Hamiltonian Monte Carlo & eigenvector continuation for *ab initio* nuclear physics

Andreas Ekström

CHALMERS
UNIVERSITY OF TECHNOLOGY

ISNET-9, 22-26 May 2023, Washington University in St. Louis

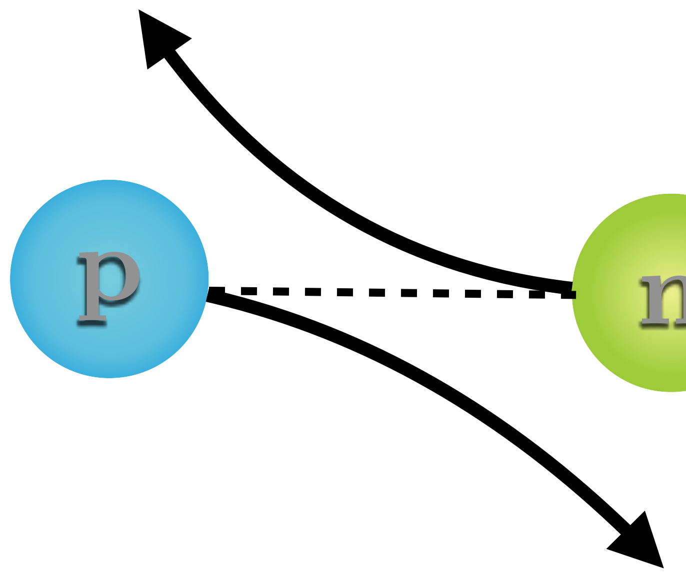
Swedish
Research Council

erc

Contents

Ab initio modelling of nuclear systems

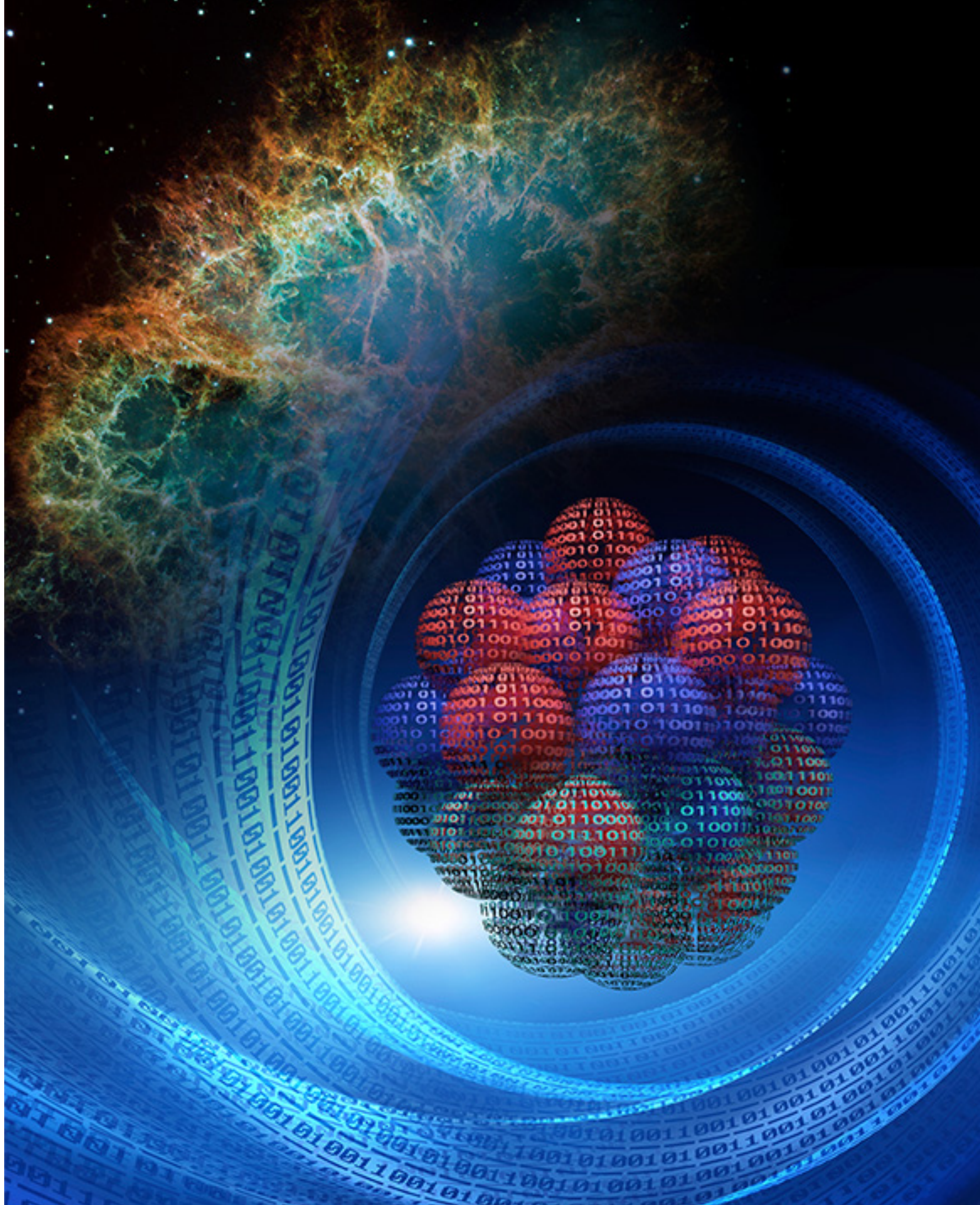
Bayesian inference of neutron-proton scattering



a prototypical system for studying the strong nuclear interaction

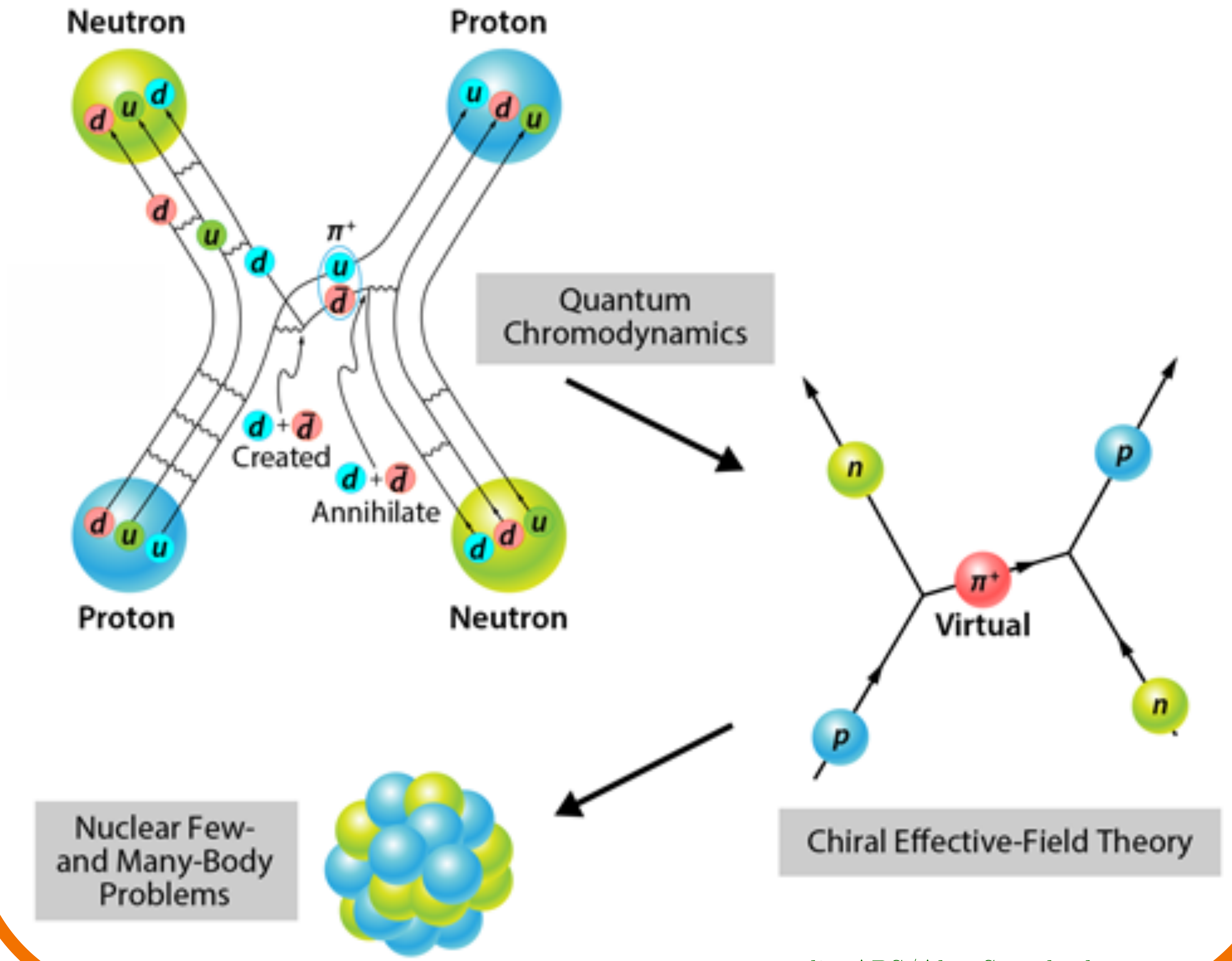
For details, see our recent preprint

Isak Svensson, Andreas Ekström, Christian Forssén arXiv:2304.02004 [nucl-th]



Ab initio modelling of nuclear systems

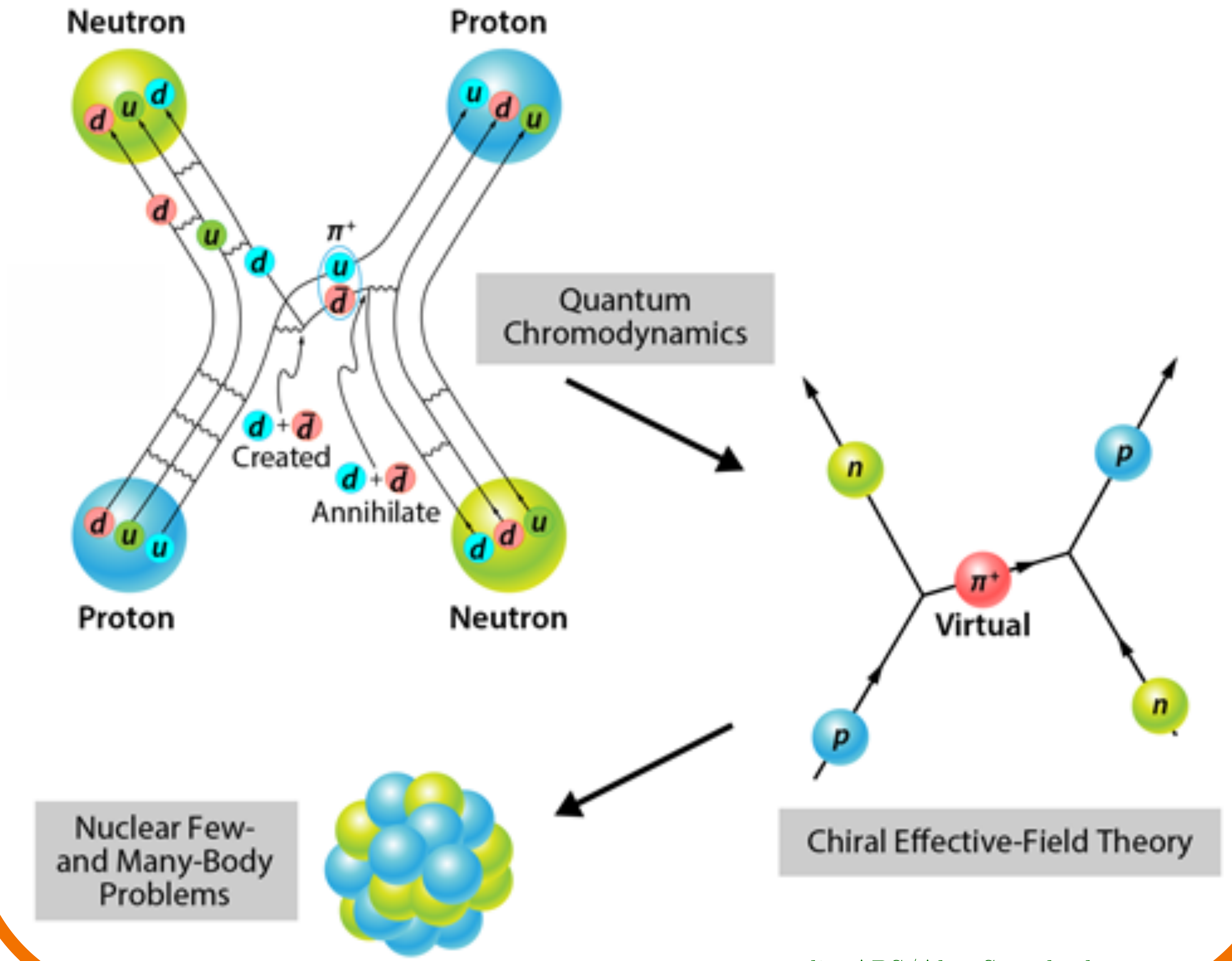
The Idea



credit: APS/Alan Stonebraker

Ab initio modelling of nuclear systems

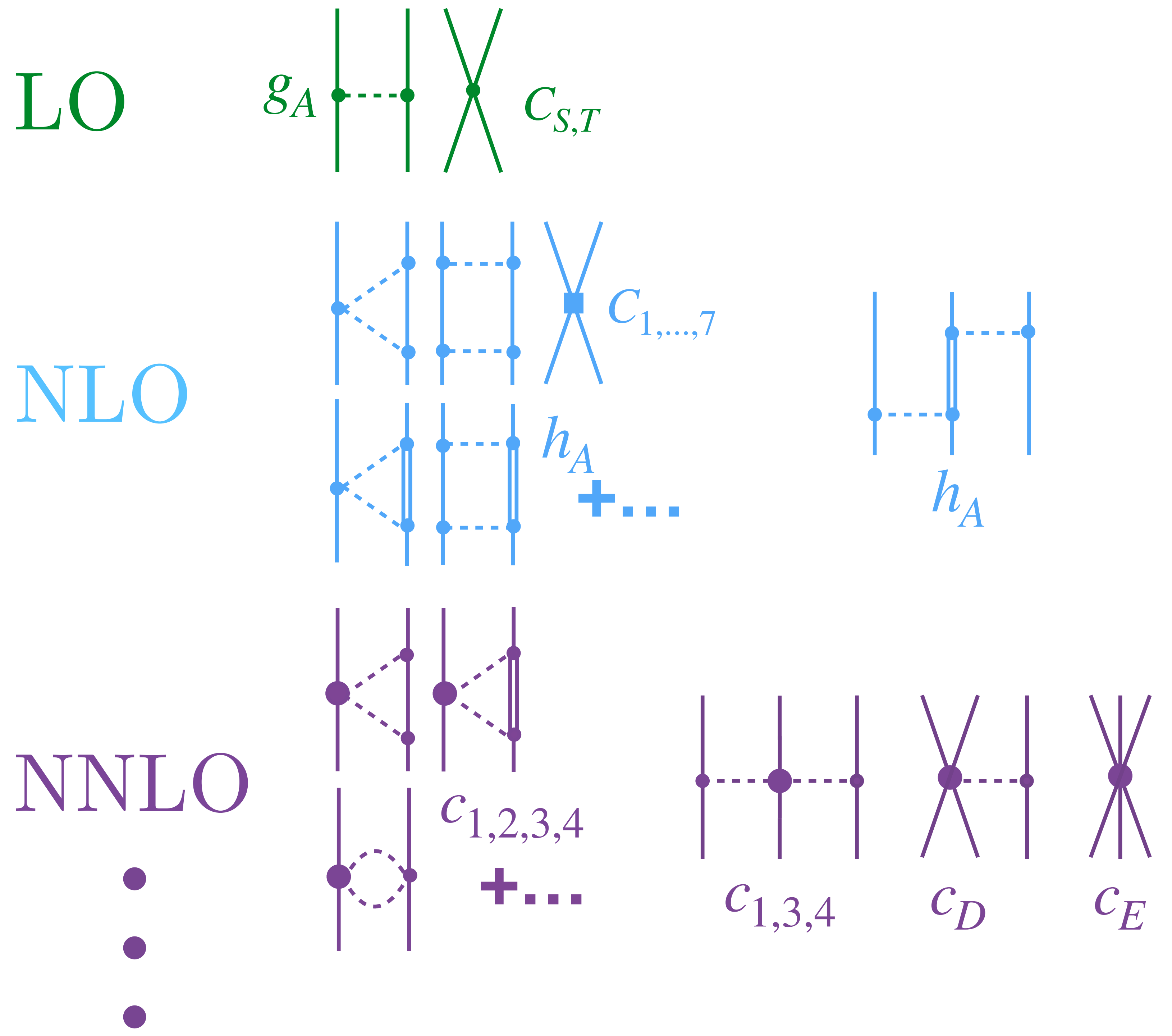
The Idea



credit: APS/Alan Stonebraker

The Model

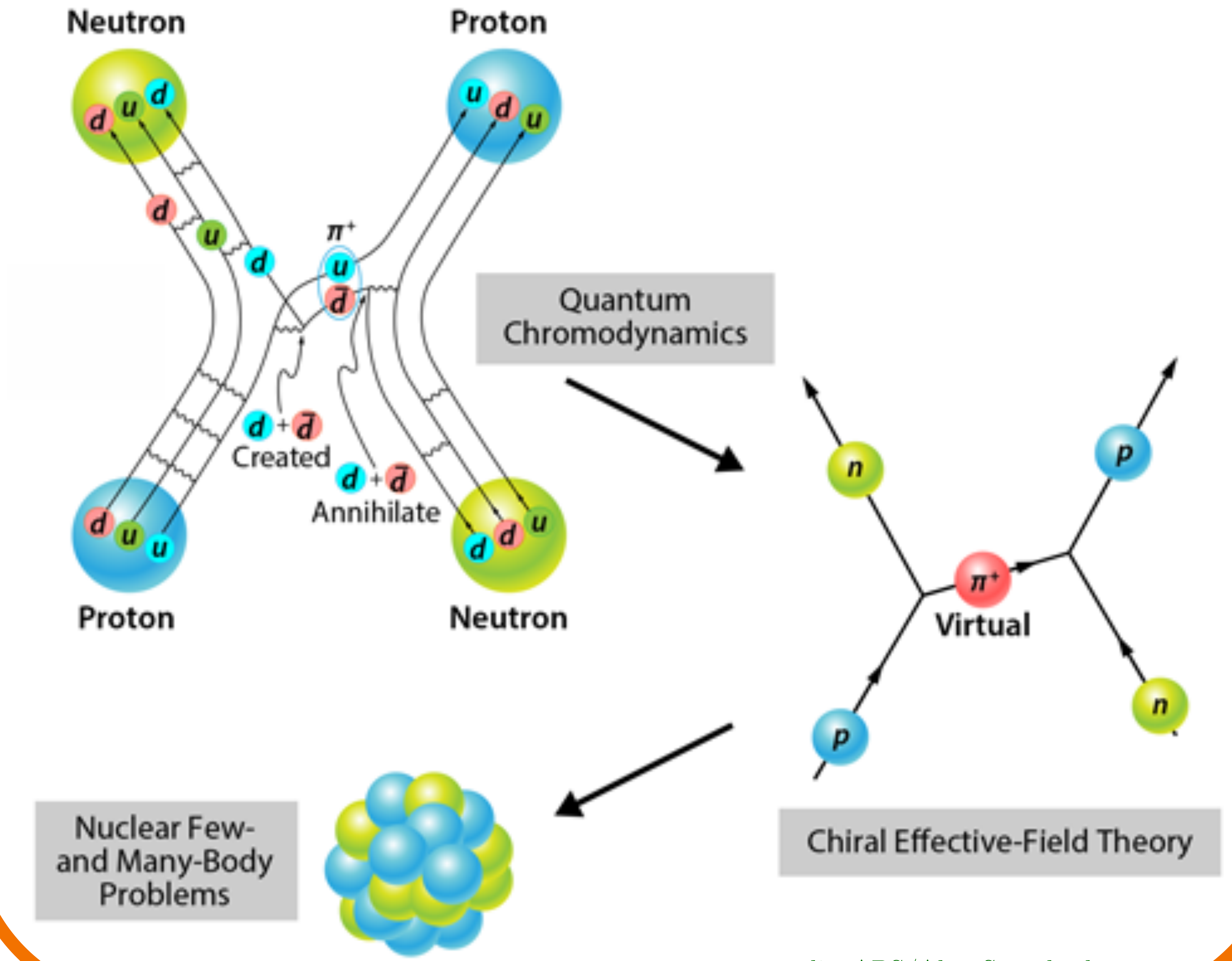
$$\widehat{H}\,|\Psi\rangle = E\,|\Psi\rangle \qquad \widehat{H}(\vec{\alpha}) = \widehat{T} + \widehat{V}(\vec{\alpha})$$



Weinberg, van Kolck, Meißner, Epelbaum, Machleidt, Entem, ...

