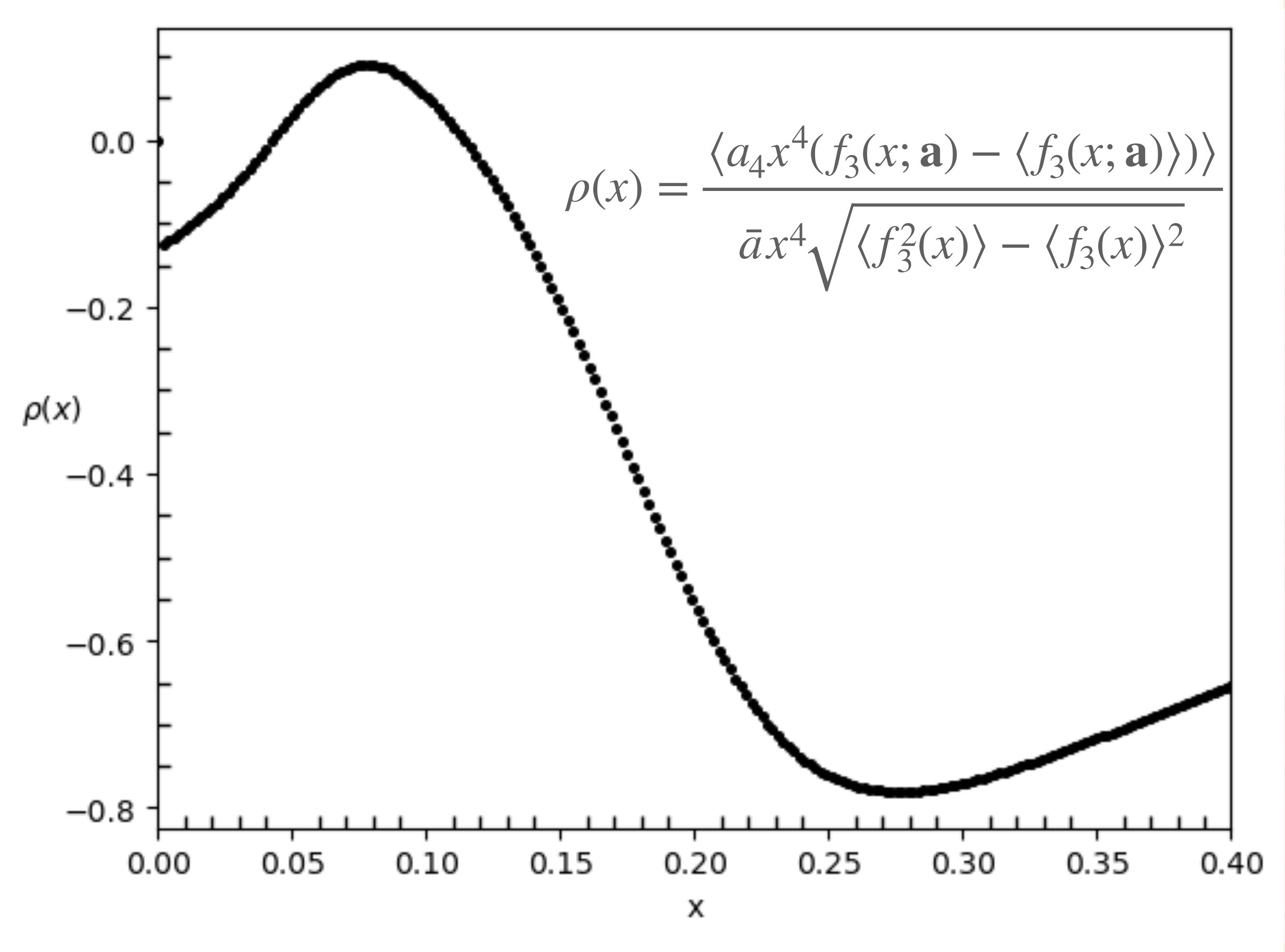
Prediction

$$g(x) = f_k(x;\boldsymbol{\theta}) + \delta y(x)$$

- Use the posterior at third order for **θ**, fit simultaneously with a fourth-order "model discrepancy", a_4x^4, to obtain the posterior for $f_k(x;\boldsymbol{\theta}) + \delta y(x)$

- But the parametric & model errors are each as large, or larger, than the total uncertainty

- Because they're anti-correlated



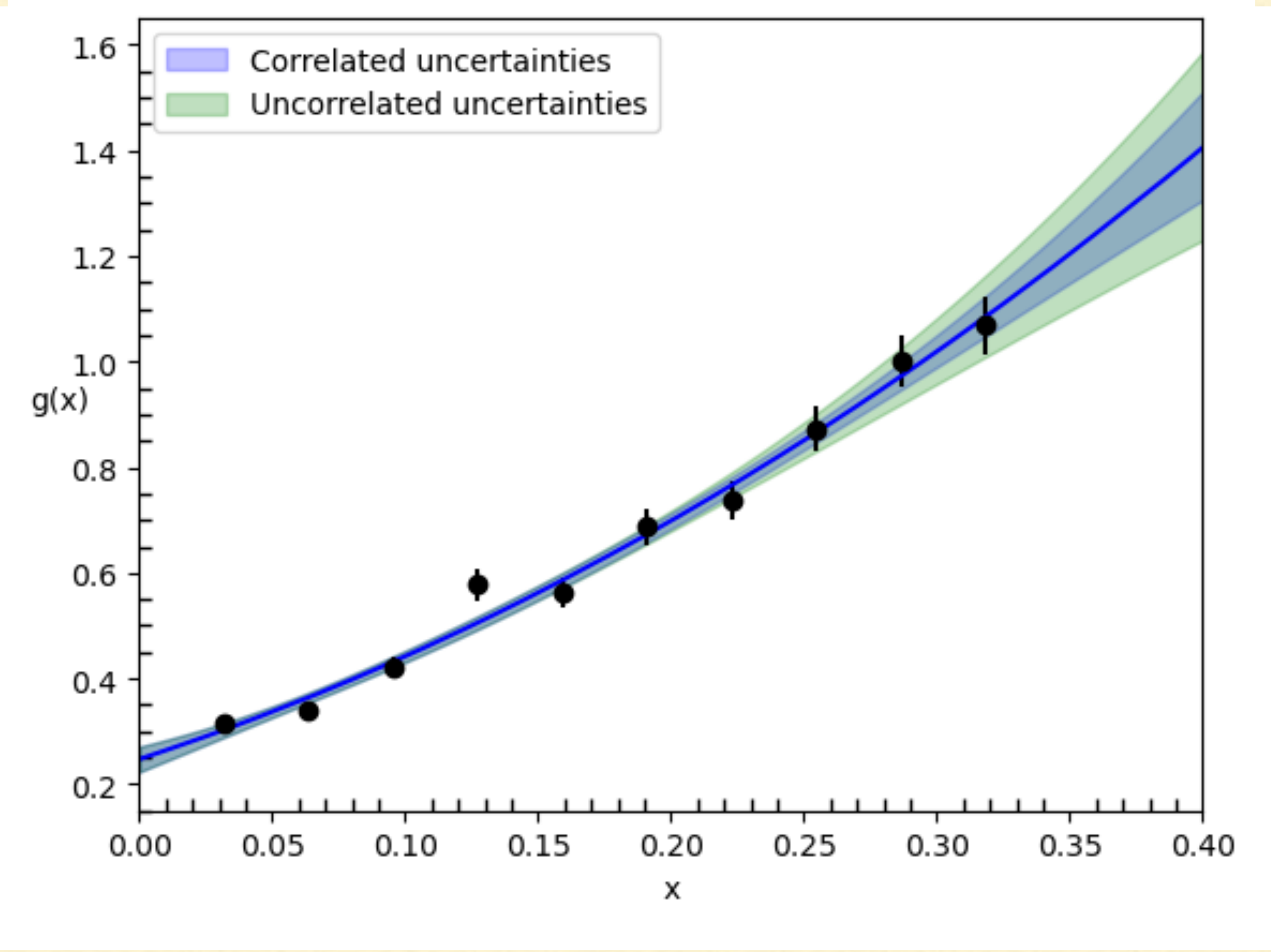
Prediction

Carter, Furnstahl, Melendez, DP, to appear

$$g(x) = f_k(x; \boldsymbol{\theta}) + \delta y(x)$$

- Use the posterior at third order for **θ**, fit simultaneously with a fourth-order "model discrepancy", a_4x^4, to obtain the posterior for $f_k(x; \boldsymbol{\theta}) + \delta y(x)$

- But the parametric & model errors are each as large, or larger, than the total uncertainty
- Because they're anti-correlated
- Neglect that and you overestimate the uncertainty of your prediction



Modeling correlations in theory uncertainties

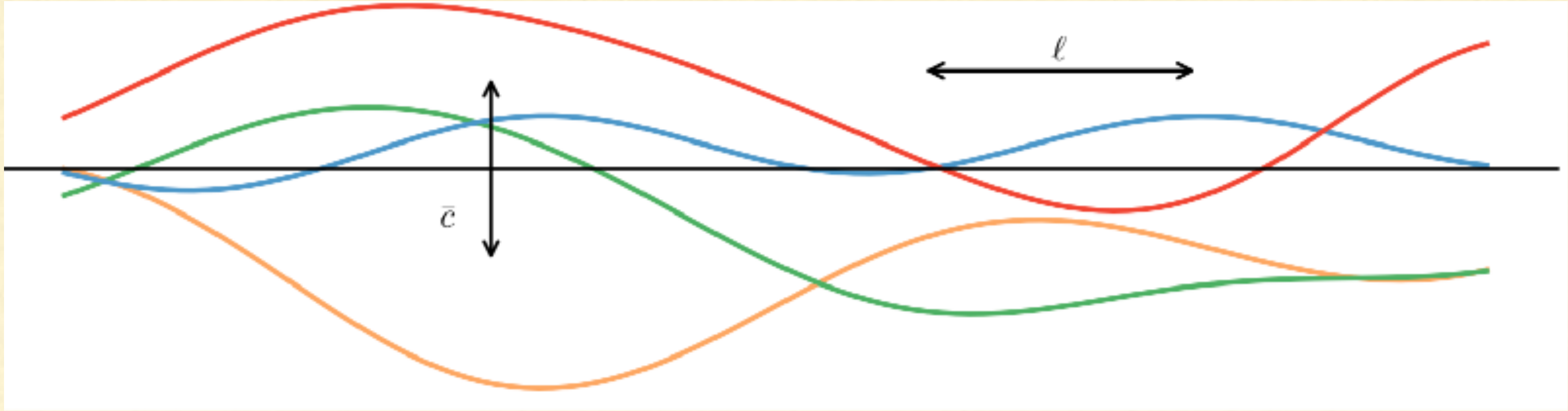
Melendez, Wesolowski, Furnstahl, DP, Pratola, PRC (2019)

$$y = y_{\mathrm{ref}} \sum_{n=0}^{k} c_n(p/m_\pi) Q^n$$

Function c_n is not a constant.
But the c_n's at different values of p aren't independent random variables either

Our hypothesis:

EFT coefficients at different orders can be modeled as independent draws from a *Gaussian Process* with a stationary kernel

