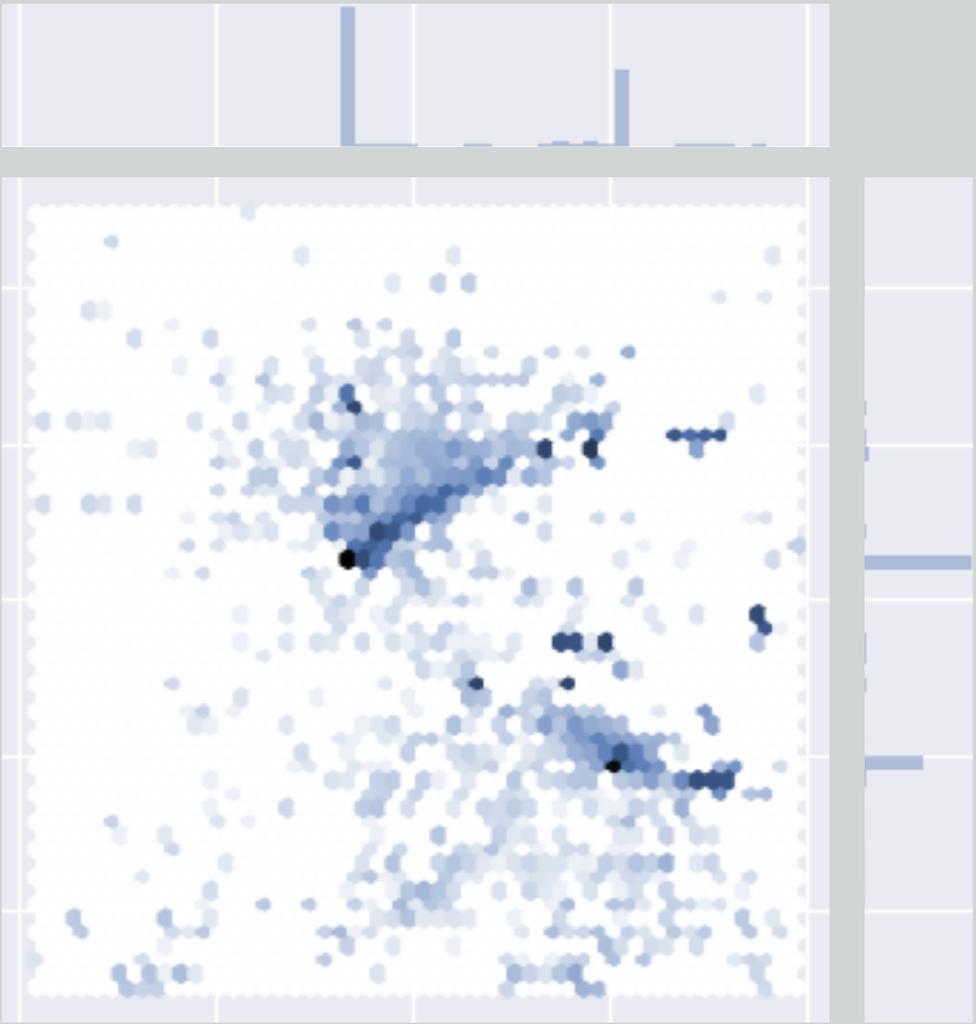


DIFFERENT STRATEGIES FOR EFT PARAMETER ESTIMATION IN THE FEW-NUCLEON SECTOR

CHRISTIAN FORSSÉN

Department of Physics, Chalmers, Sweden

ECT* workshop, Trento, June 4-8, 2018

MANY THANKS TO MY COLLABORATORS

- **Boris Carlsson**, **Andreas Ekström, Andreas Johansson** (Chalmers)
- Gustav Jansen (ORNL), Kyle Wendt (ORNL/ UT now LLNL)

And many people in the *ab initio nuclear theory* community for enlightening discussions

Research funded by:

- STINT
- European Research Council
- Swedish Research Council

INTRODUCTION

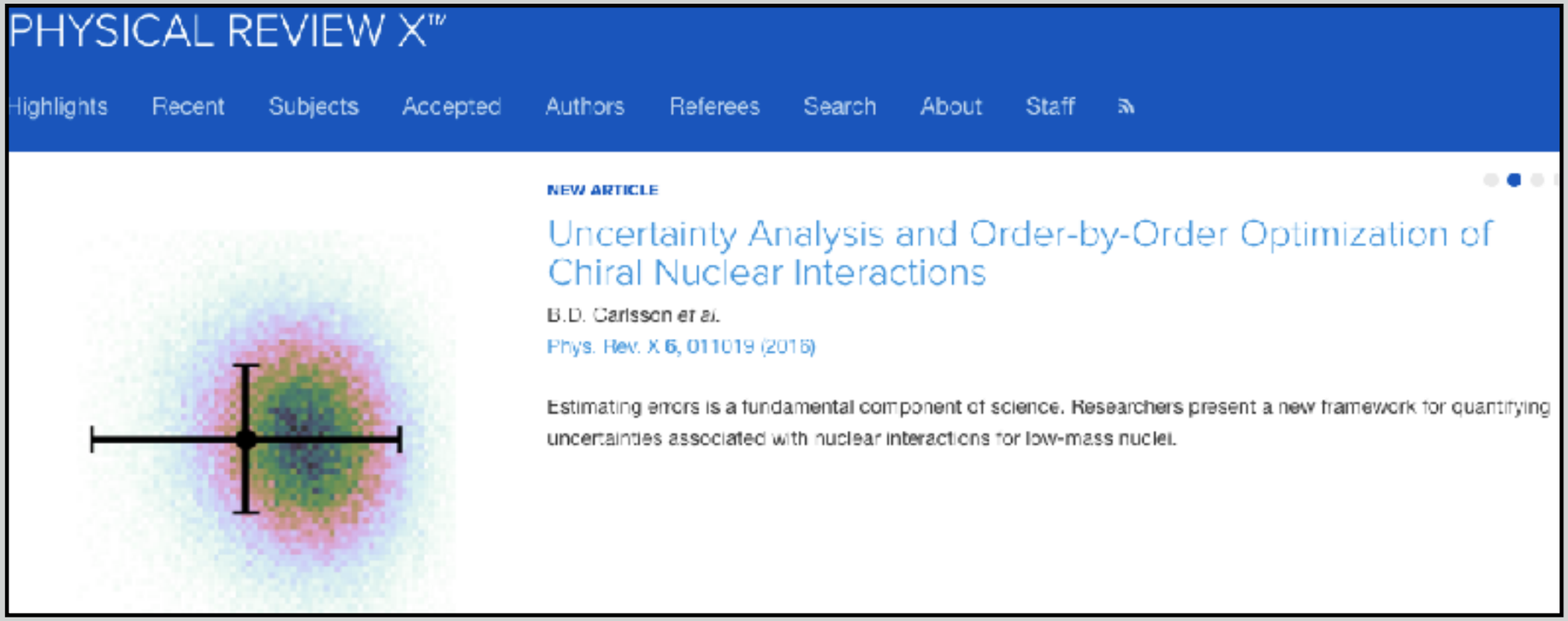
ONE-BULLET OVERVIEW OF WHAT I WILL TALK ABOUT

- What information can be inferred[1] from available few-nucleon data[2,3] to state–of-the-art "models"[4] of the strong force between nucleons?

[1] using either frequentist or bayesian approaches.
[2] Here: πN and NN scattering data
[3] ...plus A=2-3 bound-state observables
[4] effective field theories => systematically improvable

FIVE-BULLET OVERVIEW OF WHAT I WILL NOT TALK ABOUT

- What are the predictions of these *few-body constrained* Hamiltonians for *many-body observables*…
- Different *power-counting schemes* (but I will offer the Bayesian viewpoint on this)…
- Details of the *computational tools* that we are using: the nsopt code base, Automatic differentiation, MCMC algorithms, …
- The facts that Sweden eliminated both the Netherlands and Italy in the World Cup qualifier (I'm actually a bit sorry about that), plus the observation that the German national team has not won a single game this year.
- Inference from the above data on to the probability that Sweden will beat Germany in the World Cup on June 23rd.



Note: Statistical analysis revisited following the *"Schiavilla-correction"* (relation between cD in the one-pion exchange plus contact NNN potential and the LEC multiplying the contact axial-vector current)

A DATA-CENTRIC VIEW ON THE EFT DESCRIPTION OF NUCLEAR FORCES

Inference

"the act of passing from one proposition, statement, or judgment considered as true to another whose truth is believed to follow from that of the former" (Webster)

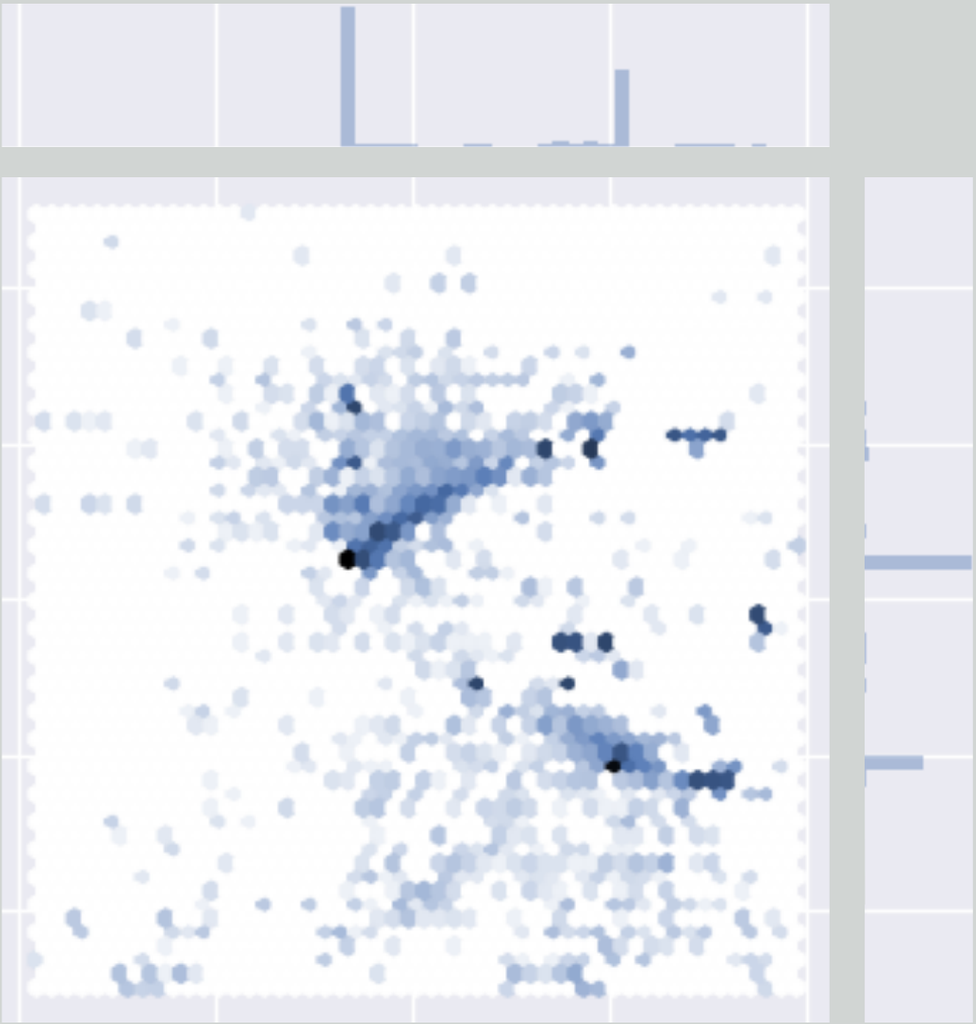
Do premises $A, B, \ldots \rightarrow$ hypothesis, H?