Ab initio modelling of nuclear systems

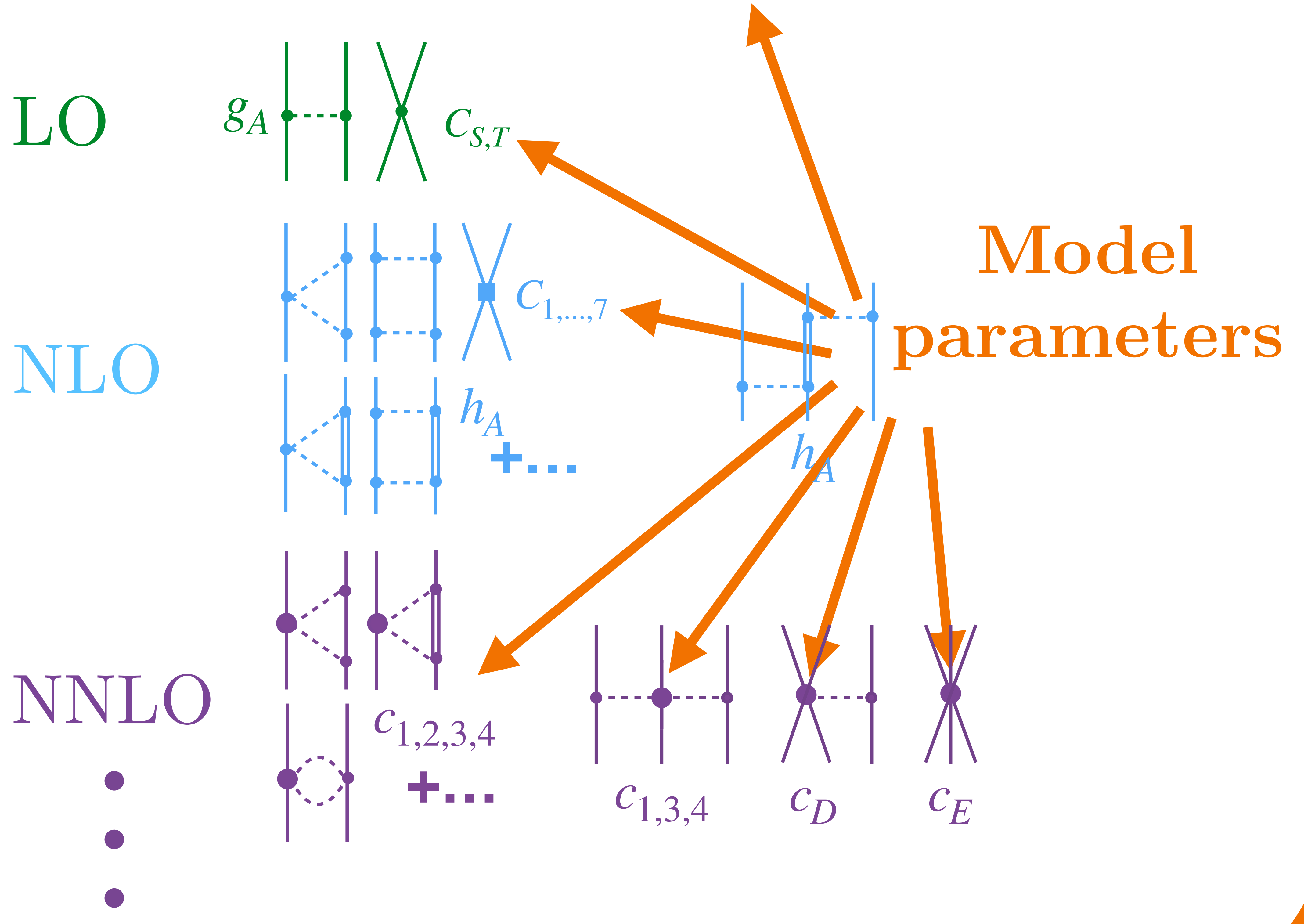
The Idea



credit: APS/Alan Stonebraker

The Model

$\widehat{H}\,|\Psi\rangle = E\,|\Psi\rangle \qquad \widehat{H}(\vec{\alpha}) = \hat{T} + \widehat{V}(\vec{\alpha})$



Weinberg, van Kolck, Meißner, Epelbaum, Machleidt, Entem, ...

Why *ab initio* nuclear theory?

Data uncertainty

$$D_i = M_i(\overrightarrow{\alpha}) + \delta D_i + \delta M_i$$

Data

Model parameters

Model discrepancy

R. J. Furnstahl, N. Klco, D. R. Phillips, S. Wesolowski Phys. Rev. **C 92,** 024005 (2015)
A. Ekström, C.Forssen, G.Hagen, G. R. Jansen, W. Jiang, and T. Papenbrock Front. Phys. **11**, 1129094 (2023)

Why *ab initio* nuclear theory?

Data uncertainty

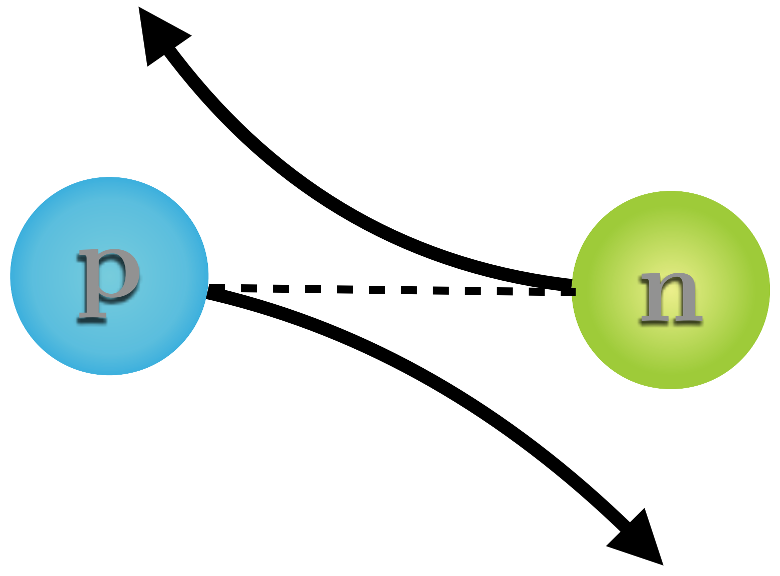
$$D_i = M_i(\vec{\alpha}) + \delta D_i + \delta M_i$$

Data

Model parameters

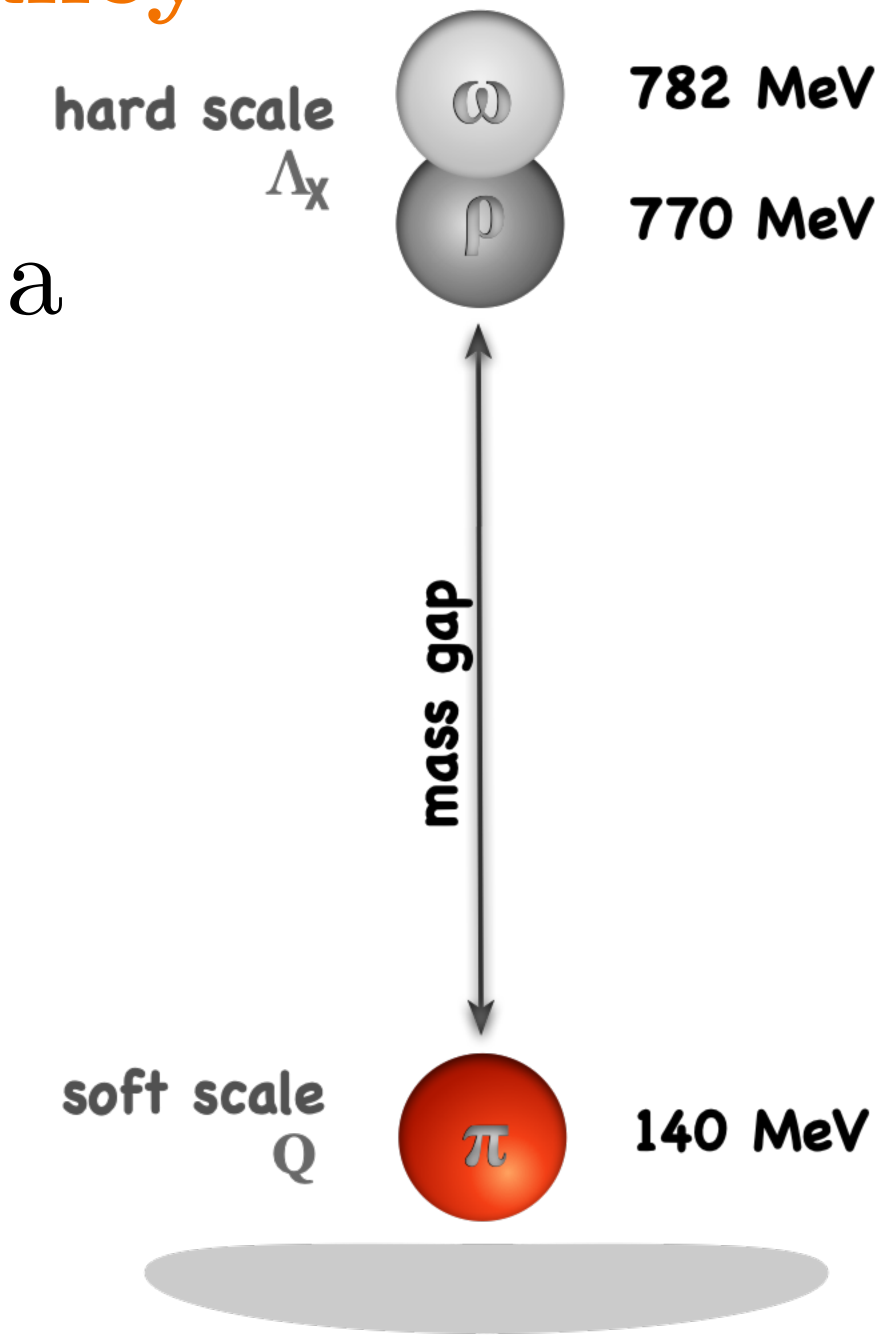
Model discrepancy

We have a model (an effective field theory) for describing data



$$M_i(\vec{\alpha}) = y^{(k)}(\vec{\alpha};\vec{x}_i) = y_{\text{ref}}\left[\gamma_0\left(\frac{Q}{\Lambda_\chi}\right)^0 + \gamma_1\left(\frac{Q}{\Lambda_\chi}\right)^1 + \gamma_2\left(\frac{Q}{\Lambda_\chi}\right)^2 + \ldots + \gamma_k\left(\frac{Q}{\Lambda_\chi}\right)^k\right]$$

LO NLO N2LO NkLO



nucleon-nucleon scattering data ($\sim$2000 of relevance)

R. J. Furnstahl, N. Klco, D. R. Phillips, S. Wesolowski Phys. Rev. **C 92,** 024005 (2015)

A. Ekström, C.Forssen, G.Hagen, G. R. Jansen, W. Jiang, and T. Papenbrock Front. Phys. **11**, 1129094 (2023)

Why *ab initio* nuclear theory?

Data uncertainty

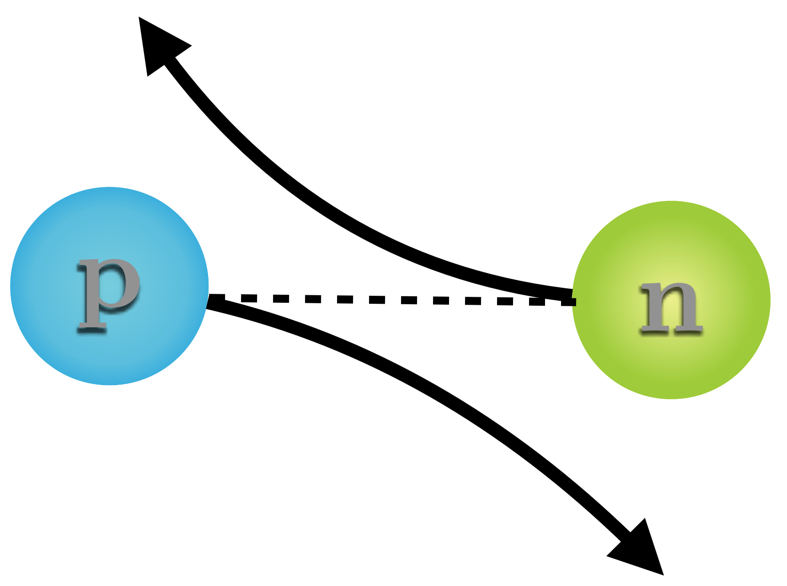
$$D_i = M_i(\vec{\alpha}) + \delta D_i + \delta M_i$$

Data

Model parameters

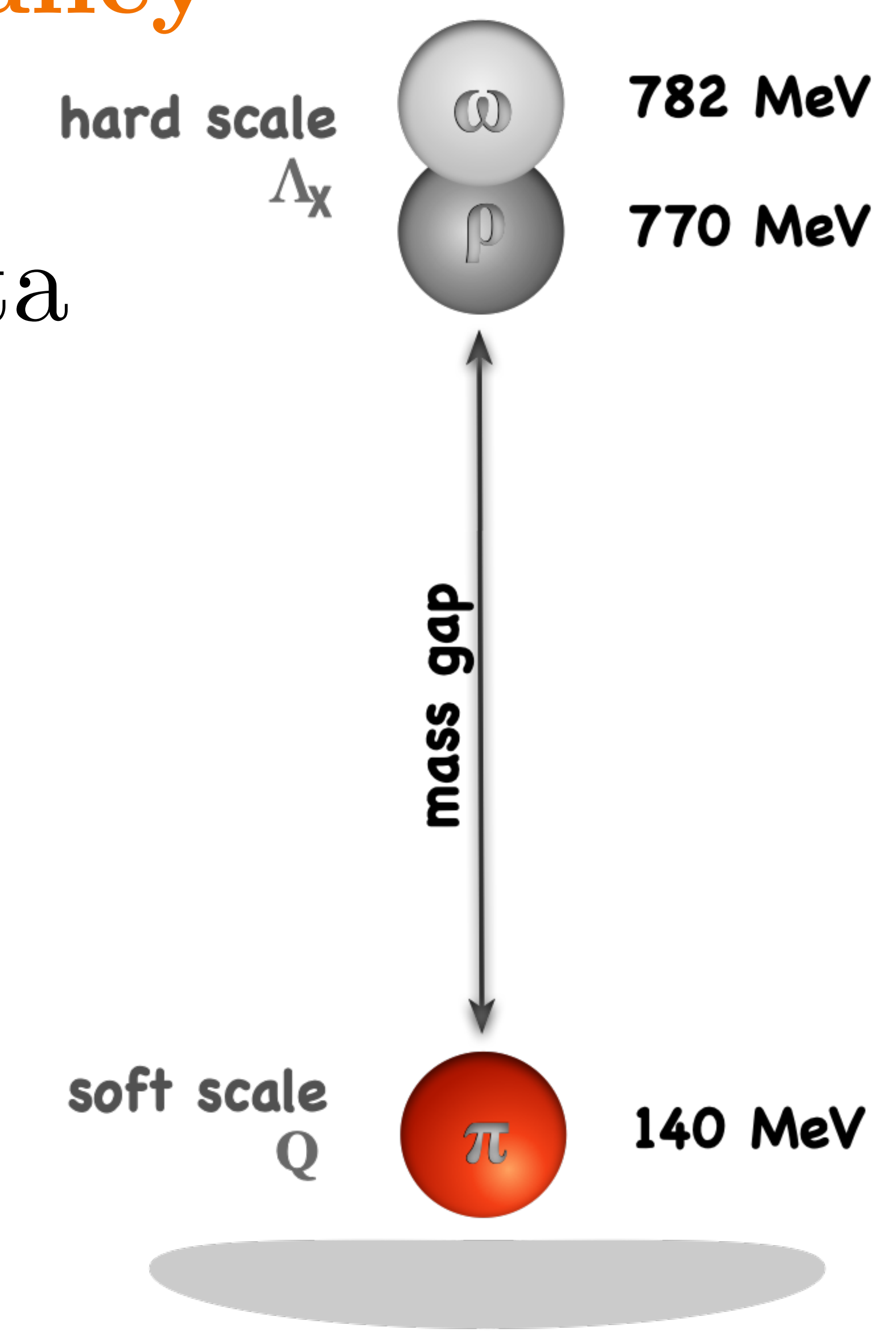
Model discrepancy

We have a model (an effective field theory) for describing data



$$M_i(\vec{\alpha}) = y^{(k)}(\vec{\alpha};\vec{x}_i) = y_{\mathrm{ref}}\left[\gamma_0\left(\frac{Q}{\Lambda_\chi}\right)^0 + \gamma_1\left(\frac{Q}{\Lambda_\chi}\right)^1 + \gamma_2\left(\frac{Q}{\Lambda_\chi}\right)^2 + \ldots + \gamma_k\left(\frac{Q}{\Lambda_\chi}\right)^k\right]$$

LO NLO N2LO NkLO



nucleon-nucleon scattering data (~2000 of relevance)

The systematicity of the *ab initio* method creates an **inferential advantage** since we claim to know something about the model discrepancy.

$$\delta M_i = y_{\mathrm{ref}} \sum_{i=k+1}^{\infty} \gamma_i \left(\frac{Q}{\Lambda_\chi}\right)^i$$

