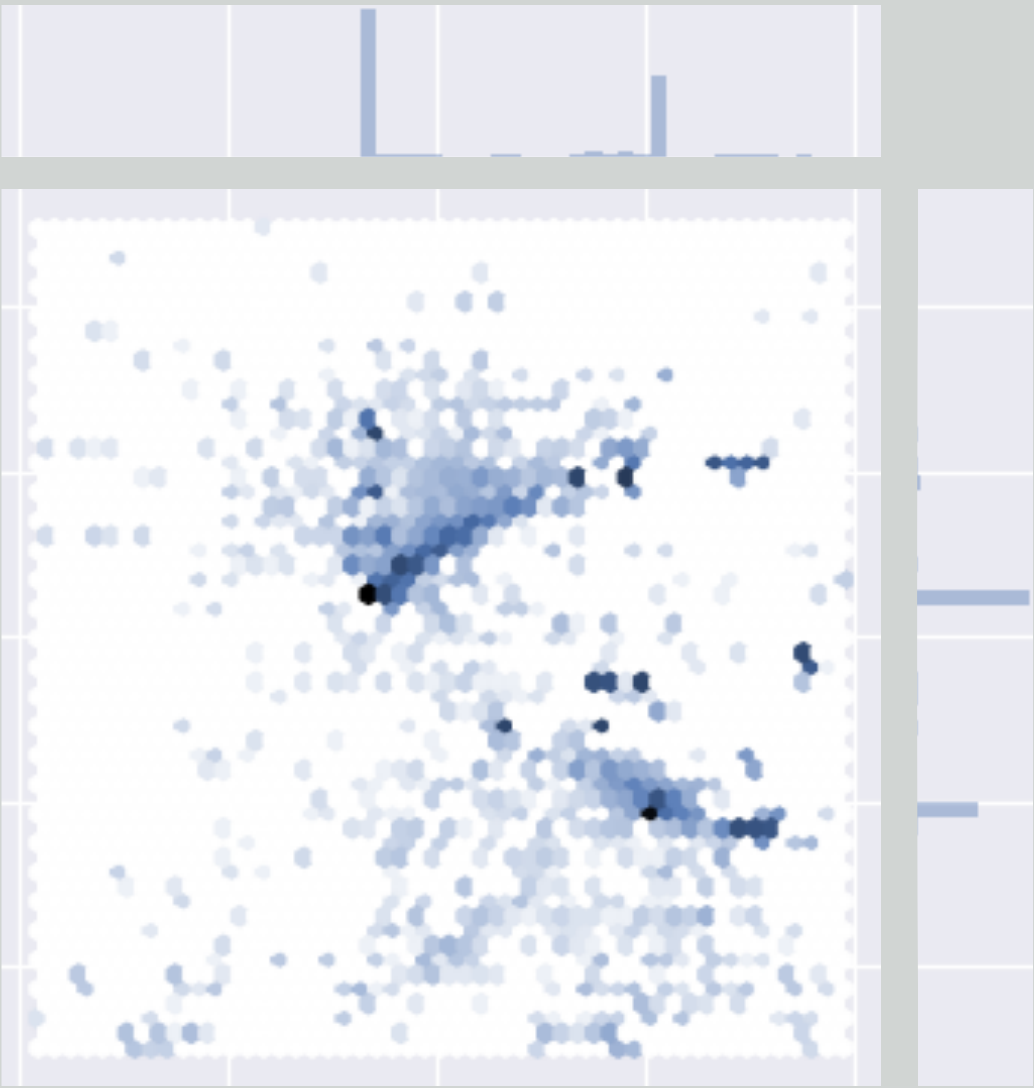


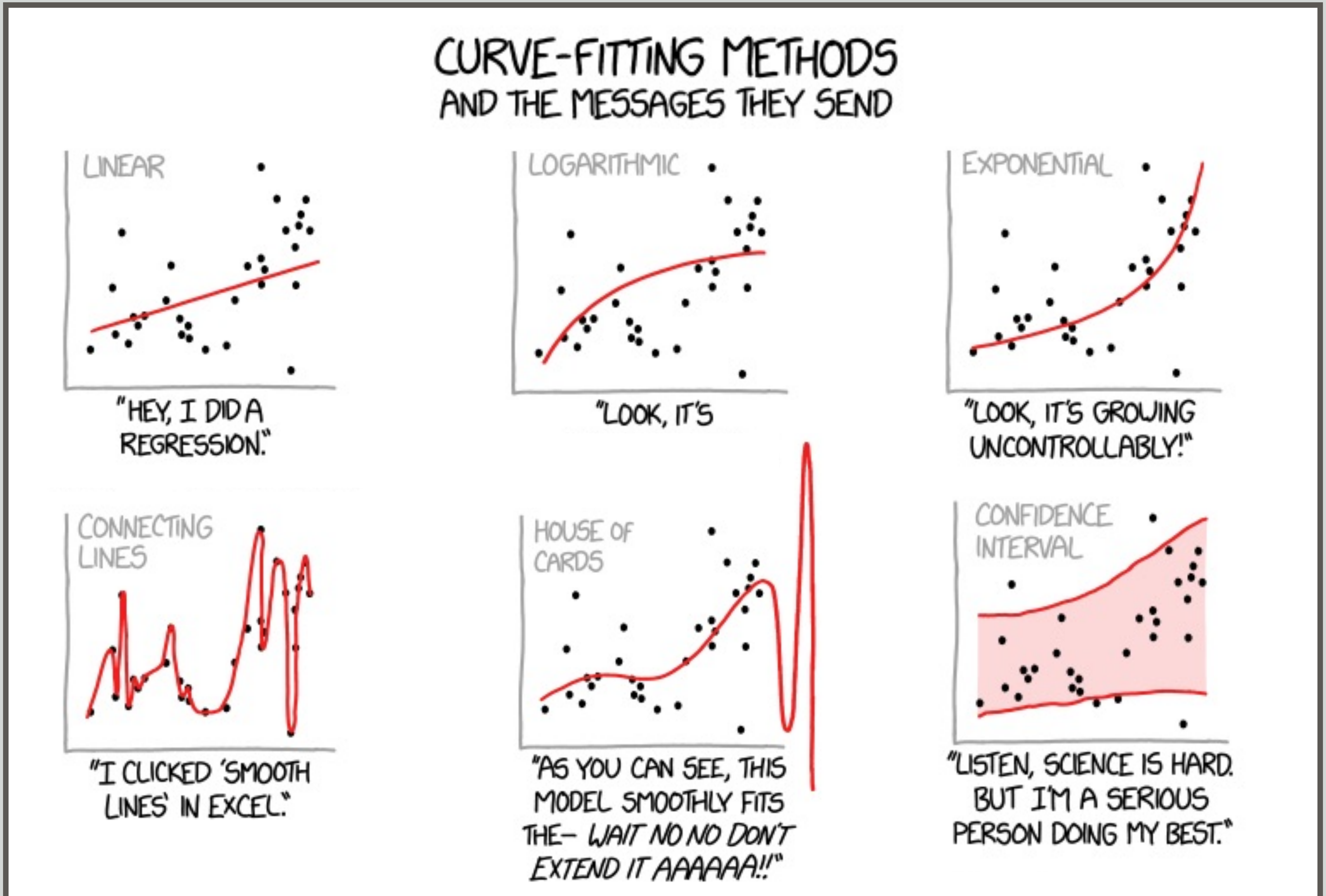
BAYESIAN PARAMETER ESTIMATION AND MODEL EVIDENCE FOR NUCLEAR EFT

CHRISTIAN FORSSÉN

Department of Physics, Chalmers, Sweden

ISNET-6, Darmstadt, Oct 9-12, 2018

XKCD


CURVE-FITTING METHODS
AND THE MESSAGES THEY SEND
LINEAR
"HEY, I DID A REGRESSION."
LOGARITHMIC
"LOOK, IT'S
EXPONENTIAL
"LOOK, IT'S GROWING UNCONTROLLABLY!"
CONNECTING LINES
"I CLICKED 'SMOOTH LINES' IN EXCEL."
HOUSE OF CARDS
"AS YOU CAN SEE, THIS MODEL SMOOTHLY FITS THE- WAIT NO NO DON'T EXTEND IT AAAAAA!!"
CONFIDENCE INTERVAL
"LISTEN, SCIENCE IS HARD. BUT I'M A SERIOUS PERSON DOING MY BEST."

MANY THANKS TO MY COLLABORATORS

- **Andreas Johansson**, Boris Carlsson, Andreas Ekström**,** Isak Svensson (Chalmers)

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

Research funded by:

- STINT
- European Research Council
- Swedish Research Council

INTRODUCTION

TWO-BULLET OVERVIEW OF THIS TALK

- What information can be inferred[1] from available data[2] to state–of-the-art "models"[3] of the strong force between nucleons?
- How can we compute the evidence for a particular "model"[3]?

[1] using a Bayesian approach.