- Assume that **hypothesis** $\boldsymbol{H_i}$ is a model M_i with parameters $\boldsymbol{\alpha}_i$.
- **Inductive inference**: Premises bear on truth/falsity of H, but don't allow its definite determination
- **Statistical Inference**: Quantify the strength of inductive inferences from data and other premises to hypotheses about the phenomena producing the data.
- Quantify via **probabilities**, or averages calculated using probabilities. *Frequentists* and *Bayesians* use probabilities very different for this.

DATA AND ITS IMPORTANCE FOR EFFECTIVE FIELD THEORIES

- Short-range physics for an EFT is encoded in LECs.
- These LECs can be inferred by confrontation with (low-energy) observables.
- Many issues for chiral EFT remain:
  - regulator artefacts,
  - convergence,
  - The choice of relevant degrees-of-freedom

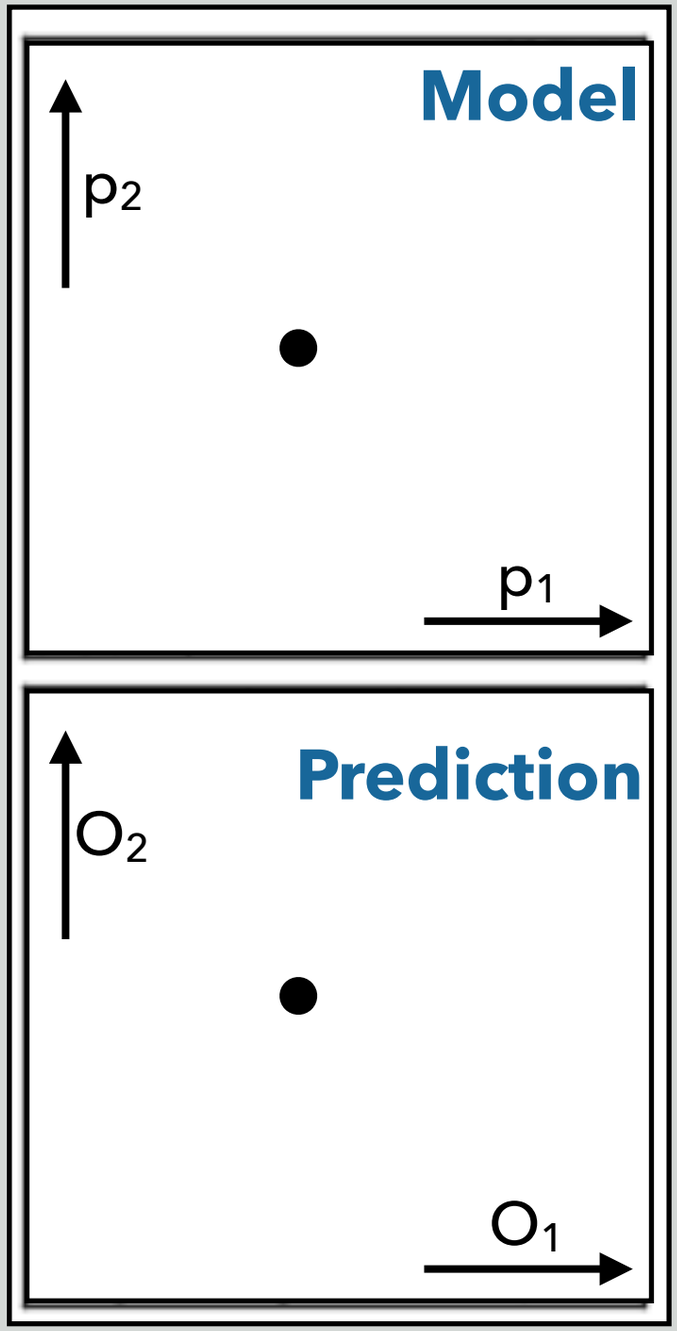
such issues can only be adressed with proper uncertainty quantification (UQ).

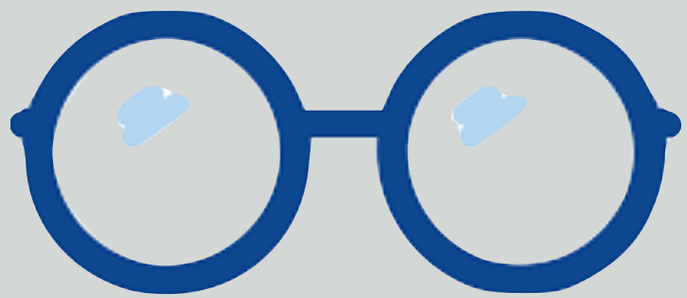


Work with A. Johansson and A. Ekström

BAYESIAN POSTERIORS IN THE NUCLEON-NUCLEON SECTOR

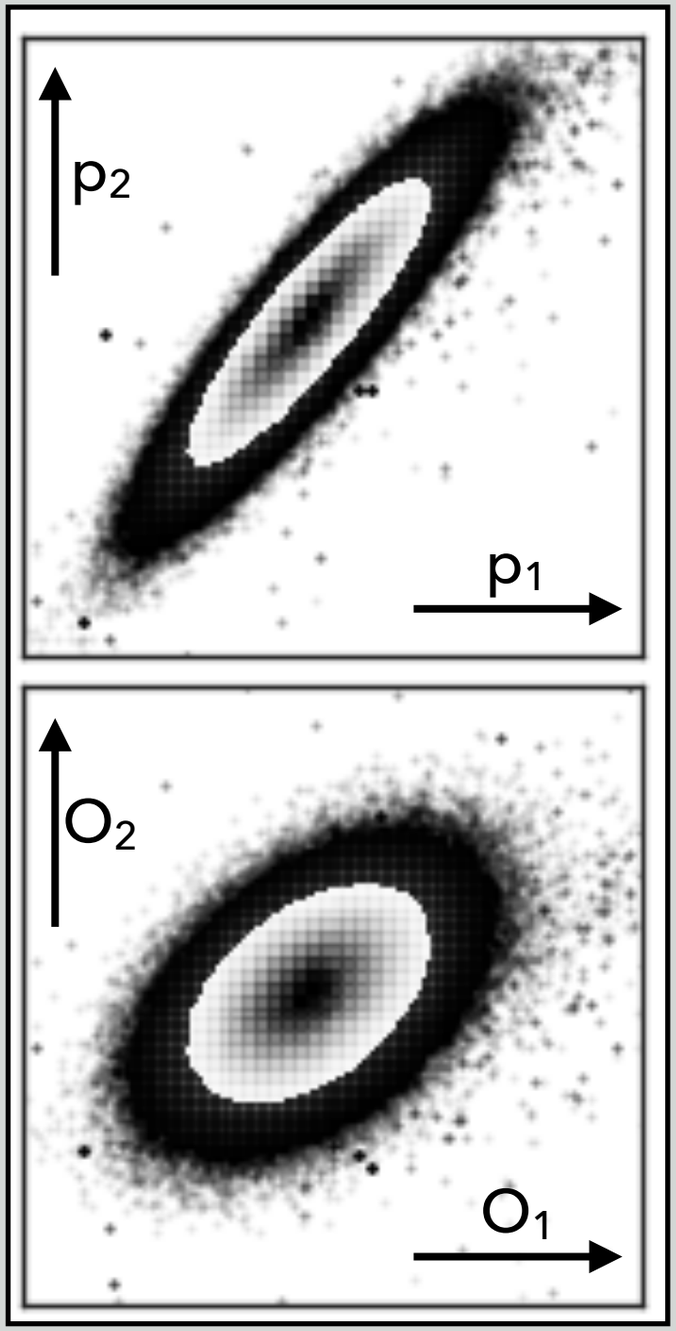
THE BAYESIAN VIEWPOINT





Data,
Prior information

$$p(\boldsymbol{p}|D, I)$$



Parametric models

- Assume that hypothesis H_i is a model M_i with parameters $\boldsymbol{p}_i$.
- In **Bayesian statistics** we assess the hypotheses by calculating their probabilities $p(H_i| \ldots)$ conditional on known and/or presumed information using the rules of probability theory.
- **Parameter estimation**:
  Assume that the model M_i is true;
  Compute: $p(\boldsymbol{p}_i \,|\, D_{\text{obs}}, M_i, I)$
- **Model comparison**:
  Compute ratio: $p(D_{\text{obs}} \,|\, M_i, I) \,/\, p(D_{\text{obs}} \,|\, M_j, I)$

Bayesian parameter estimation

Bayes' theorem (follows from probability product rule):

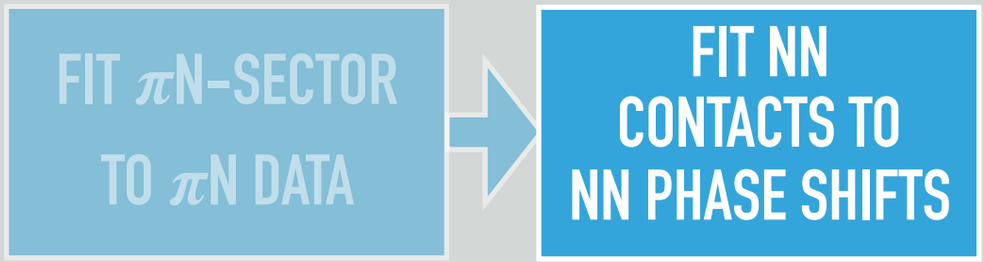
$$\underbrace{p(\boldsymbol{p}|D,I)}_{\text{posterior}} = \frac{\overbrace{p(D|\boldsymbol{p},I)}^{\text{likelihood}}\overbrace{p(\boldsymbol{p}|I)}^{\text{prior}}}{\underbrace{p(D|I)}_{\text{Bayesian evidence}}}$$

Marginalization: $$p(p_1|D,I) = \int dp_2 \ldots dp_k p(\boldsymbol{p}|D,I)$$

- For many lessons and suggestions on the use of Bayesian methods in Effective Field Theories, see work by the BUQEYE collaboration (and talk by Sarah).
- Here we report on progress in implementing Bayesian methods for parameter estimation in Chiral EFT (up to N3LO) using *NN* scattering data (phase shifts).