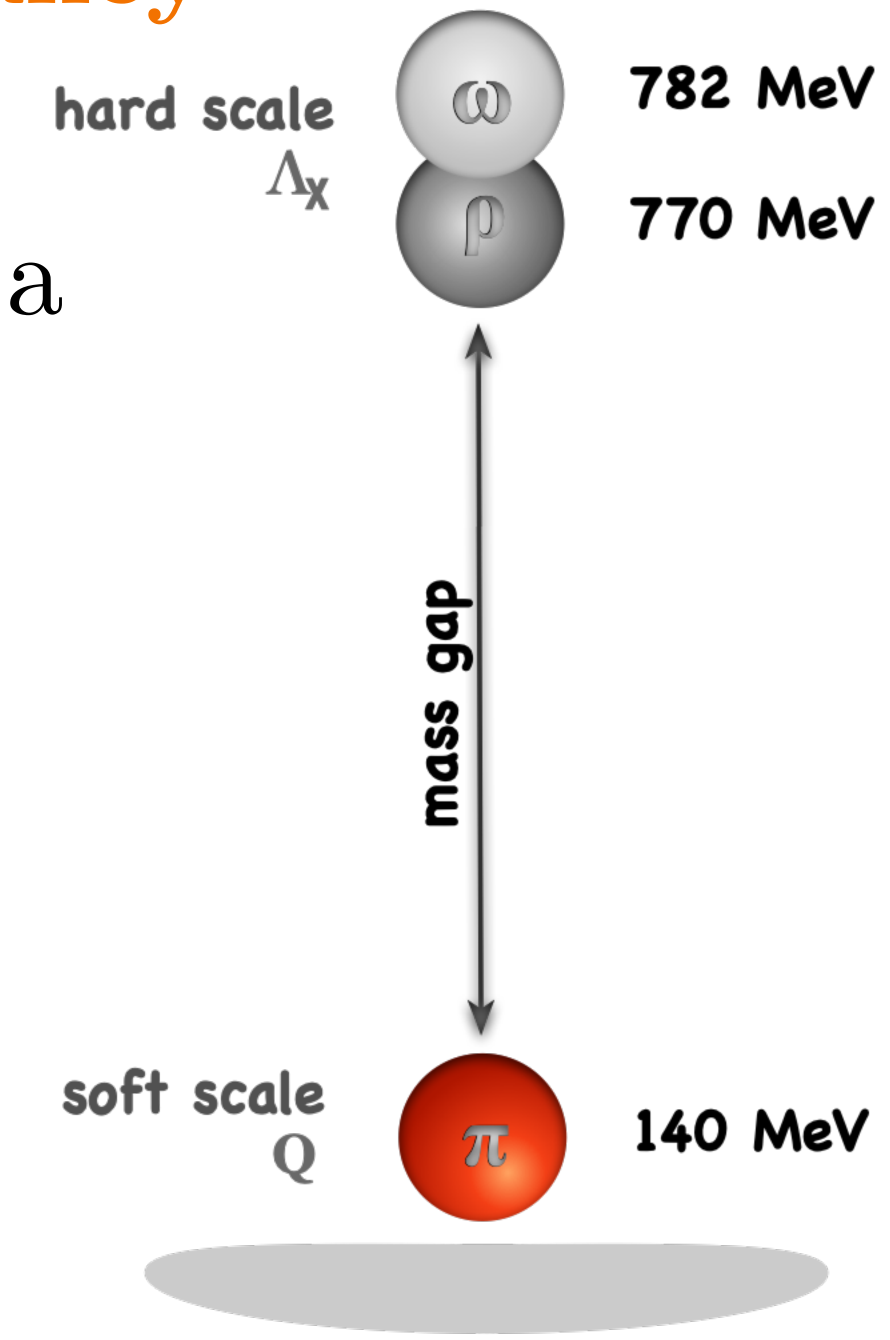
R. J. Furnstahl, N. Klco, D. R. Phillips, S. Wesolowski Phys. Rev. **C 92,** 024005 (2015)
A. Ekström, C.Forssen, G.Hagen, G. R. Jansen, W. Jiang, and T. Papenbrock Front. Phys. **11**, 1129094 (2023)

Why *ab initio* nuclear theory?

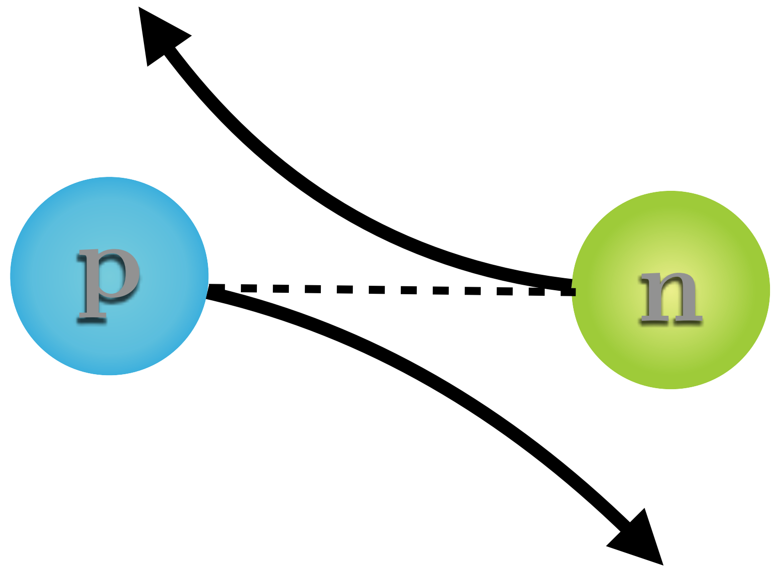
Data uncertainty

$$D_i = M_i(\vec{\alpha}) + \delta D_i + \delta M_i$$

Data — Model parameters — Model discrepancy



We have a model (an effective field theory) for describing data



$$M_i(\vec{\alpha}) = y^{(k)}(\vec{\alpha};\vec{x}_i) = y_{\text{ref}}\left[\gamma_0\left(\frac{Q}{\Lambda_\chi}\right)^0 + \gamma_1\left(\frac{Q}{\Lambda_\chi}\right)^1 + \gamma_2\left(\frac{Q}{\Lambda_\chi}\right)^2 + \ldots + \gamma_k\left(\frac{Q}{\Lambda_\chi}\right)^k\right]$$

LO NLO N2LO NkLO

nucleon-nucleon scattering data (~2000 of relevance)

The systematicity of the *ab initio* method creates an **inferential advantage** since we claim to know something about the model discrepancy.

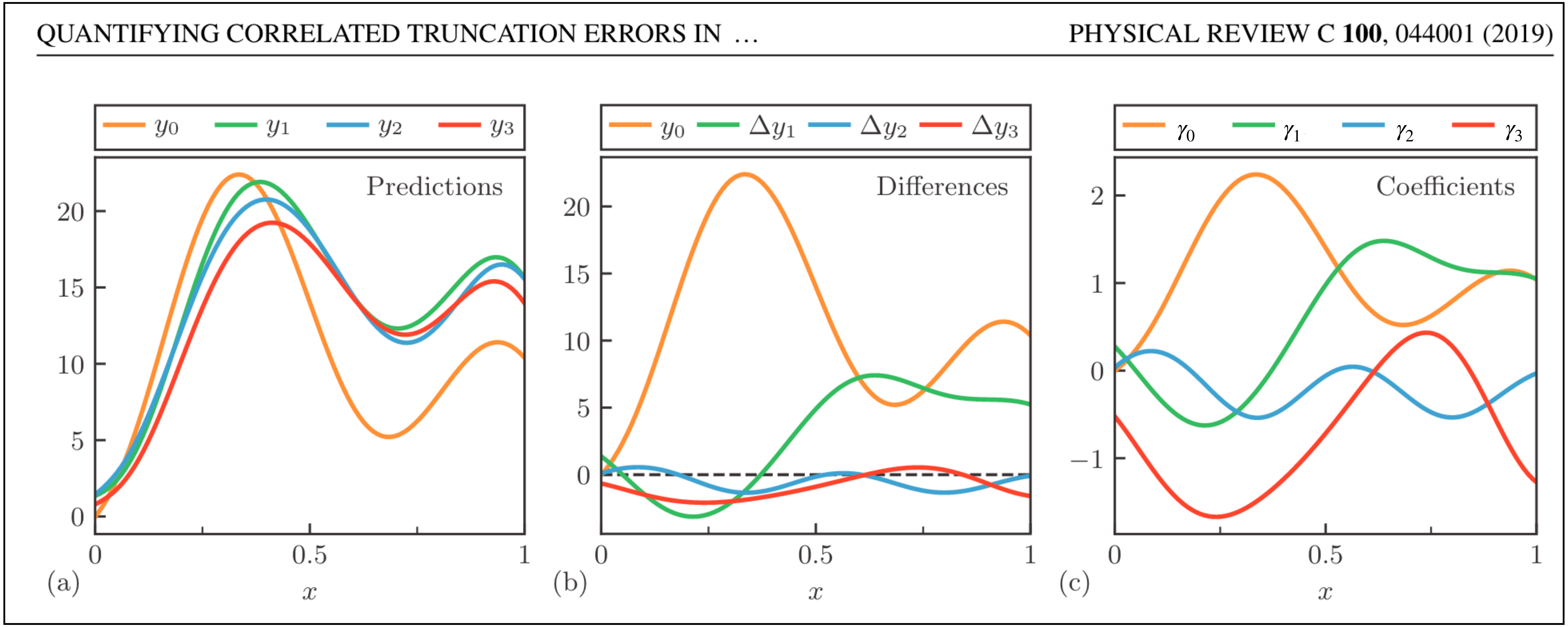
$$\delta M_i = y_{\text{ref}} \sum_{i=k+1}^{\infty} \gamma_i \left(\frac{Q}{\Lambda_\chi}\right)^i$$

We are uncertain about the expansion coefficients

R. J. Furnstahl, N. Klco, D. R. Phillips, S. Wesolowski Phys. Rev. **C 92,** 024005 (2015)
A. Ekström, C.Forssen, G.Hagen, G. R. Jansen, W. Jiang, and T. Papenbrock Front. Phys. **11**, 1129094 (2023)

Predictions $y^{(k)}$ correlated across scattering energy & angle

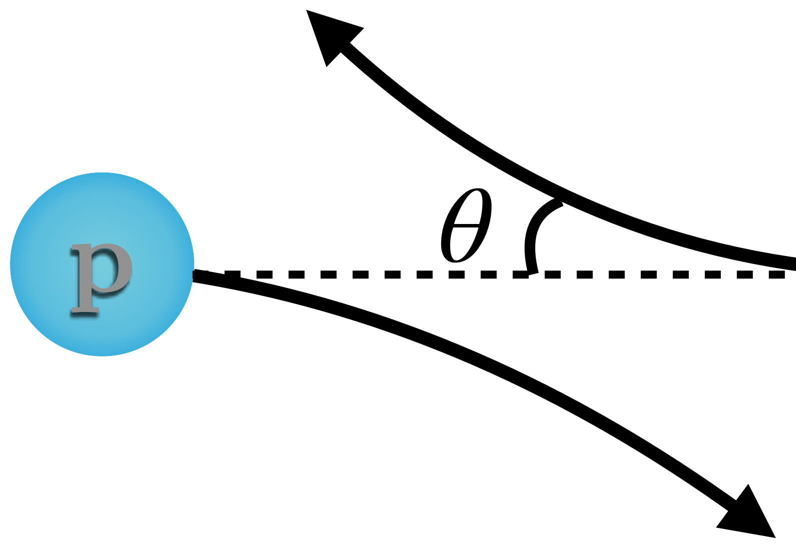
adapted from J. A. Melendez, R. J. Furnstahl, D. R. Phillips, M. T. Pratola, S. Wesolowski Phys. Rev. **C 100** (4), 044001 (2019)



we use $y_{\mathrm{ref}} = 0.5$ (Idaho-N3LO) for polarization (σ_{tot} & $\sigma(\theta)$)

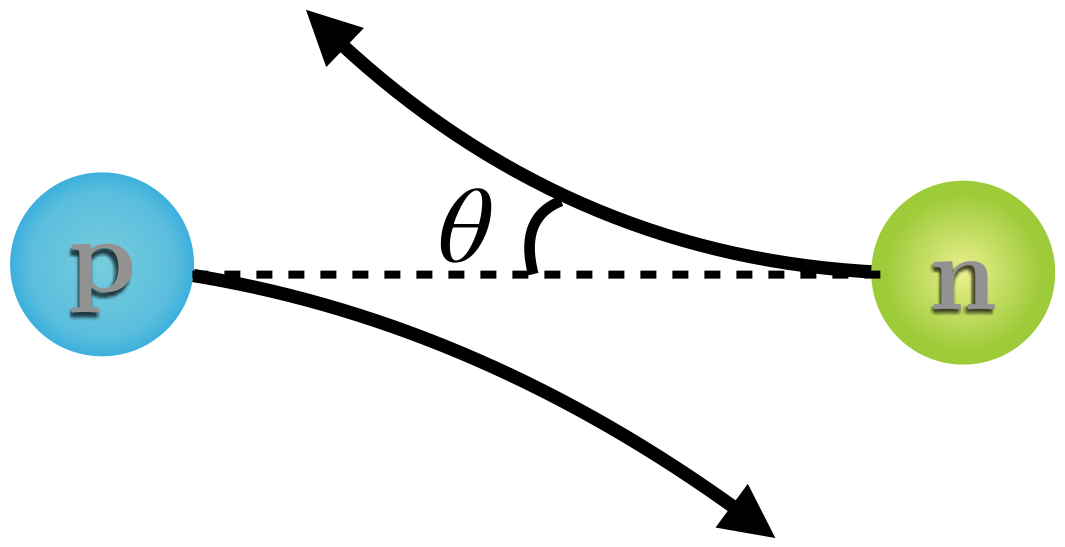
$$\gamma^{(k)}(\vec{x}_j) = \frac{y_{\mathrm{th}}^{(k)}(\vec{\alpha}^{\star};\vec{x}_j) - y_{\mathrm{th}}^{(k-1)}(\vec{\alpha}^{\star};\vec{x}_j)}{y_{\mathrm{ref}}\,(Q/\Lambda_\chi)^k}$$

A Gaussian-process model for correlated δM



Quantify expansion coefficients by evaluating the model at consecutive chiral orders for a 2D grid of $\vec{x}_j$ values (energies T_{lab} and angles θ).
n.b. we do this for a sensible choice of $\vec{\alpha} = \vec{\alpha}^{\star}$

A Gaussian-process model for correlated δM



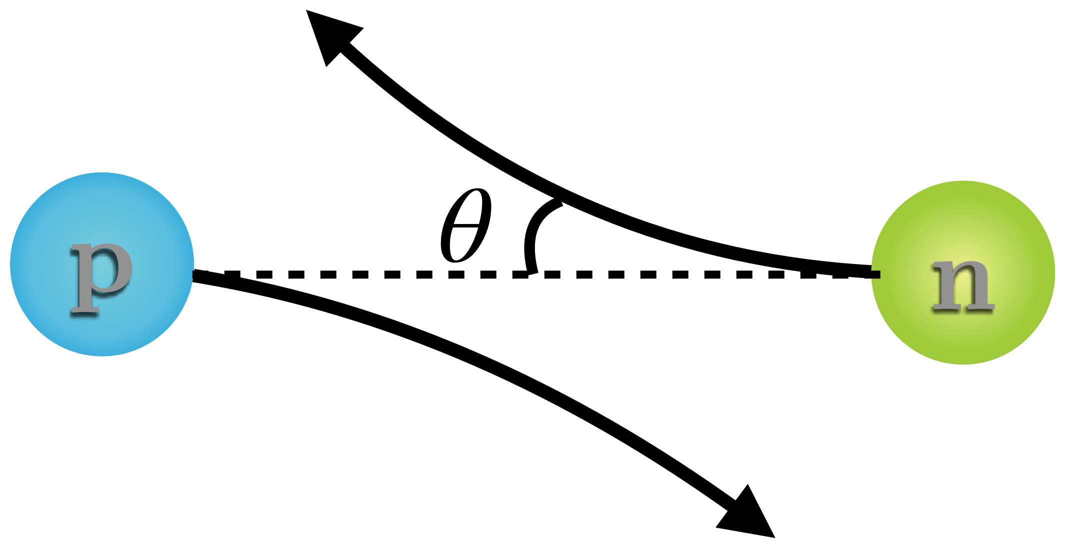
Quantify expansion coefficients by evaluating the model at consecutive chiral orders for a 2D grid of $\vec{x}_j$ values (energies T_{lab} and angles θ).
n.b. we do this for a sensible choice of $\vec{\alpha} = \vec{\alpha}^\star$

We model the correlated expansion parameters as a two-feature Gaussian process using a SQE kernel

$$\gamma(\boldsymbol{x})\,|\,m, \vec{\ell}, \bar{c}^2 \sim \mathcal{GP}[m, \bar{c}^2 k(\boldsymbol{x}', \boldsymbol{x}; \vec{\ell})]$$

Isak Svensson, Andreas Ekström, Christian Forssén arXiv:2304.02004 [nucl-th]

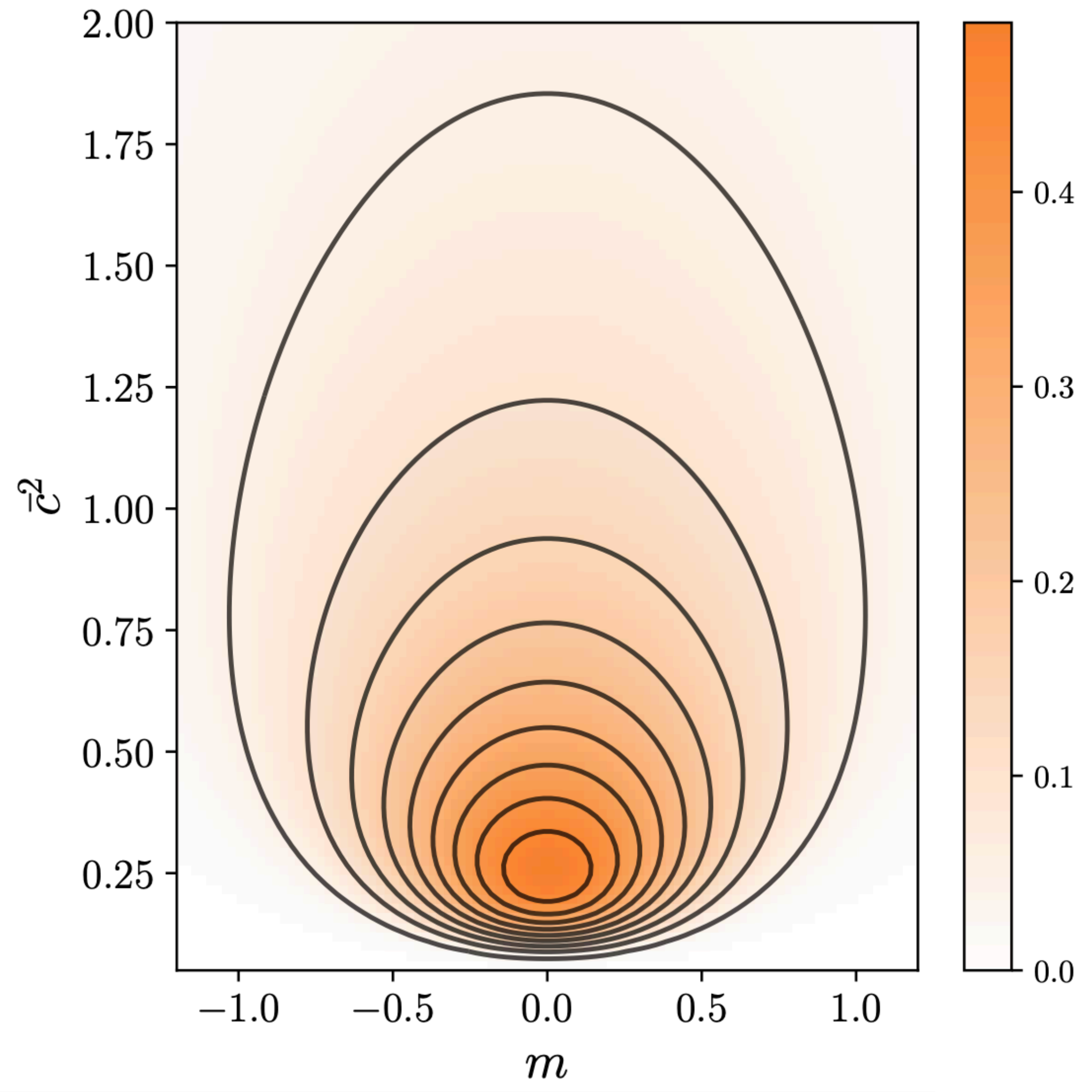
A Gaussian-process model for correlated δM