[2] Here: πN and NN scattering data

[3] effective field theories => systematically improvable

QUESTIONS

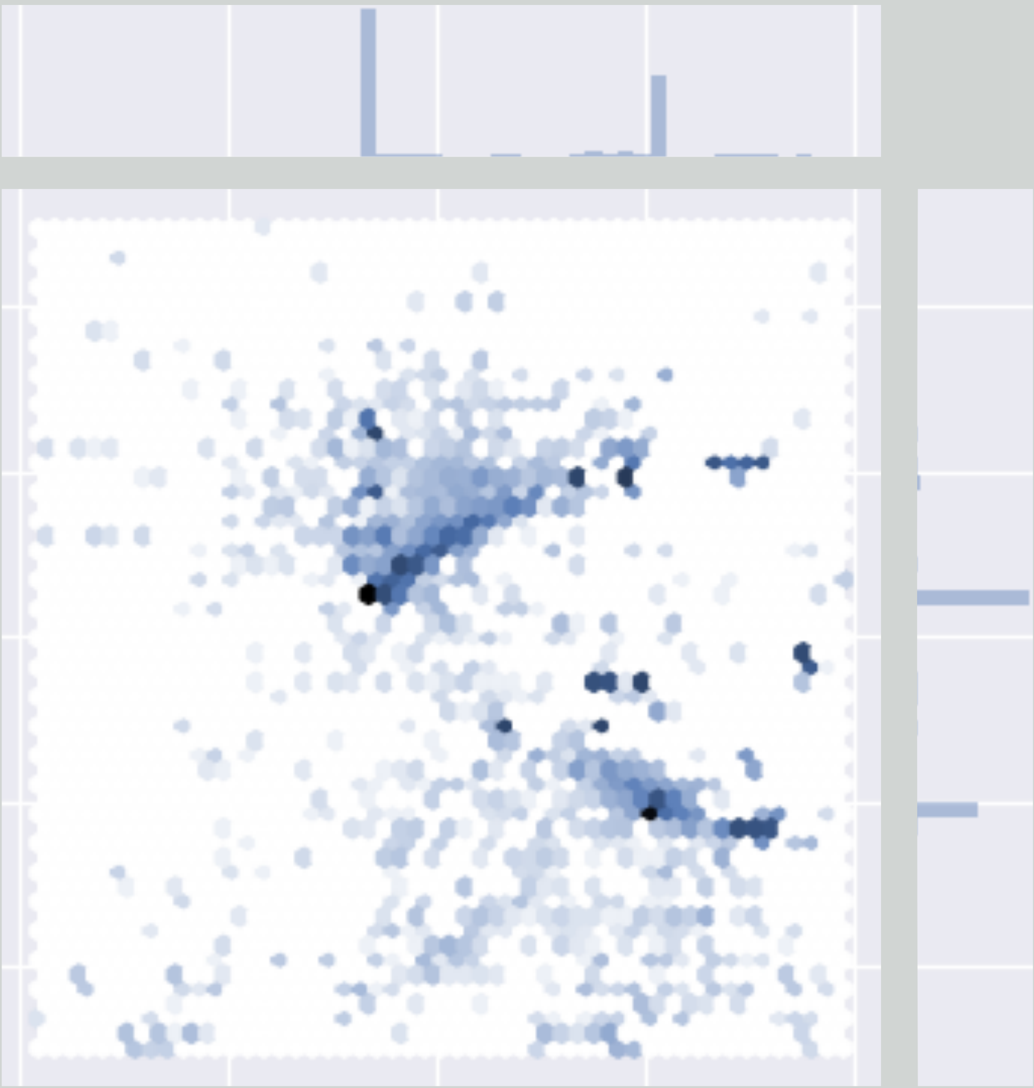
- Best-practice for evidence calculation (norm of posterior pdf from MCMC samples)
- Feasible (approximate) Bayes for multi-dimensional posterior distributions, expensive likelihood evaluations, and sub-sequent error propagation?
- Feasible ways to include EFT truncation errors (see also Sarah’s and Jordan’s presentations)?
- Diagnostics of MCMC convergence?
- Physics: What are the most relevant tests of various power counting schemes given the Bayesian framework?

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*, . . . → hypothesis, *H*?

- Assume that **hypothesis $\boldsymbol{H_i}$** is a model M_i with parameters $\boldsymbol{\theta}_i$.
- **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.



Work with A. Johansson and A. Ekström

BAYESIAN POSTERIORS IN THE NUCLEON-NUCLEON SECTOR

Parametric models

- Assume that hypothesis H_i is a model M_i with parameters $\boldsymbol{\theta}_i$.
- 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: $p(\boldsymbol{\theta}_i \mid D_{\mathrm{obs}}, M_i, I)$
- **Model comparison**:
  Compute ratio: $p(D_{\mathrm{obs}} \mid M_i, I) / p(D_{\mathrm{obs}} \mid M_j, I)$

Bayesian parameter estimation

Bayes' theorem (follows from probability product rule):

posterior likelihood prior

$$p\left(\boldsymbol{\theta} \,|\, D, I\right) = \frac{p(D \,|\, \boldsymbol{\theta}, I)\, p(\boldsymbol{\theta} \,|\, I)}{p(D \,|\, I)}$$

Bayesian evidence

Marginalization: $$p\left(\theta_1 \,|\, D, I\right) = \int d\theta_2 \ldots d\theta_k\, p\left(\boldsymbol{\theta} \,|\, D, I\right)$$

- 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.

Bayesian parameter estimation

Bayes' theorem (follows from probability product rule):

posterior likelihood prior

$$p\left(\boldsymbol{\theta}\,|\,D,I\right) = \frac{p(D\,|\,\boldsymbol{\theta},I)p(\boldsymbol{\theta}\,|\,I)}{p(D\,|\,I)}$$

Bayesian evidence

posterior $p(\boldsymbol{\theta}) \equiv p\left(\boldsymbol{\theta}\,|\,D,I\right)$

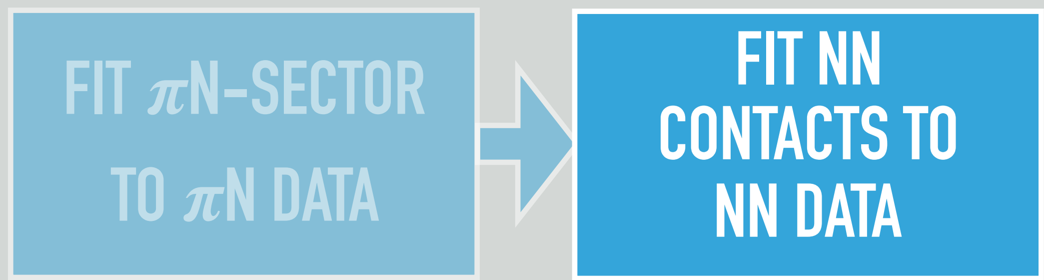
likelihood $L(\boldsymbol{\theta}) \equiv p\left(D\,|\,\boldsymbol{\theta},I\right)$

prior $\pi(\boldsymbol{\theta}) \equiv p\left(\boldsymbol{\theta}\,|\,I\right)$

Bayesian evidence $Z \equiv p\left(D\,|\,I\right) = \int d\boldsymbol{\theta}\, L\left(\boldsymbol{\theta}\right)\pi\left(\boldsymbol{\theta}\right)$

LO EFT AND NN SCATTERING DATA

Let us consider the task of determining a set of low-energy constants (LECs) by confrontation with experimental NN scattering data.

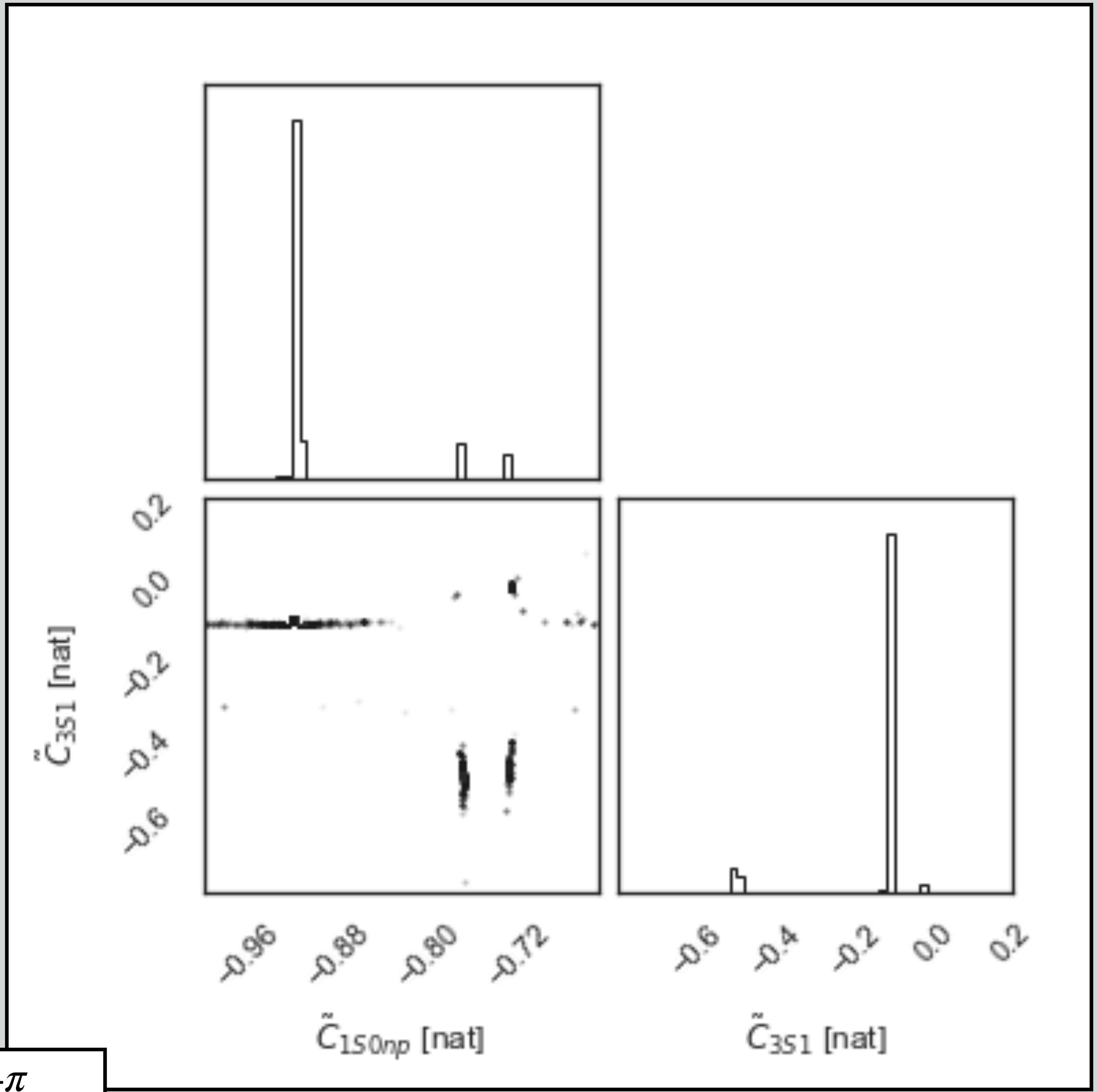


(Leading order)

$$p\left(\tilde{C}_{1S0}, \tilde{C}_{3S1} \,|\, D, I\right) \propto p\left(D \,|\, \tilde{C}_{1S0}, \tilde{C}_{3S1}, I\right) p\left(\tilde{C}_{1S0}, \tilde{C}_{3S1} \,|\, I\right)$$

posterior — likelihood — prior

NN SCATTERING DATA OPTIMIZATION



Leading order. X-sec data up to 10 MeV without (!) truncation error

How do we best sample a multimodal posterior pdf?