NN PHASE SHIFT OPTIMIZATION

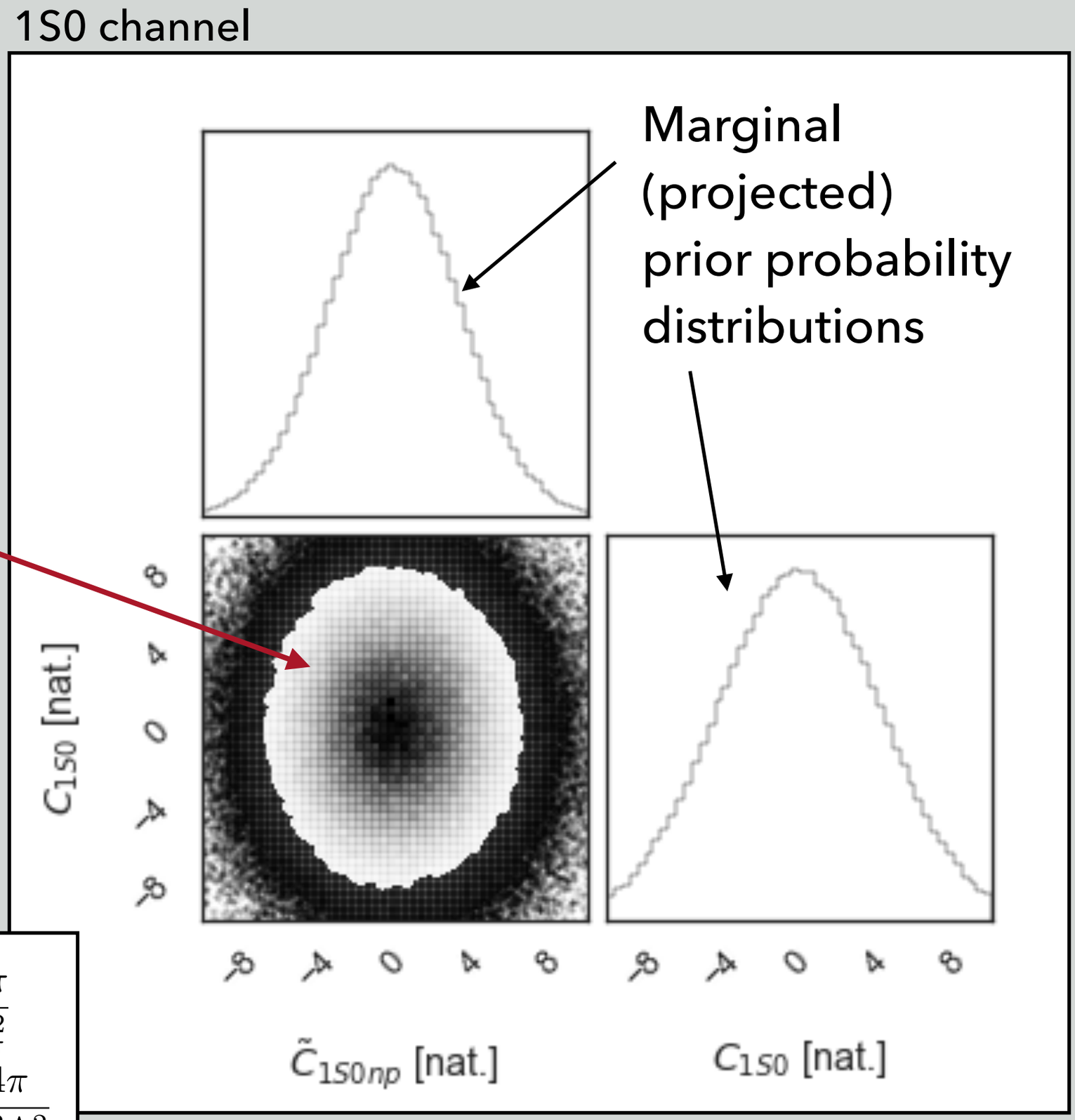
Let us consider the task of determining a set of low-energy constants (LECs) by fitting to phase shifts.



$$p(\tilde{C}_{1S0}, C_{1S0}|D, I) \propto p(D|\tilde{C}_{1S0}, C_{1S0}, I)p(\tilde{C}_{1S0}, C_{1S0}|I)$$

posterior likelihood prior

The 1S0(np) channel @ NLO/N2LO: Prior distribution

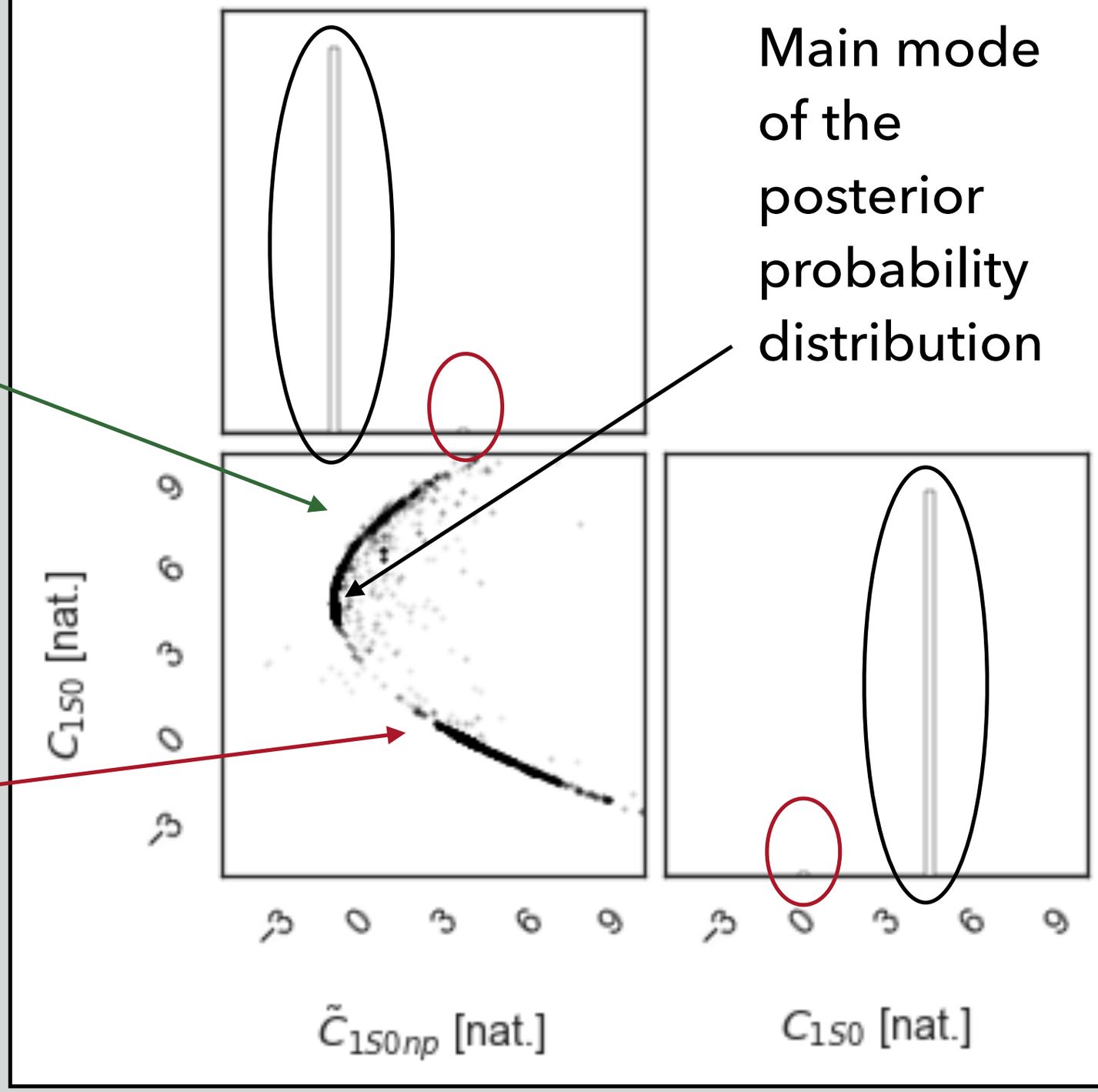


$$\tilde{C}_{1S0} = 1[\text{nat}] = \frac{4\pi}{f_\pi^2}$$

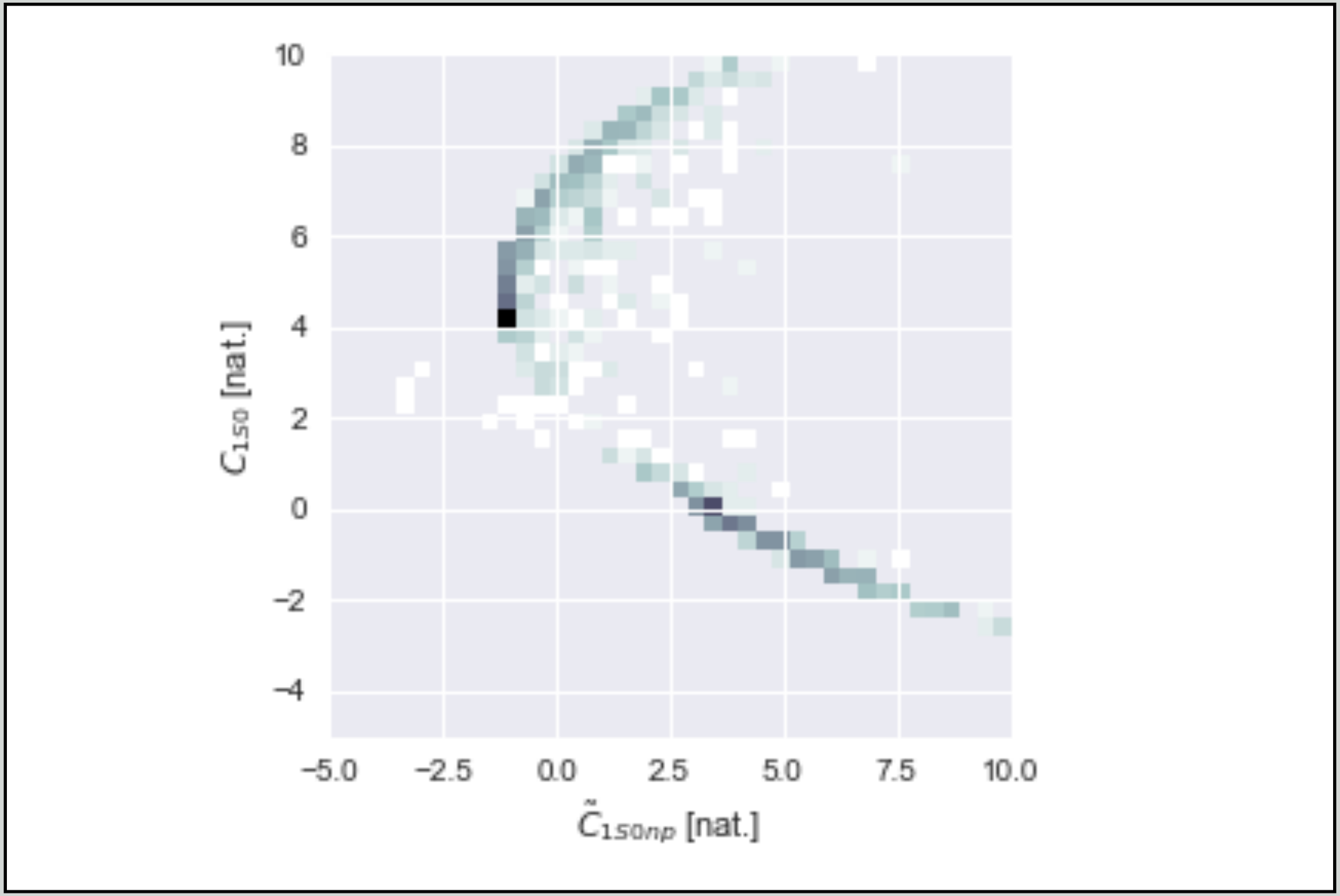
$$C_{1S0} = 1[\text{nat}] = \frac{4\pi}{f_\pi^2 \Lambda^2}$$