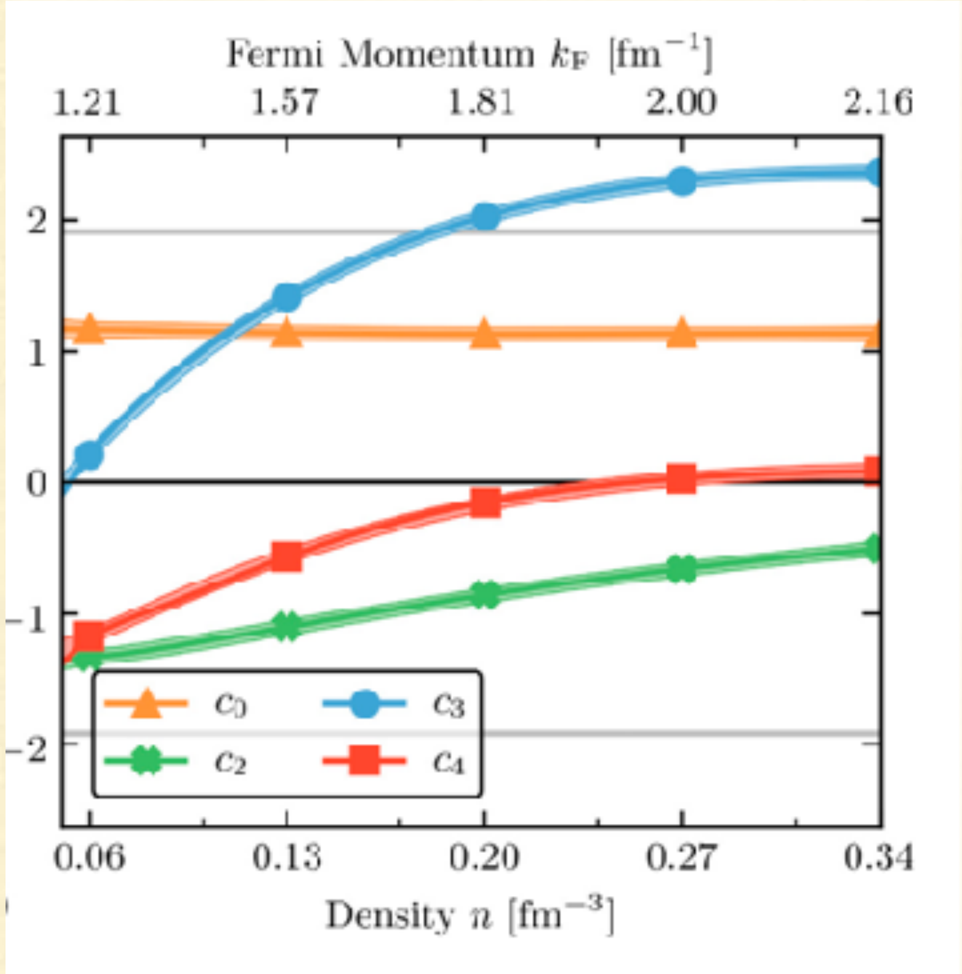


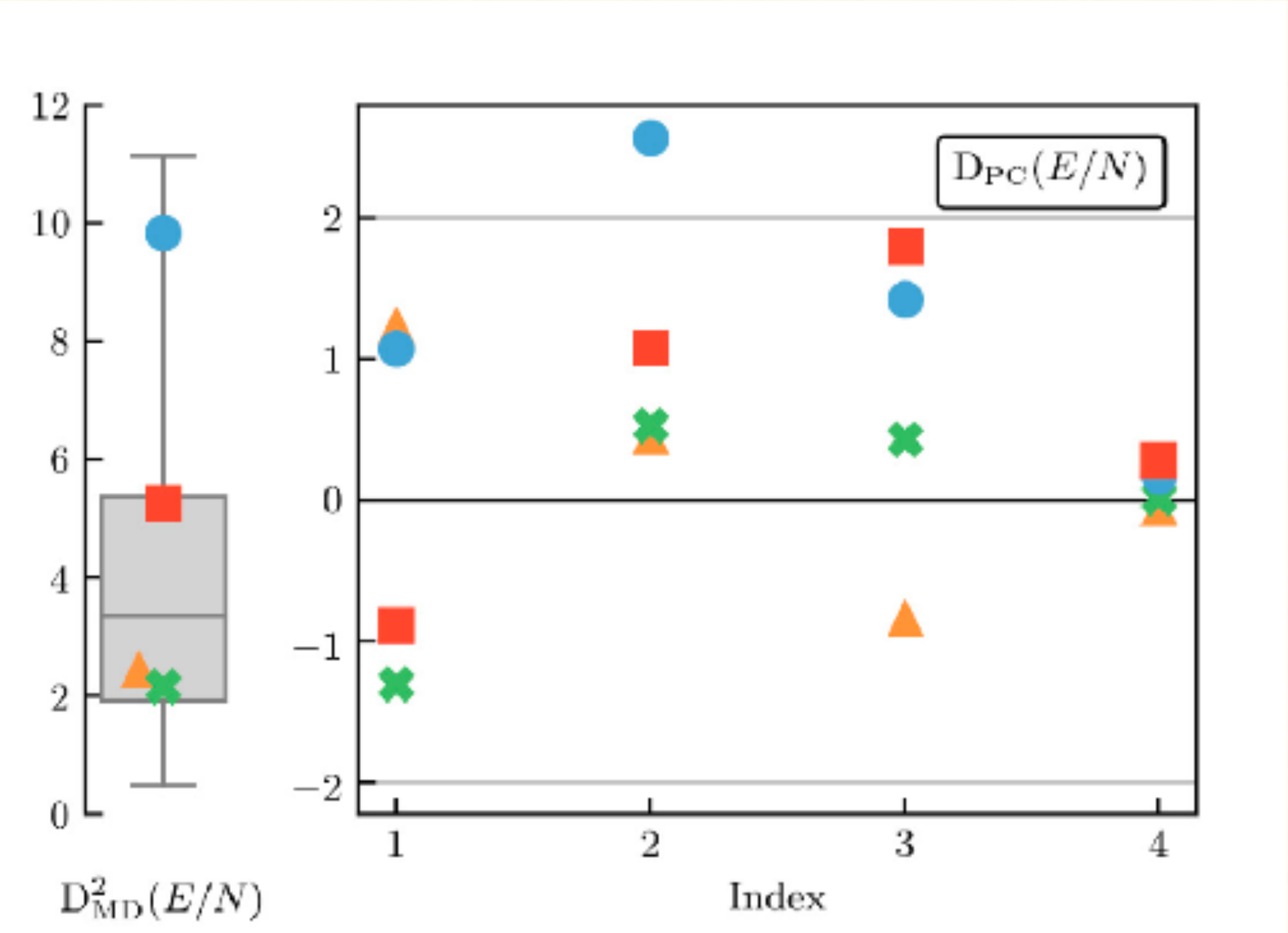
- Gaussian distribution at each point
- With correlation structure parameterized by a single $\bar{c}^2$ and ℓ at all orders

Example: E/N for pure neutron matter

Drischler, Melendez, Furnstahl, DP, PRL, PRC (2020)

- Order-by-order uncertainties for pure neutron matter
- Obtained by applying BUQEYE™ approach to truncation errors

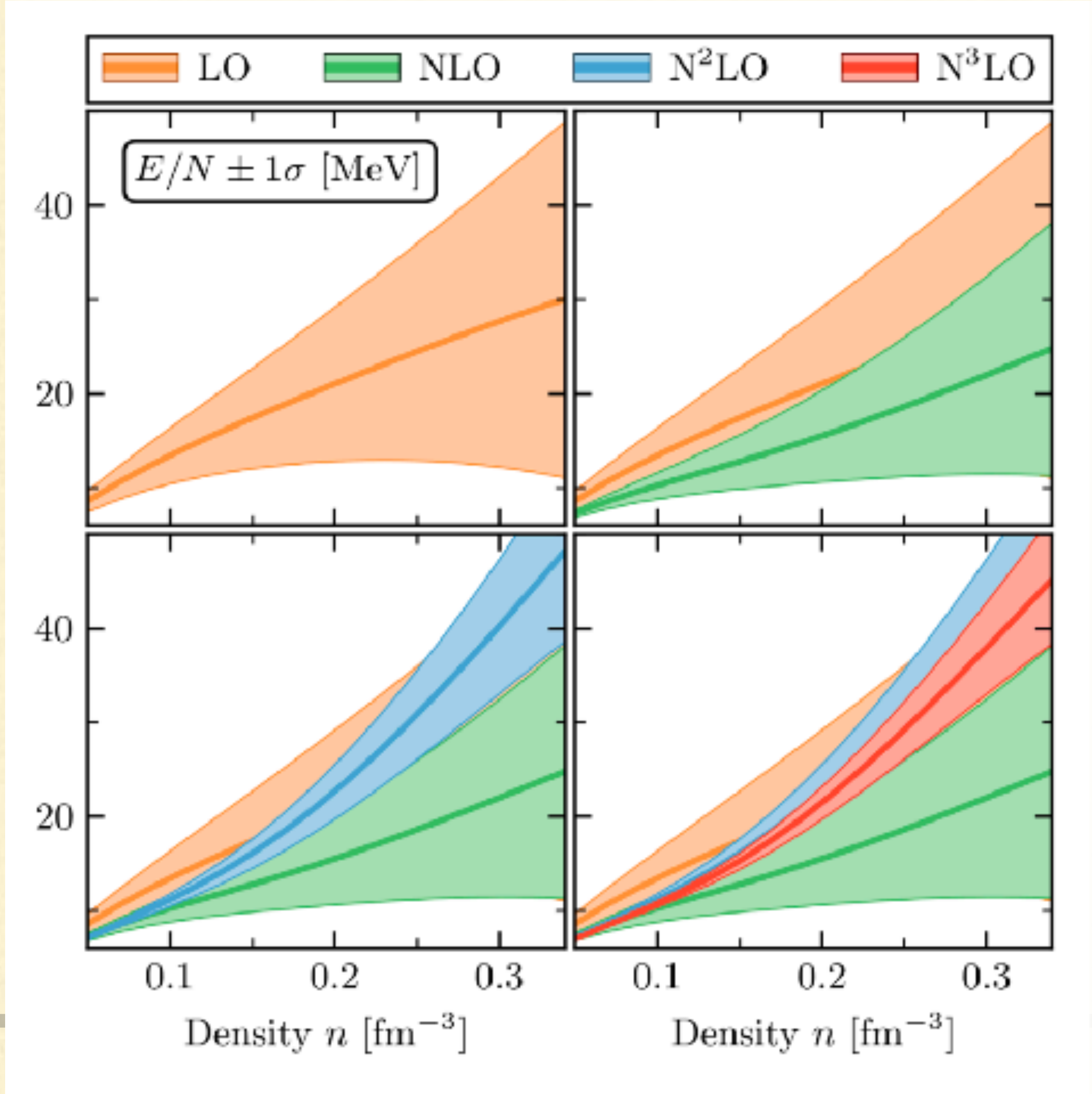




Example: E/N for pure neutron matter

Drischler, Melendez, Furnstahl, DP, PRL, PRC (2020)

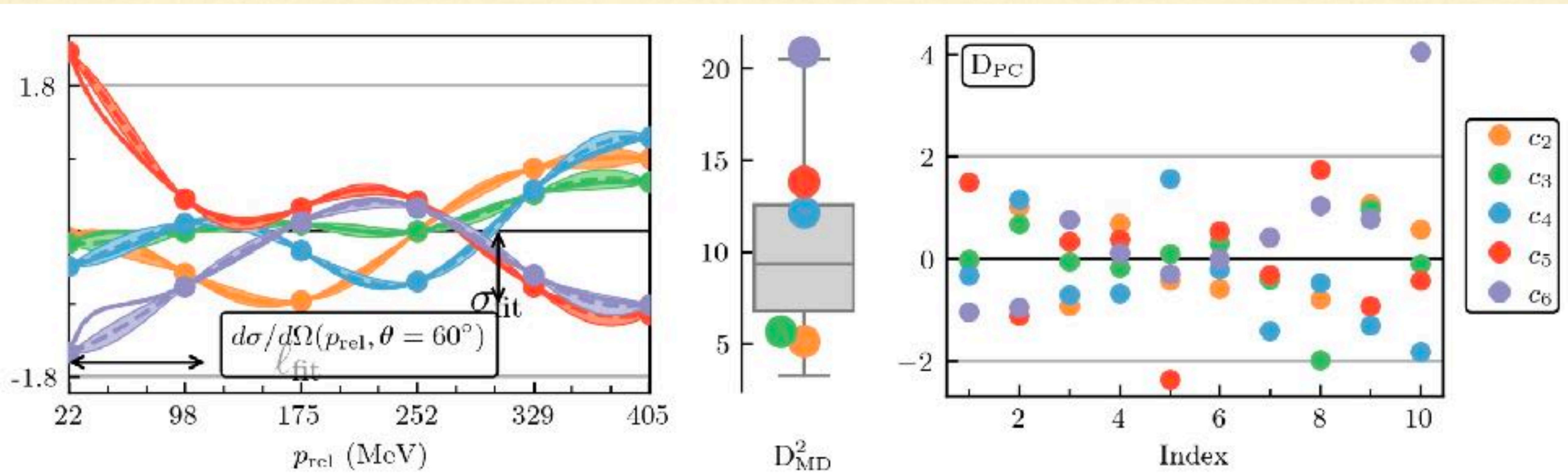
- Order-by-order uncertainties for pure neutron matter
- Obtained by applying BUQEYE™ approach to truncation errors



Example: NN differential cross section

Millican, Furnstahl, Melendez, DP, Pratola (2024)