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

MCMC SAMPLING WITH PARALLEL TEMPERING (PT-MCMC)

- Perform MCMC sampling in parallel from tempered versions of the posterior distribution

$$p(\boldsymbol{\theta}) \equiv p\left(\boldsymbol{\theta} \,|\, D, I\right)$$

$$L(\boldsymbol{\theta}) \equiv p\left(D \,|\, \boldsymbol{\theta}, I\right)$$

$$\pi(\boldsymbol{\theta}) \equiv p\left(\boldsymbol{\theta} \,|\, I\right)$$

$$p_T(\boldsymbol{\theta}) \equiv L(\boldsymbol{\theta})^{1/T} \pi(\boldsymbol{\theta})$$

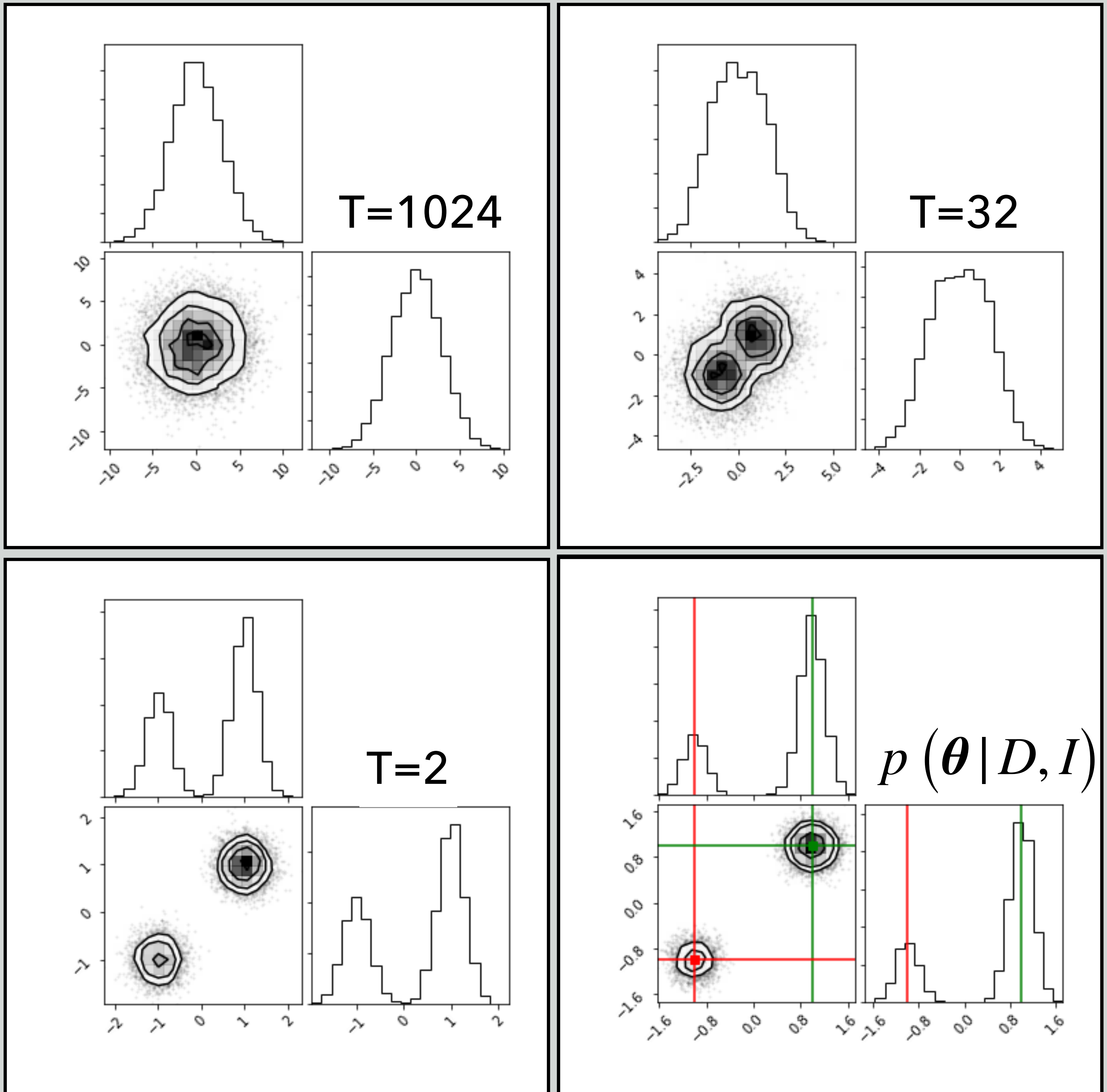
- Employ a well-chosen(?) temperature ladder $T_1 < T_2 < \ldots < T_N$, where $T_1 = 1$ corresponds to the target distribution.

- Swaps $\boldsymbol{\theta}_i \leftrightarrow \boldsymbol{\theta}_j$ can occur at pre-determined intervals with probability

$$\beta \equiv 1/T$$

$$A_{i,j} = \min\left\{ \left(\frac{L(\boldsymbol{\theta}_i)}{L(\boldsymbol{\theta}_j)} \right)^{\beta_j - \beta_i}, 1 \right\}$$

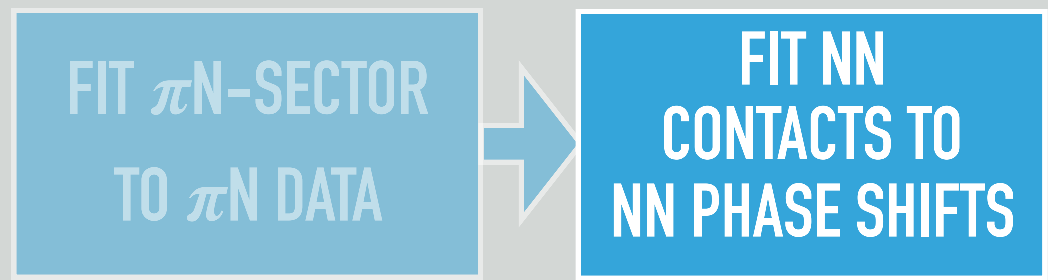
PT-MCMC MULTIMODAL EXAMPLE



$$p_T(\boldsymbol{\theta}) \equiv L(\boldsymbol{\theta})^{1/T} \pi(\boldsymbol{\theta})$$