The 1S0(np) channel @ N2LO: Posterior distribution

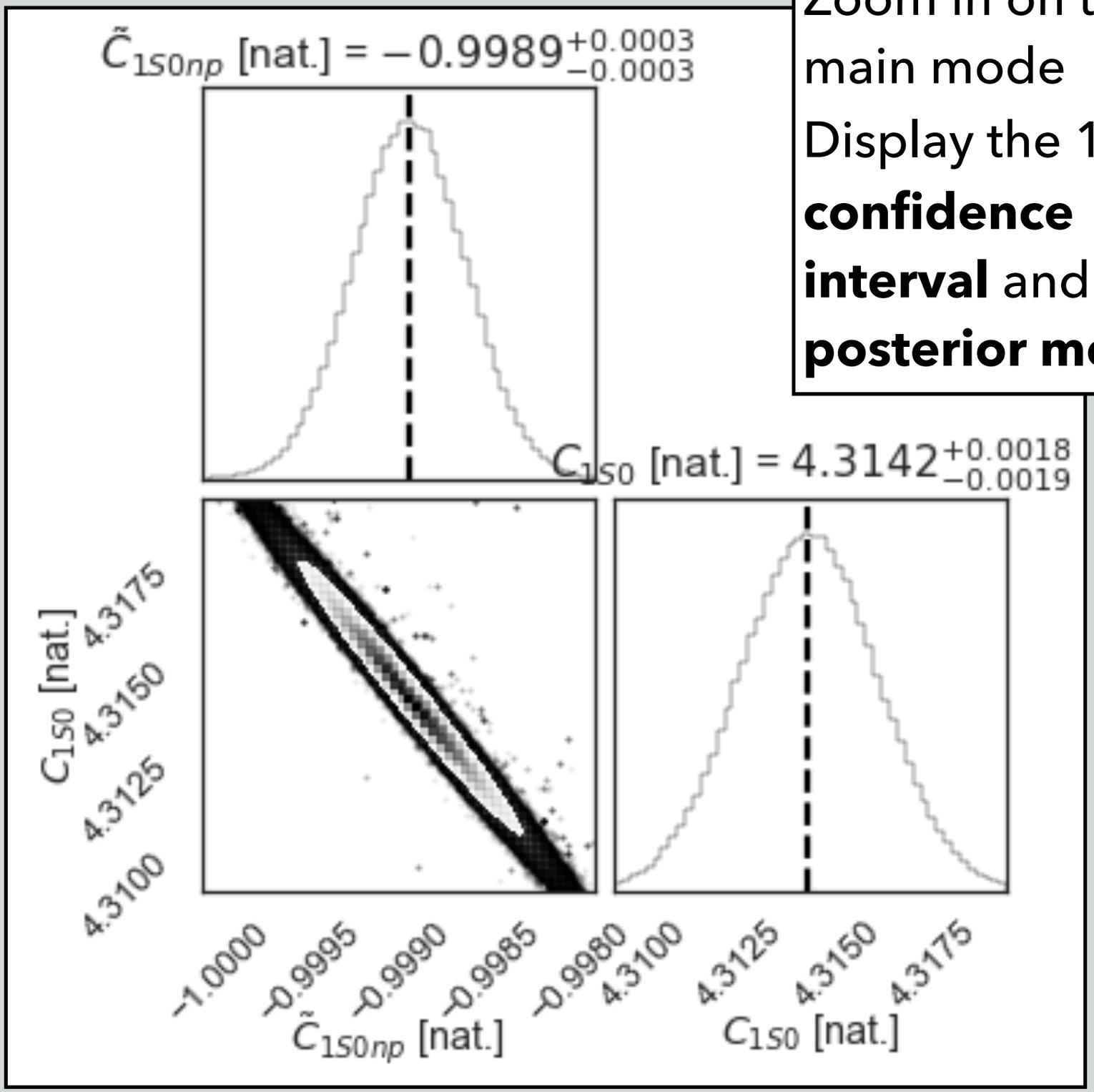
$$p(\tilde{C}_{1S0}, C_{1S0}|D, I) \propto p(D|\tilde{C}_{1S0}, C_{1S0}, I)p(\tilde{C}_{1S0}, C_{1S0}|I)$$



The 1S0(np) channel @ N2LO: Posterior distribution



The 1S0(np) channel @ N2LO: Posterior distribution



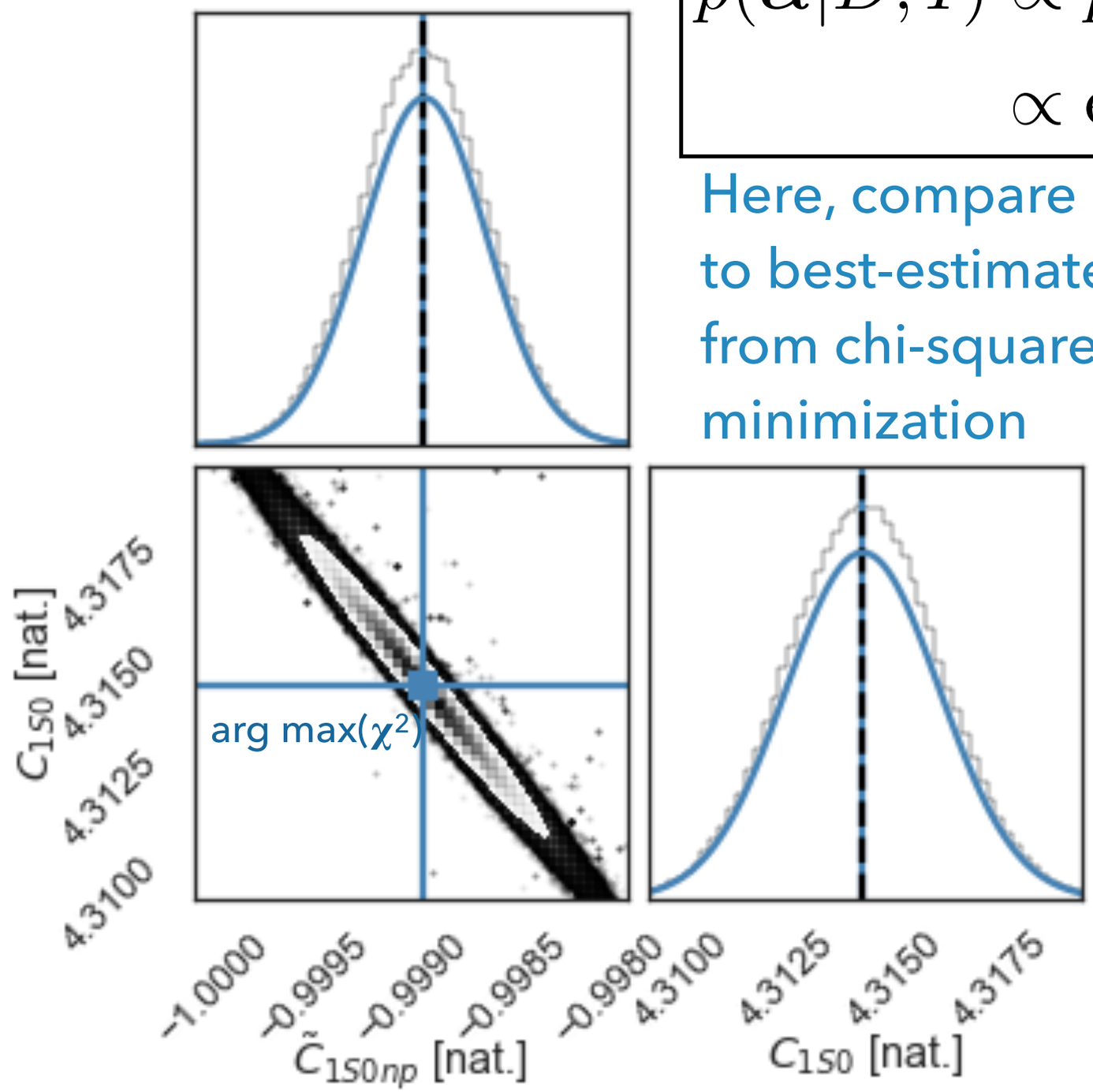
Zoom in on the main mode

Display the 1σ **confidence interval** and the **posterior mean**

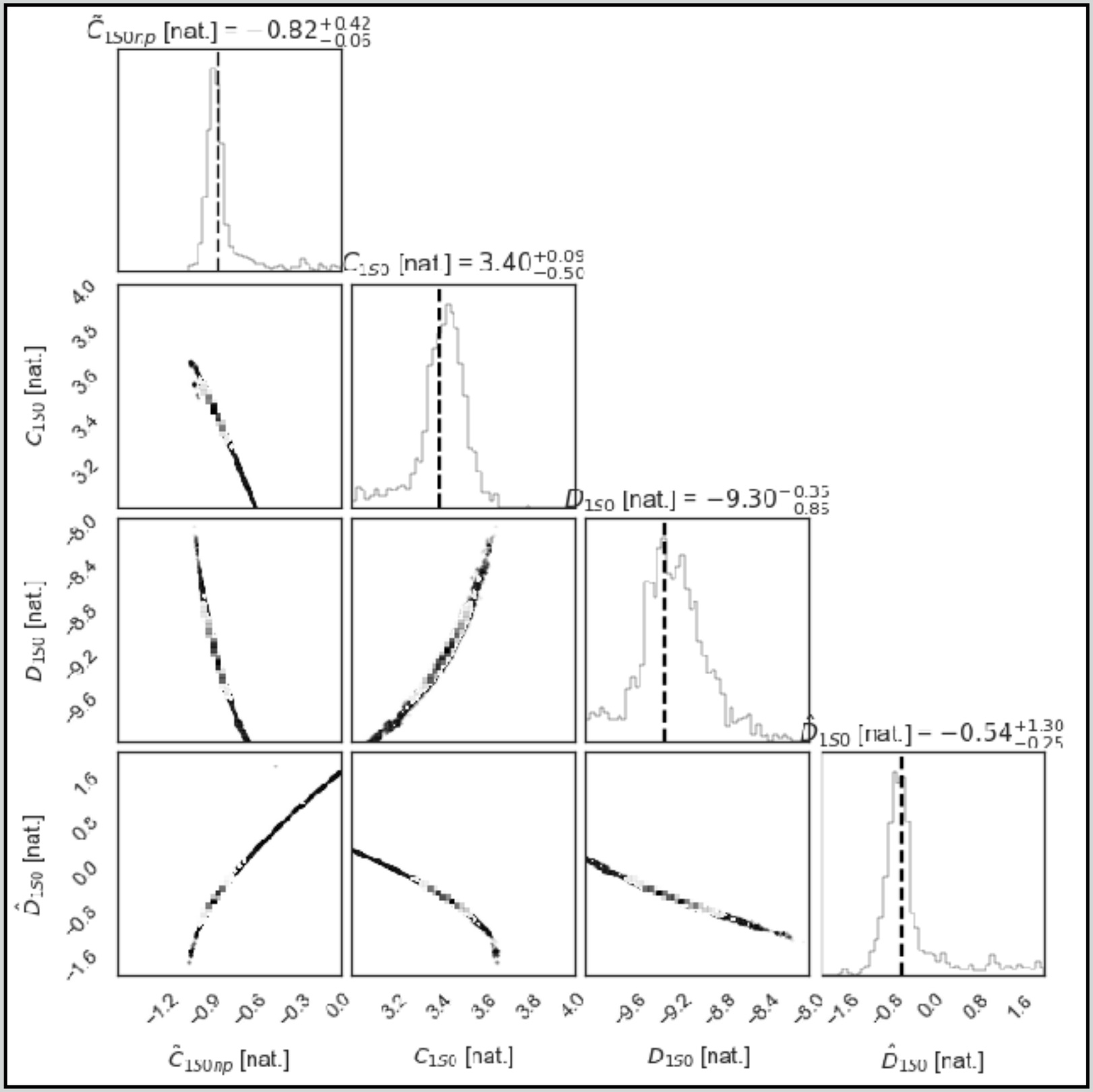
The 1S0(np) channel @ N2LO: Posterior distribution

$$p(\boldsymbol{\alpha}|D,I) \propto p(D|\boldsymbol{\alpha},I)$$
$$\propto \exp(-\chi^2/2)$$