NN differential cross section at 60° for SMS 500 MeV

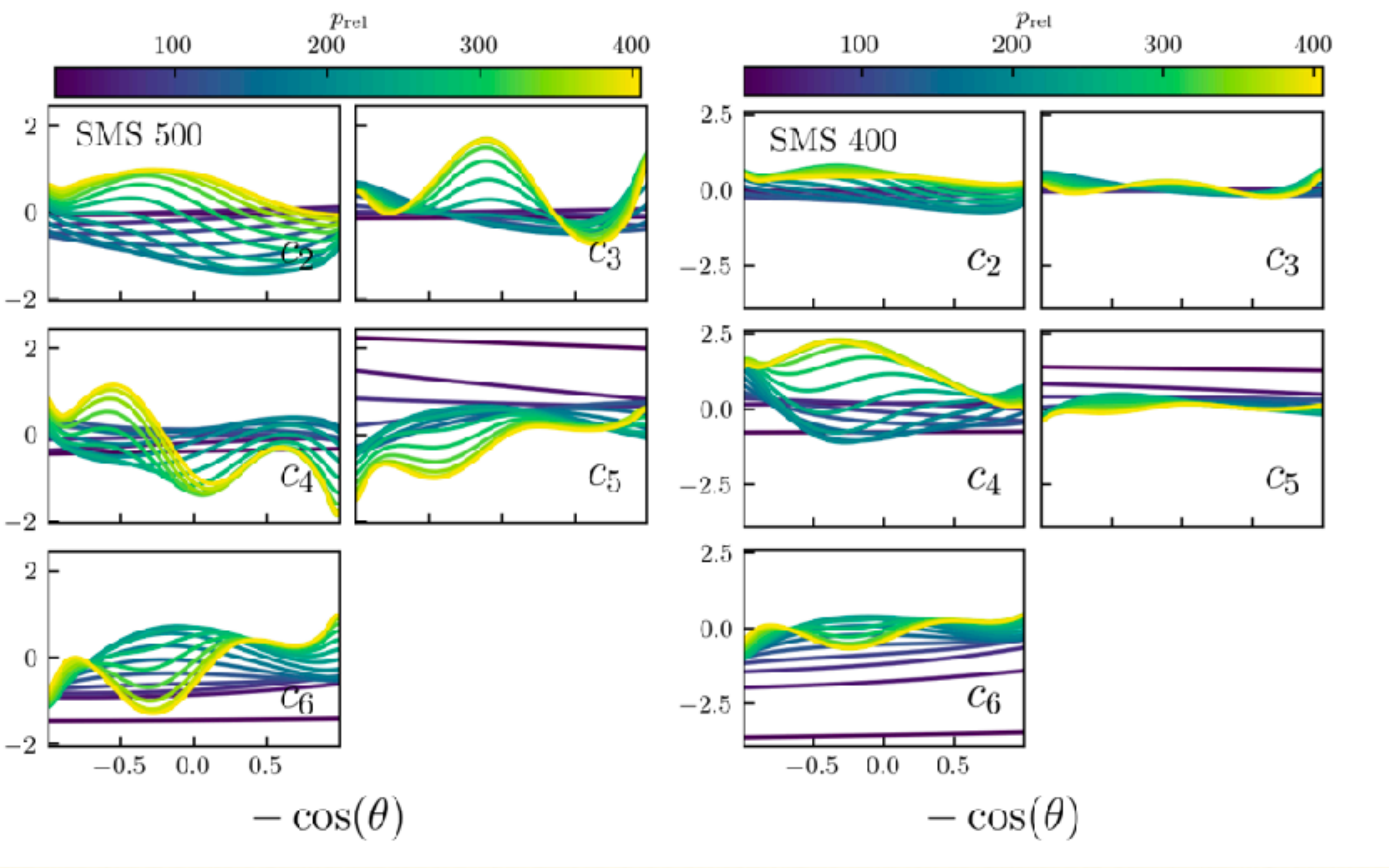


$m_\pi^{\rm eff} = 138$ MeV, Λ$_b$=570 MeV, $Q = \dfrac{m_\pi^{\rm eff} + p}{m_\pi^{\rm eff} + \Lambda_b}$

But….

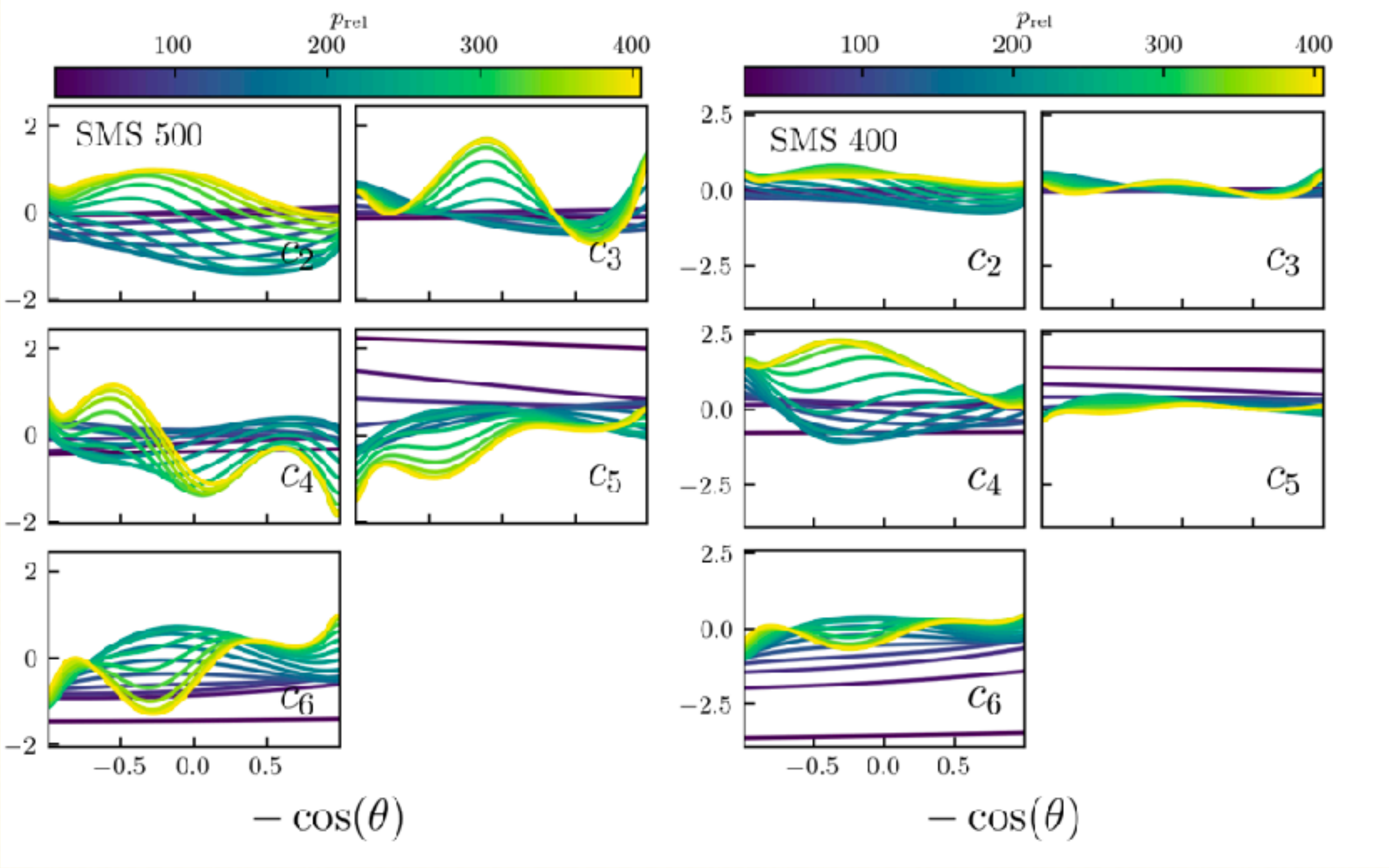
You can see a lot by looking

Millican et al., in preparation (2025)



You can see a lot by looking

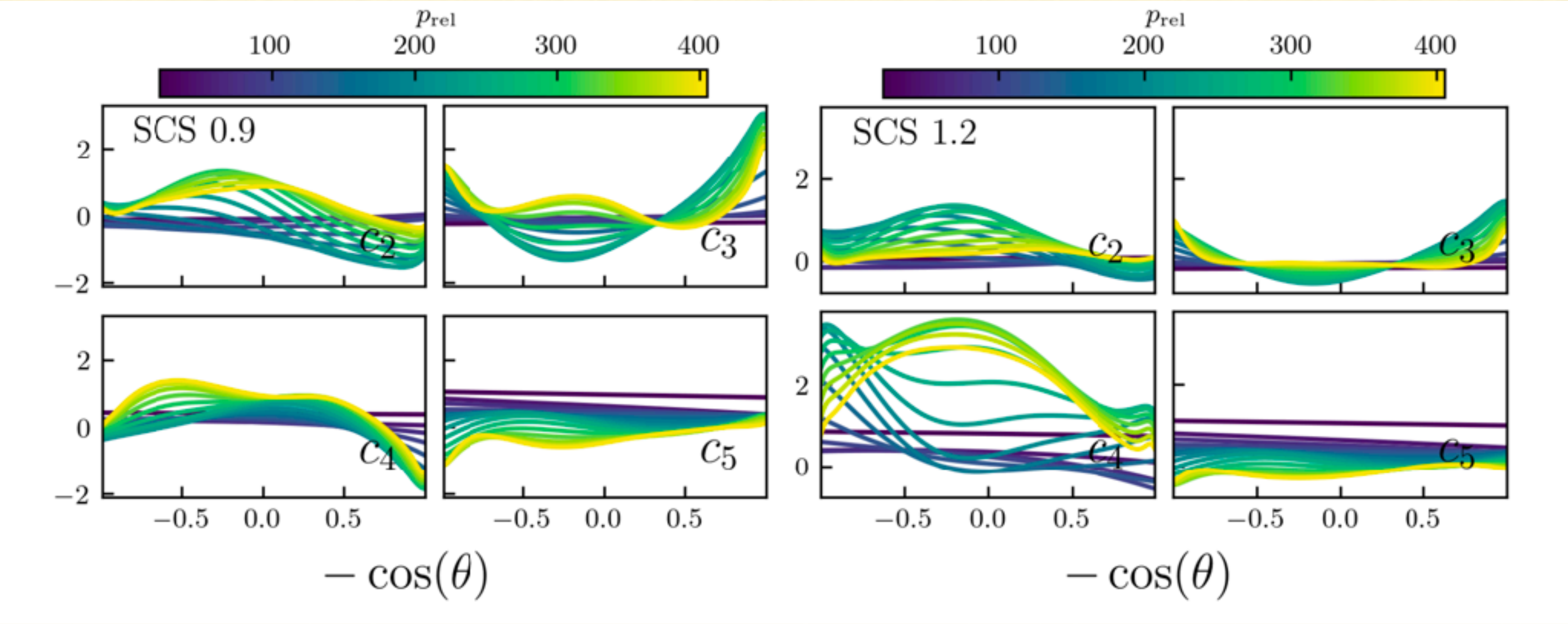
Millican et al., in preparation (2025)



Even/odd orders have different sizes for soft potentials

You can see a lot by looking

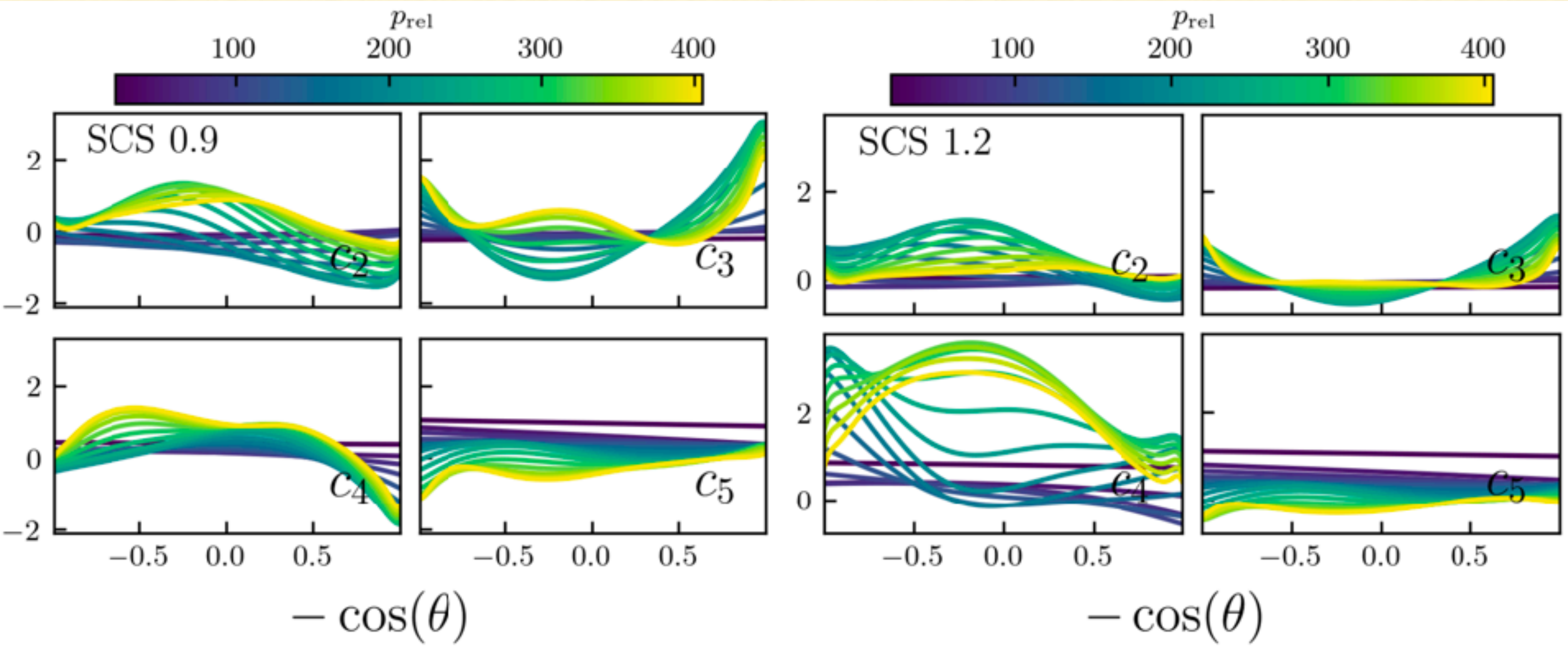
Millican et al., in preparation (2025)



Even/odd orders have different sizes for soft potentials

You can see a lot by looking

Millican et al., in preparation (2025)