Quantify expansion coefficients by evaluating the model at consecutive chiral orders for a 2D grid of $\vec{x}_j$ values (energies T_{lab} and angles θ).
n.b. we do this for a sensible choice of $\vec{\alpha} = \vec{\alpha}^\star$

We model the correlated expansion parameters as a two-feature Gaussian process using a SQE kernel

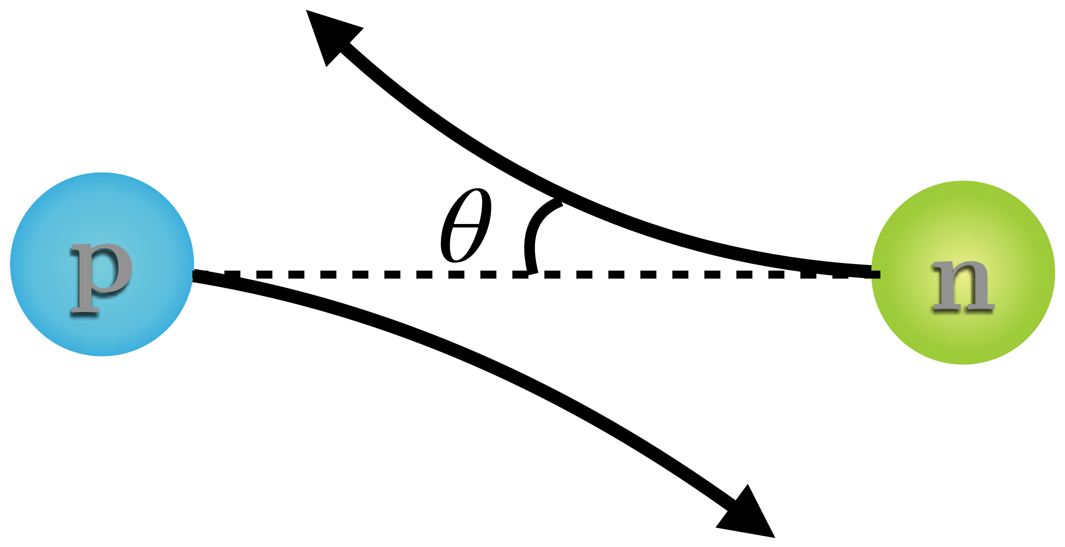
$$\gamma(\boldsymbol{x}) \,|\, m, \vec{\ell}, \bar{c}^2 \sim \mathcal{GP}[m, \bar{c}^2 k(\boldsymbol{x}', \boldsymbol{x}; \vec{\ell})]$$



Our prior for the GP hyperparameters

uniform distribution for the lengthscales $\vec{\ell}$

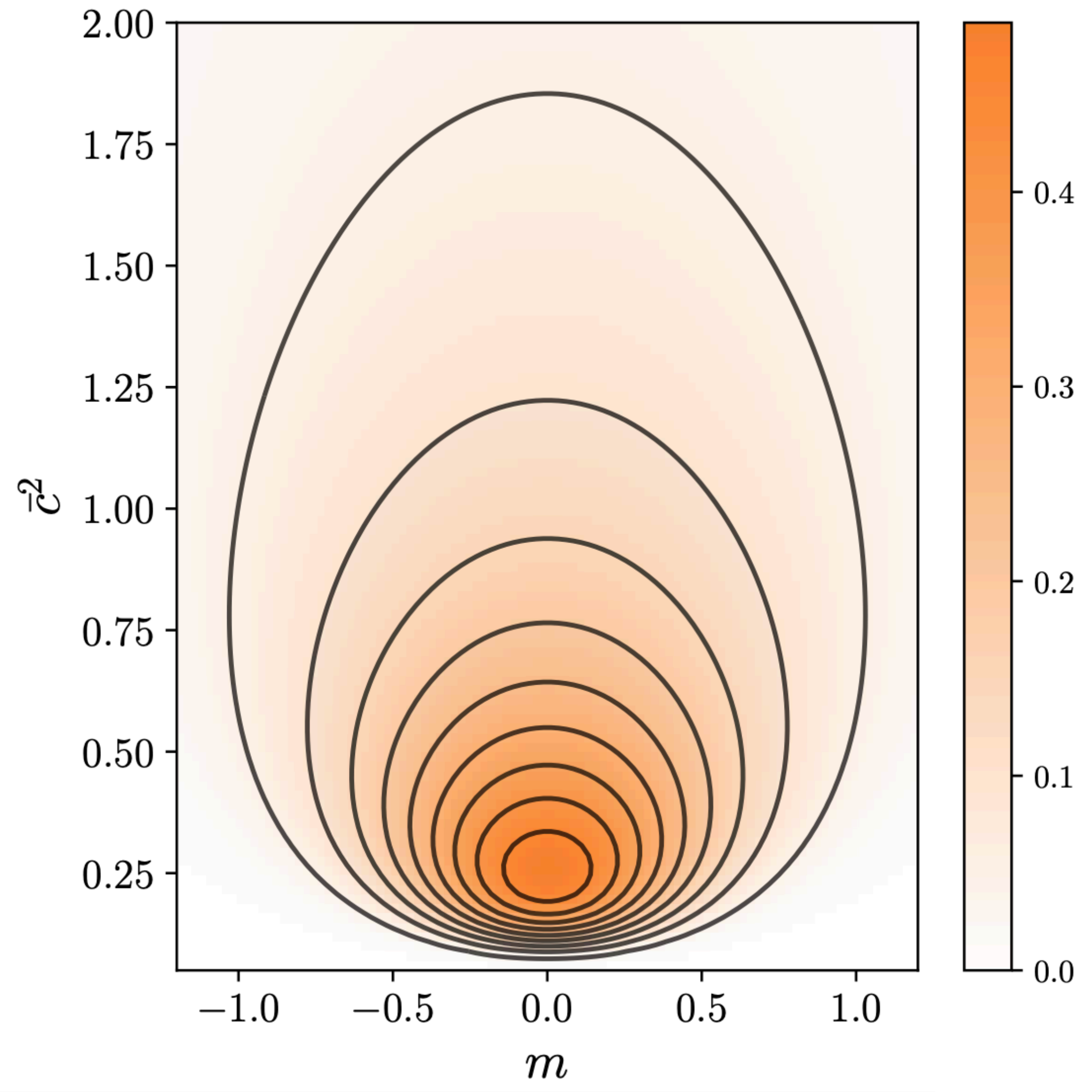
A Gaussian-process model for correlated δM



Quantify expansion coefficients by evaluating the model at consecutive chiral orders for a 2D grid of $\vec{x}_j$ values (energies T_{lab} and angles θ).
n.b. we do this for a sensible choice of $\vec{\alpha} = \vec{\alpha}^\star$

We model the correlated expansion parameters as a two-feature Gaussian process using a SQE kernel

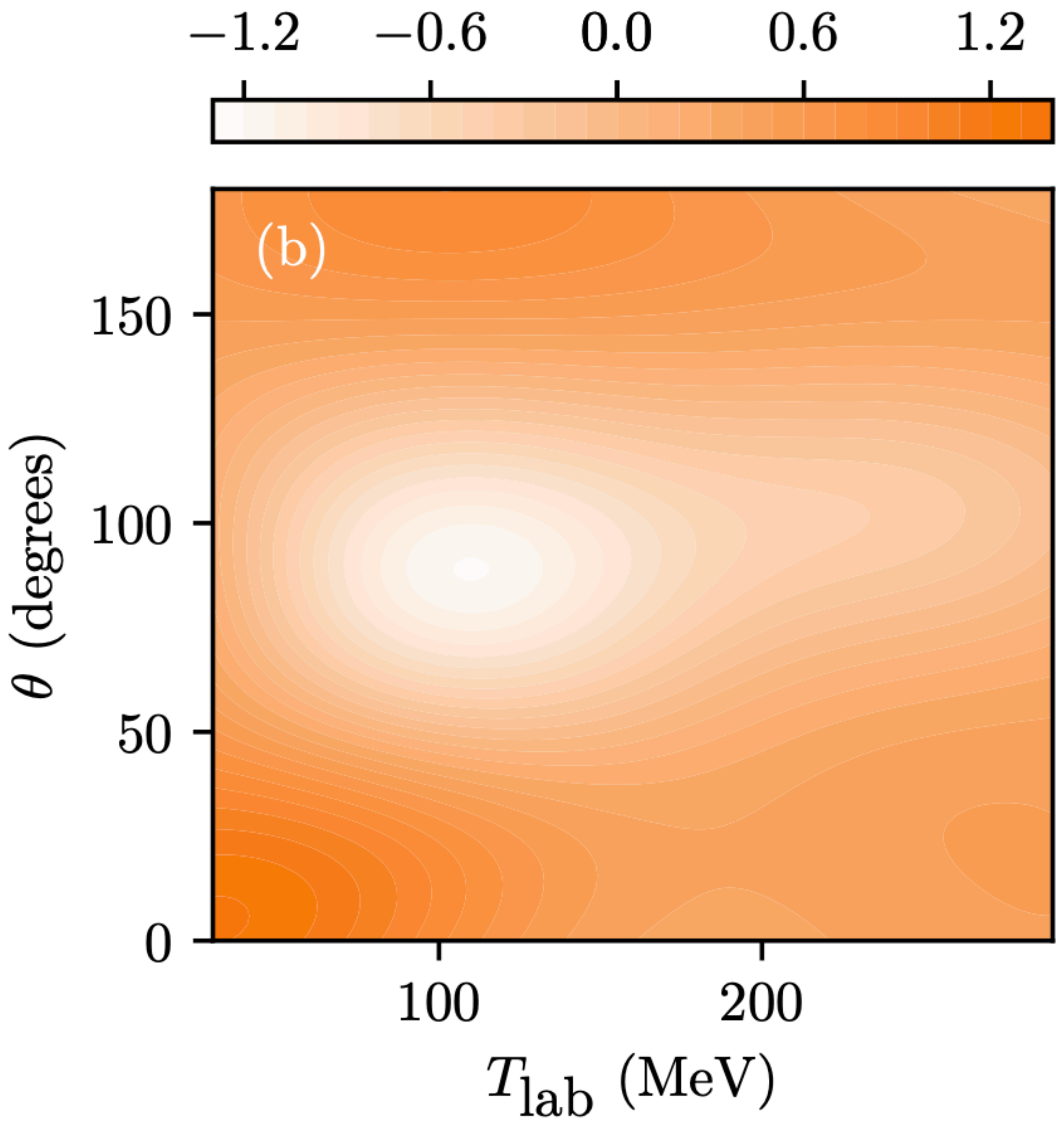
$$\gamma(\boldsymbol{x})\,|\,m, \vec{\ell}, \bar{c}^2 \sim \mathcal{GP}[m, \bar{c}^2 k(\boldsymbol{x}', \boldsymbol{x}; \vec{\ell})]$$



Our prior for the GP hyperparameters

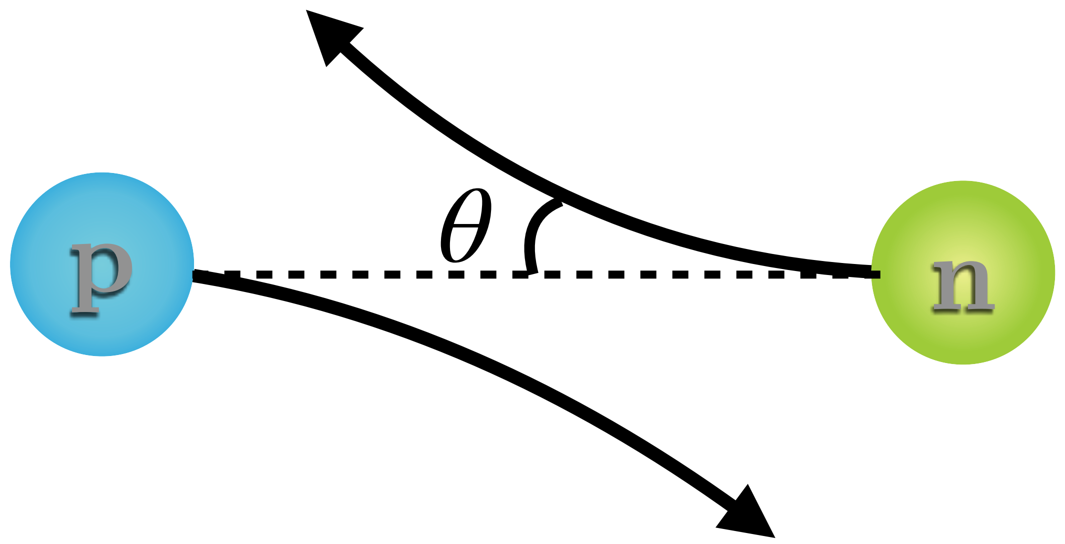
uniform distribution for the lengthscales $\vec{\ell}$

Our posterior for the EFT expansion coefficients



Isak Svensson, Andreas Ekström, Christian Forssén arXiv:2304.02004 [nucl-th]

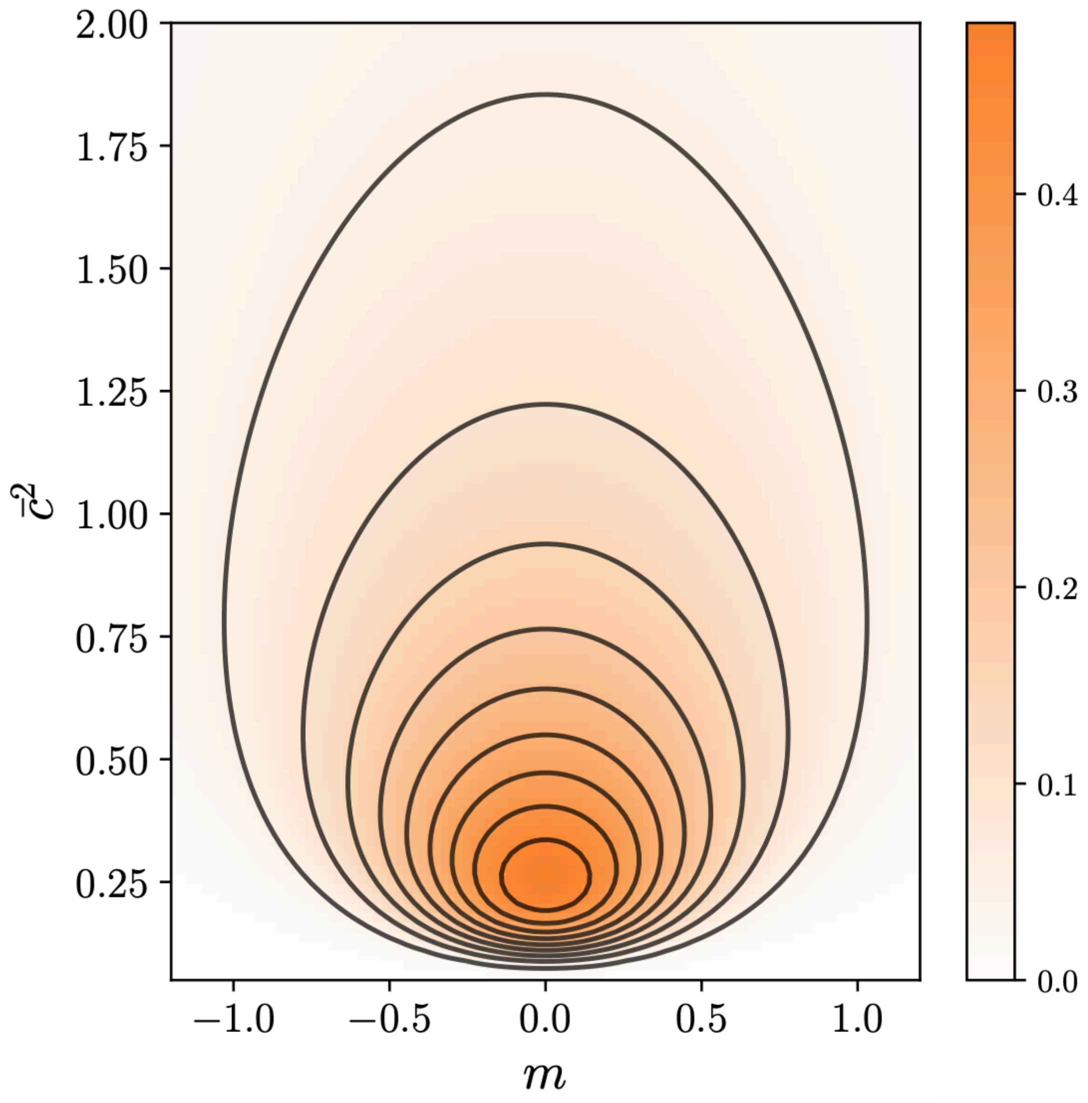
A Gaussian-process model for correlated δM



Quantify expansion coefficients by evaluating the model at consecutive chiral orders for a 2D grid of $\vec{x}_j$ values (energies $T_{\rm lab}$ and angles θ).
n.b. we do this for a sensible choice of $\vec{\alpha} = \vec{\alpha}^\star$

We model the correlated expansion parameters as a two-feature Gaussian process using a SQE kernel