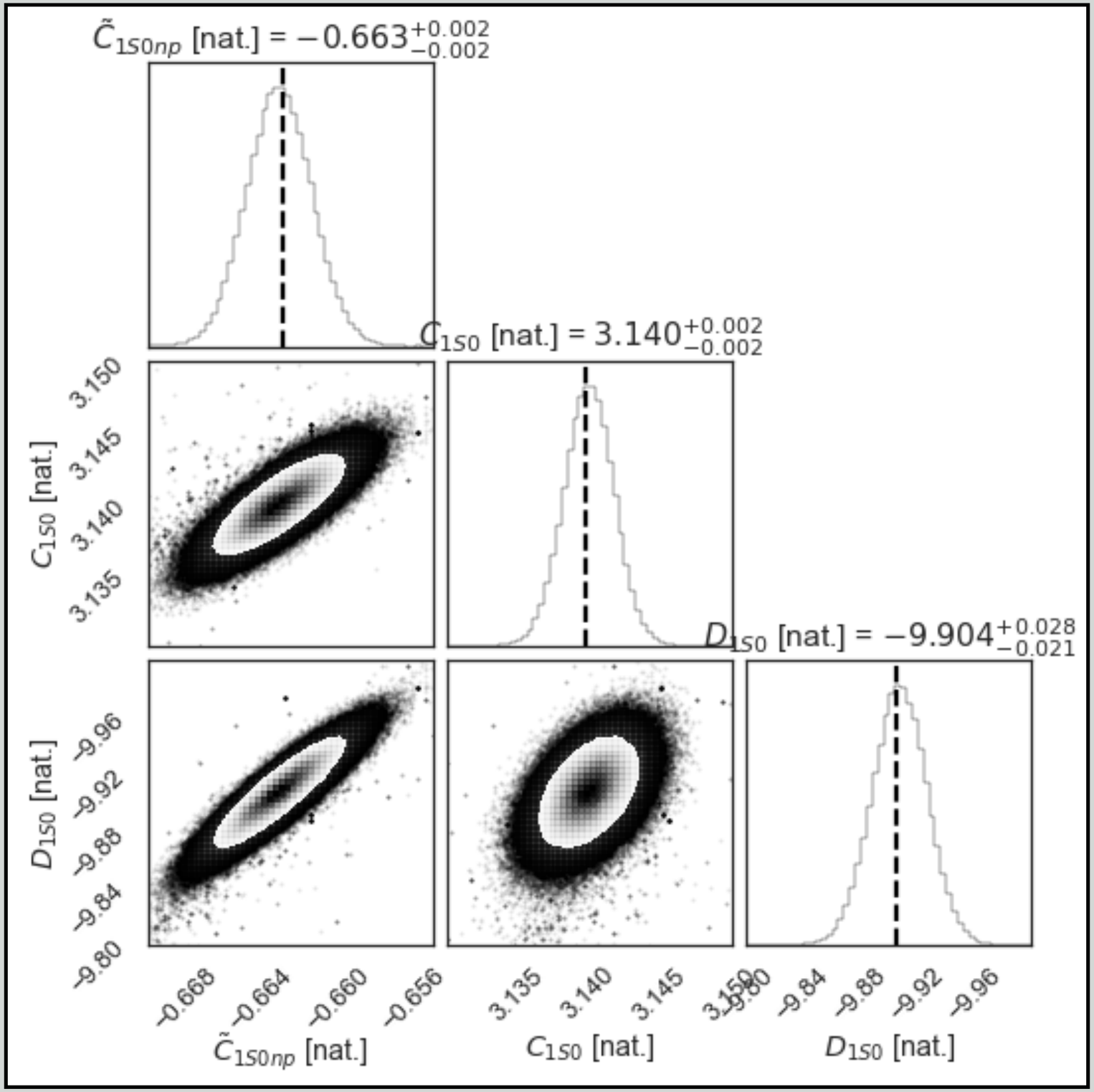
Here, compare to best-estimate from chi-squared minimization



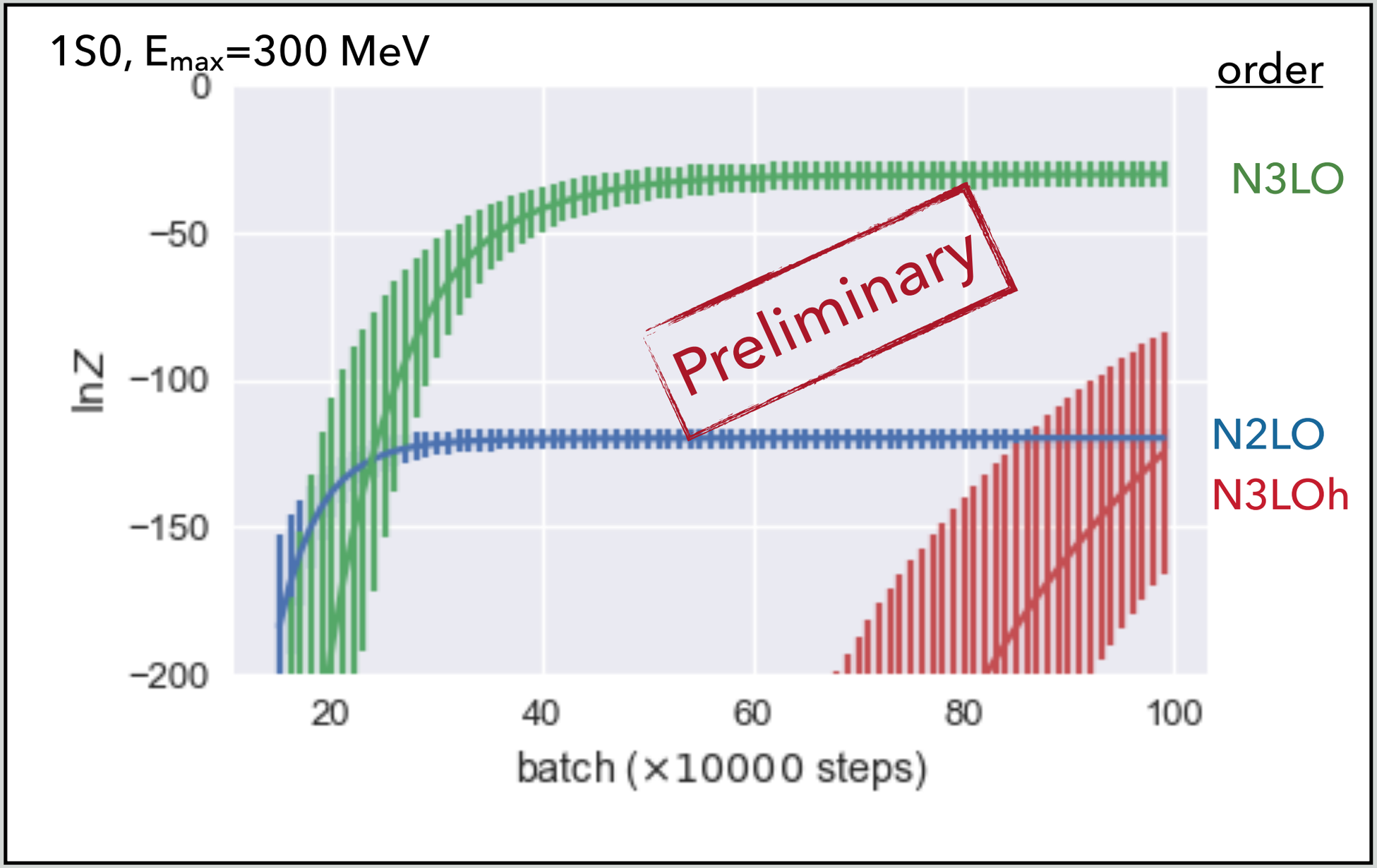
The 1S0(np) channel @ N3LO with redundant parameter