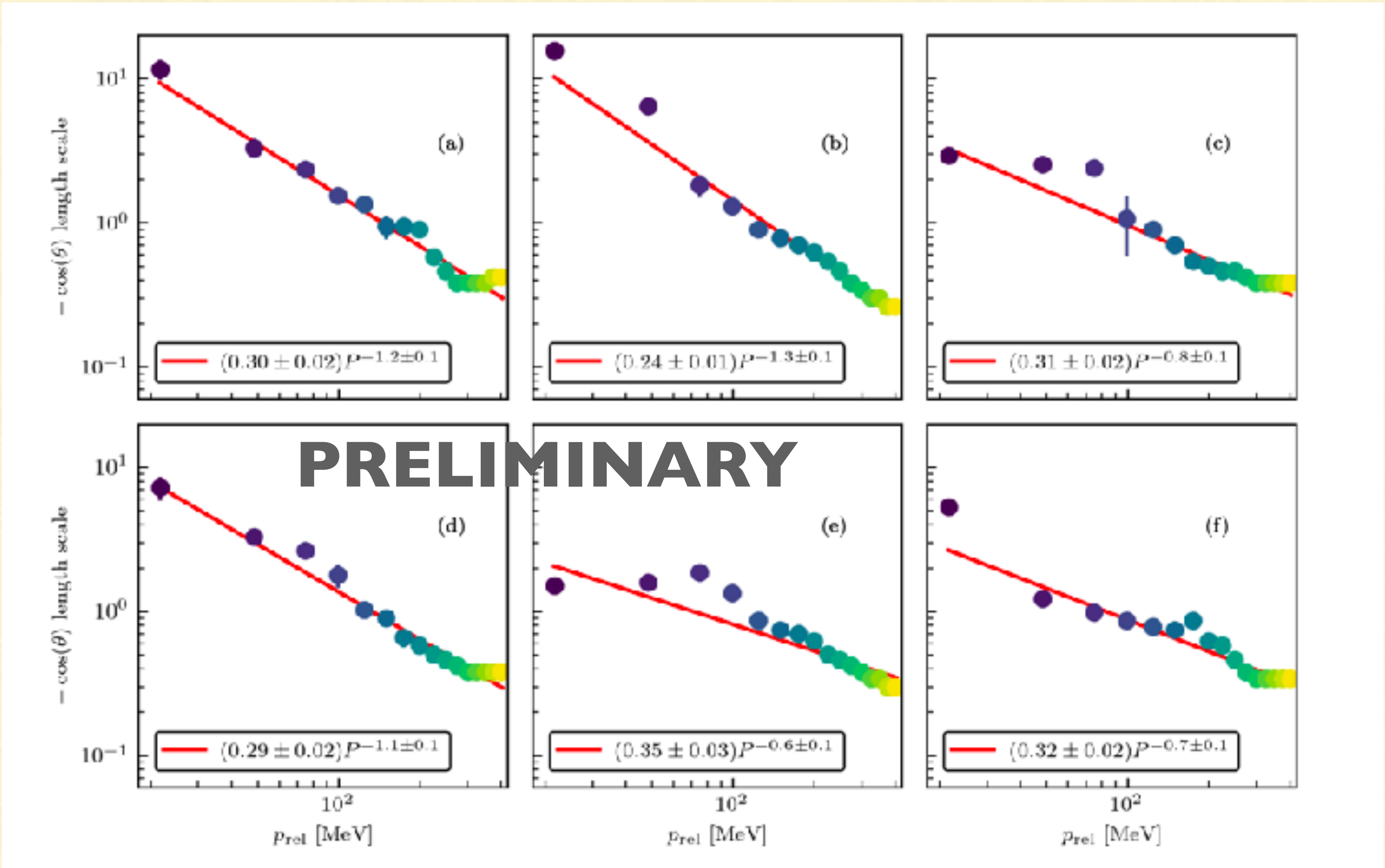
Length scale of curves gets shorter as momentum gets higher

Even/odd orders have different sizes for soft potentials

The GP is not 2D stationary in (p,θ_{cm})

Millican et al., in preparation (2025)

SMS 500 MeV



- $\ell_\theta \sim 1/p$ to within uncertainties
- "Warp" input space to account for 1/p effect

$$\ell_\theta(p) = \ell_\theta \left(\frac{405 \text{ MeV}}{p} \right)$$

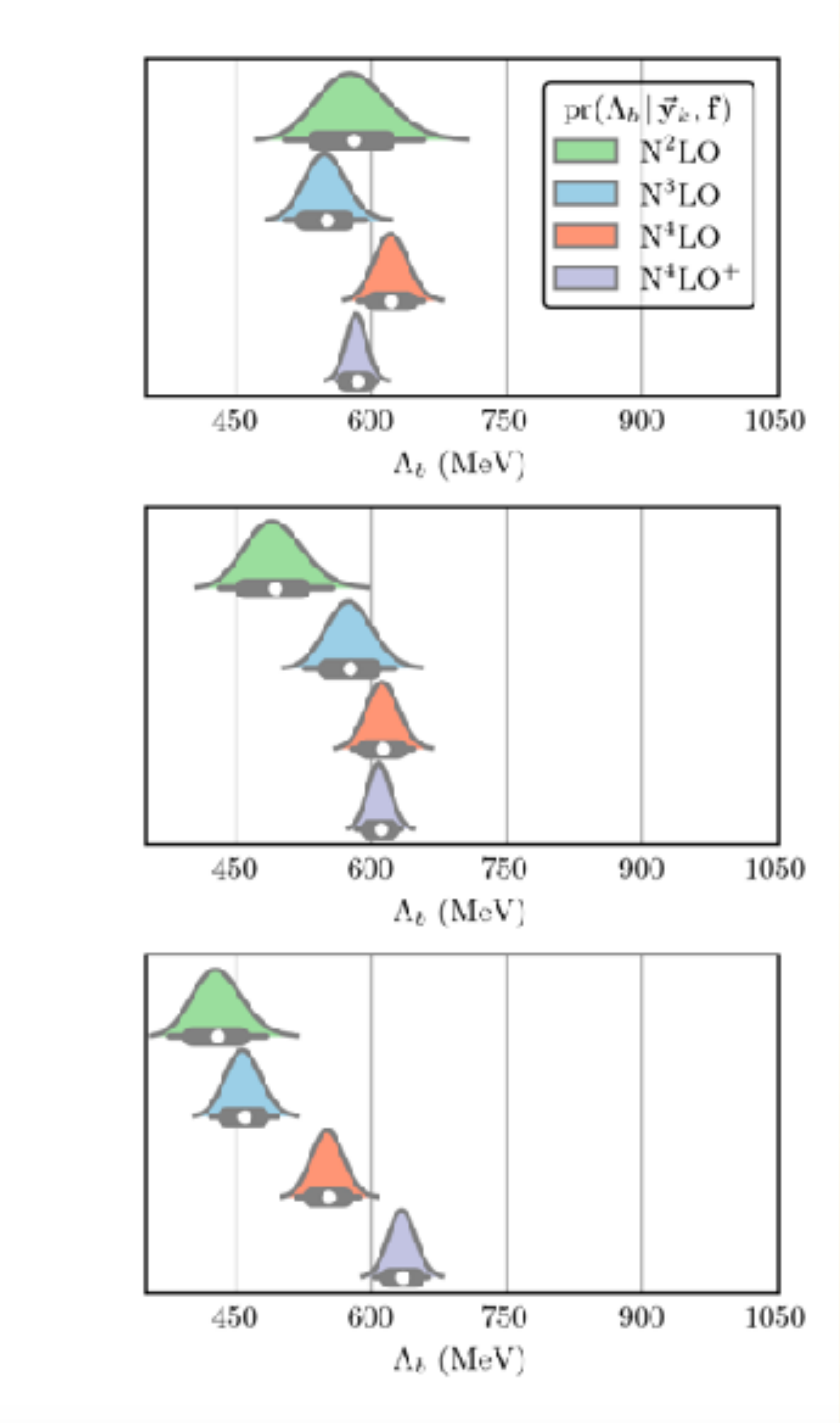
Results for Λ_b : SMS potentials

Millican et al., in preparation (2025)

SMS 450 MeV

SMS 500 MeV

SMS 550 MeV



PRELIMINARY

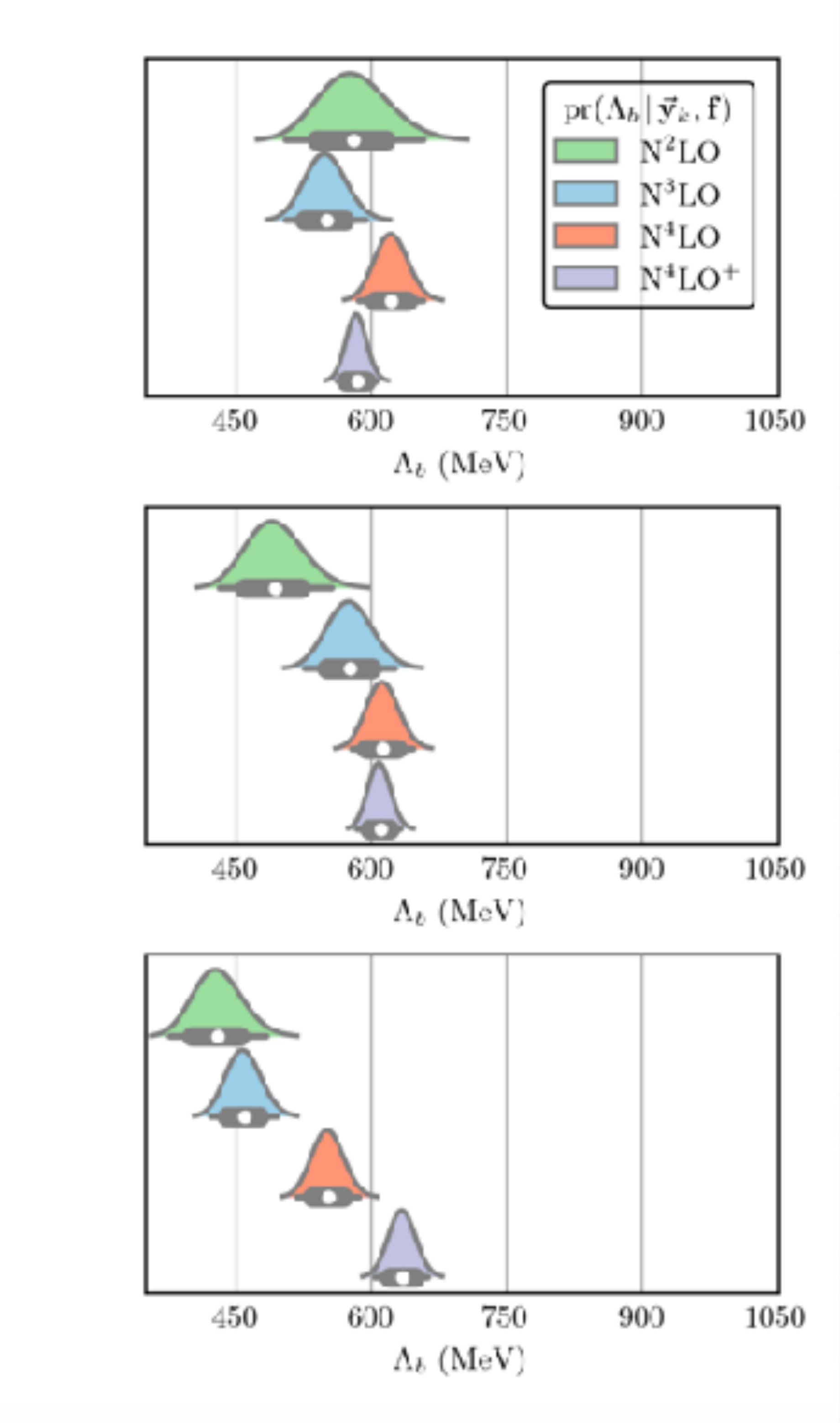
Results for Λ_b : SMS potentials

Millican et al., in preparation (2025)

SMS 450 MeV

SMS 500 MeV

SMS 550 MeV



- $m_\pi^{\mathrm{eff}} = 138$ MeV, GP diagnostics are not better (or worse) for $m_\pi^{\mathrm{eff}} = 200$ MeV
- "Downsampling" to stop over-representation of coefficients at small momenta