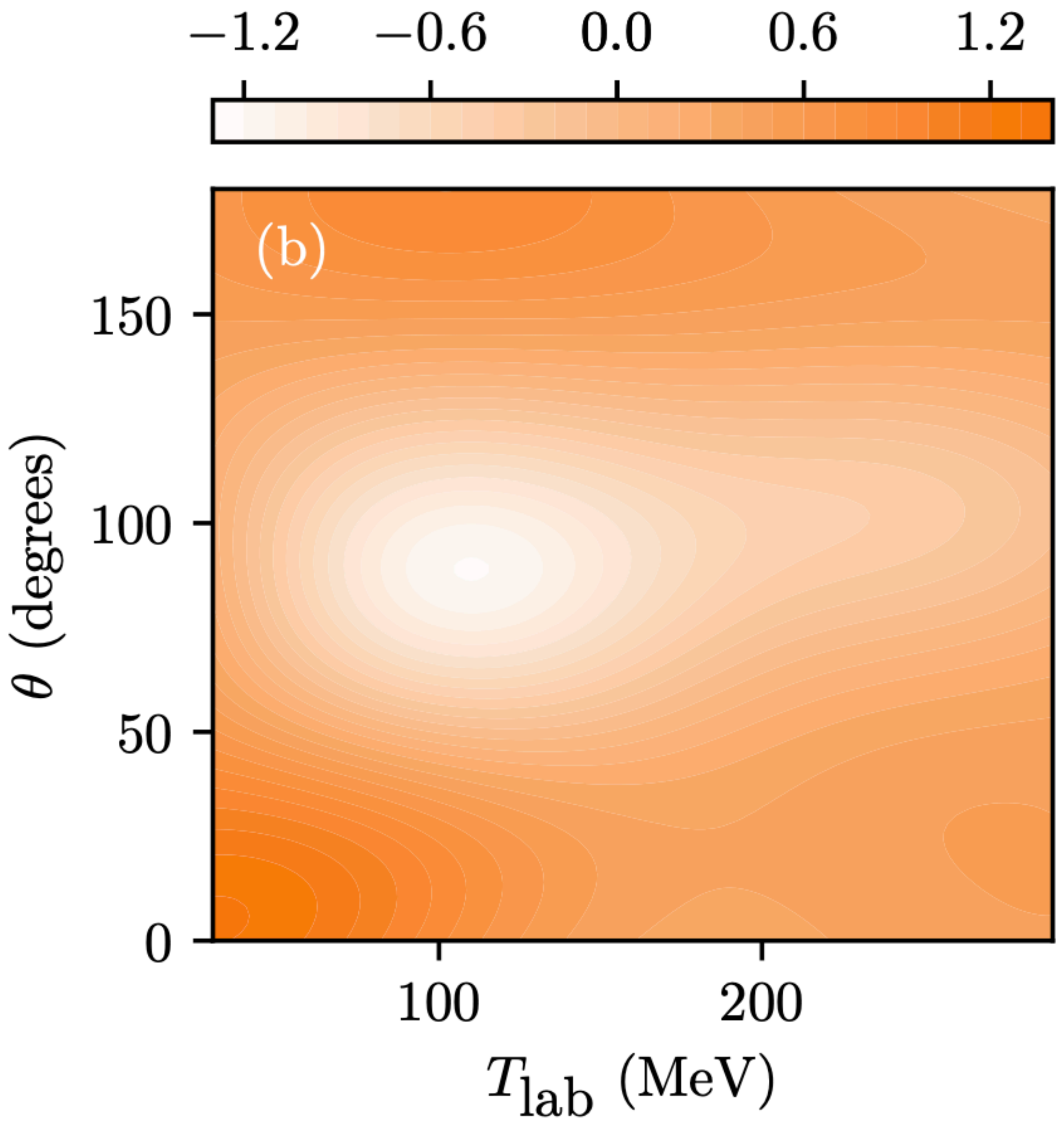
$$\gamma(\boldsymbol{x})\,|\,m, \vec{\ell}, \bar{c}^2 \sim \mathcal{GP}[m, \bar{c}^2 k(\boldsymbol{x}', \boldsymbol{x}; \vec{\ell})]$$



uniform distribution for the lengthscales $\vec{\ell}$

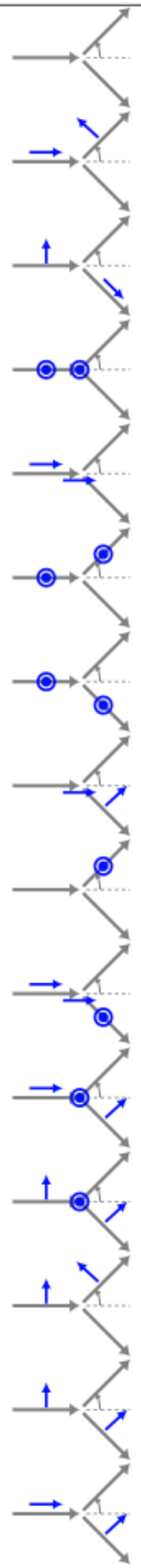
Our prior for the GP hyperparameters

Our posterior for the EFT expansion coefficients



$$\delta M \sim \mathcal{GP}[M, K]$$

Notation	Definition		Acronym	$N_{d,y}$	$N_{T_{\text{lab}},y}$	n_{eff}	$\widehat{\ell_{T_{\text{lab}}}}$ (MeV)		$\widehat{\bar{c}}^2$
σ_{tot}	total cross section		SGT	119	113	26.7	61		0.61^2
σ_T	$\sigma_{\text{tot}}(\uparrow\downarrow) - \sigma_{\text{tot}}(\uparrow\uparrow)$		SGTT	3	3	—	—		—
σ_L	$\sigma_{\text{tot}}(\leftrightarrows) - \sigma_{\text{tot}}(\rightrightarrows)$		SGTL	4	4	3.7	59		2.28^2
Notation	**Tensor**	**Illustration**	**Acronym**	$N_{d,y}$	$N_{T_{\text{lab}},y}$	n_{eff}	$\widehat{\ell_{T_{\text{lab}}}}$ (MeV)	$\widehat{\ell_\theta}$ (deg)	$\widehat{\bar{c}}^2$
$\sigma(\theta)$	I_{0000}		DSG	1207	68	383.8	73	39	0.56^2
$A(\theta)$	D_{s0k0}		A	5	1	5.0	70	37	0.66^2
$A_t(\theta)$	K_{0ks0}		AT	30	2	23.0	60	36	0.61^2
$A_{yy}(\theta)$	A_{00nn}		AYY	58	4	19.4	60	33	0.97^2
$A_{zz}(\theta)$	A_{00kk}		AZZ	45	2	12.1	120	37	1.45^2
$D(\theta)$	D_{n0n0}		D	13	1	4.7	68	26	0.6^2
$D_t(\theta)$	K_{0nn0}		DT	36	5	33.6	41	45	0.52^2
$D_x^z(\theta)$	D_{0s0k}		D0SK	8	1	3.1	59	29	0.52^2
$A_y(\theta)$	P_{n000}		P	503	28	269.6	65	32	0.35^2
$N_{zz}^y(\theta)$	N_{0nkk}		NNKK	8	1	8.0	45	25	0.22^2
$N_{zy}^x(\theta)$	N_{0skn}		NSKN	12	1	11.0	83	36	0.85^2
$N_{xy}^x(\theta)$	N_{0ssn}		NSSN	4	1	4.0	78	31	0.62^2
$R(\theta)$	D_{s0s0}		R	5	1	5.0	74	32	0.67^2
$R_t(\theta)$	K_{0ss0}		RT	29	3	24.2	87	28	0.59^2
$R_t'(\theta)$	K_{0sk0}		RPT	1	1	1.0	62	35	0.54^2



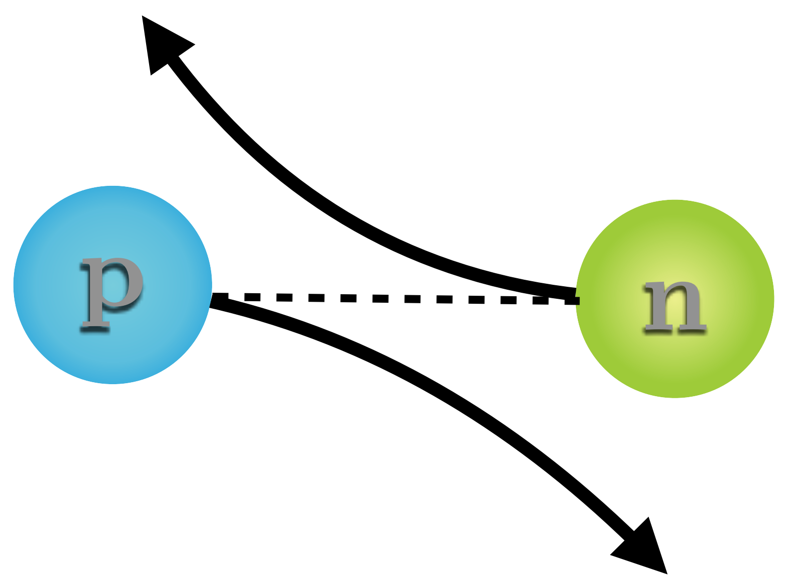
We construct a model discrepancy for each observable type (y) we condition our inference on.

correlation lengths for *np* scattering energies and angles in the ranges **40–120 MeV and 25–45 degrees.**

small marginal variances

Isak Svensson, Andreas Ekström, Christian Forssén arXiv:2304.02004 [nucl-th]

The likelihood



$$\text{pr}(\boldsymbol{\mathcal{D}}|\vec{\alpha}, I) = \prod_{y=1}^{N_y} \text{pr}(\boldsymbol{\mathcal{D}}_y|\vec{\alpha}, I_y),$$

$$\text{pr}(\boldsymbol{\mathcal{D}}_y|\vec{\alpha}, I_y) \propto \exp\left[-\frac{\boldsymbol{r}_y^T(\vec{\alpha})(\boldsymbol{\Sigma}_{\text{exp},y} + \boldsymbol{\Sigma}_{\text{th},y})^{-1}\boldsymbol{r}_y(\vec{\alpha})}{2}\right]$$

$$r_{y,j}(\vec{\alpha}) = [y_{\text{exp}}(\vec{x}_j) - y_{\text{th}}^{(k)}(\vec{\alpha}; \vec{x}_j)]$$

$$\boldsymbol{\Sigma}_{\text{th}} = \begin{bmatrix} \boldsymbol{\Sigma}_{\text{th},1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\text{th},2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Sigma}_{\text{th},N_y} \end{bmatrix}$$

$$(\boldsymbol{\Sigma}_{\text{th},y})_{mn} = \text{cov}[\delta y_{\text{th}}^{(k)}(\vec{x}_m), \delta y_{\text{th}}^{(k)}(\vec{x}_n)]$$

The likelihood

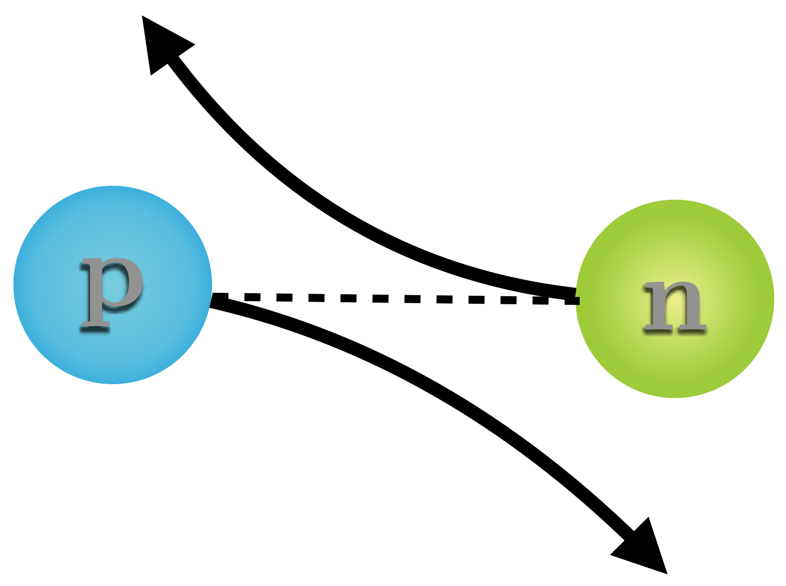
$$\mathrm{pr}(\boldsymbol{\mathcal{D}}|\vec{\alpha}, I) = \prod_{y=1}^{N_y} \mathrm{pr}(\boldsymbol{\mathcal{D}}_y|\vec{\alpha}, I_y),$$

$$\mathrm{pr}(\boldsymbol{\mathcal{D}}_y|\vec{\alpha}, I_y) \propto \exp\left[-\frac{\boldsymbol{r}_y^T(\vec{\alpha})(\boldsymbol{\Sigma}_{\mathrm{exp},y} + \boldsymbol{\Sigma}_{\mathrm{th},y})^{-1}\boldsymbol{r}_y(\vec{\alpha})}{2}\right]$$

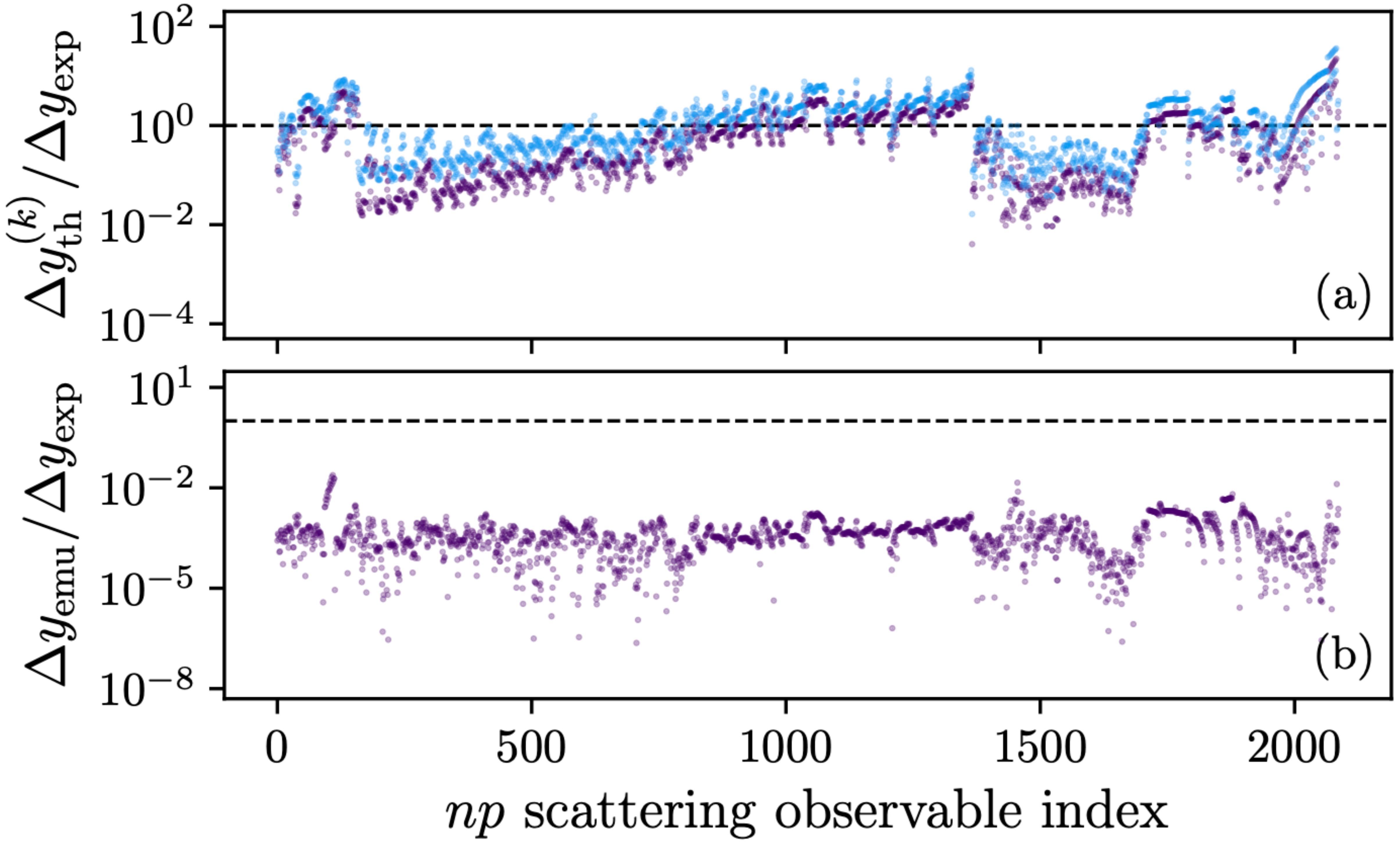
$$r_{y,j}(\vec{\alpha}) = [y_{\mathrm{exp}}(\vec{x}_j) - y_{\mathrm{th}}^{(k)}(\vec{\alpha}; \vec{x}_j)]$$

$$\boldsymbol{\Sigma}_{\mathrm{th}} = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathrm{th},1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\mathrm{th},2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Sigma}_{\mathrm{th},N_y} \end{bmatrix}$$

$$(\boldsymbol{\Sigma}_{\mathrm{th},y})_{mn} = \mathrm{cov}[\delta y_{\mathrm{th}}^{(k)}(\vec{x}_m), \delta y_{\mathrm{th}}^{(k)}(\vec{x}_n)]$$



We emulate all ~2000 cross sections in $\mathscr{D}$. Takes 1s (non-threaded + Google JAX) to evaluate likelihood and its derivatives.