HIGHER ORDER EFT AND NN PHASE SHIFT

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

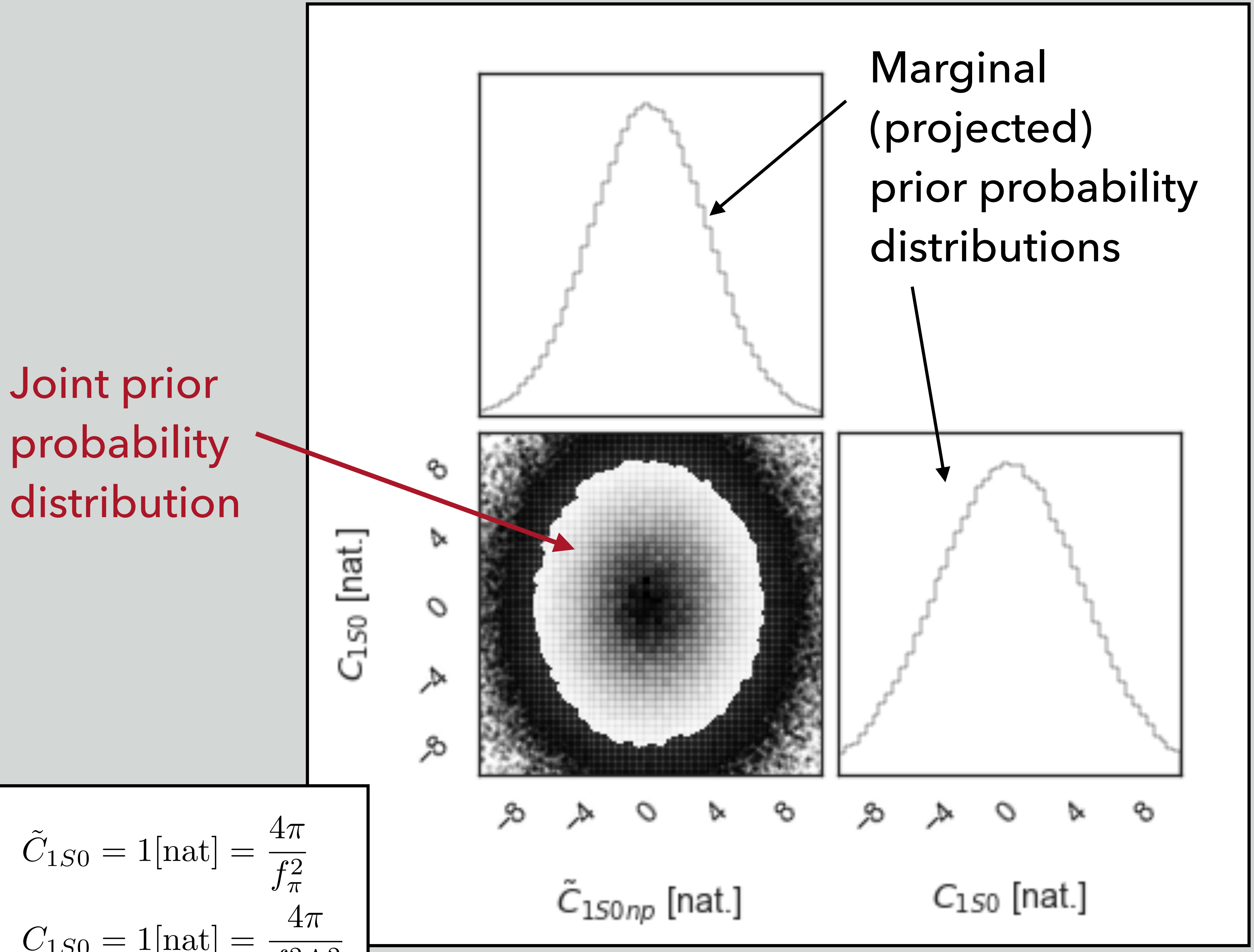


$$p(\tilde{C}_{1S0}, C_{1S0}|D, I) \propto p(D|\tilde{C}_{1S0}, C_{1S0}, I)p(\tilde{C}_{1S0}, C_{1S0}|I)$$

posterior likelihood prior

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

1S0 channel



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

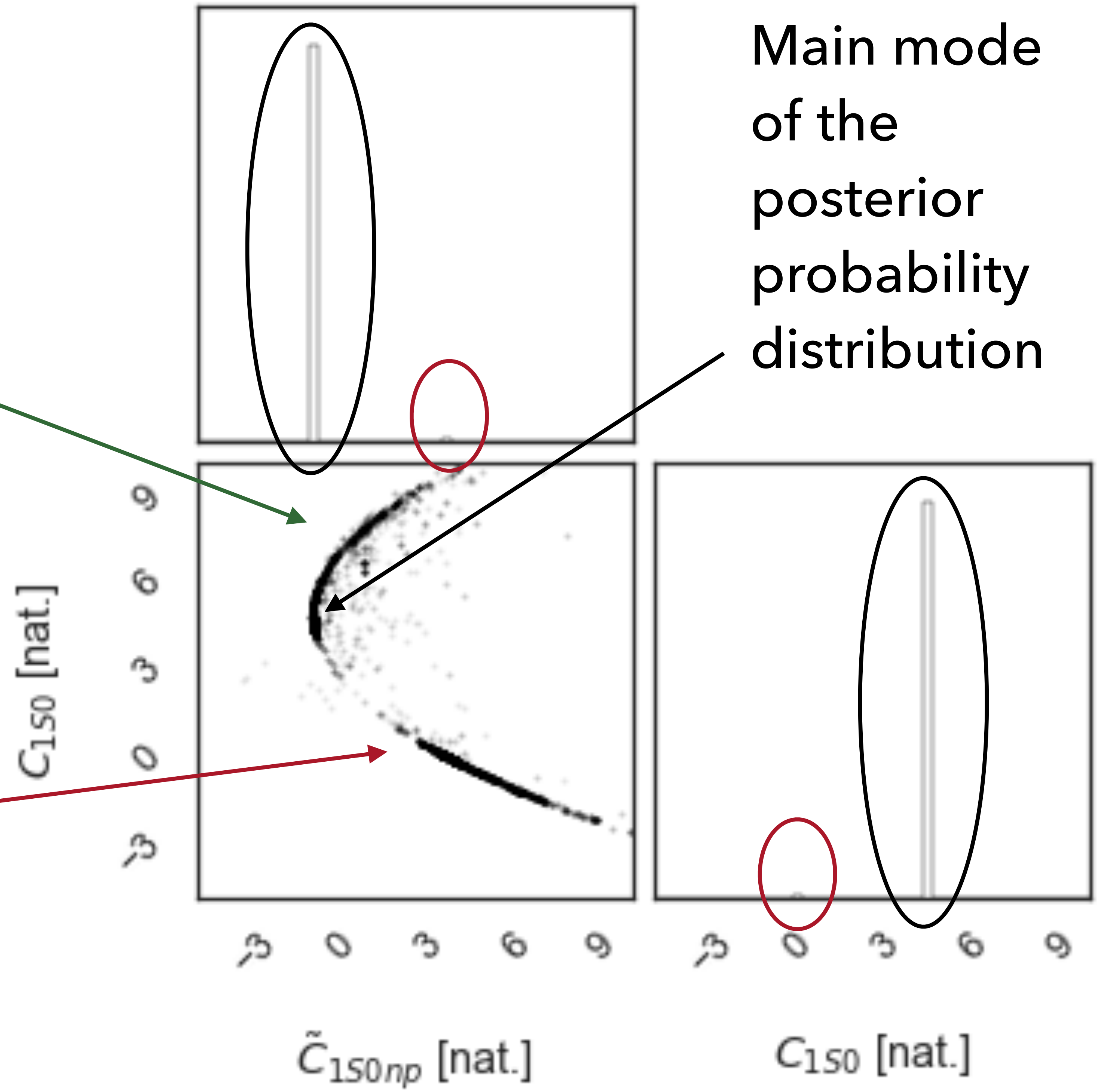
$$C_{1S0} = 1[\text{nat}] = \frac{4\pi}{f_\pi^2 \Lambda^2}$$