PRELIMINARY

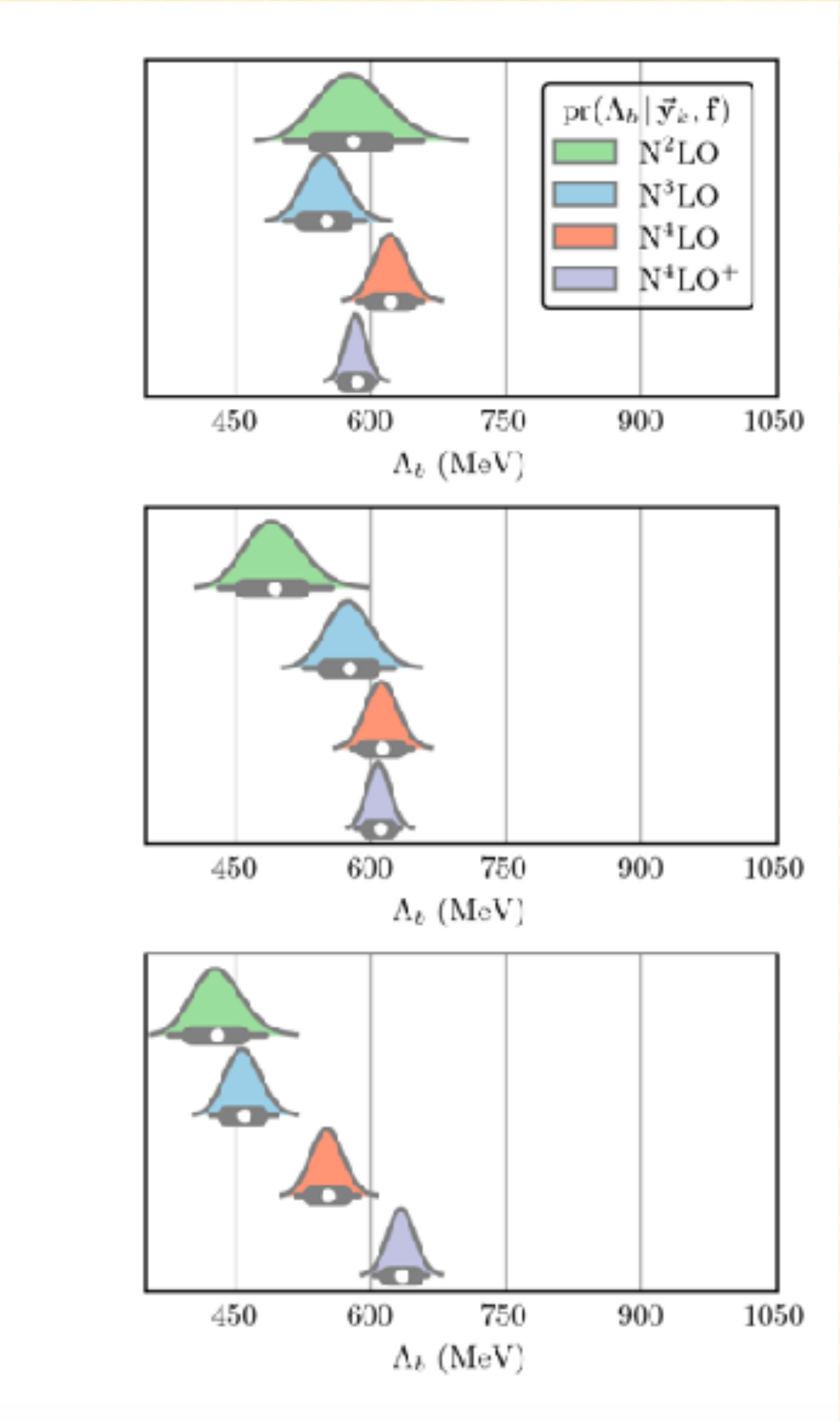
Results for Λ_b : SMS potentials

Millican et al., in preparation (2025)

SMS 450 MeV

SMS 500 MeV

SMS 550 MeV



- $m_\pi^{\text{eff}} = 138$ MeV, GP diagnostics are not better (or worse) for $m_\pi^{\text{eff}} = 200$ MeV
- "Downsampling" to stop over-representation of coefficients at small momenta
- 450 MeV & 500 MeV potentials have Λ_b ≈600 MeV consistent across orders
- 550 MeV shows Λ_b increasing with order