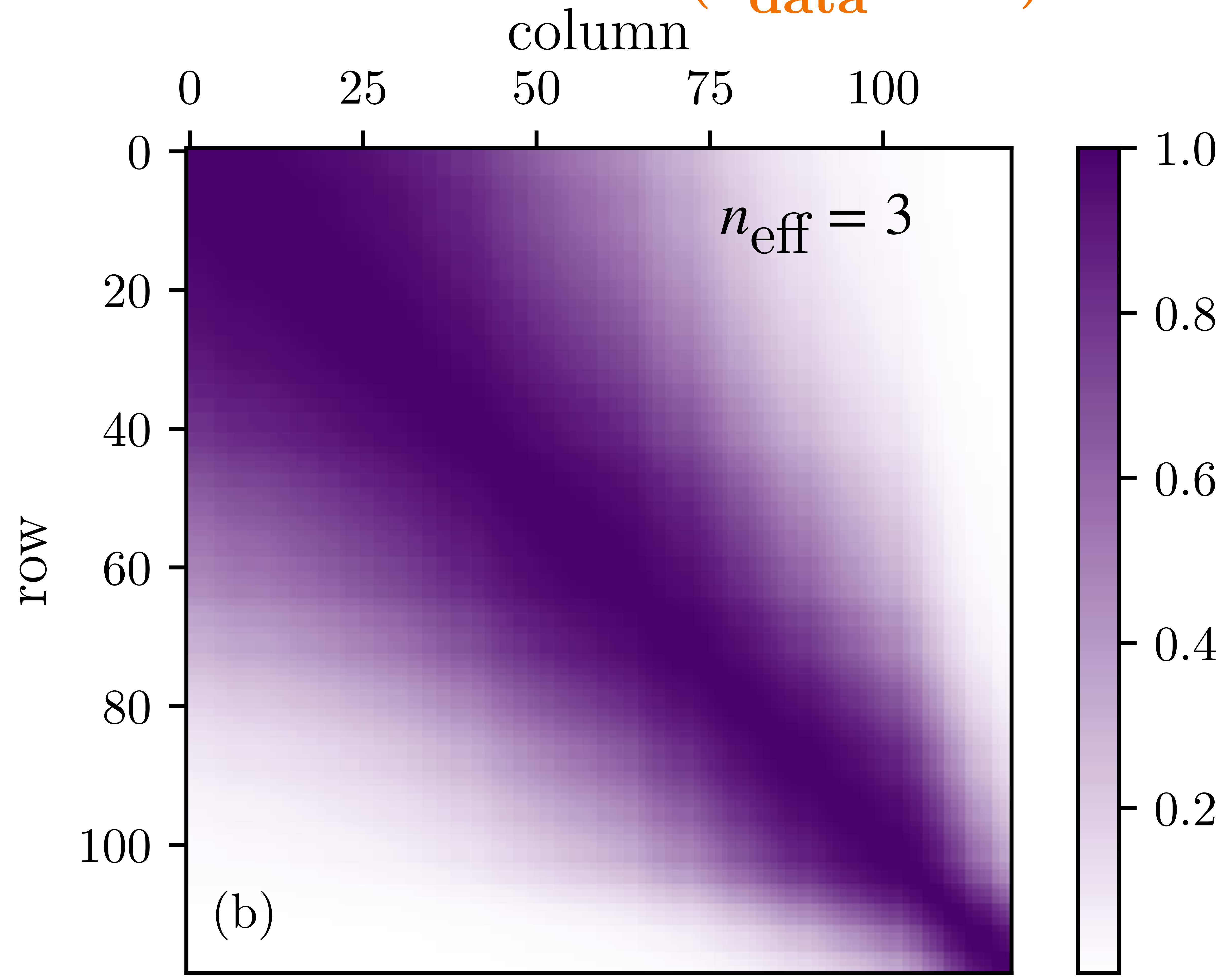


Isak Svensson, Andreas Ekström, Christian Forssén arXiv:2304.02004 [nucl-th]

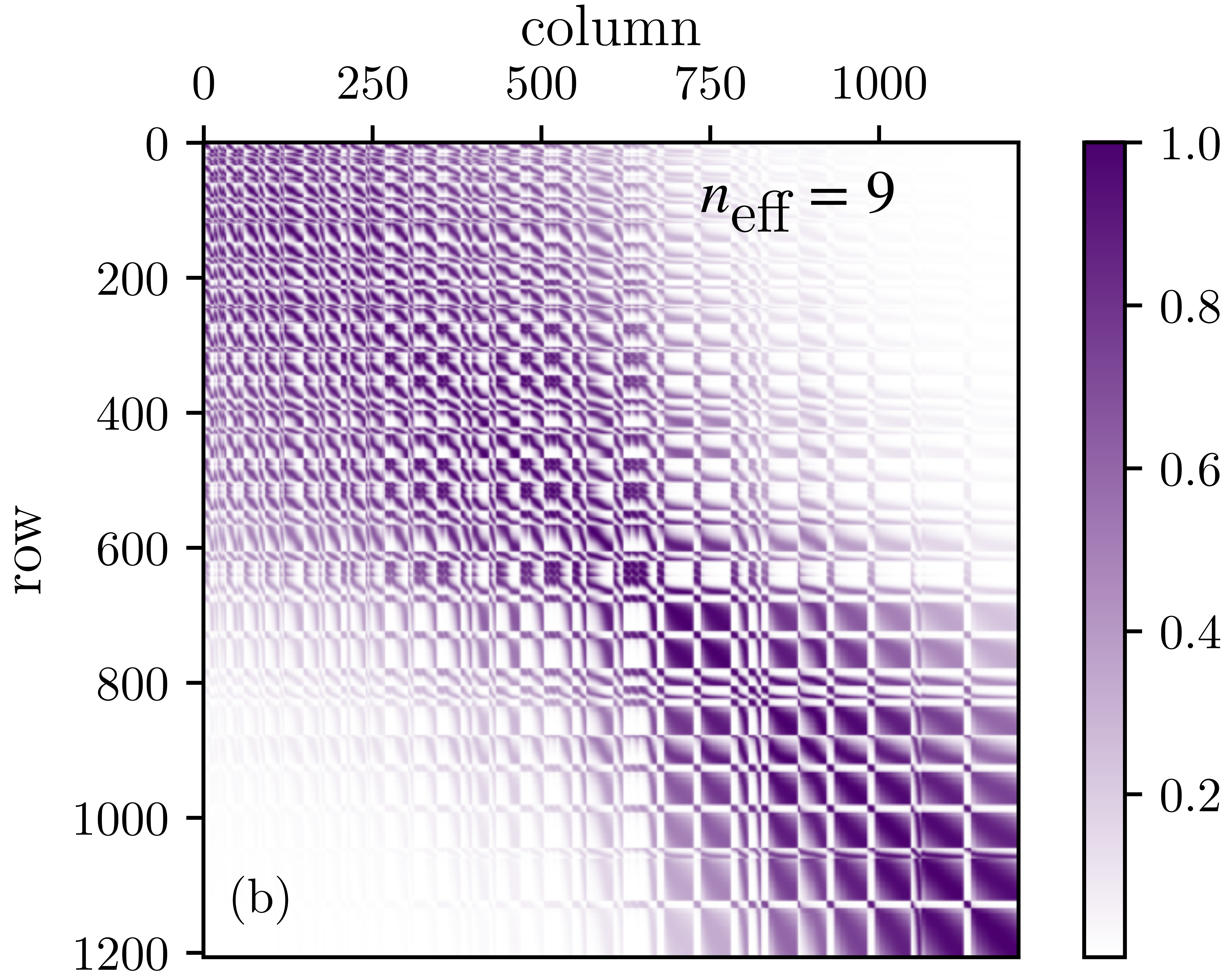
See R. Furnstahl's talk for more on emulators

Effects of a small marginal variance in the model discrepancy

total cross section ($N_{\mathbf{data}} = 119$)

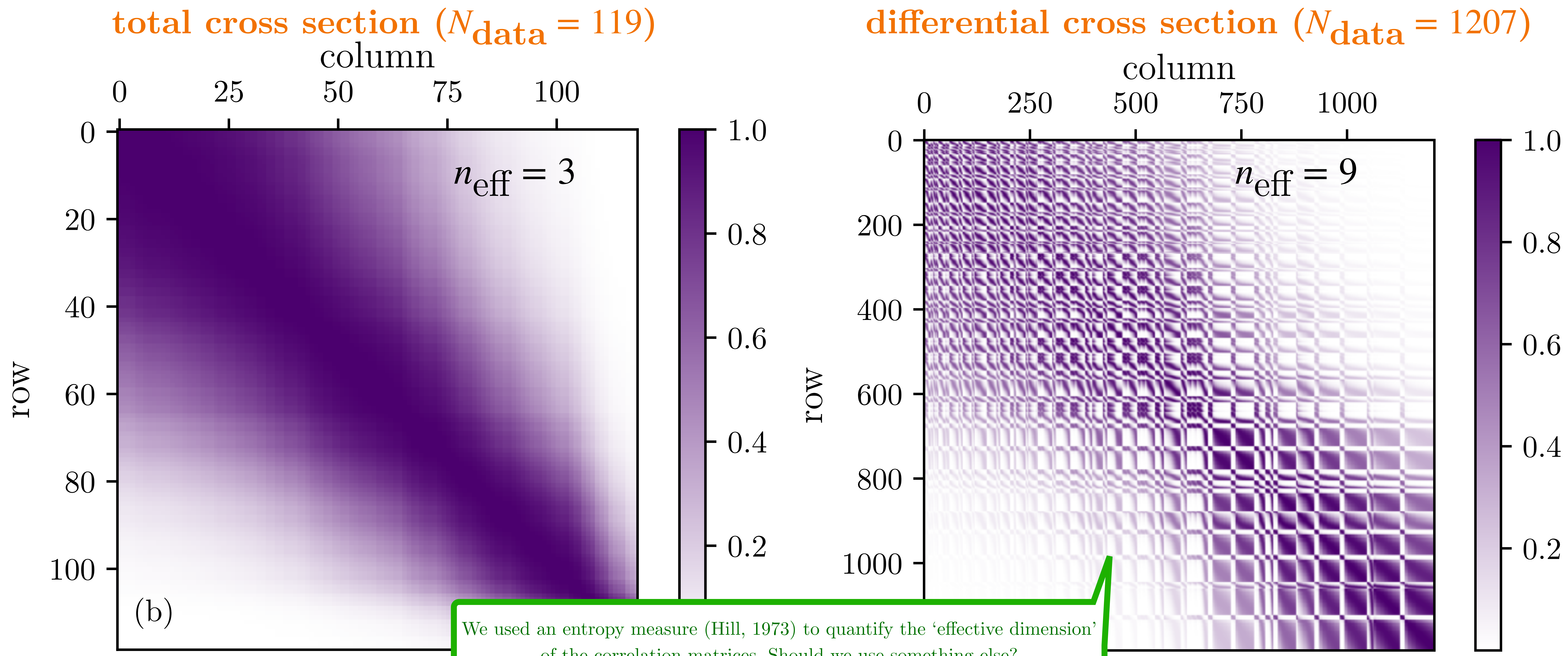


differential cross section ($N_{\mathbf{data}} = 1207$)



Isak Svensson, Andreas Ekström, Christian Forssén arXiv:2304.02004 [nucl-th]

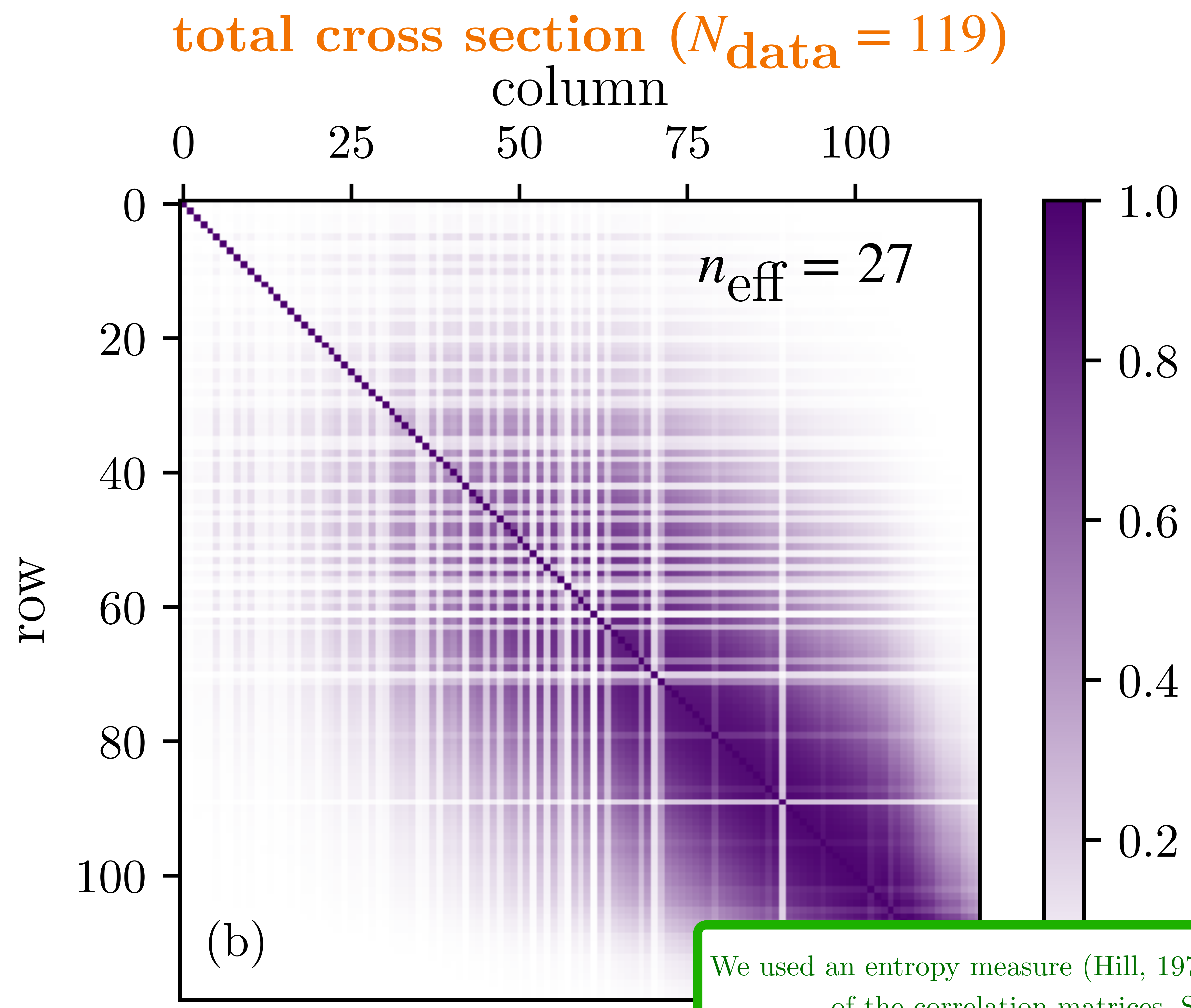
Effects of a small marginal variance in the model discrepancy

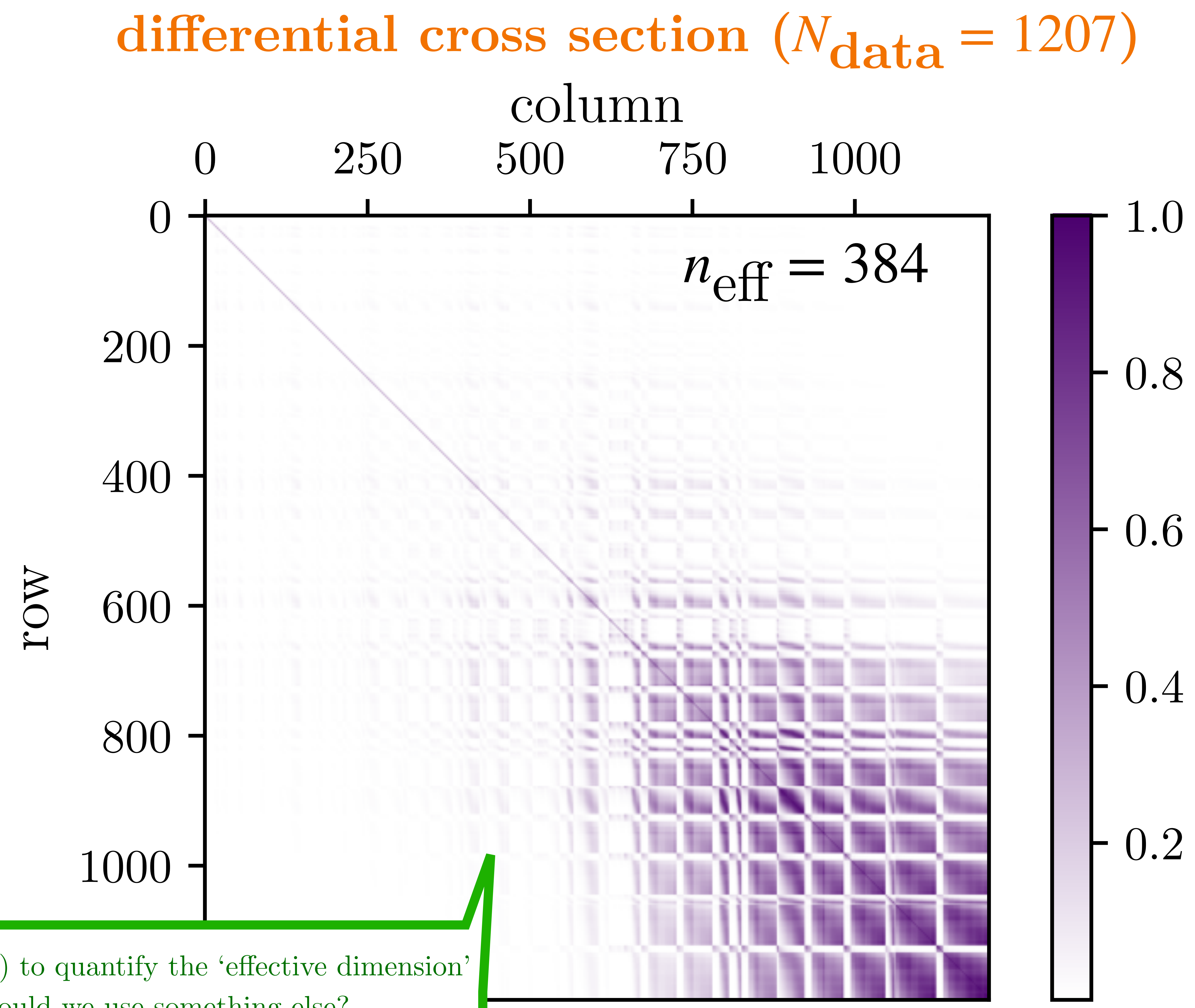


We used an entropy measure (Hill, 1973) to quantify the 'effective dimension' of the correlation matrices. Should we use something else?

Isak Svensson, Andreas Ekström, Christian Forssén arXiv:2304.02004 [nucl-th]

Effects of a small marginal variance in the model discrepancy





We used an entropy measure (Hill, 1973) to quantify the 'effective dimension' of the correlation matrices. Should we use something else?