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

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

PT-MCMC samples

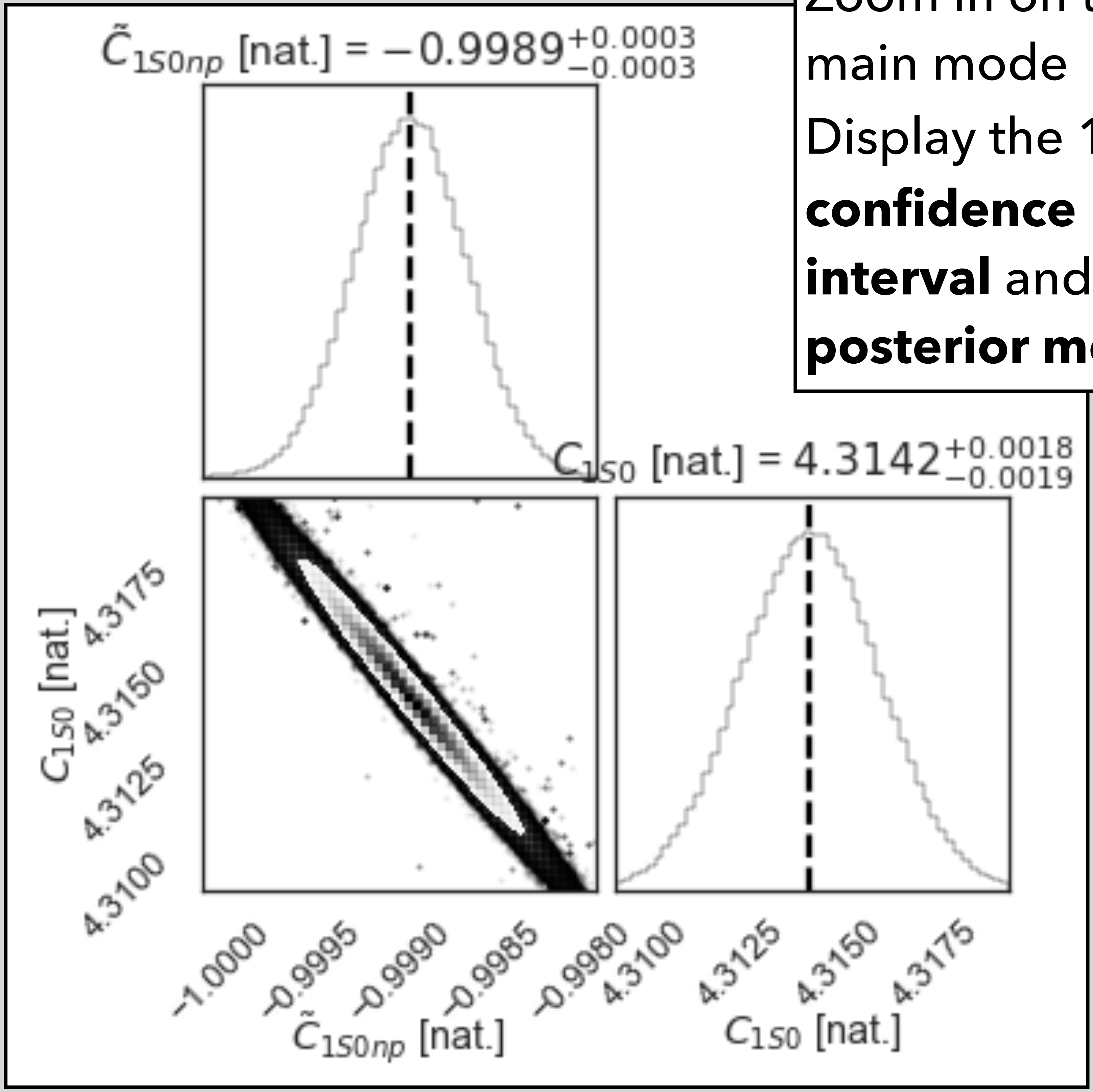
emcee The MCMC Hammer



Main mode of the posterior probability distribution

Second mode

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

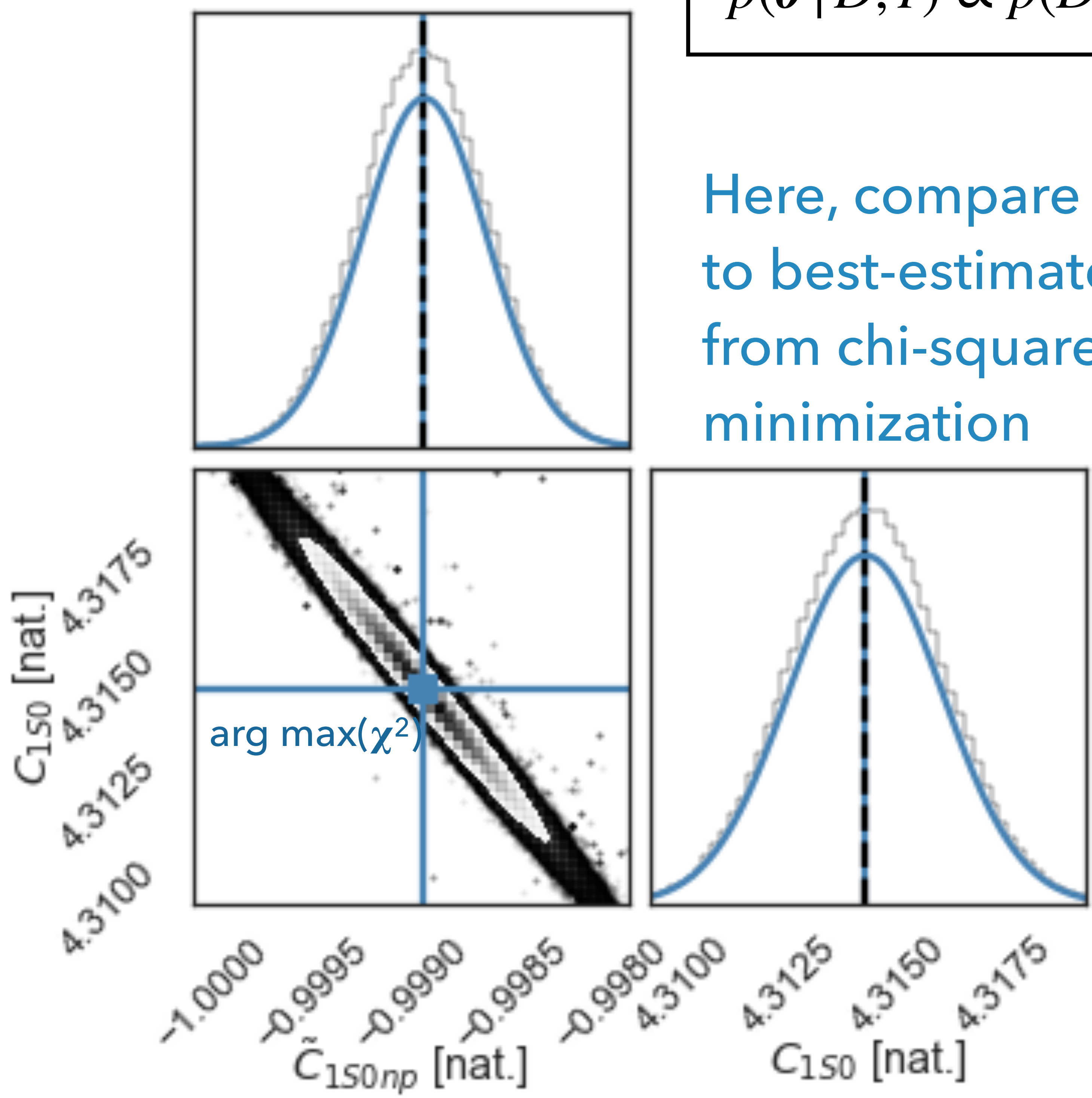


Zoom in on the main mode

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

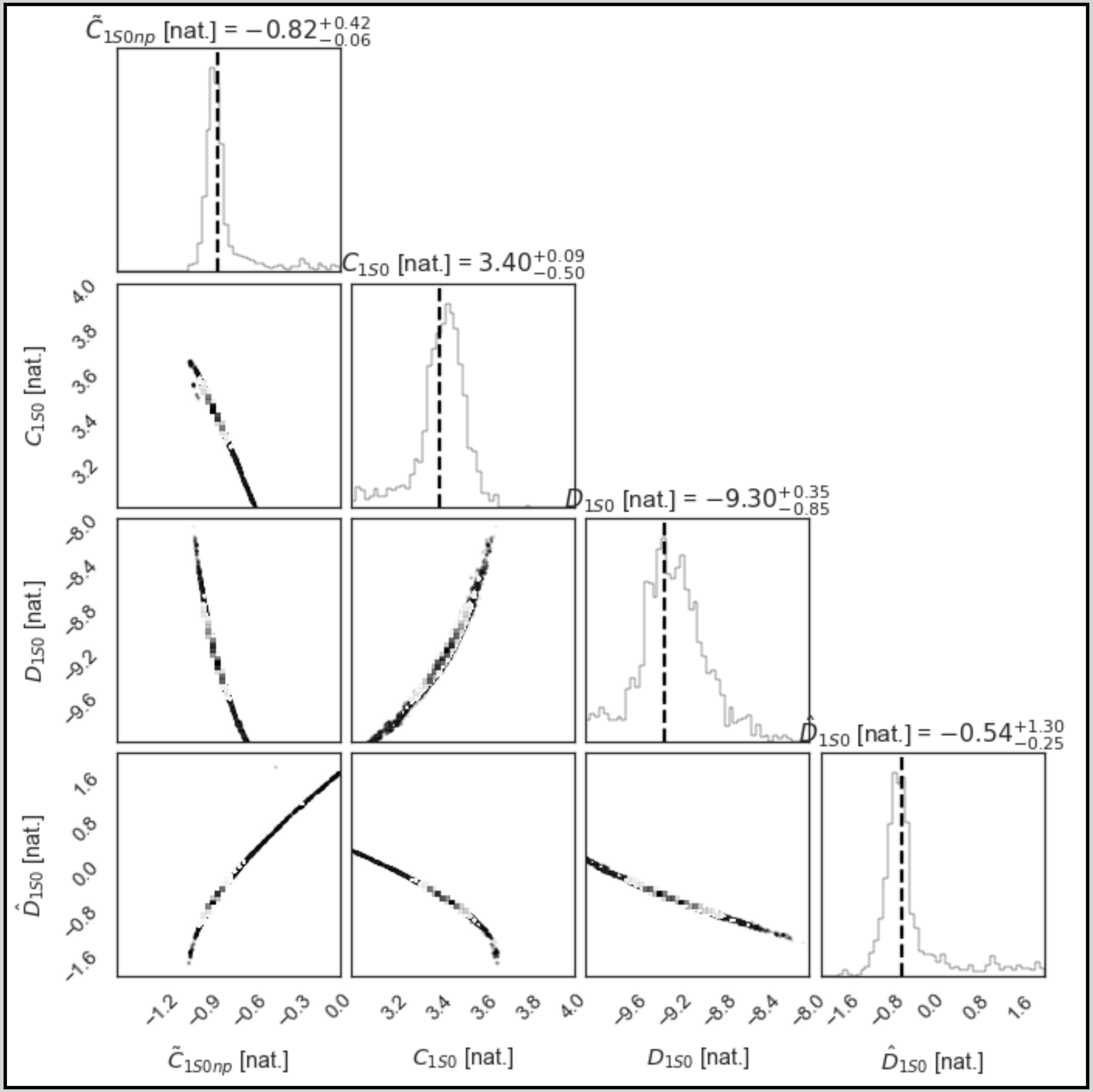
The 1S0(np) channel @ N2LO: Posterior distribution