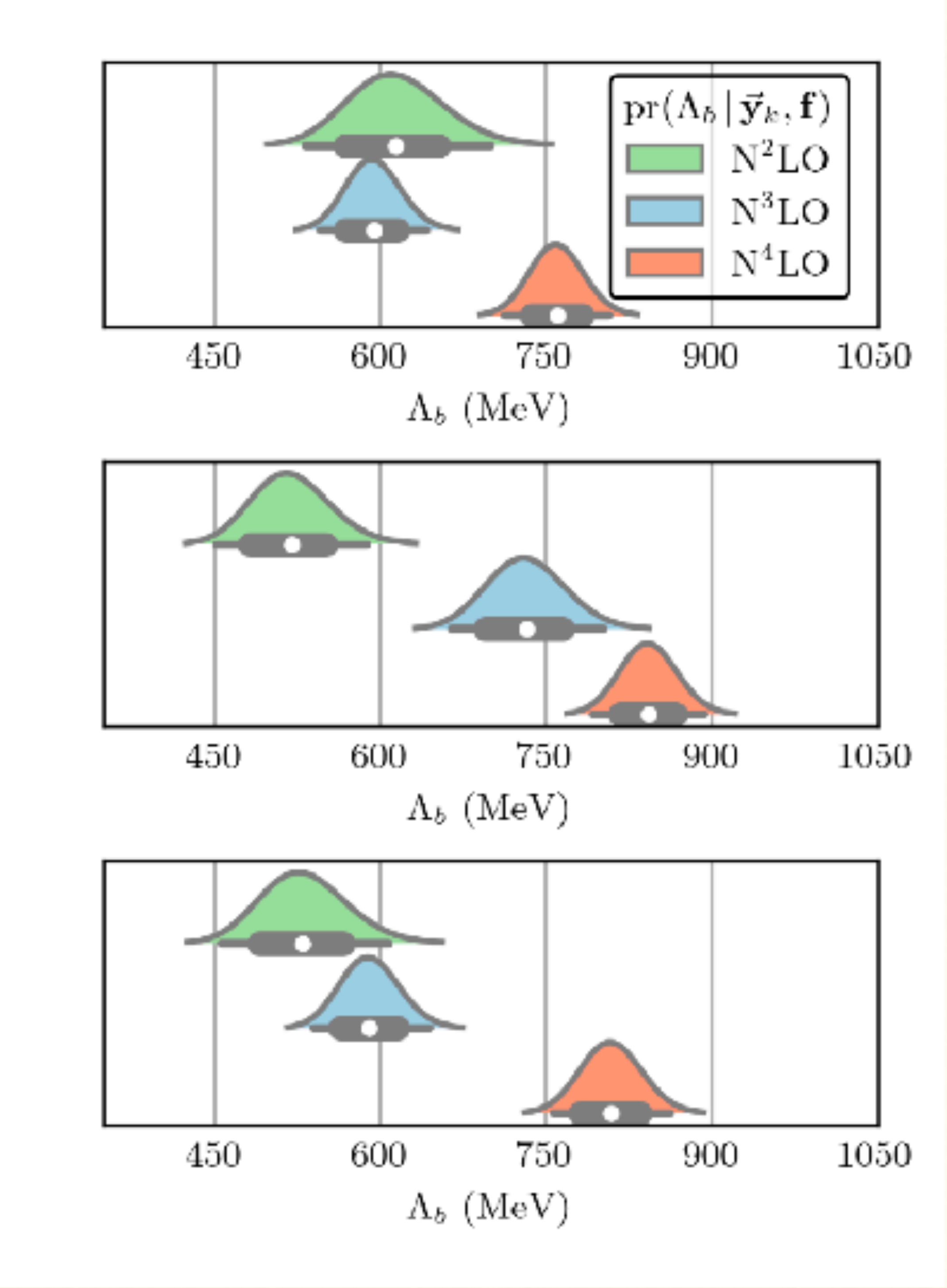
PRELIMINARY

Other non-soft potentials

SCS
1.0
fm

SCS
0.9
fm

EMN
500
MeV



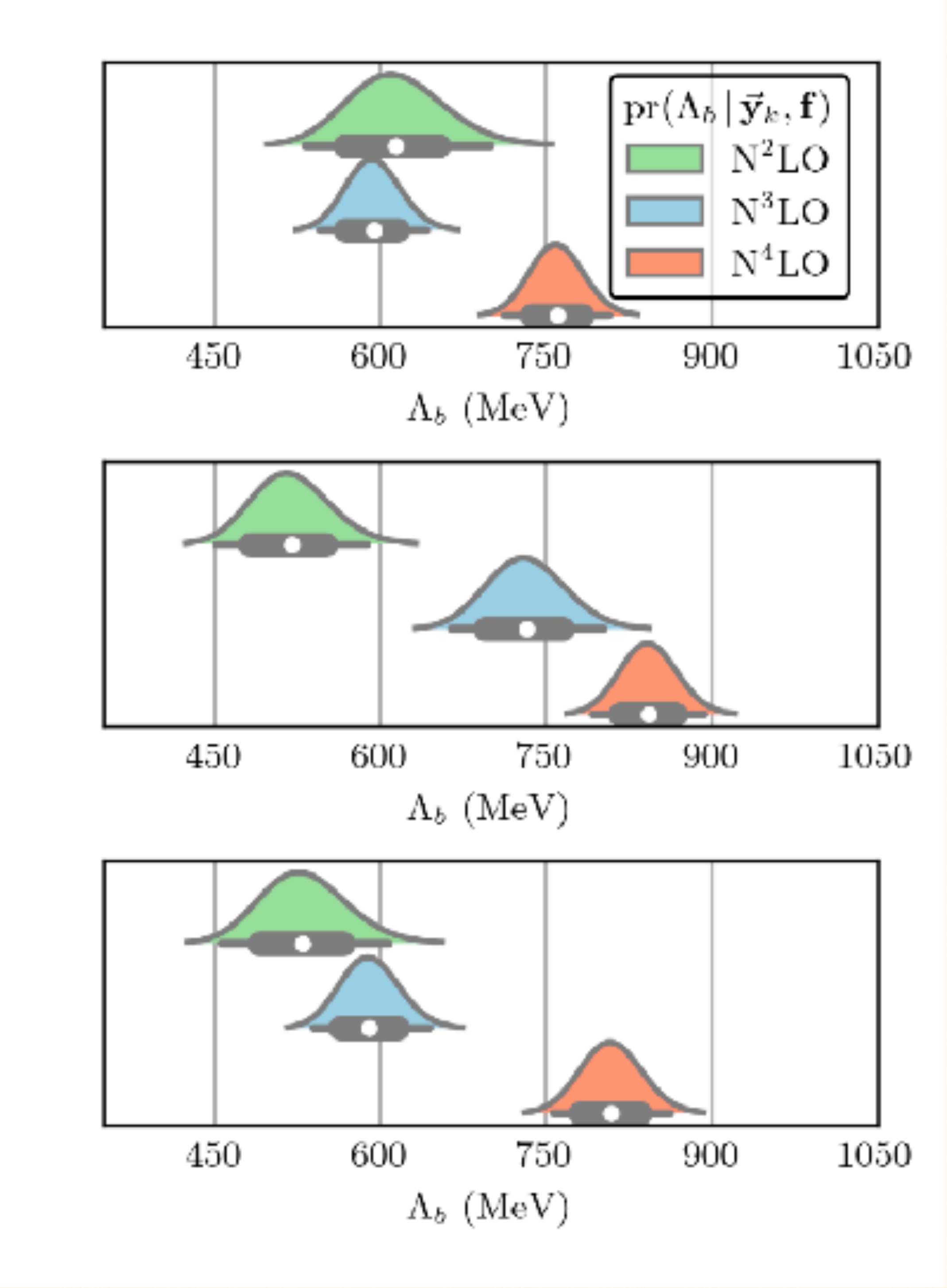
PRELIMINARY

Other non-soft potentials

SCS 1.0 fm

SCS 0.9 fm

EMN 500 MeV



- $m_\pi^{\text{eff}} = 138$ MeV
- EMN has fantastic GP diagnostics

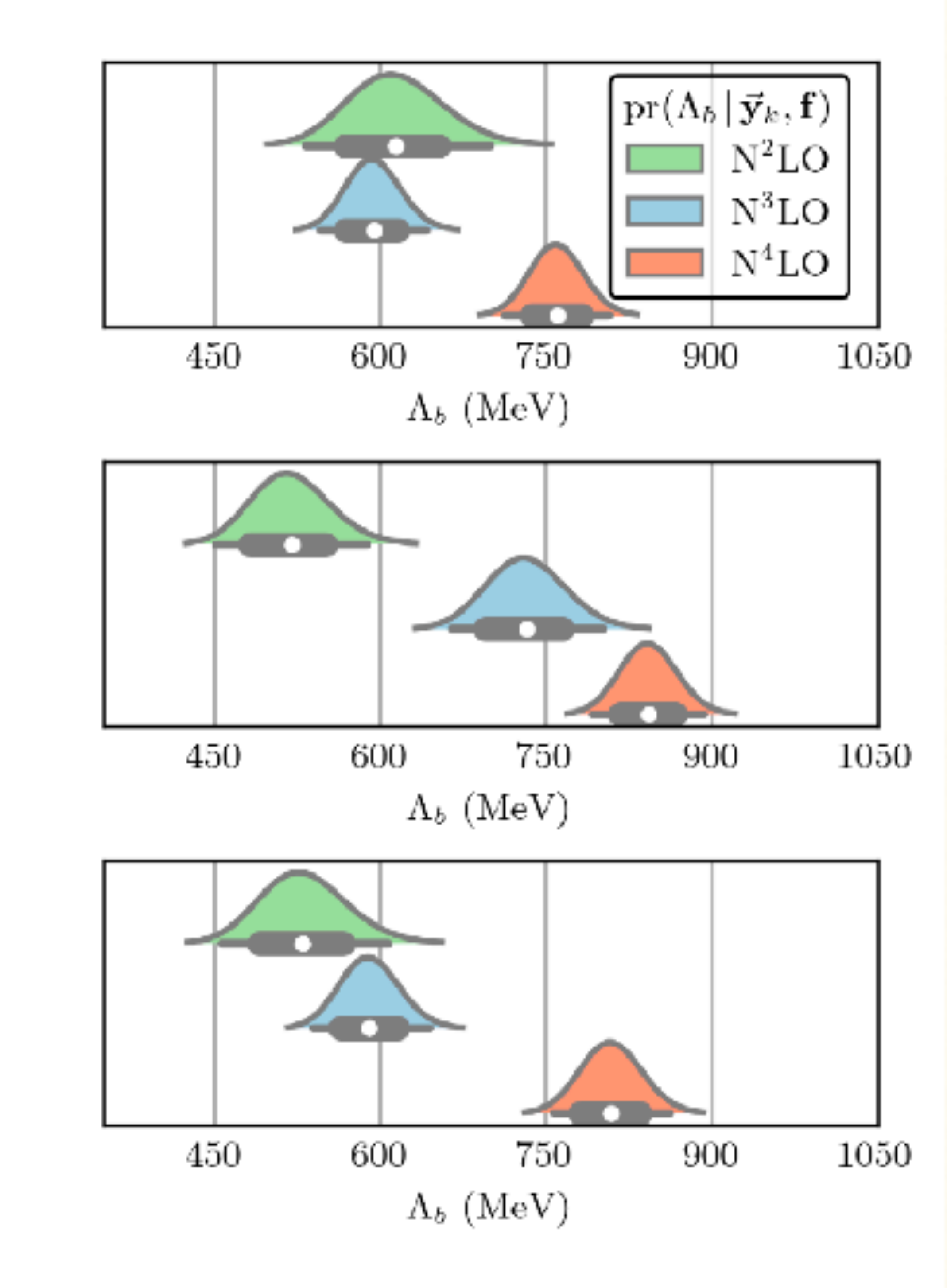
PRELIMINARY

Other non-soft potentials

SCS
1.0
fm

SCS
0.9
fm

EMN
500
MeV



- $m_\pi^{\mathrm{eff}} = 138$ MeV
- EMN has fantastic GP diagnostics
- Is the increase from N^{3}LO to N^{4}LO due to overfitting at N^{3}LO?
- UQ can still be done on potentials at a particular order, physical meaning of Λ_b then unclear
- Is that okay as long as Λ_b goes up with order?

PRELIMINARY

Global model combination

- Bayesian Model Averaging involves combining probability distributions according to: $\mathrm{pr}(y_{N+1} | D, I) = \sum_k \mathrm{pr}(y_{N+1} | M_k, D, I)\mathrm{pr}(M_k | D, I)$

- But one can, of course, also combine pdfs using other weights, e.g., $\mathrm{pr}(y_{N+1} | D, I) = \sum_k w_{M_k} \mathrm{pr}(y_{N+1} | M_k, D, I)$

- Weights can be adjusted to optimize predictive performance, e.g., "stacking"

  Yao et al., Bayesian Analysis 13 (2018), 917-1007

- In such an approach the weights do not have a rigorous probabilistic interpretation

  Höge, Guthke, Nowak, Journal of Hydrology 572 (2019) 96–107

Combined pdf from many many-body methods

Cirigliano et al., J. Phys. G (2022)

For a set of models M_k: $$y_{N+1}^{\text{true}} = y_{N+1}(\boldsymbol{\theta}, \eta; \lambda) + \delta y_{N+1,\text{th}}(\lambda)$$

parameters of nuclear Hamiltonian, calibrated to data

LEC that may be needed if this is not a binding energy. η determined using different reaction/nucleus

Control parameters associated with many-body method

Combined pdf from many many-body methods

Cirigliano et al., J. Phys. G (2022)

For a set of models M_k: $$y_{N+1}^{\text{true}} = y_{N+1}(\boldsymbol{\theta}, \eta; \lambda) + \delta y_{N+1,\text{th}}(\lambda)$$

- $\boldsymbol{\theta}$: parameters of nuclear Hamiltonian, calibrated to data
- η: LEC that may be needed if this is not a binding energy. η determined using different reaction/nucleus
- λ: Control parameters associated with many-body method

- We then marginalize over ***θ***, *η*, and the model discrepancy to generate a pdf for each model $p(y_{N+1} \,|\, y, y_\eta, M_k)$. Here y is the data set used to calibrate the Hamiltonian and y_η is the data set used to obtain η.

Combined pdf from many many-body methods

Cirigliano et al., J. Phys. G (2022)

For a set of models M_k: $$y_{N+1}^{\text{true}} = y_{N+1}(\boldsymbol{\theta}, \eta; \lambda) + \delta y_{N+1,\text{th}}(\lambda)$$

$\boldsymbol{\theta}$: parameters of nuclear Hamiltonian, calibrated to data

η: LEC that may be needed if this is not a binding energy. η determined using different reaction/nucleus

λ: Control parameters associated with many-body method

- We then marginalize over $\boldsymbol{\theta}$, η, and the model discrepancy to generate a pdf for each model $p(y_{N+1} | y, y_\eta, M_k)$. Here y is the data set used to calibrate the Hamiltonian and y_η is the data set used to obtain η.

- Assess ability of model M_k to describe y_{N+1} using another set of observables y_{ev}

Combined pdf from many many-body methods

Cirigliano et al., J. Phys. G (2022)

For a set of models M_k: $y_{N+1}^{\text{true}} = y_{N+1}(\boldsymbol{\theta}, \eta; \lambda) + \delta y_{N+1,\text{th}}(\lambda)$

parameters of nuclear Hamiltonian, calibrated to data

LEC that may be needed if this is not a binding energy. η determined using different reaction/nucleus

Control parameters associated with many-body method

- We then marginalize over $\boldsymbol{\theta}$, η, and the model discrepancy to generate a pdf for each model $p(y_{N+1} \,|\, y, y_\eta, M_k)$. Here y is the data set used to calibrate the Hamiltonian and y_η is the data set used to obtain η.

- Assess ability of model M_k to describe y_{N+1} using another set of observables y_{ev}

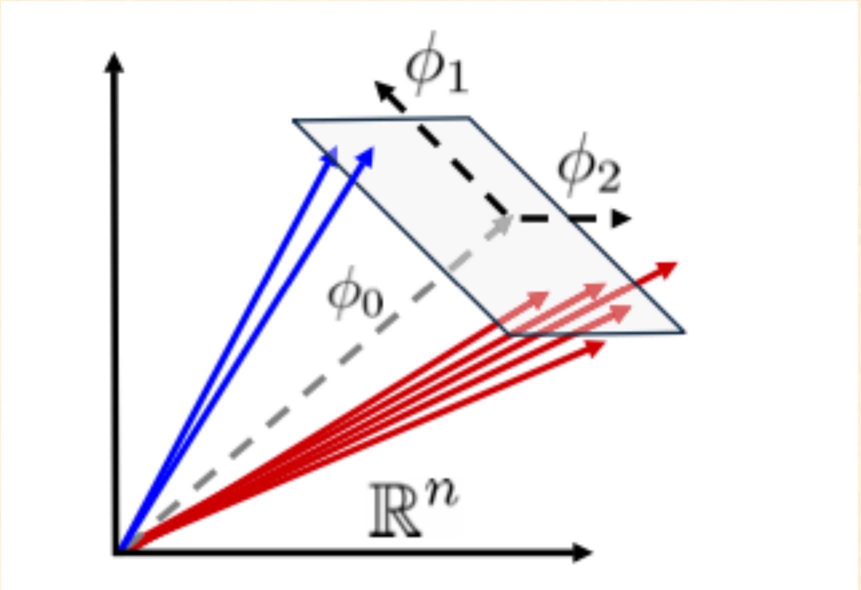
- Combine model pdfs: $p(y_{N+1} \,|\, y_{\text{ev}}, y) = \sum_{k=1}^{K} w_k(y_{\text{ev}}) p(y_{N+1} \,|\, y, y_\eta, M_k)$. Could derive weights from model evidence $p(y_{\text{ev}} \,|\, M_k)$. But there may be better options…

Diagnosing & removing model correlations

Giuliani, Godbey, Kejzlar, Nazarewicz, Phys. Rev. Res. (2024)

- How to ensure that there are not multiple copies of the same model in the combination? I.e., want to try and combine models that are actually independent, and not keep adding redundant (degenerate) models
- Consider predictions of models for N observables, indexed by i: $y_i^{(k)}$
- These define an N-dimensional vector $Y^{(k)}$ for each model. We subtract the average of that vector over the models, and then break the rest into principal components

$$y_i^{(k)} - \bar{y}_i = U_{N\times N} S_{N\times k} V^T_{k\times k} \approx U_{N\times p} S_{p\times p} V^T_{p\times p}$$



- Columns of U contain p principal components of models for the data Y, $\phi_i^{(k)}$;k=1,…p
- It is more efficient (less degenerate) to mix $\phi^{(k)}(x)$

Summary

- Discrepancy modeling provides a tool to account for "model discrepancies" aka "theory uncertainties"
- χEFT prescribes how the model discrepancy should increase with Q
- But what is Q? Information on Λ_b in order-by-order behavior and through likelihood
- Correlations matter; need to learn↔model those correlations across input space and between different observables
- Posterior for parameters broadened by χEFT uncertainty, but that of predicted observable may not be if prediction is highly correlated with calibration data set
- Combining predictions from different many-body methods could improve UQ for nuclear observable: best choice of scoring scheme is a research frontier
- Need to also ensure absence of degenerate models (and degenerate modelers)

Modeling correlated truncation errors

Consider χEFT, where we have two light scales, p and m_π

- General χEFT series for observable to order k: $y = y_{\rm ref} \sum_{n=0}^{k} c_n(p_{\rm typ}/m_\pi) Q^n$:

$$Q = \frac{(p_{\rm typ}, m_\pi)}{\Lambda_b}; \quad \Lambda_b \approx 600 \text{ MeV}$$

- Then c_n are "order 1"

Higher-order uncertainties

- Exist
- Have a characteristic size $\sim Q^{k+1}$
- Are correlated across the input space
- Have a characteristic correlation length of order the light scale
- Can be modeled statistically

A bit more on Gaussian Processes

- Non-parametric, probabilistic model for a function
- Specify how f(y) is correlated with $f(x_1)$, $f(x_2)$, …..; don't specify underlying functional form.
- But value of f(y) is not deterministic: it's given by a (Gaussian) probability distribution.
- Correlation decreases as points get further away from each other.
- Specify correlation matrix of f at x and y, e.g.:

$$k(f(x), f(y)) = \bar{c}^2 \exp\left(-\frac{(x-y)^2}{2\ell^2}\right)$$

- $\mathrm{pr}(\bar{c}^2 \,|\, I) \sim \chi^{-2}(\nu_0, \tau_0^2)$; $\mathrm{pr}(\ell \,|\, I)$ uniform

A bit more on Gaussian Processes

- Non-parametric, probabilistic model for a function
- Specify how f(y) is correlated with $f(x_1)$, $f(x_2)$, …..; don't specify underlying functional form.
- But value of f(y) is not deterministic: it's given by a (Gaussian) probability distribution.
- Correlation decreases as points get further away from each other.
- Specify correlation matrix of f at x and y, e.g.:

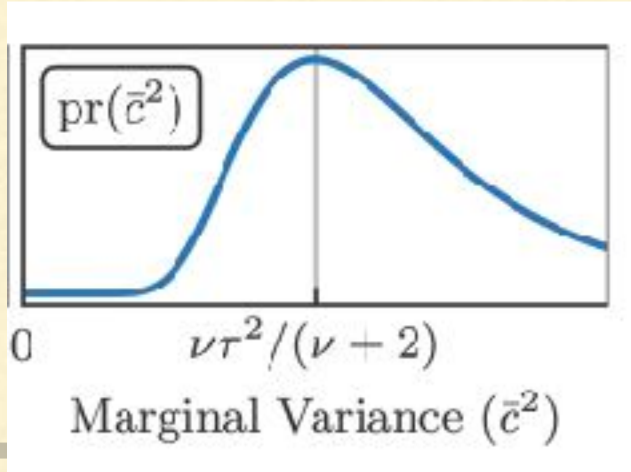
$$k(f(x), f(y)) = \bar{c}^2 \exp\left(-\frac{(x-y)^2}{2\ell^2}\right)$$