The 1S0(np) channel @ N3LO



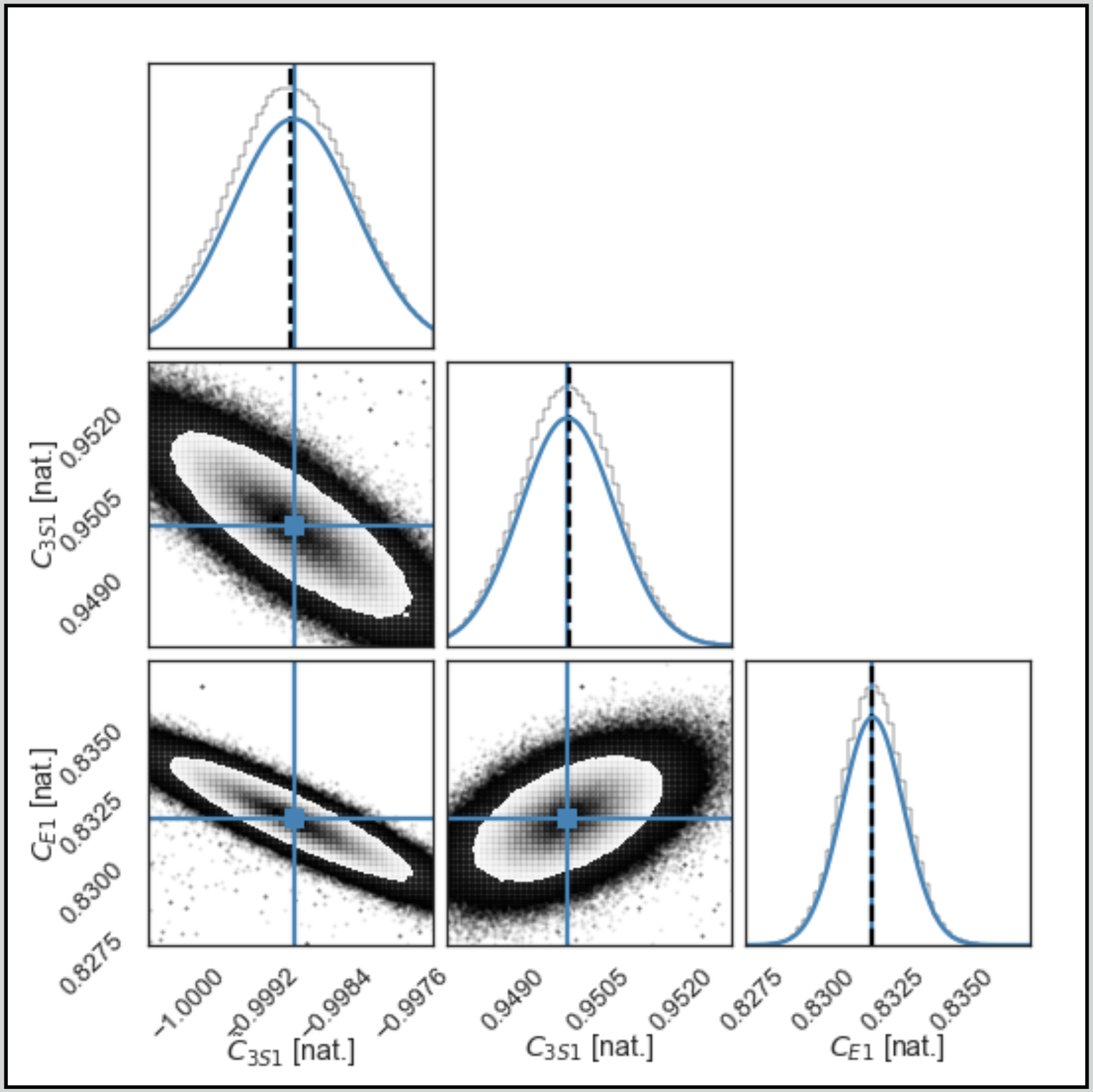
BAYESIAN EVIDENCE



$$p(\boldsymbol{\alpha}|D,I) = \frac{p(D|\boldsymbol{\alpha},I)p(\boldsymbol{\alpha}|I)}{p(D|I)}$$

Bayesian evidence (=Z)

The deuteron channel



Expectation integrals, error propagation

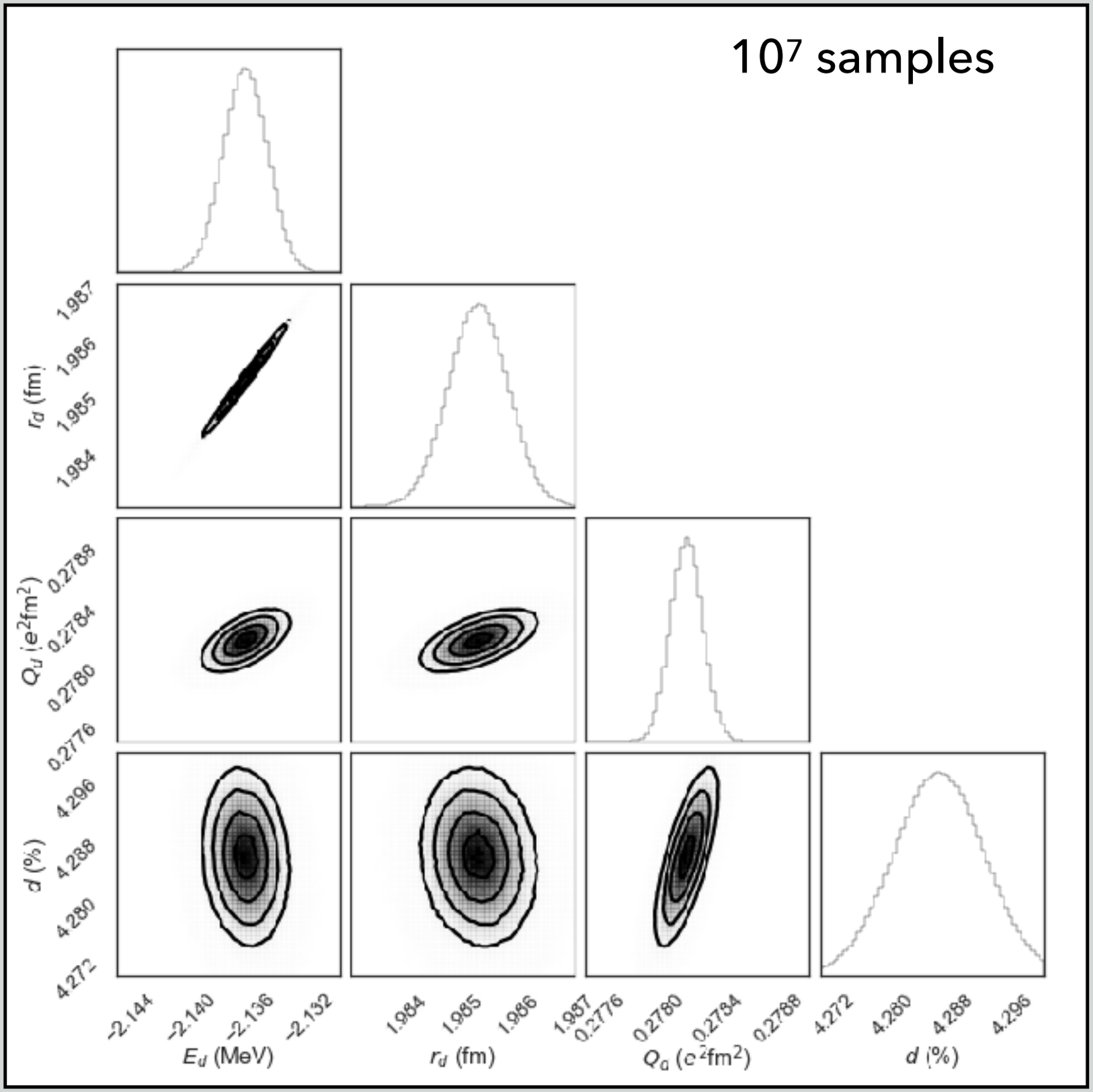
- Expectation integrals for observables can be performed using the posterior pdf

$$\langle O(\boldsymbol{\alpha})\rangle = \int d\boldsymbol{\alpha} p(\boldsymbol{\alpha}|D,I)O(\boldsymbol{\alpha})$$

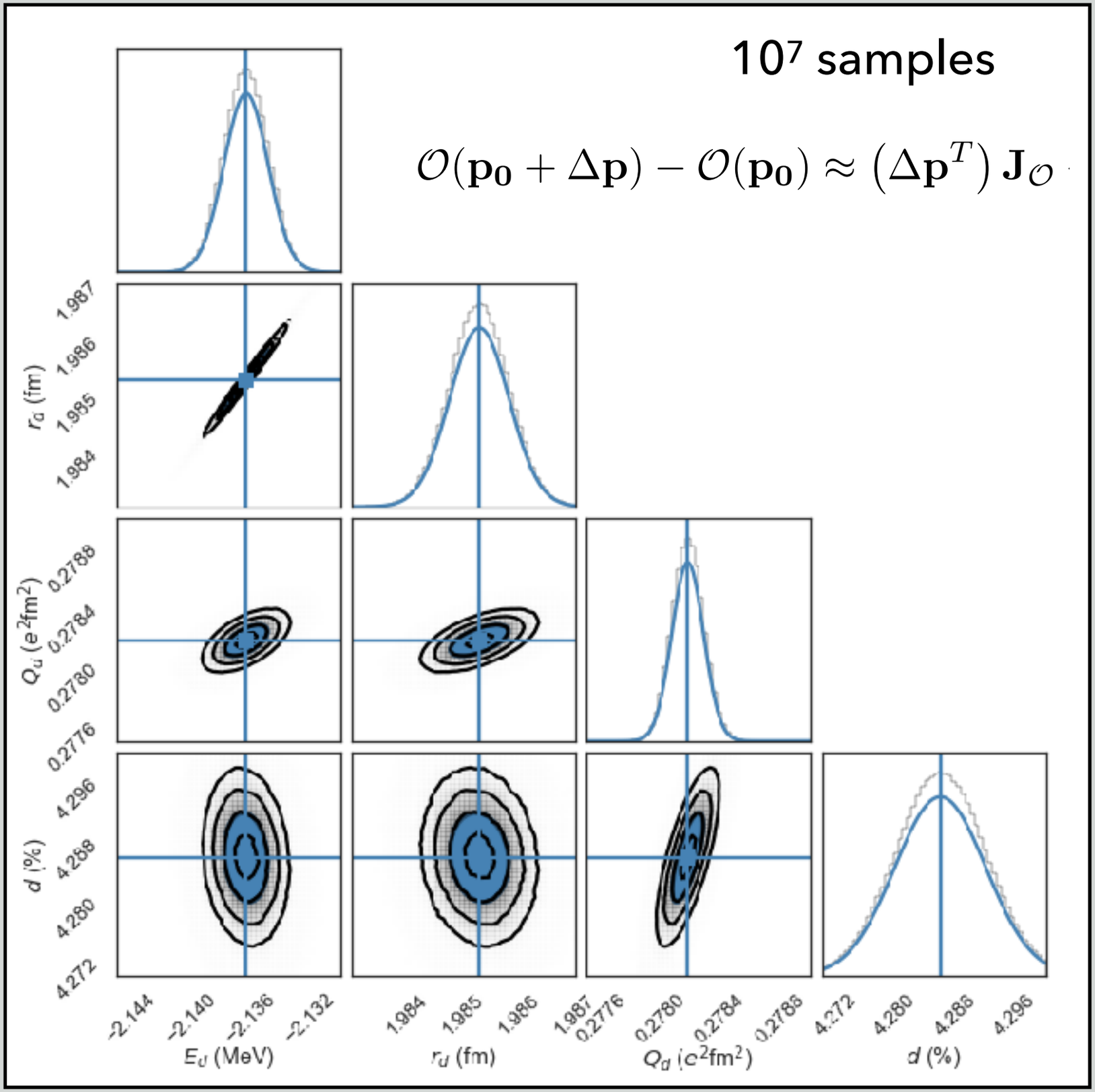
$$\approx \frac{1}{N}\sum_{j=1}^{N} O(\boldsymbol{\alpha}_j)$$