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

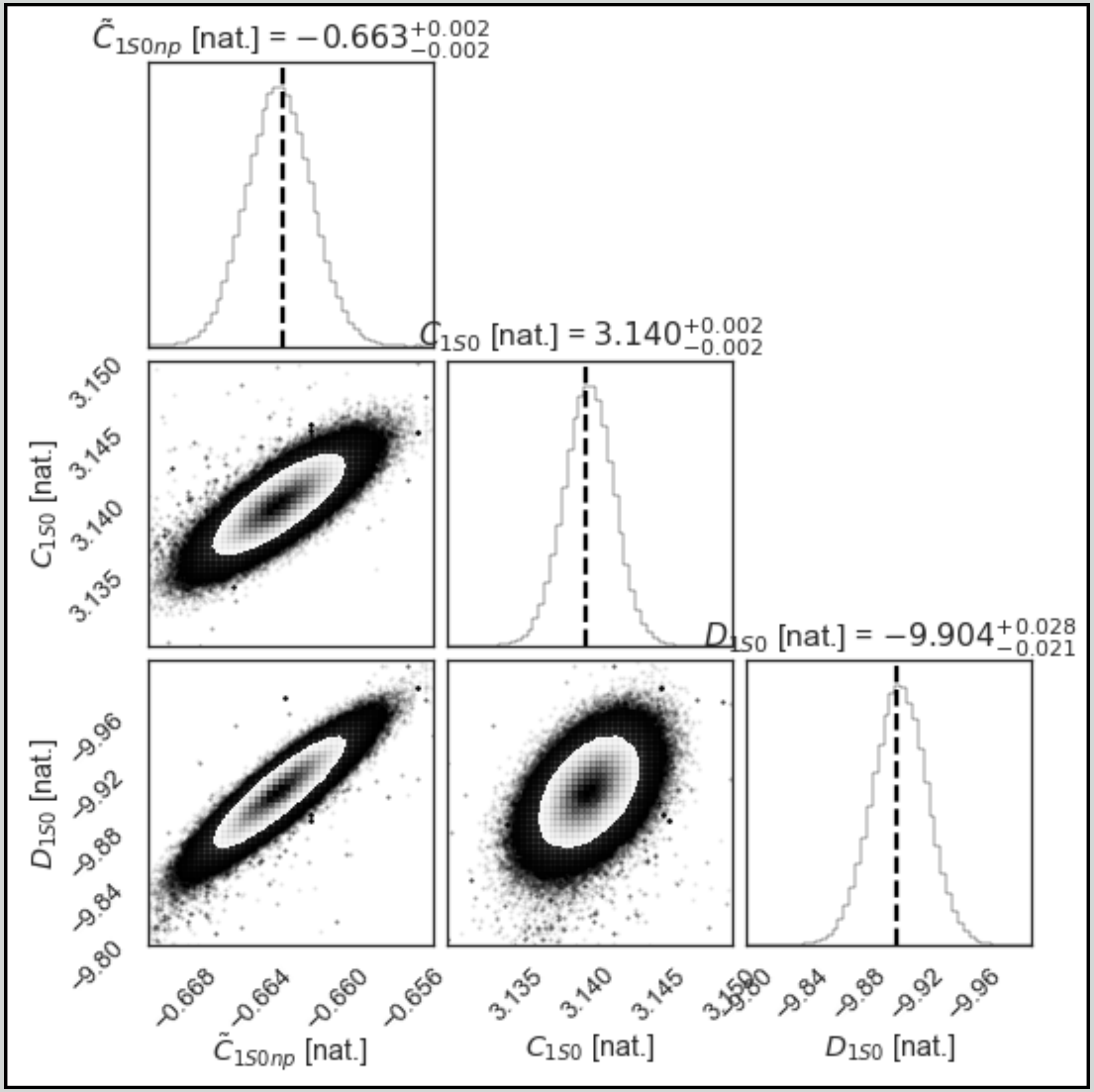


Here, compare to best-estimate from chi-squared minimization

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



The 1S0(np) channel @ N3LO without redundant parameter



Expectation integrals, error propagation

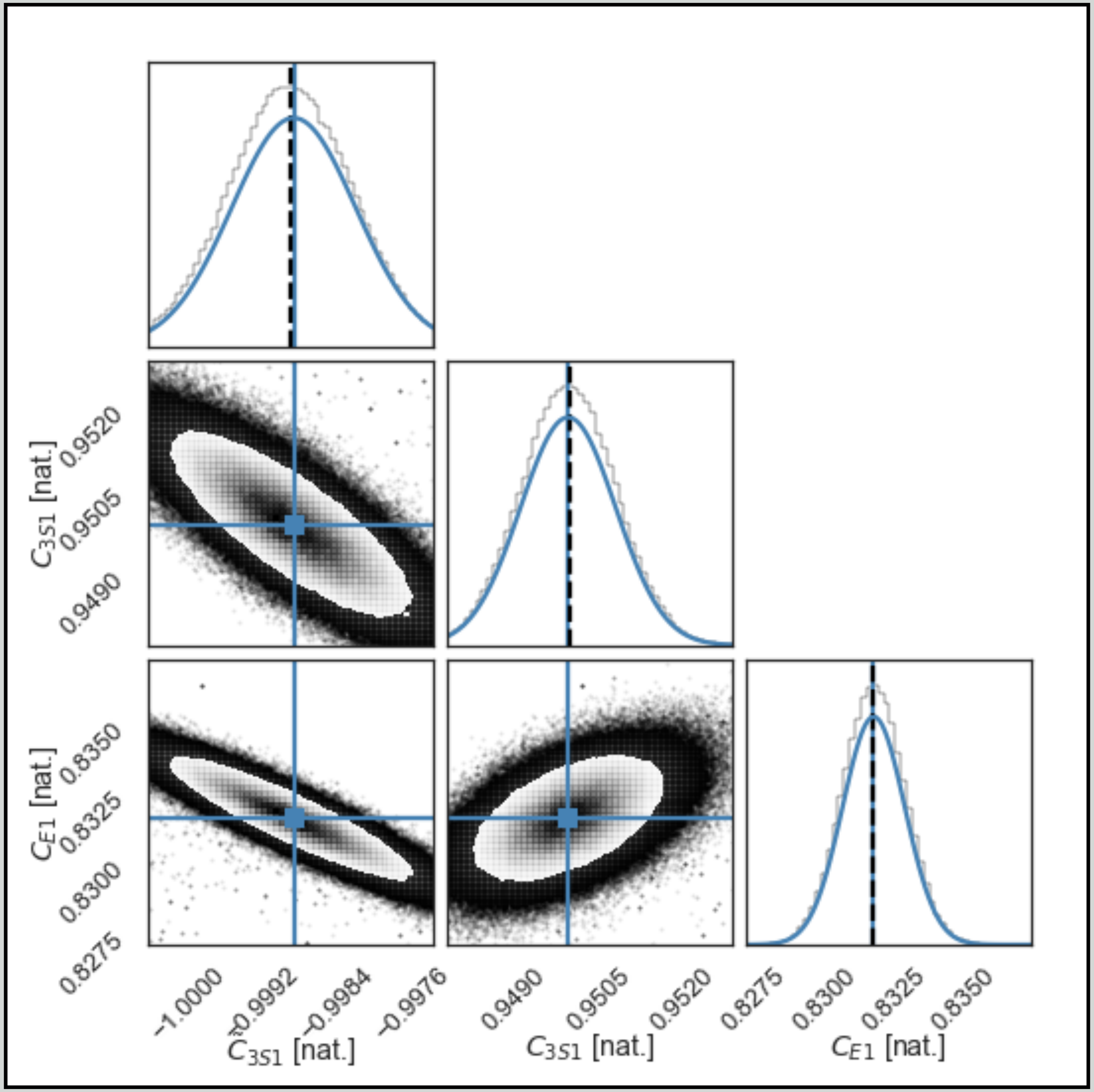
- Expectation integrals for observables can be performed using the posterior pdf

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

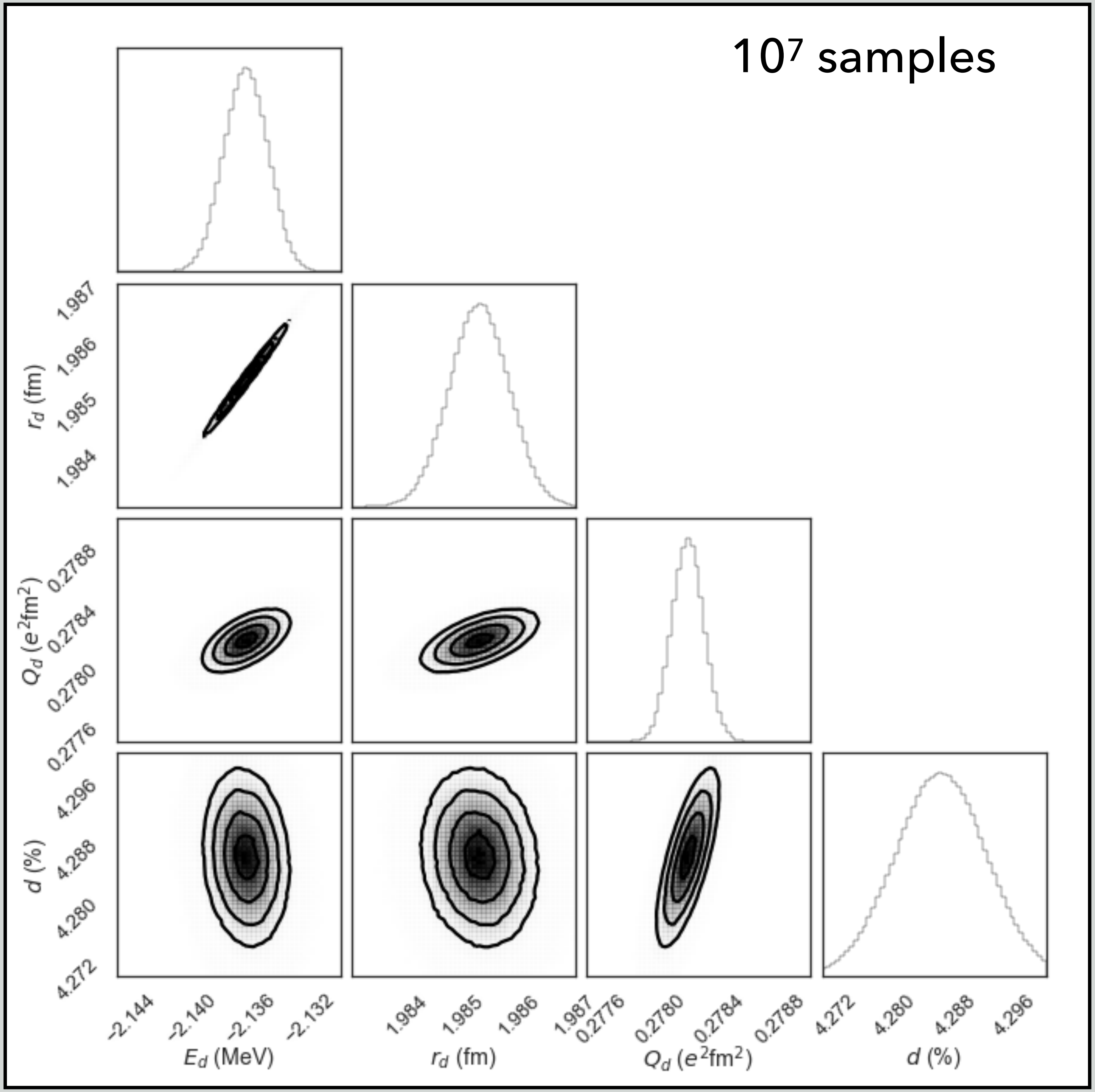
$$\approx \frac{1}{N} \sum_{j=1}^{N} O(\boldsymbol{\theta}_j)$$