Bayes' rule: from likelihood & prior to posterior

- Collect N data points that we gather in a data vector D
- To explain the data, propose some model M, depending on parameters $\vec{\alpha}$
- Apply Bayes' rule

Posterior **Likelihood** **Prior**

$$p(\vec{\alpha}\,|\,D,M,I) = \frac{p(D\,|\,\vec{\alpha},M,I)\cdot p(\vec{\alpha}\,|\,M,I)}{p(D\,|\,M,I)}$$

Marginal likelihood

most likely not Rev. T. Bayes

- The **prior** encodes our knowledge about the parameter values before analyzing the data
- The **likelihood** is the probability of the data given a set of parameters
- The **marginal likelihood** (or model evidence) provides normalization of the posterior
- The **posterior** is the complete inference and resulting probability density for the parameters $\vec{\alpha}$

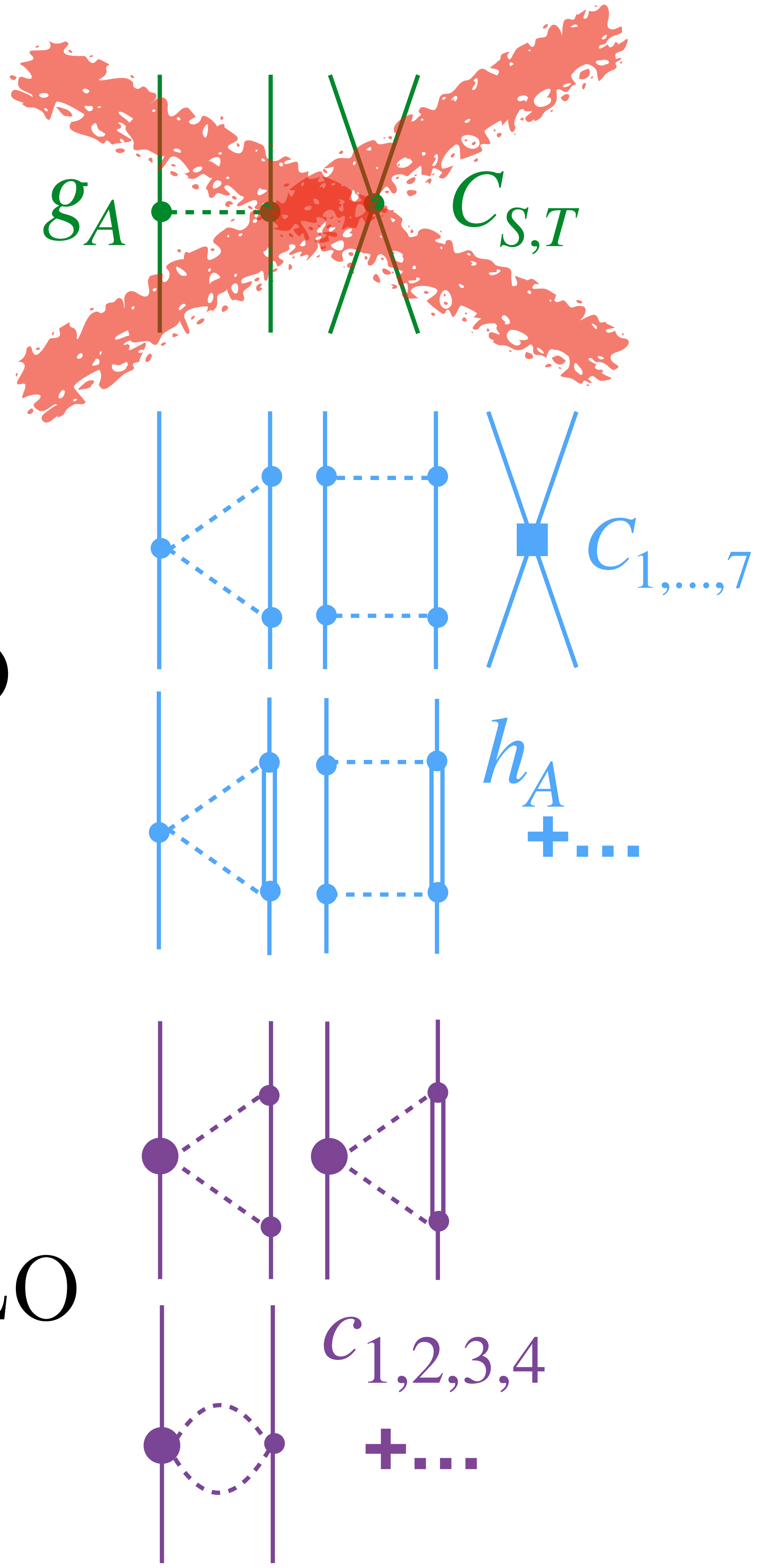
Setting up the prior for the model params'

$$p(\vec{\alpha}\,|\,I) = p(\vec{\alpha}_{NN}\,|\,I) \cdot p(\vec{\alpha}_{\pi N}\,|\,I)$$

LO

NLO

NNLO

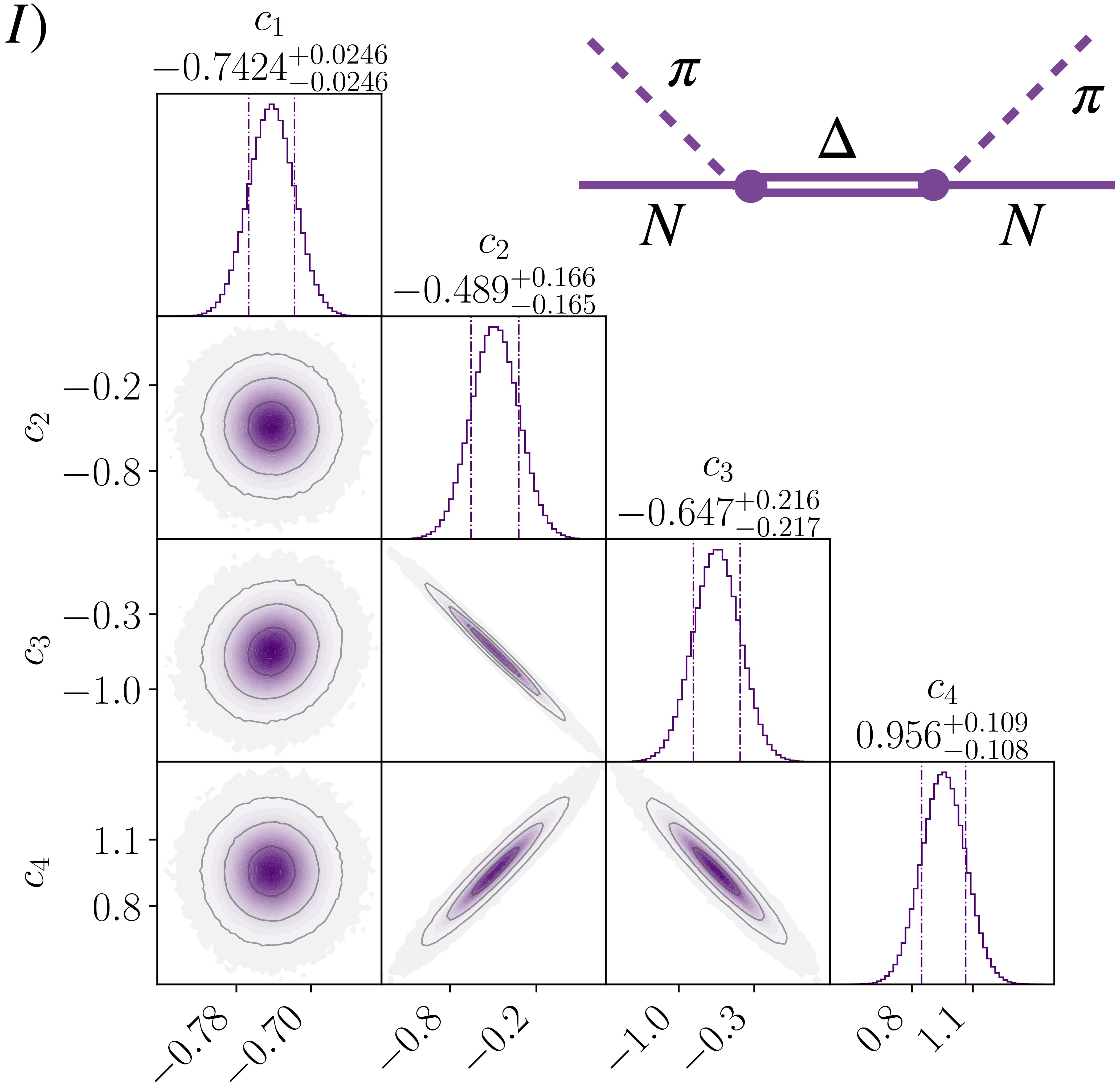


***NN* contacts: iid normal**

$$p(\vec{\alpha}\,|\,I) = p(\vec{\alpha}_{NN}\,|\,I)$$
$$= \mathcal{N}(\mathbf{0}, \mathbf{1} \cdot 5^2)$$

***πN* exchange: Roy-Steiner**

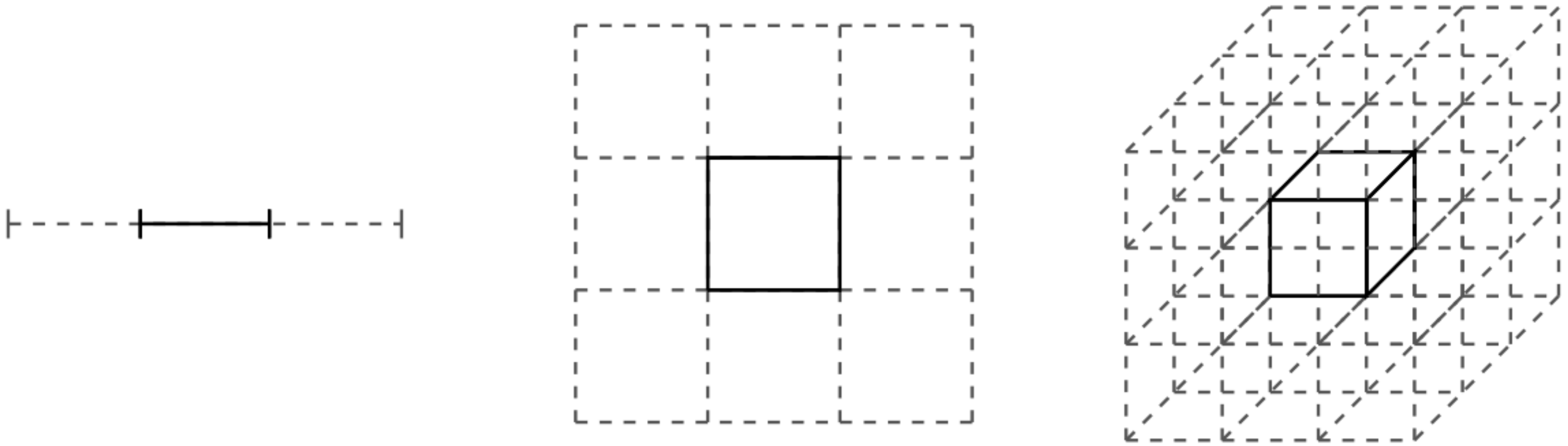
$$p(\vec{\alpha}\,|\,I) = p(\vec{\alpha}_{\pi N}\,|\,I)$$



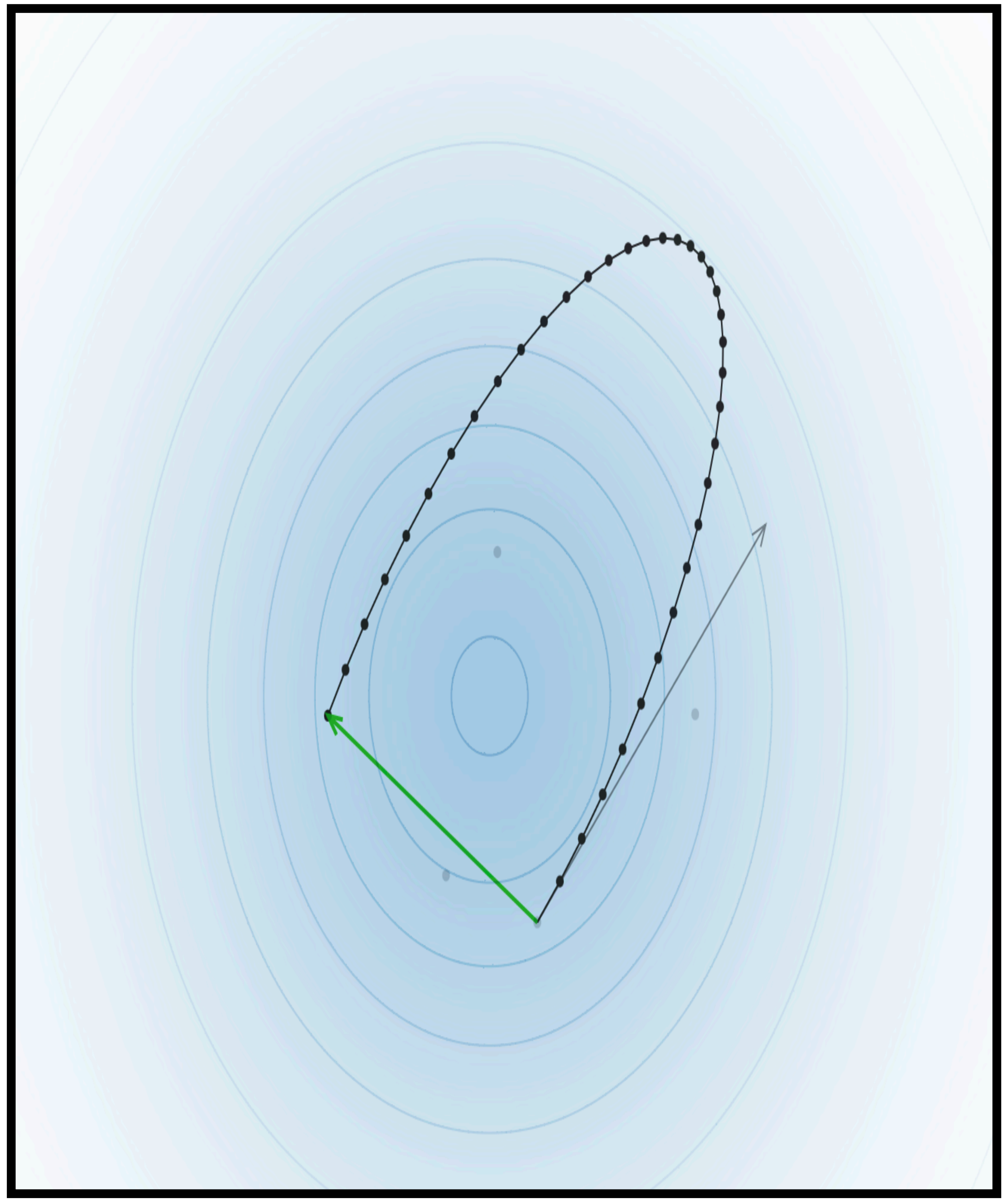
A. Ekström, et al. Phys. Rev **C 97**, 024332 (2018)
W. Jiang, et al. Phys Rev **C 102,** 054301 (2020)

D. Siemens et al, Phys. Lett **B 770**, 27-34 (2017)

Hamiltonian Monte Carlo



credit: M. Betancourt arXiv:1701.02434



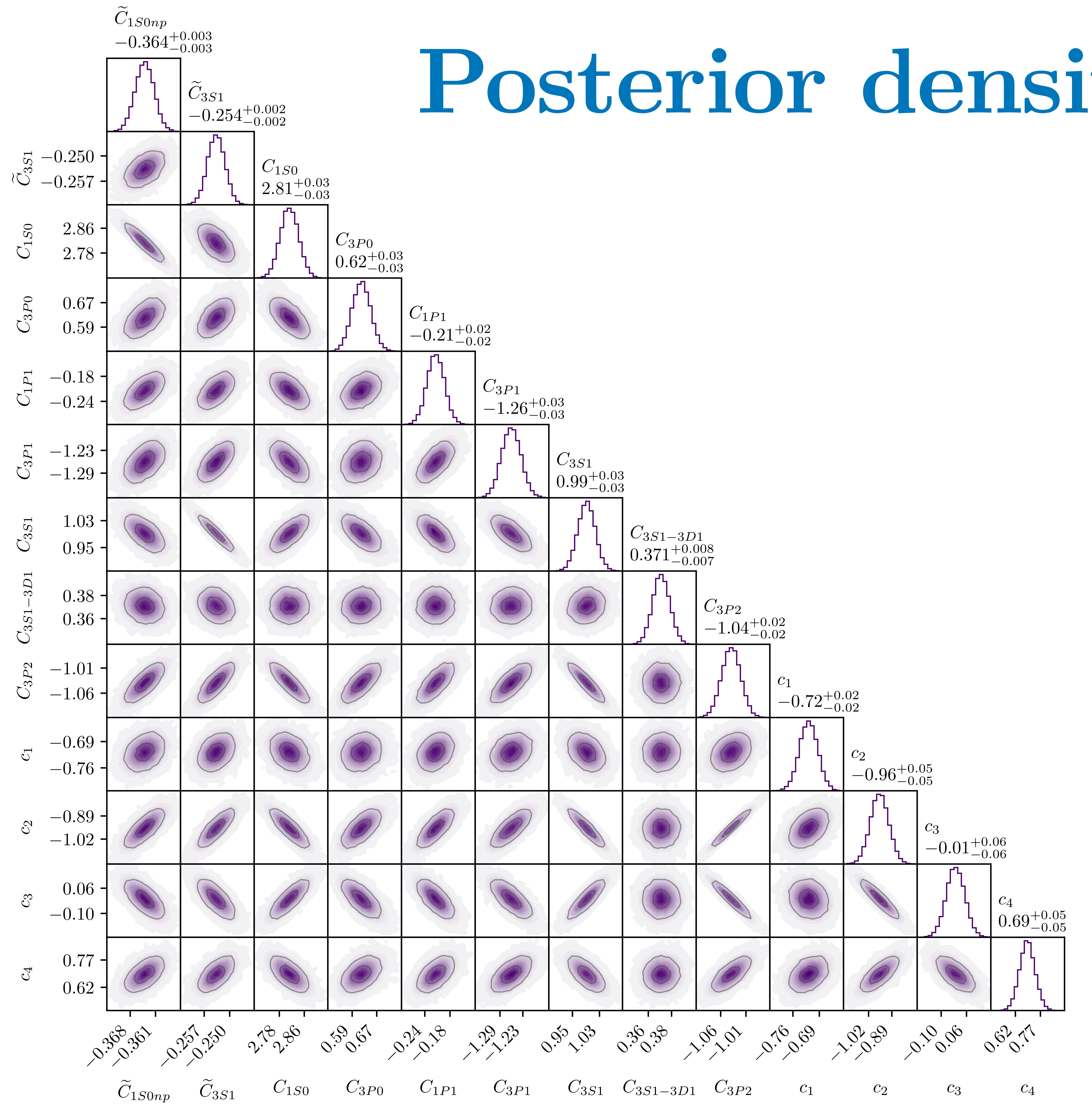
credit: https://chi-feng.github.io/mcmc-demo/

The LEC posterior is often multivariate (30 np/pp LECs at N3LO). Naive "guess and check" (random walk metropolis) will fail exponentially. We use Hamiltonian Monte Carlo (HMC) to take long jumps in parameter space while staying in regions with high probability mass.