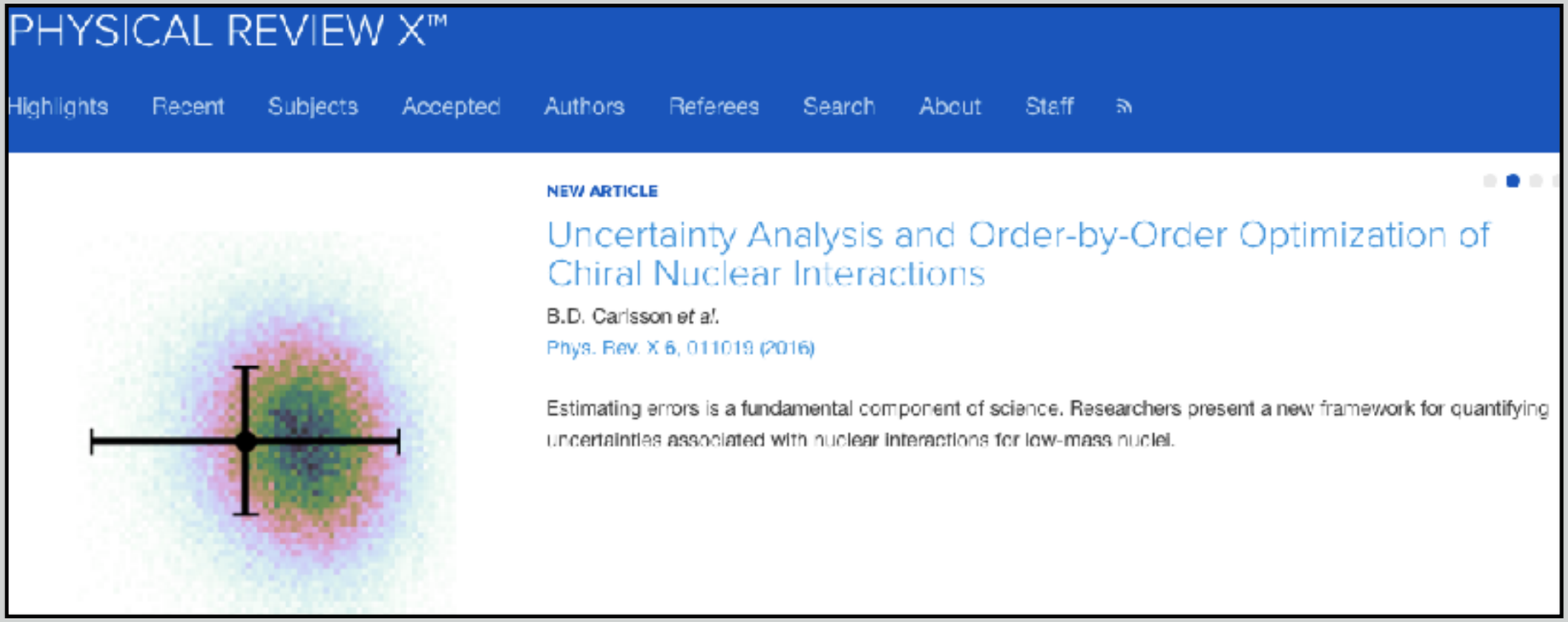
The MCMC algorithm generates *N* samples $\{\boldsymbol{\alpha}_j\}$ according to the posterior pdf

Deuteron observables



Deuteron observables





Note: Statistical analysis revisited following the *"Schiavilla-correction"* (relation between cD in the one-pion exchange plus contact NNN potential and the LEC multiplying the contact axial-vector current)

A DATA-CENTRIC VIEW ON NUCLEAR FORCES: CHI-SQUARED MINIMIZATION

OPTIMIZATION STRATEGY

Low-energy constants (LECs) need to be fitted to experimental data.

$$\chi^2(\vec{p}) \equiv \sum_i r_i^2(\vec{p}) = \boxed{\sum_{j\in NN} r_j^2(\vec{p}) + \sum_{k\in \pi N} r_k^2(\vec{p}) + \sum_{l\in 3N} r_l^2(\vec{p})}$$

 # parameters that are allowed to vary:

13	(1-3 per channel)	10	2+1=3	26: N2LO
13	(1-8 per channel)	25	2+1=3	41: N3LO
FIT πN-SECTOR TO πN DATA	FIT NN CONTACTS TO NN PHASE SHIFTS	FIT NN CONTACTS TO NN DATA	FIT C_D,C_E (+ Ct^{nn}_{1S0}) TO A=3 DATA	SIMULTANEOUS OPTIMIZATION

INPUT AND TECHNOLOGY

πN scattering

- WI08 database
- T_{lab} between 10-70 MeV
- N_{data} = 1347
- R. Workman et al. (2012)

NN scattering

- Granada '13 database
- T_{lab} between 0-290 MeV
- N_{data} = 4753 (np + pp)
- R. Navarro Pérez et al. (2013)

All 6000 residuals computed on 1 node in ~90 sec.

A=2,3 bound states

- ^{2}H,^{3}H,^{3}He [binding energy, radius, Q(^{2}H), ^{3}H half life]

On 1 node in ~10 sec

+ derivatives! (×2-20 cost)

Alternatively... theoretical analysis of data

πN scattering

- Roy-Steiner analysis
  M. Hoferichter et al. (2015)

NN scattering

- Phase shifts from partial wave analysis

OPTIMIZATION STRATEGY

Low-energy constants (LECs) need to be fitted to experimental data.

$$\chi^2(\vec{p}) \equiv \sum_i \left(\frac{O_i^{\text{theo}}(\vec{p}) - O_i^{\text{expr}}}{\sigma_{\text{tot},i}} \right)^2 \equiv \sum_i r_i^2(\vec{p})$$

- Derivative-free optimization using POUNDerS was used in our earliest works
- More efficient minimization algorithms (Levenberg-Marquardt, Newton), and statistical error analysis require **derivatives**

$$\frac{\partial r_i}{\partial p_j} \quad \text{and} \quad \frac{\partial^2 r_i}{\partial p_j \partial p_k}$$

- Numerical derivation using **finite differences** is plagued by **low numerical precision** and is **computationally costly.**
- Instead, we use **Automatic Differentiation** (AD)

TOTAL ERROR BUDGET

- The total error budget is

E.g. , NCSM Neglected

$$\sigma_{\rm tot}^2 = \sigma_{\rm exp}^2 + \sigma_{\rm method}^2 + \sigma_{\rm numerical}^2 + \sigma_{\rm model}^2$$

- At a given chiral order v, the omitted diagrams should be of order

$$\mathcal{O}\left((Q/\Lambda_\chi)^{\nu+1}\right)$$

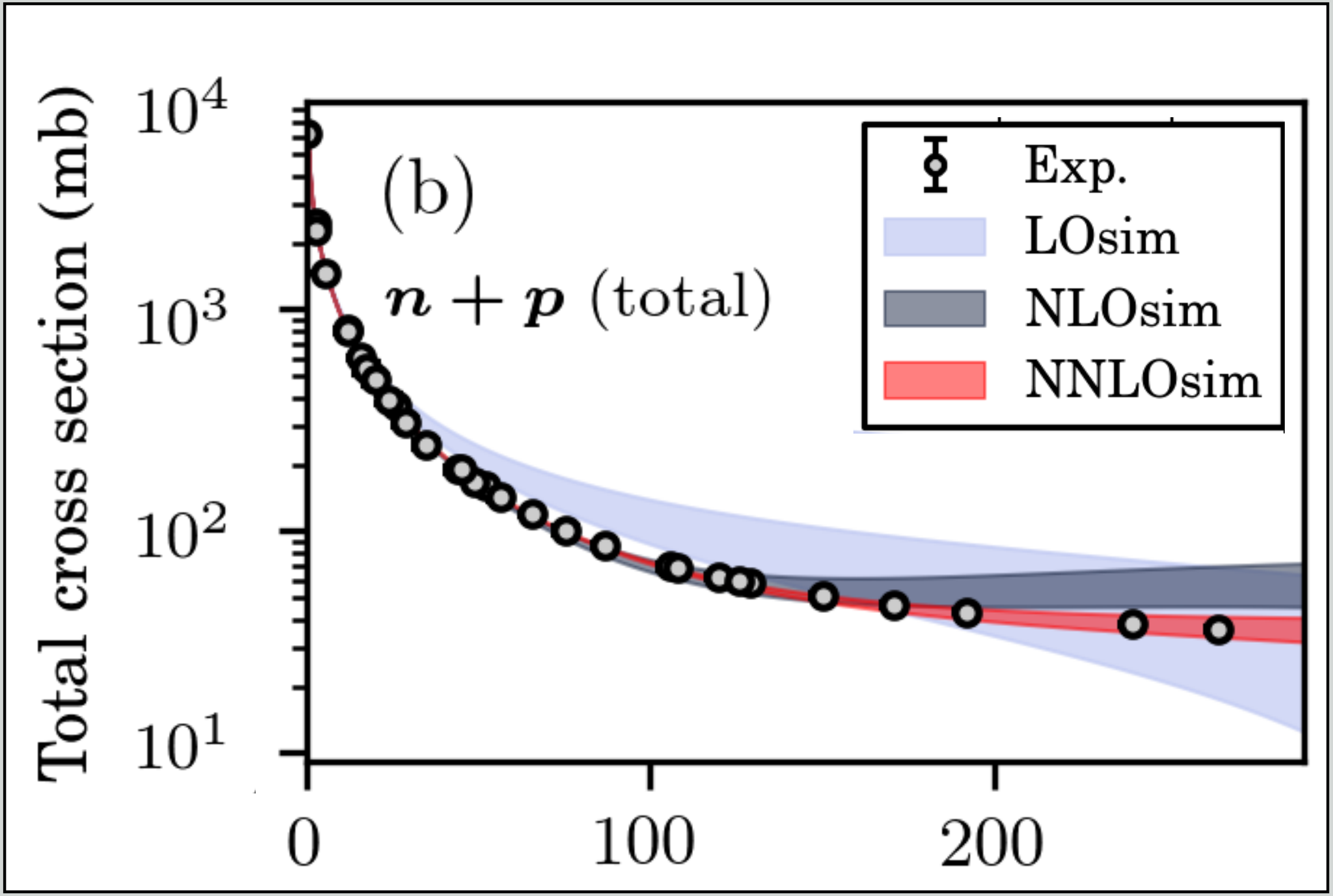
- Still needs to be converted to actual numbers $\sigma_{\rm model}$

- We translate this EFT knowledge into an error in the scattering amplitudes

$$\sigma_{\rm model,x}^{\rm (amp)} = C_x \left(\frac{Q}{\Lambda_\chi}\right)^{\nu+1}, \quad x \in \{NN, \pi N\}$$