Statistical model choices

- $\mathrm{pr}(\bar{c}^2 \,|\, I) \sim \chi^{-2}(\nu_0, \tau_0^2)$; $\mathrm{pr}(\ell \,|\, I)$ uniform

$$\nu = \nu_0 + n_c;$$

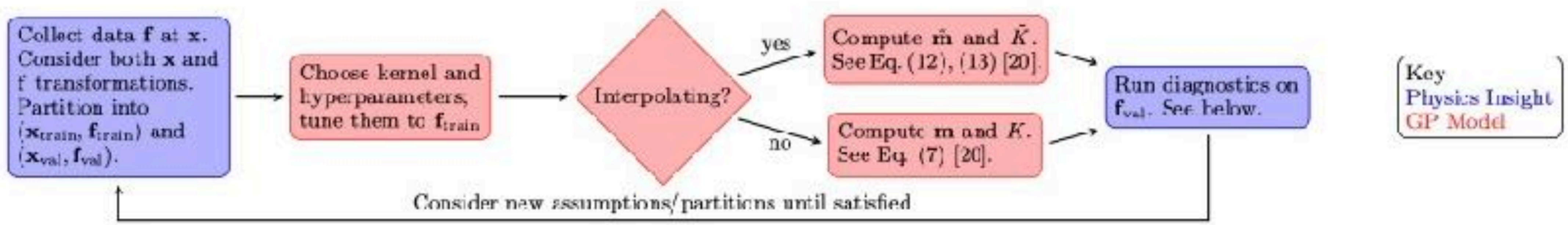
$$\nu\tau^2 = \nu_0\tau_0^2 + \vec{c}_k^{\,2}$$



Model checking

https://github.com/buqeye/gsum

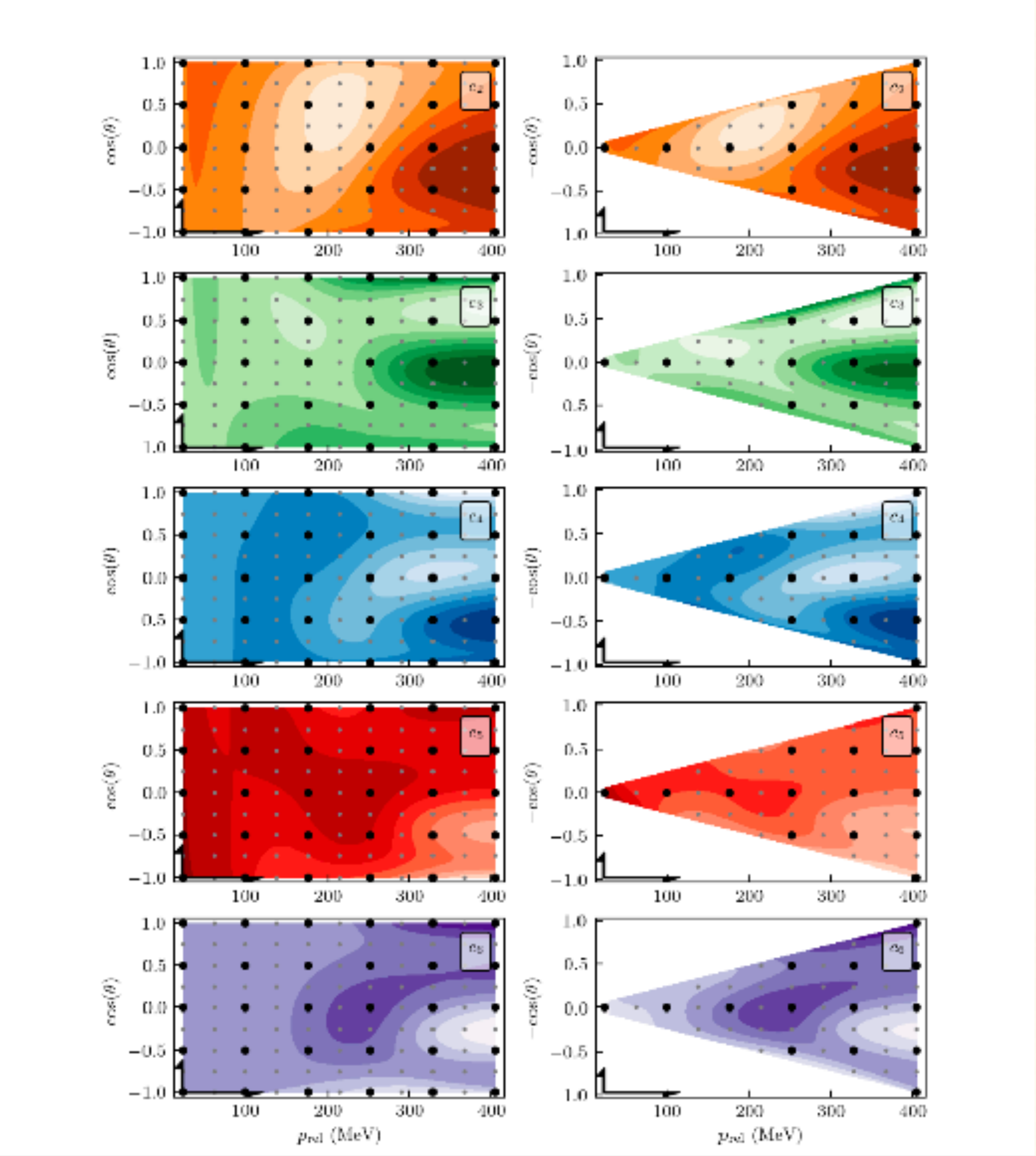
Melendez et al. (2019), Millican et al. (2024), Bastos & O'Hagan (2009)



Diagnostic	Formula	Motivation	Success	Failure
Visualise the function	—	Does $\mathbf{f}_{\text{val}}$ look like a draw from a GP? What kind of GP?	$\mathbf{f}_{\text{val}}$ "looks similar" to draws from a GP	$\mathbf{f}_{\text{val}}$ "stands out" compared to GP draws
Mahalanobis Distance D^2_{MD}	$(\mathbf{f}_{\text{val}} - \mathbf{m})^\intercal K^{-1}(\mathbf{f}_{\text{val}} - \mathbf{m})$	Can we *quantify* how much the $\mathbf{f}_{\text{val}}$ looks like a GP?	D^2_{MD} follows its theoretical distribution (χ^2_M)	D^2_{MD} lies too far away from the expected value of M
Pivoted Cholesky $\mathbf{D}_{\text{PC}}$	$G^{-1}(\mathbf{f}_{\text{val}} - \mathbf{m})$	Can we understand why D^2_{MD} is failing?	At each index, points follow standard Gaussian	Many cases (see below)
Credible Interval $\mathrm{D}_{\text{CI}}(P)$ for $P \in [0, 1]$	$\frac{1}{M}\sum_{i=1}^{M} \mathbf{1}[\mathbf{f}_{\text{val},i} \in \mathrm{CI}_i(P)]$	Do $100P\%$ credible intervals capture data roughly $100P\%$ of the time?	Plot $\mathrm{D}_{\text{CI}}(P)$ for $P \in [0, 1]$; the curve should be within errors of $\mathrm{D}_{\text{CI}}(P) = P$.	$\mathrm{D}_{\text{CI}}(P)$ is far from $100P\%$, particularly for large $100P\%$ (e.g., 68% and 95%).

Variance	Length Scale	Observed Pattern in $\mathbf{D}_{\text{PC}}$
$\sigma_{\text{est}} = \sigma_{\text{true}}$	$\ell_{\text{est}} = \ell_{\text{true}}$	Points are distributed as a standard Gaussian, with no pattern across index (e.g., only $\approx 5\%$ of points outside 2σ lines).
$\sigma_{\text{est}} = \sigma_{\text{true}}$	$\ell_{\text{est}} > \ell_{\text{true}}$	Points look well distributed at small index but expand to a too-large range at high index.
$\sigma_{\text{est}} = \sigma_{\text{true}}$	$\ell_{\text{est}} < \ell_{\text{true}}$	Points look well distributed at small index but shrink to a too-small range at high index.
$\sigma_{\text{est}} > \sigma_{\text{true}}$	$\ell_{\text{est}} = \ell_{\text{true}}$	Points are distributed in a too-small range at all indices.
$\sigma_{\text{est}} < \sigma_{\text{true}}$	$\ell_{\text{est}} = \ell_{\text{true}}$	Points are distributed in a too-large range at all indices.

Unwarped vs warped coefficients



What about amplitudes?

McClung, Elster, DP, PRC (2025)

Wolfenstein amplitudes

Wolfenstein & Ashkin (1952)

$$\overline{M}(q,\theta) = A(q,\theta)\mathbb{1} + iC(q,\theta)(\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_2)\cdot\hat{\mathbf{n}} + M(q,\theta)(\boldsymbol{\sigma}_1\cdot\hat{\mathbf{n}})(\boldsymbol{\sigma}_2\cdot\hat{\mathbf{n}}) + [G(q,\theta) - H(q,\theta)](\boldsymbol{\sigma}_1\cdot\hat{\mathbf{q}})(\boldsymbol{\sigma}_2\cdot\hat{\mathbf{q}}) + [G(q,\theta) + H(q,\theta)](\boldsymbol{\sigma}_1\cdot\hat{\mathcal{K}})(\boldsymbol{\sigma}_2\cdot\hat{\mathcal{K}}).$$

$$\mathbf{K} = \frac{1}{2}(\mathbf{p} + \mathbf{p}'); \mathbf{q} = \mathbf{p}' - \mathbf{p}; \mathbf{n} = \mathbf{p}\times\mathbf{p}'$$