See D. Mondal's talk for more on HMC

S. Duane, et al. Phys. Lett B **195,** 216 (1986)
I. Svensson, et al Phys. Rev. C **105**, 014004(2022)
I. Svensson, et al Phys. Rev. C **107**, 014001(2023)

Posterior density (correlated δM)

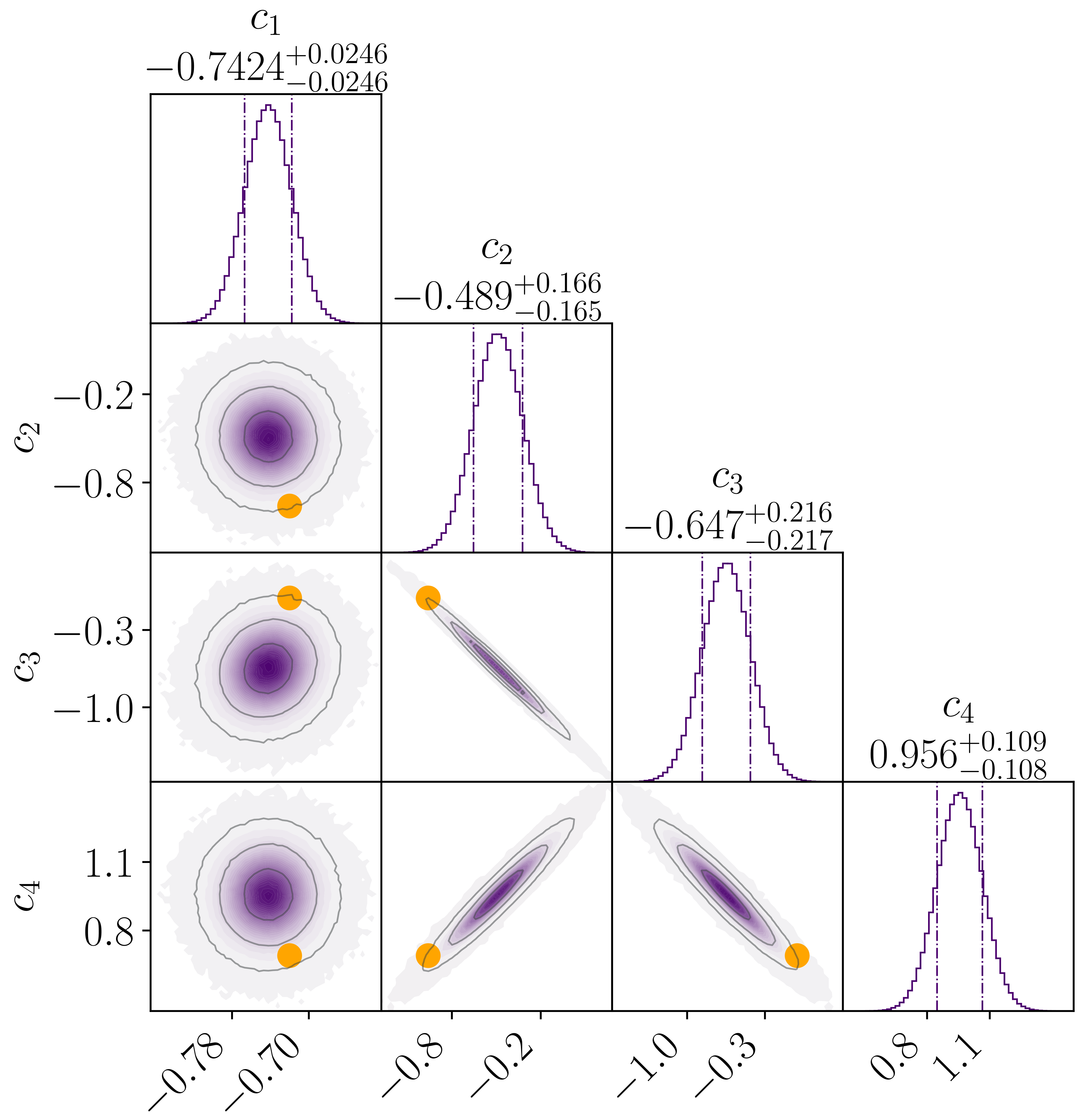


credible ranges for the model parameters increase by a factor 1.5 when accounting for correlations in the model discrepancy

np data reduces the prior credible ranges for the πN parameters c_1, c_2, c_3, c_4 by a factor 3-5

Isak Svensson, Andreas Ekström, Christian Forssén arXiv:2304.02004 [nucl-th]

The πN prior vs posterior MAP

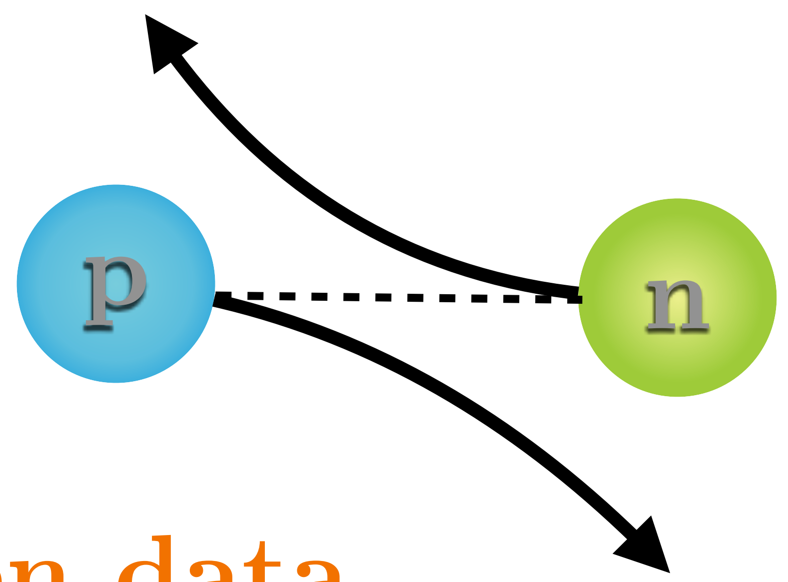


The Mahalanobis distance ($D^2_{MD} = 9.95$) between the πN posterior MAP and prior (normal) mean is far enough to be outside 95% of the prior probability mass.

We need a better understanding of the model discrepancy and the underlying EFT.

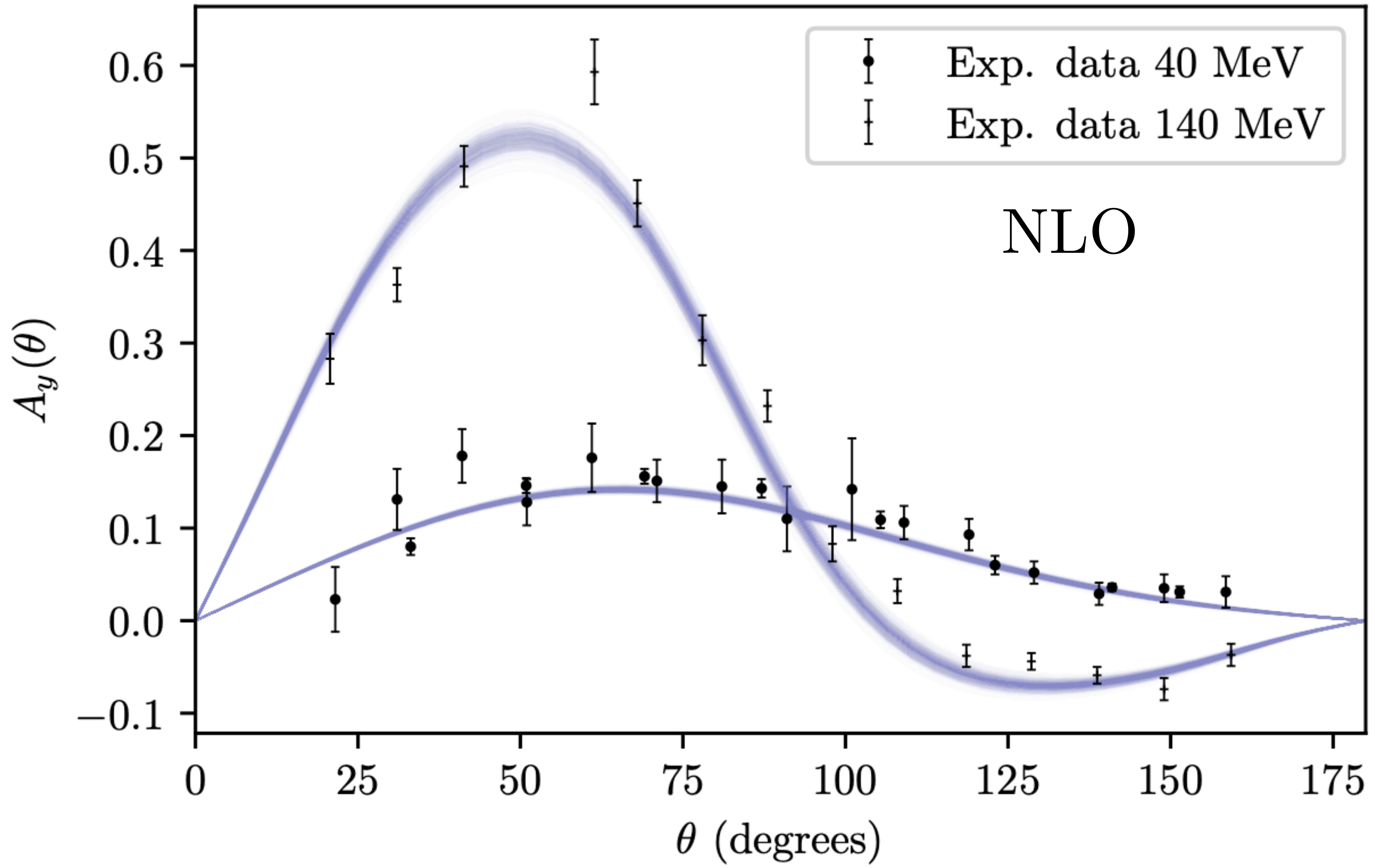
Isak Svensson, Andreas Ekström, Christian Forssén arXiv:2304.02004 [nucl-th]

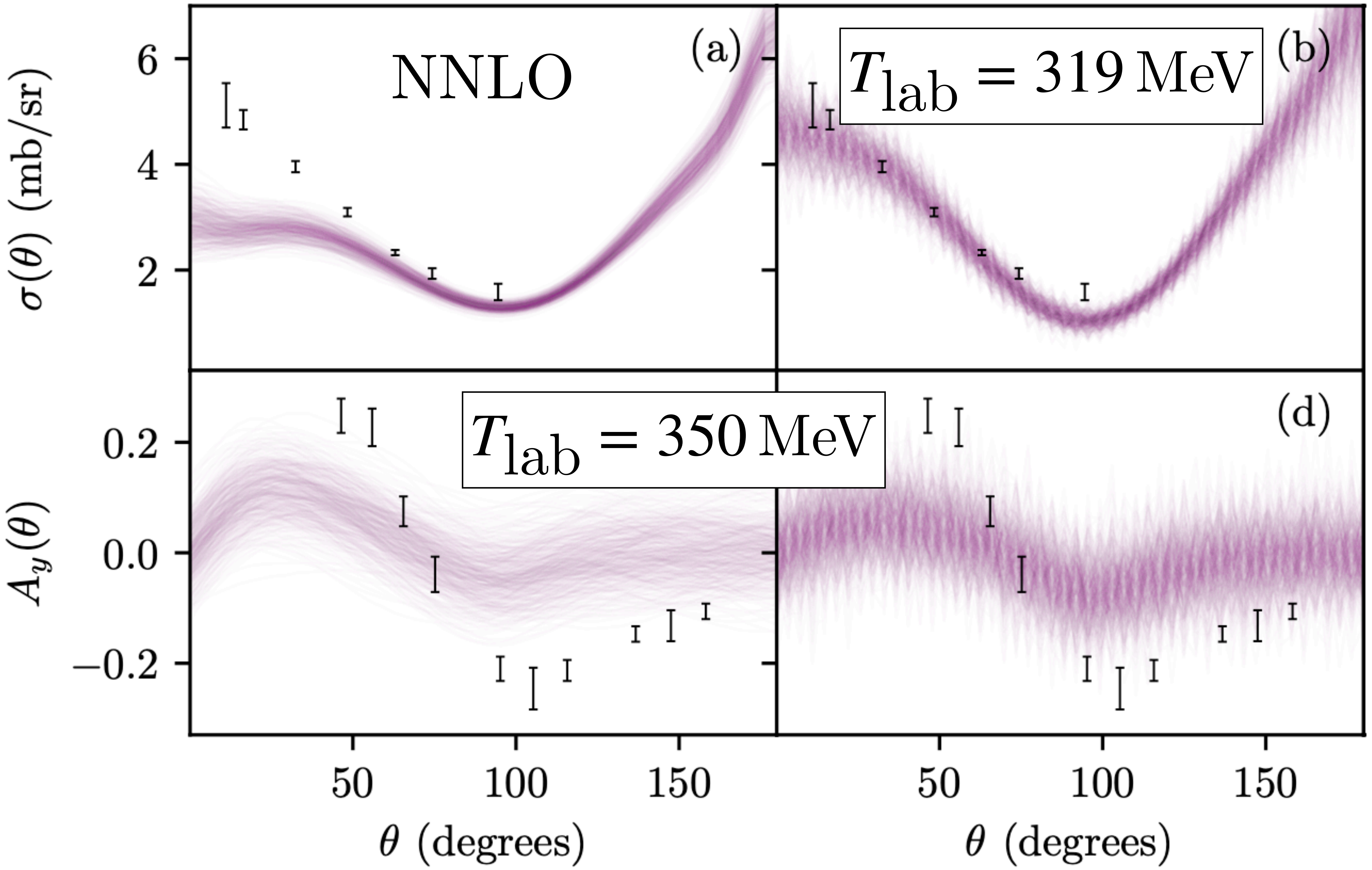
Posterior predictive distributions



seen data

unseen data
(some of the 'bad ones')

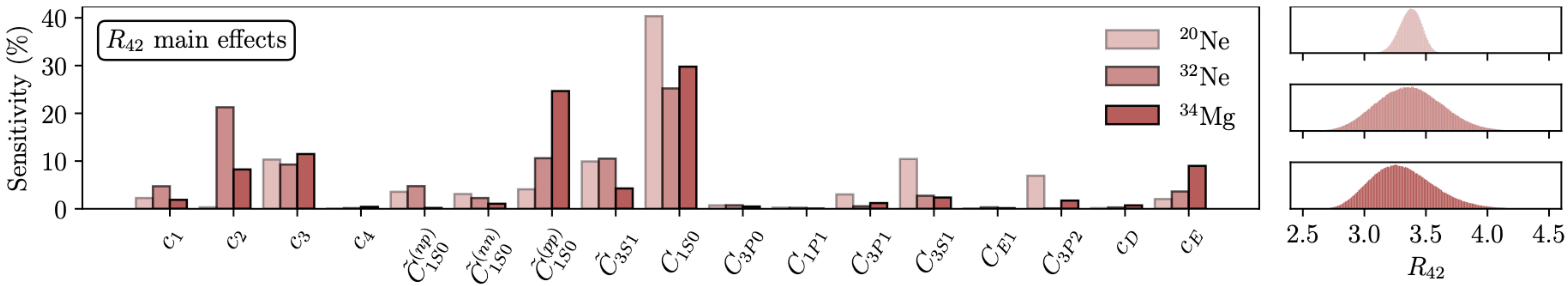


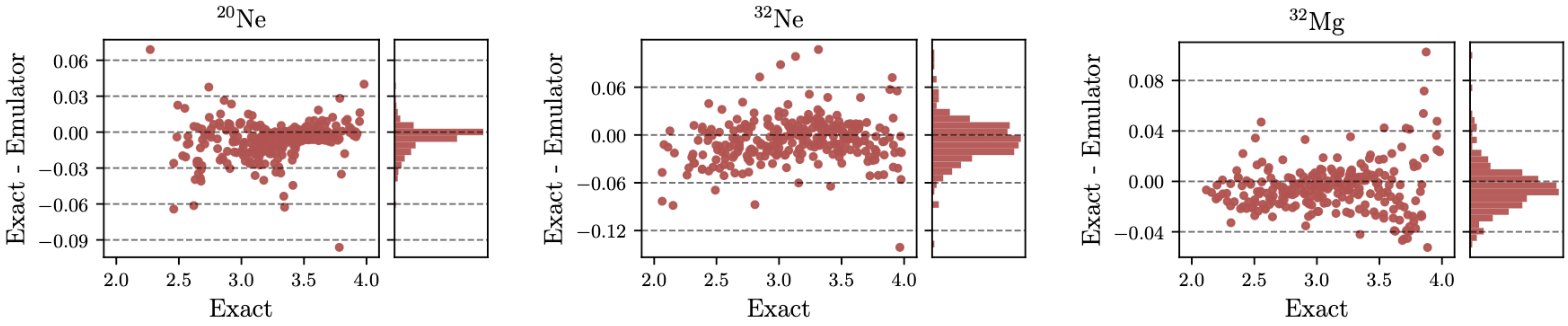


Isak Svensson, Andreas Ekström, Christian Forssén arXiv:2304.02004 [nucl-th]

Global sensitivity analysis

Emulating excited state energies of deformed atomic nuclei





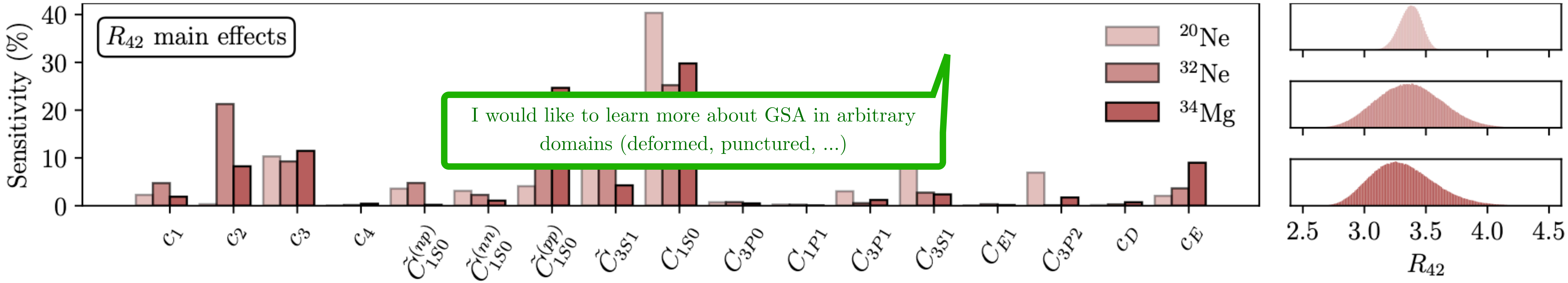
See K. Becker's talk for more on GSA

A Ekström, C Forssén, G Hagen, GR Jansen, T Papenbrock, Z. H. Sun arXiv:2305.06955

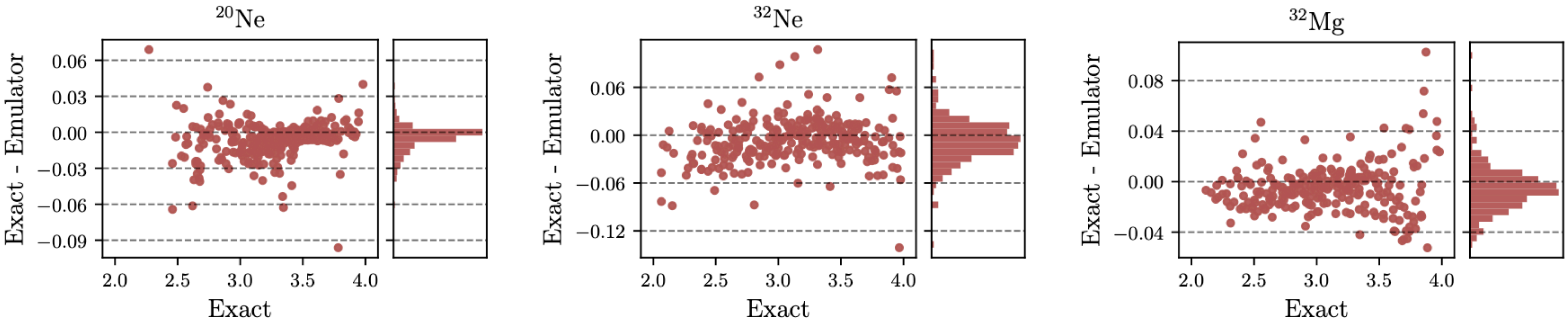
Global sensitivity analysis

Nuclear Few- and Many-Body Problems

Emulating excited state energies of deformed atomic nuclei



I would like to learn more about GSA in arbitrary domains (deformed, punctured, ...)



See K. Becker's talk for more on GSA

A Ekström, C Forssén, G Hagen, GR Jansen, T Papenbrock, Z. H. Sun arXiv:2305.06955

- **we developed** a Gaussian process to quantify the correlated model discrepancy (truncation error) of Δ-full chiral effective field theory at NLO and NNLO in *np* scattering. Much of the essential machinery is in place (open source!).

- **we found** that the *correlated* discrepancy has a small impact in Δ-full EFT up to NNLO. Small difference between NLO and NNLO -> underestimated prediction errors.

- **problems abound:**
  - tension between πN and NN sectors. (maybe just misspecified discrepancy)
  - sampling in high-dimensional spaces is always challenging.
  - do we have a sensible chiral EFT expansion? (likely not)
  - always a challenge to solve the many-body Schrödinger equation & quantify the associated method errors.

- **emulators** work like a charm and open up for many exciting and useful Bayesian analyses in *ab initio* nuclear physics!

Thanks for your attention!