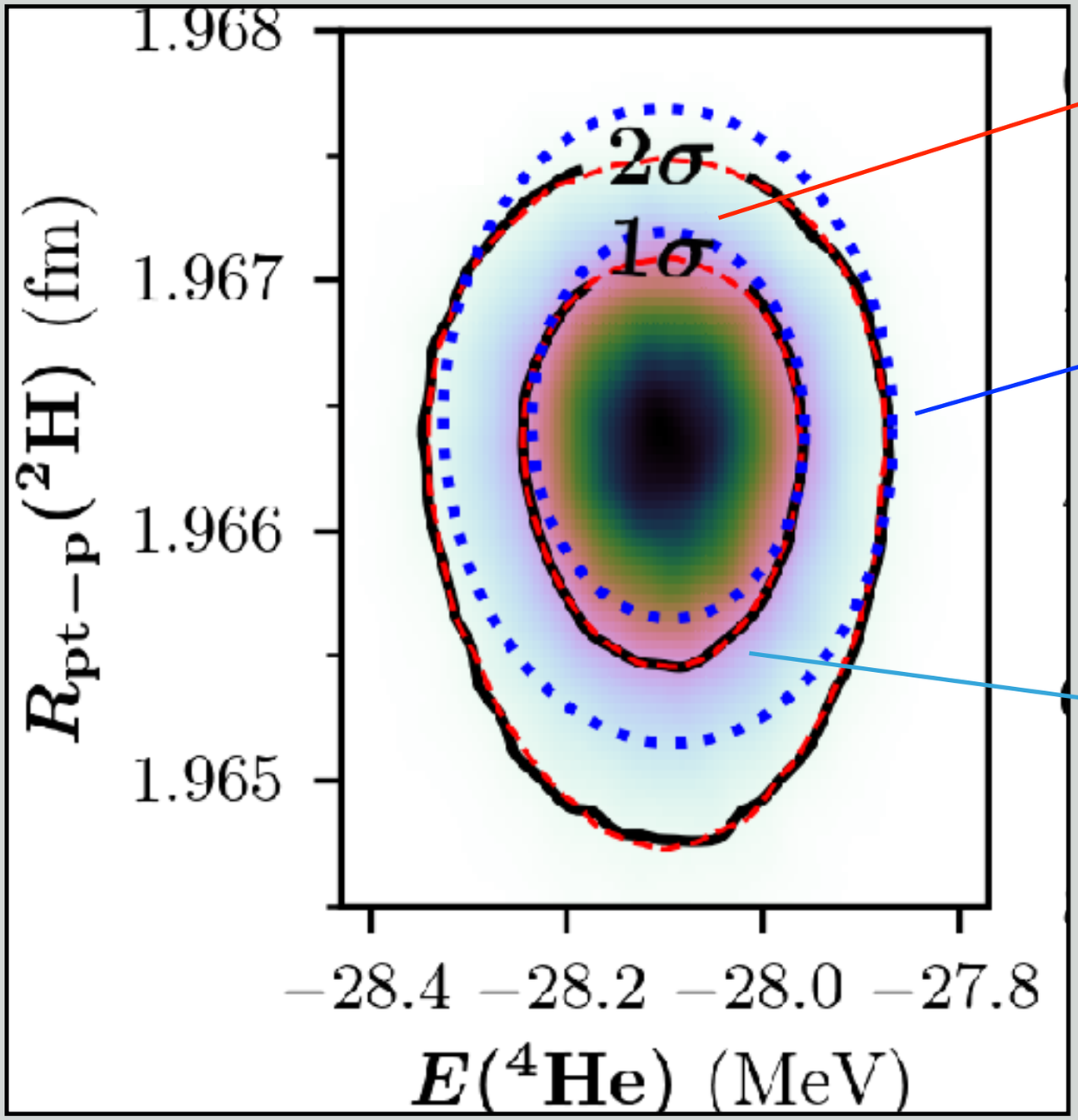
- which is then propagated to an error in the observable.

TOTAL NP CROSS SECTION



Quadratic error propagation vs Brute force sampling

$$O(\mathbf{p}) \approx O(\mathbf{p}_0) + J_O \Delta\mathbf{p} + \frac{1}{2}\Delta\mathbf{p}^T H_O \Delta\mathbf{p}$$



Quadratic error prop.

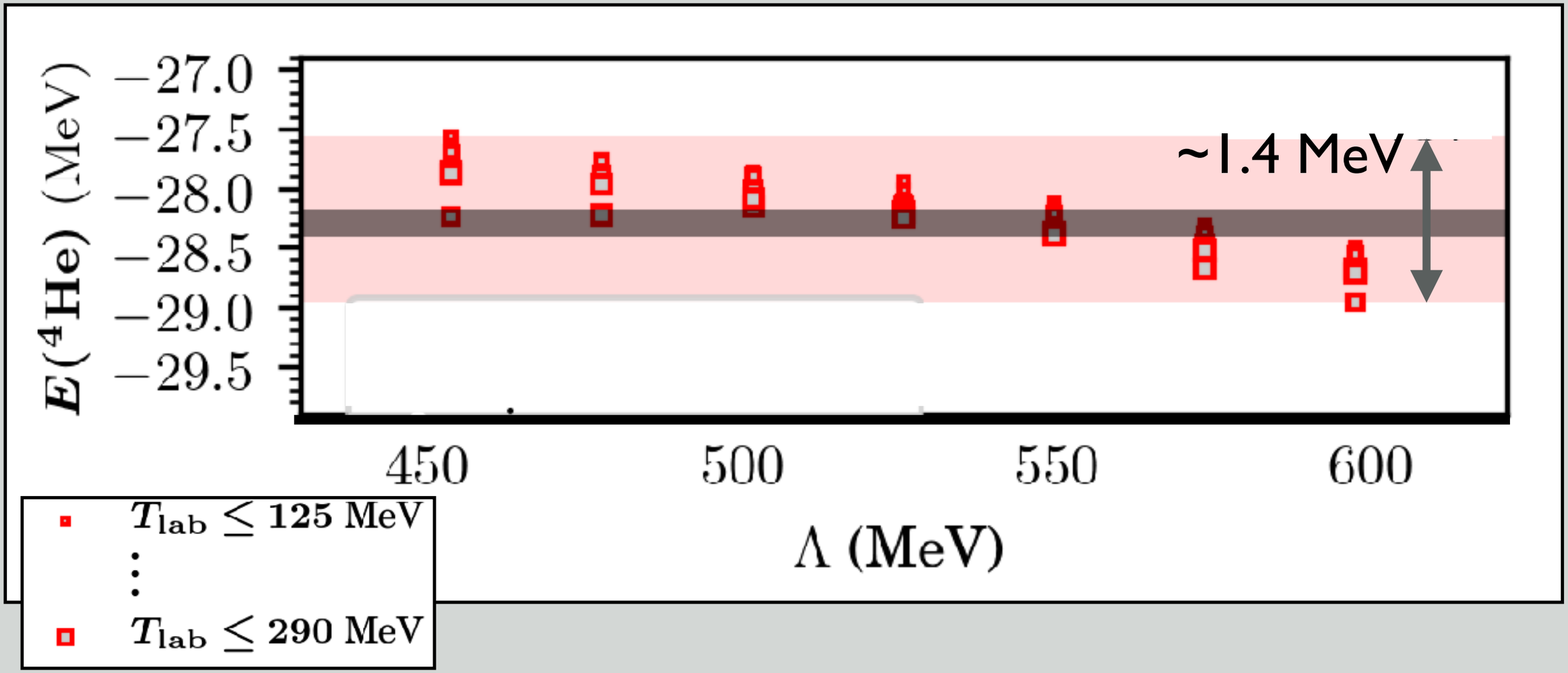
Linear error prop.

Monte Carlo prob. dens.

$E(^4He) = -28.10^{+8}_{-10}$ (MeV)

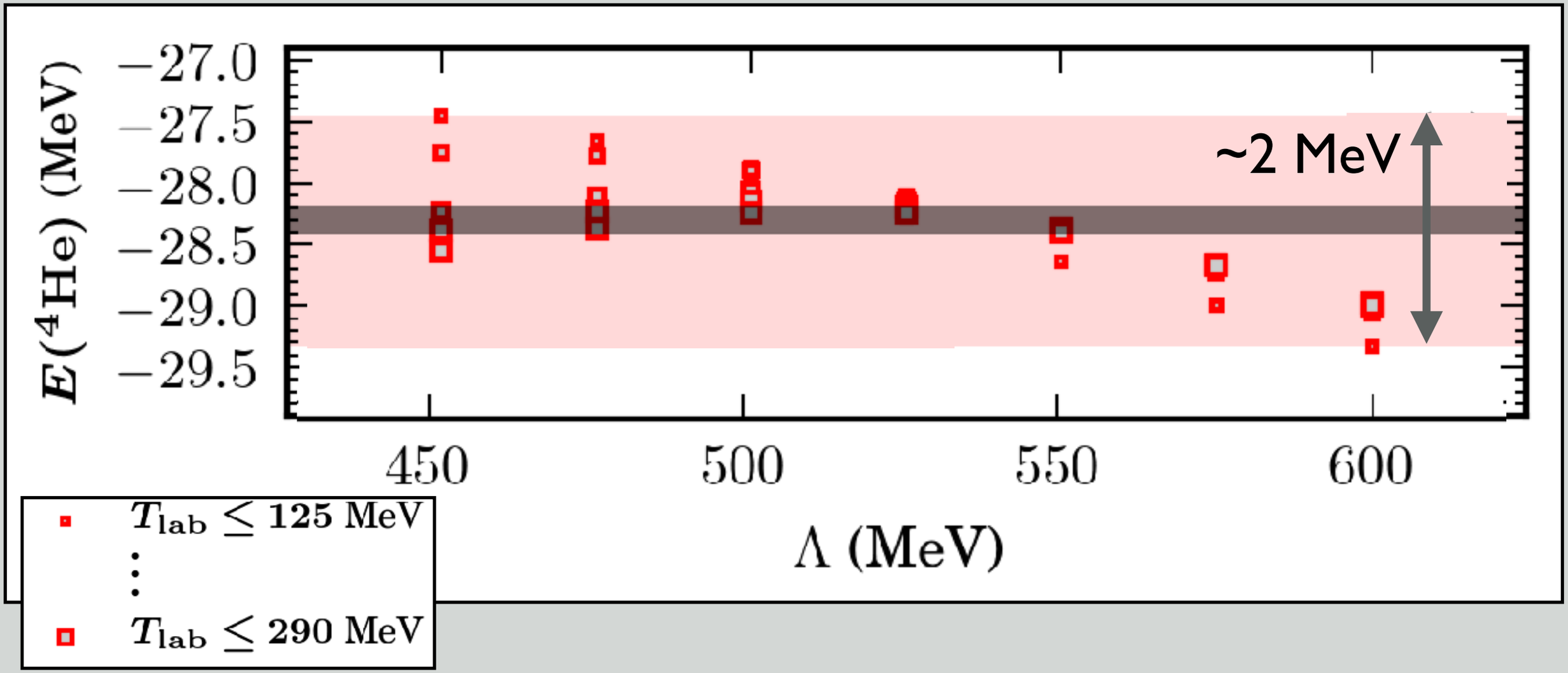
Statistical uncertainty

Systematic uncertainties: input data, regulator cutoff



- 7 different regulator cutoffs: Λ=450, 475, …, 575, 600 MeV
- 6 different *NN*-scattering datasets $T_{lab} \in [0, T_{lab,max}]$, with $T_{lab,max}$=125, …, 290 MeV

Systematic uncertainties: input data, regulator cutoff



- 7 different regulator cutoffs: Λ=450, 475, …, 575, 600 MeV
- 6 different *NN*-scattering datasets $T_{lab} \in [0, T_{lab,max}]$, with $T_{lab,max}$=125, …, 290 MeV

Previous version

CONCLUSION

OUTLOOK

- The inclusion of more data in the objective function requires other approaches to the optimization problem needed. (See also Andreas' presentation.)
- The frequentist approach does not offer an easy and transparent method for handling systematic uncertainties or imposing prior knowledge.
- Bayesian parameter estimation is advantageous, but costly.
  - avoiding the need to 'judge', a priori, what data can be included in order to safely avoid overfitting.
  - offers a viable approach to include prior knowledge of certain parameters from Roy-Steiner analysis.
- Investigate other chiral EFT power-counting schemes via Bayesian evidence.

QUESTIONS

- How to best combine different kinds of experimental data (such as NN scattering and A=3 bound-state observables)?
- And how to combine this with (prior) information from a theoretical analysis (RS eqs., sub-threshold parameters)?
- Feasible ways to evaluate multi-dimensional posterior distributions, and perform sub-sequent error propagation?
- What are the most relevant tests of various power counting schemes given the Bayesian framework?