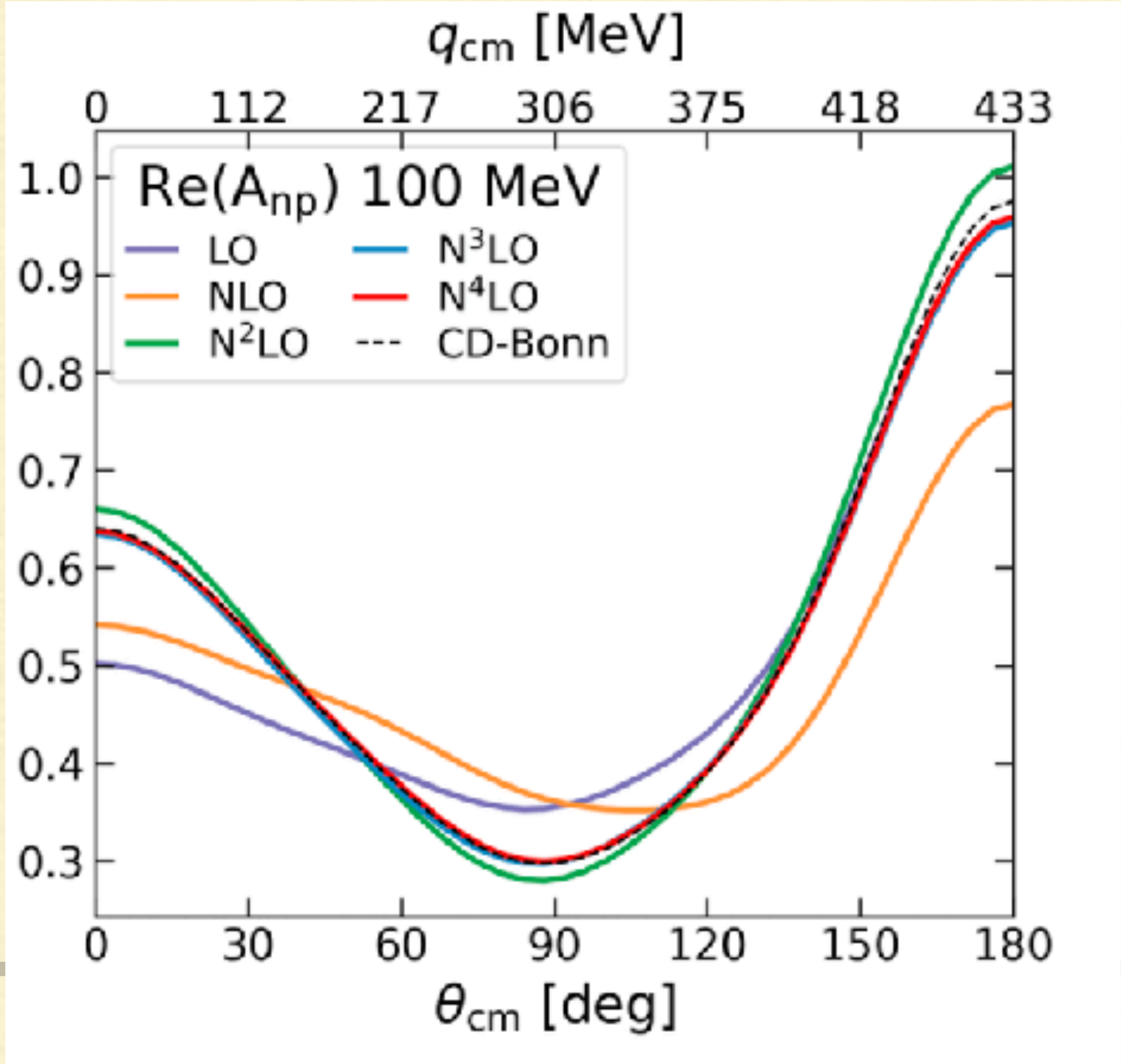
A: central part

C: spin-orbit

M, G, and H: tensor effects

Works well for amplitudes at 100 MeV

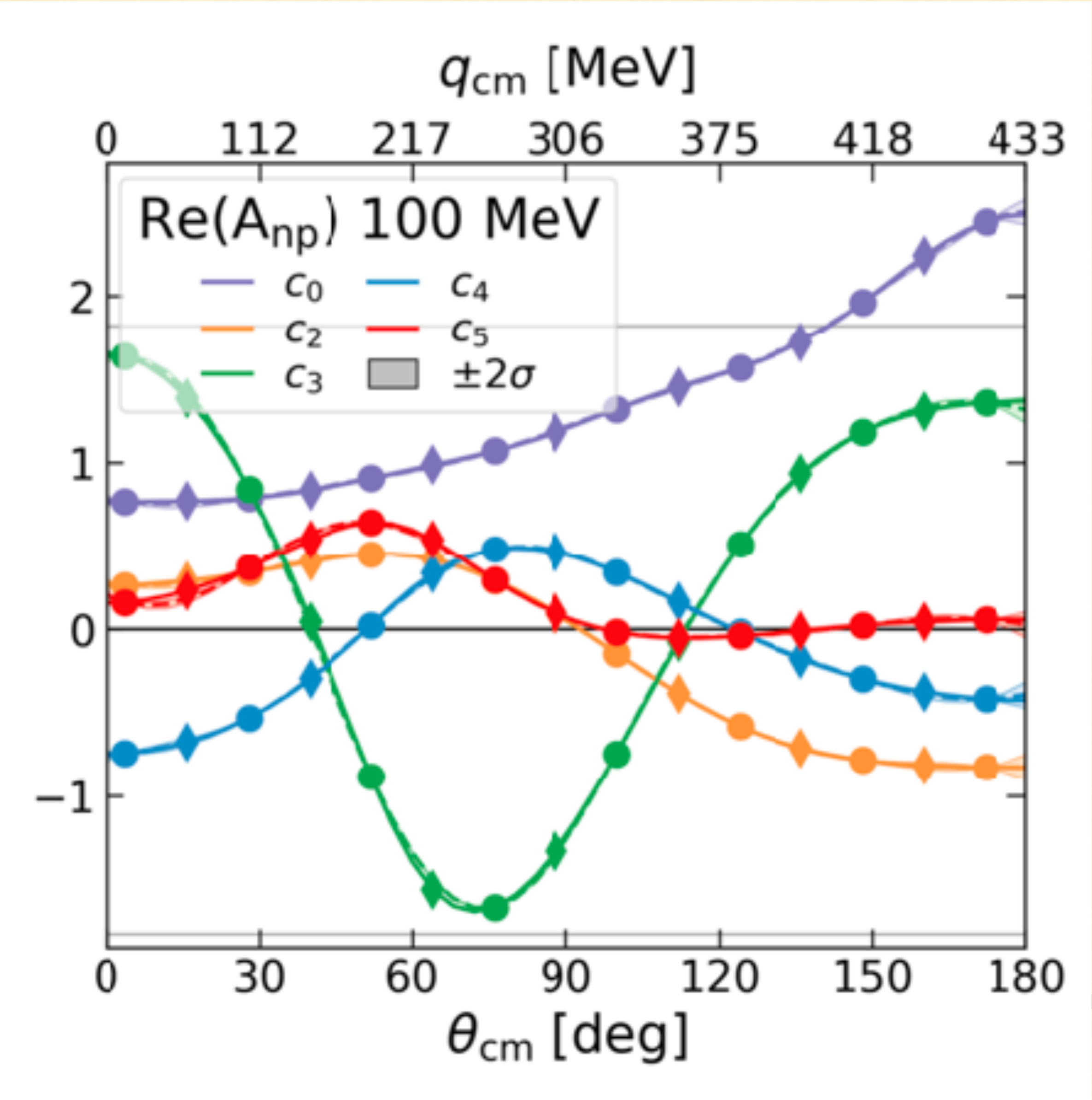
- y_{ref}=Im(A)
- $$Q = \frac{\max(p, q) + m_\pi}{\Lambda_b + m_\pi}$$

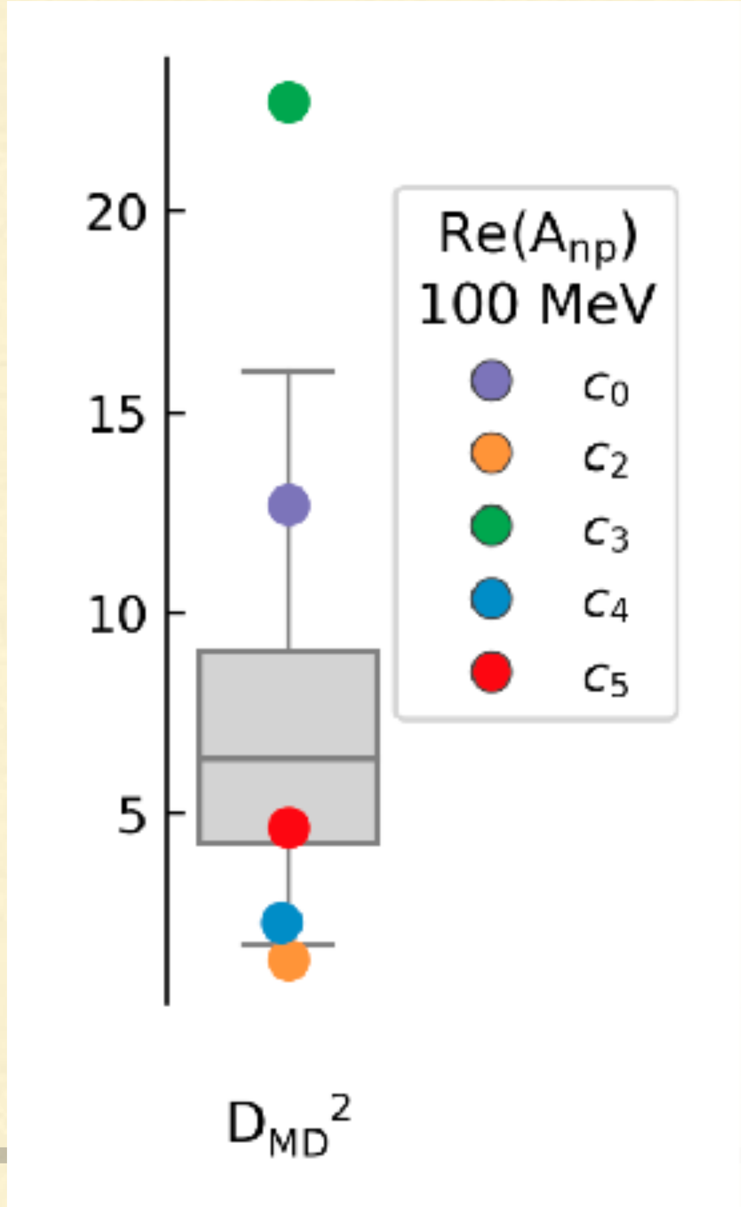


Works well for amplitudes at 100 MeV

- y_{ref}=Im(A)
- $Q = \dfrac{\max(p,q) + m_\pi}{\Lambda_b + m_\pi}$

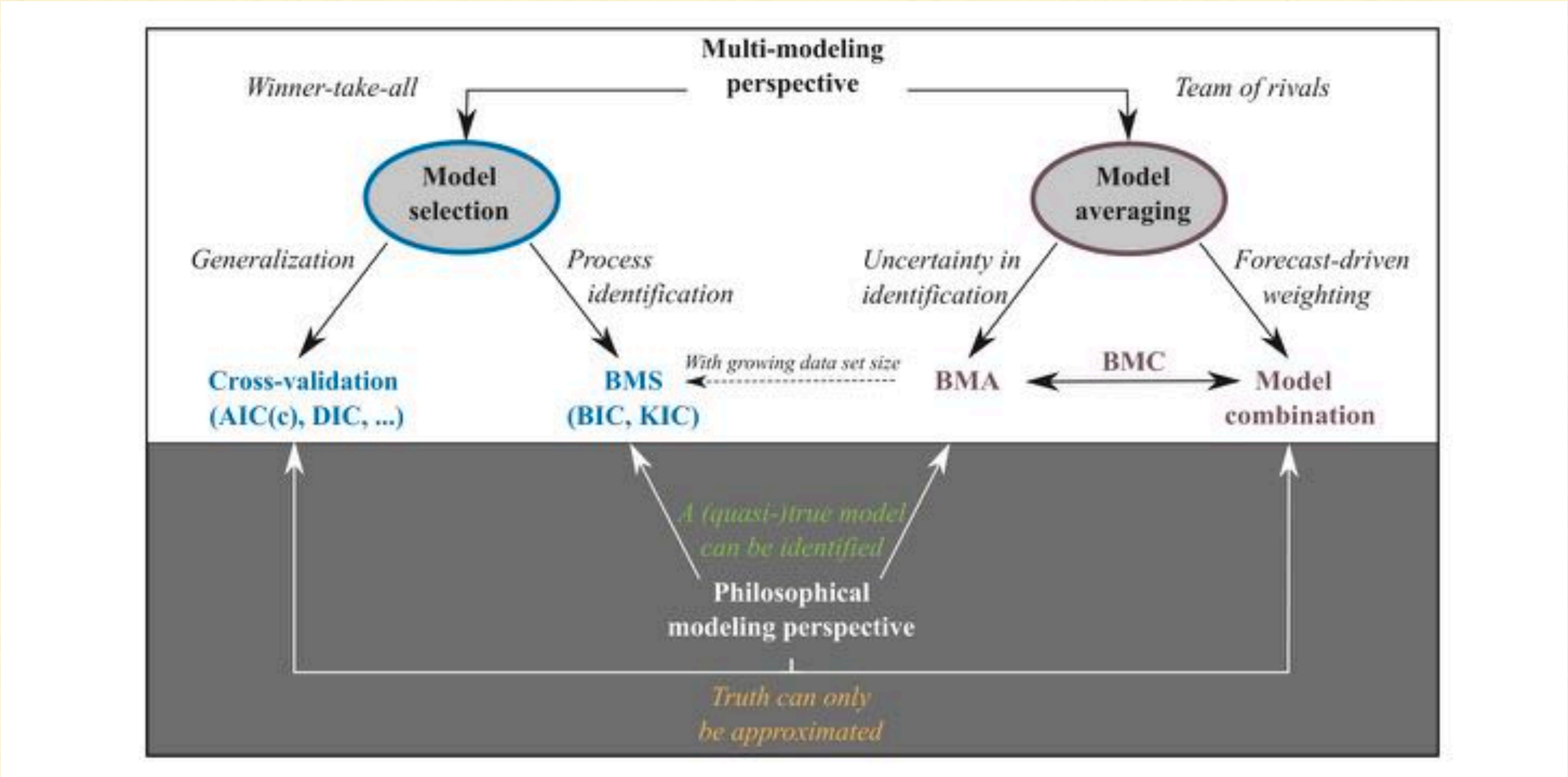
See ℓ_q is constant with energy





The M-open/M-closed distinction

- BMA converges to true model in $\mathfrak{M}$-closed situation
- And in this situation weights can indeed be interpreted as probability of model M_k being true



Höge, Guthke, Nowak, Journal of Hydrology 572 (2019) 96–107

Mixing moments locally

- $y(x_i) \sim \mathcal{N}\left(\sum_{k=1}^{p} w_k(x_i)\hat{f}_k(x_i), c(\,\cdot\,)\right)$ "mean mixing"
- Where $\hat{f}_k$ is the (possibly discrepancy corrected) mean of the kth model
- "No single model in [the set] adequately describes $\mathfrak{M}_\dagger$" [the underlying model]
- "a locally weighted combination of models in the set will provide a better description of $\mathfrak{M}_\dagger$ than any single model in the set"
- "Dirichlet mixing": $y(x_i) \sim \mathcal{N}\left(\sum_{k=1}^{p} w_k(x_i)\hat{f}_k(x_i), \sigma^2\right)$

Kejzlar, Neufcourt, Nazarewicz, Scientific Reports (2023), 19600; Gneiting et al., Mon. Weather Review (2005)

- The (uncorrelated) uncertainty σ is fit to data at the same time as the hyper parameters of the Dirichlet distributions from which the weights are drawn
- BART: $y(x_i) \sim \mathcal{N}\left(\sum_{k=1}^{p} w_k(x_i)\hat{f}_k(x_i), \sigma^2\right)$; $w_k(x)$ not simplexed and modeled as sum of trees

Evaluating $w_k(y_{ev})$

- We expect a significant portion of the error is due to systematic method deficiencies and so will be correlated across observable space
- Want observables closely related to $\mathscr{M}_{0\nu}$. Don't want to dilute weights, or spend time evaluating things that don't tell us which method has highest evidence
- Candidates:
  - Single beta-decay rates in neighboring nuclei
  - β-strength distributions
  - Known 2νββ rates
  - Magnetic moments and B(M1) rates in the three nuclei involved in a particular 0νββ decay
  - Energies of lowest $J^{\pi}=2^+$ states and $B(E2, 2^+ \rightarrow 0^+)$ rates in initial and final nuclei
  - Charge radii
  - Observables probing a 100 MeV momentum-transfer scale, e.g., in muon capture

Emulators will play a crucial role

- Assess correlation between members of y_{ev} set and 0νββ matrix element. (Significant work along these lines already done using lower-resolution methods for nuclei.) Correlations incorporated into scoring criteria.