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

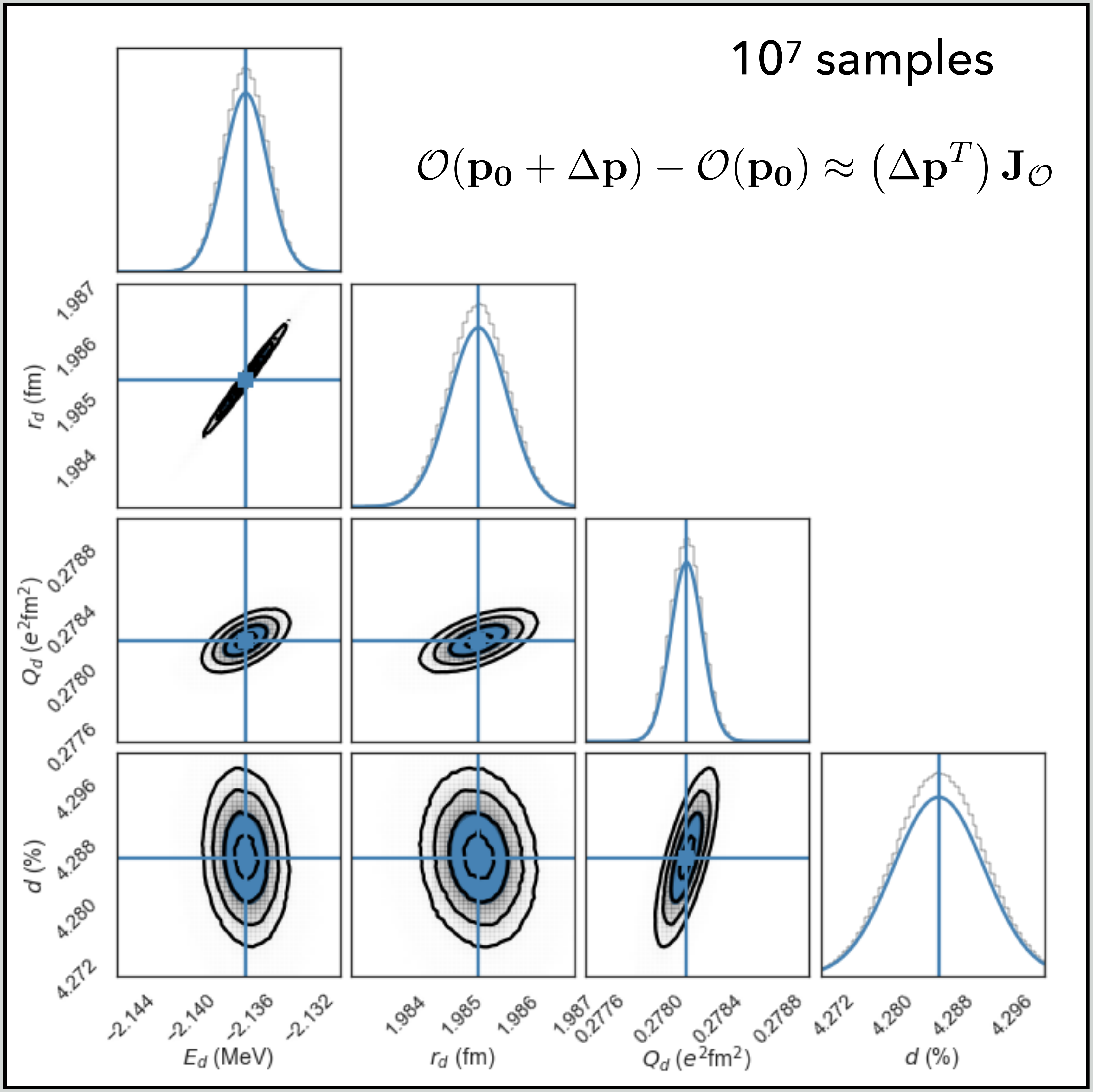
The deuteron channel



Deuteron observables



Deuteron observables



EVIDENCE COMPUTATION WITH PT-MCMC

- For comparison of different "models" = hypothesis *I*:
  - EFTs at different orders,
  - different EFTs,
  - different power counting schemes, ...)

- Compute the Bayesian evidence for the hypothesis *I*

$$Z \equiv p\left(D \,|\, I\right) = \int d\boldsymbol{\theta}\, L\left(\boldsymbol{\theta}\right) \pi\left(\boldsymbol{\theta}\right)$$

- It is hard to compute this norm when your access to the posterior pdf is only via MCMC samples

THERMODYNAMIC INTEGRATION

Goggans & Chi (2004)

- Introduce the evidence for the tempered (β=1/T) distribution

$$Z(\beta) \equiv \int d\boldsymbol{\theta}\, L^{\beta}\left(\boldsymbol{\theta}\right) \pi\left(\boldsymbol{\theta}\right)$$

- Note that *Z*(1) is the target quantity, while

$$Z(0) = \int d\boldsymbol{\theta}\, \pi\left(\boldsymbol{\theta}\right) = 1$$

- *Z*(β) satisfies the differential equation

$$\frac{d \ln Z(\beta)}{d\beta} = \frac{1}{Z(\beta)} \int d\boldsymbol{\theta}\; \ln L\left(\boldsymbol{\theta}\right) L^{\beta}\left(\boldsymbol{\theta}\right) \pi\left(\boldsymbol{\theta}\right) = \langle \ln L \rangle_{\beta}$$

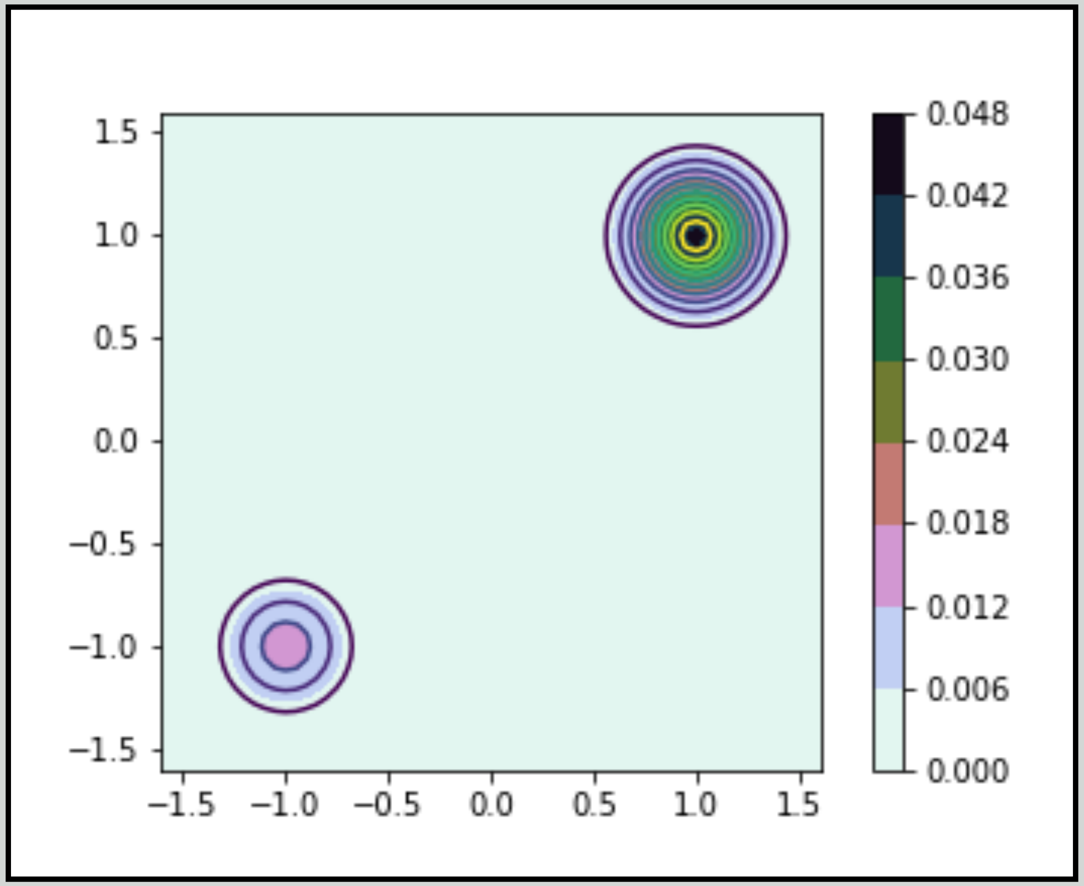
average over the posterior at temperature *T*.

- Integrate:

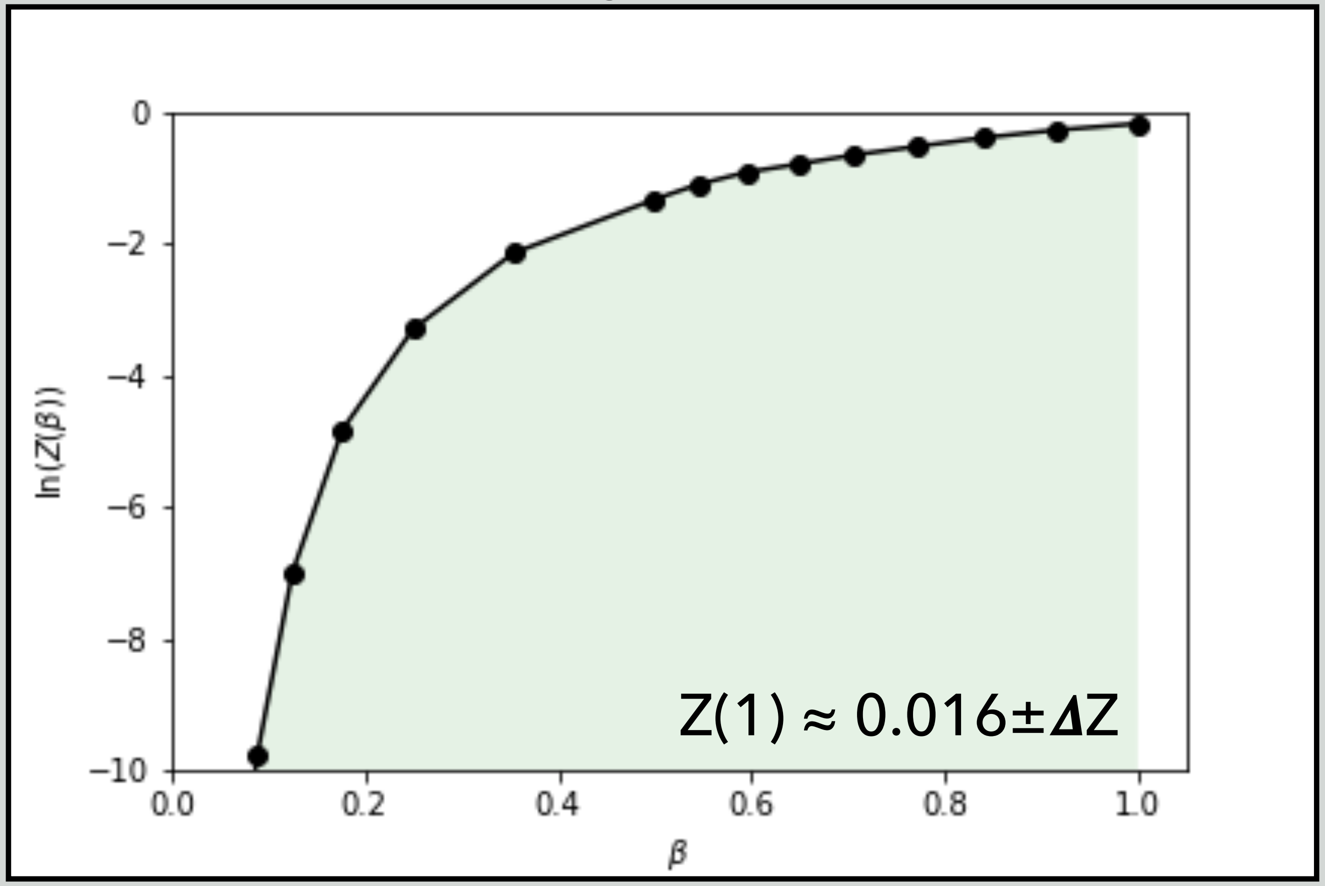
$$\ln Z(1) = \ln Z(0) + \int_0^1 d\beta\; \langle \ln L \rangle_{\beta}$$

Output from MCMC sampling

MULTIMODAL EXAMPLE



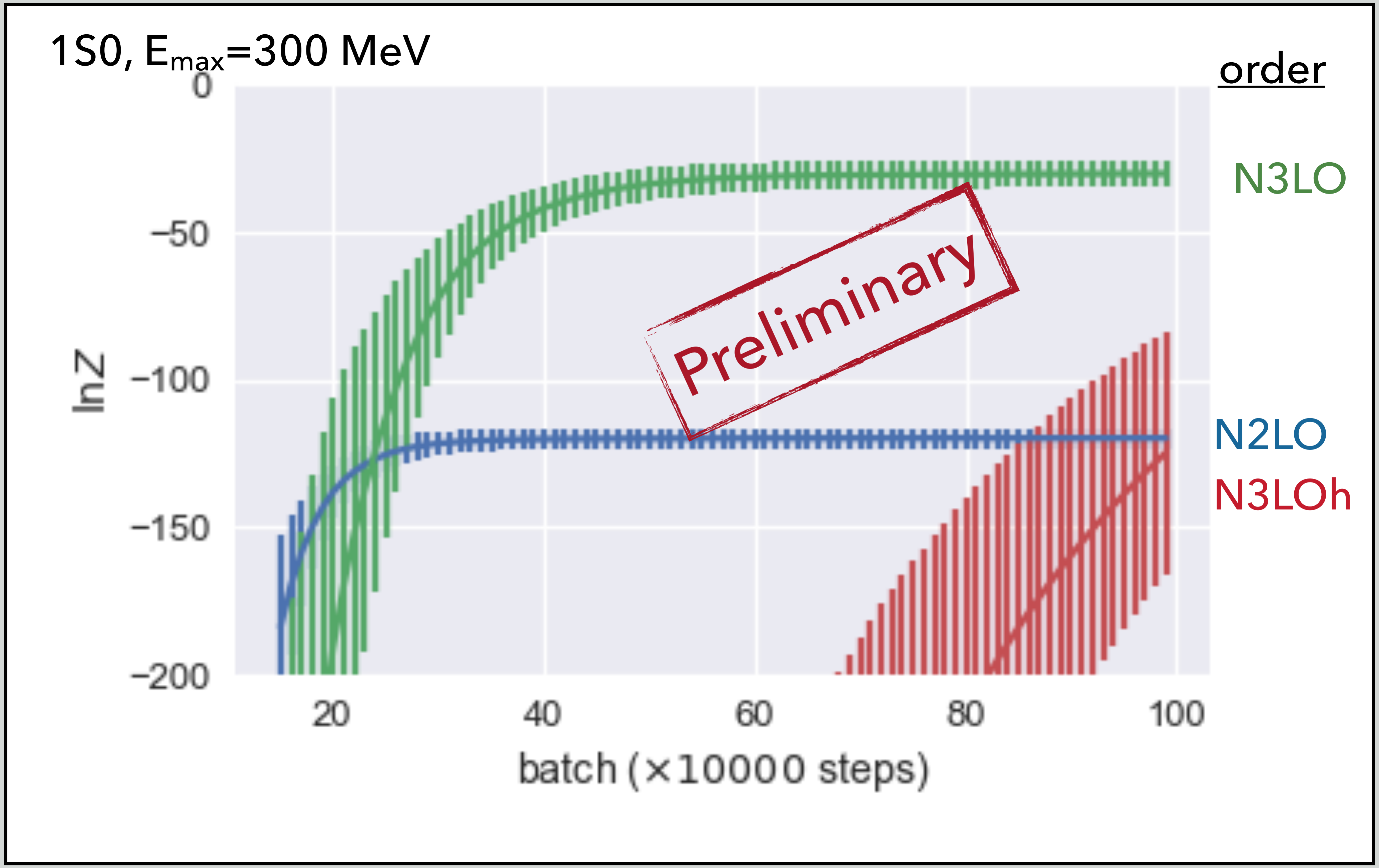
Thermodynamic integration



Two Gaussian posterior modes with norms: 0.0108 and 0.0036

The total evidence is = 0.0144

BAYESIAN EVIDENCE



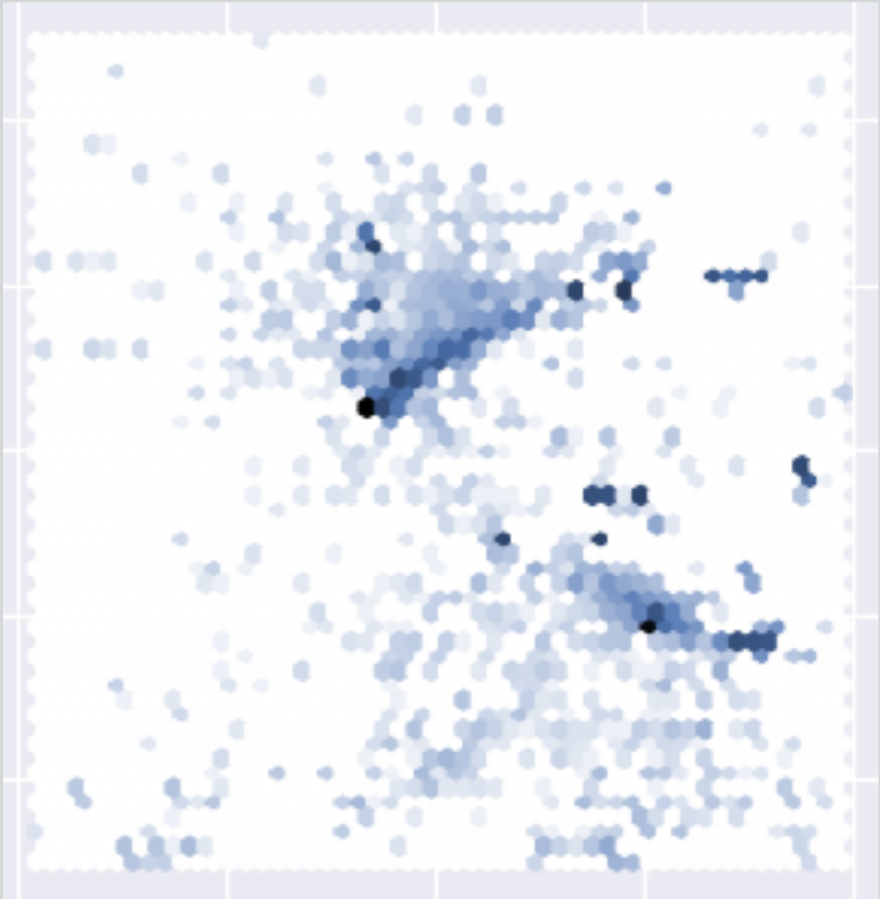
$$p(\boldsymbol{\alpha}|D,I) = \frac{p(D|\boldsymbol{\alpha},I)p(\boldsymbol{\alpha}|I)}{p(D|I)}$$

Bayesian evidence (=Z)

CONCLUSION

Future work

- **Parameter inference**: Quantify the strength of inductive inferences from data.
- **Model comparison**: Compute the evidence. Is there evidence for higher orders in the data?
- **A > 2 systems**: Very challenging computations. Approximations?
- **MCMC sampling**: What can be improved?



QUESTIONS

- Feasible (approximate) Bayes for multi-dimensional posterior distributions, expensive likelihood evaluations, and sub-sequent error propagation?
- Feasible ways to include EFT truncation errors (see also Sarah's and Jordan's presentations)?
- Diagnostics of MCMC convergence?
- Best-practice for evidence calculation (norm of pdf from MCMC samples)
- What are the most relevant tests of various power counting schemes given the Bayesian framework?