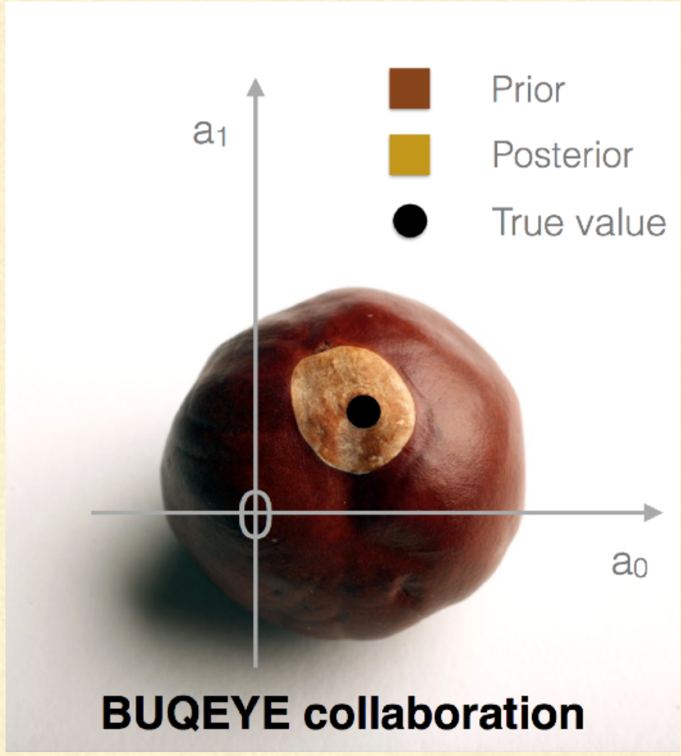
MODELING EFT TRUNCATION ERRORS USING GAUSSIAN PROCESSES

Daniel Phillips
Ohio University
TU Darmstadt
ExtreMe Matter Institute

for the BUQEYE collaboration
(Bayesian Uncertainty Quantification: Errors for Your EFT)

R. J. Furnstahl, J. Melendez (Ohio State University)
DP (Ohio University)
S. Wesolowski (Salisbury University)



RESEARCH SUPPORTED BY THE US DOE AND NSF AND EMMI

What I’m going to say

What I'm going to say

- Truncation errors can be calculated, provided we assume the persistence of EFT coefficient behavior across orders

What I’m going to say

- Truncation errors can be calculated, provided we assume the persistence of EFT coefficient behavior across orders
- The resulting theory error bar is not just something we do to make ourselves look more scientific*: it makes “model checking” possible

*Although yes, it does do that too.

What I'm going to say

- Truncation errors can be calculated, provided we assume the persistence of EFT coefficient behavior across orders
- The resulting theory error bar is not just something we do to make ourselves look more scientific*: it makes "model checking" possible
- EFT truncation errors are correlated across different observables

*Although yes, it does do that too.

What I'm going to say

- Truncation errors can be calculated, provided we assume the persistence of EFT coefficient behavior across orders
- The resulting theory error bar is not just something we do to make ourselves look more scientific*: it makes "model checking" possible
- EFT truncation errors are correlated across different observables
- For NN observables, Gaussian Process models of the truncation error can capture this correlation structure

*Although yes, it does do that too.

What I’m going to say

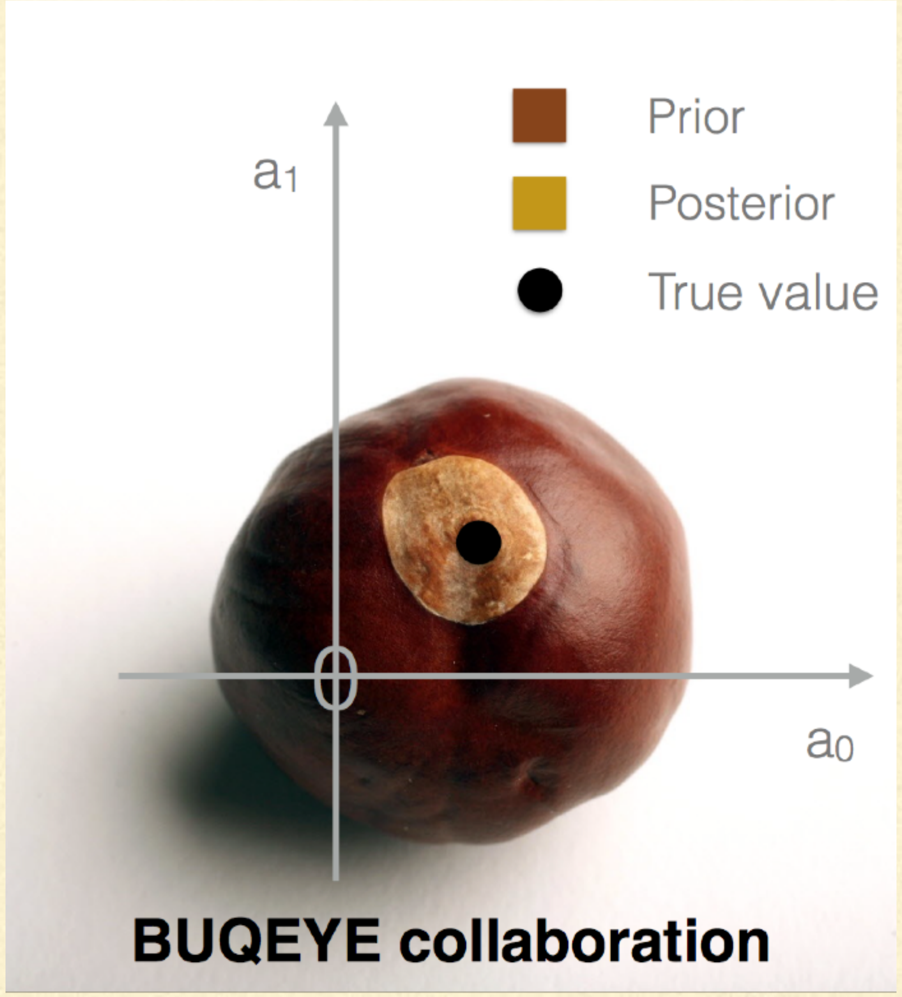
- Truncation errors can be calculated, provided we assume the persistence of EFT coefficient behavior across orders
- The resulting theory error bar is not just something we do to make ourselves look more scientific*: it makes “model checking” possible
- EFT truncation errors are correlated across different observables
- For NN observables, Gaussian Process models of the truncation error can capture this correlation structure

Check two different aspects of “model”

*Although yes, it does do that too.

Outline

- Calculating rigorous truncation errors
- Diagnosing ChiEFT NN expansions
- Gaussian processes for truncation errors
- Preliminary results from analysis of ChiEFT NN potentials with GP truncation errors
- Summary



R. J. Furnstahl, DP, and S. Wesolowski, J. Phys. G **42**, 034028 (2015)

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

S. Wesolowski, N. Klco, R. J. Furnstahl, DP, and A. Thapaliya, J. Phys. G. **43,** 074001 (2016)

J. Melendez, S. Wesolowski, R. J, Furnstahl, Phys. Rev. C **96**, 024003 (2017)

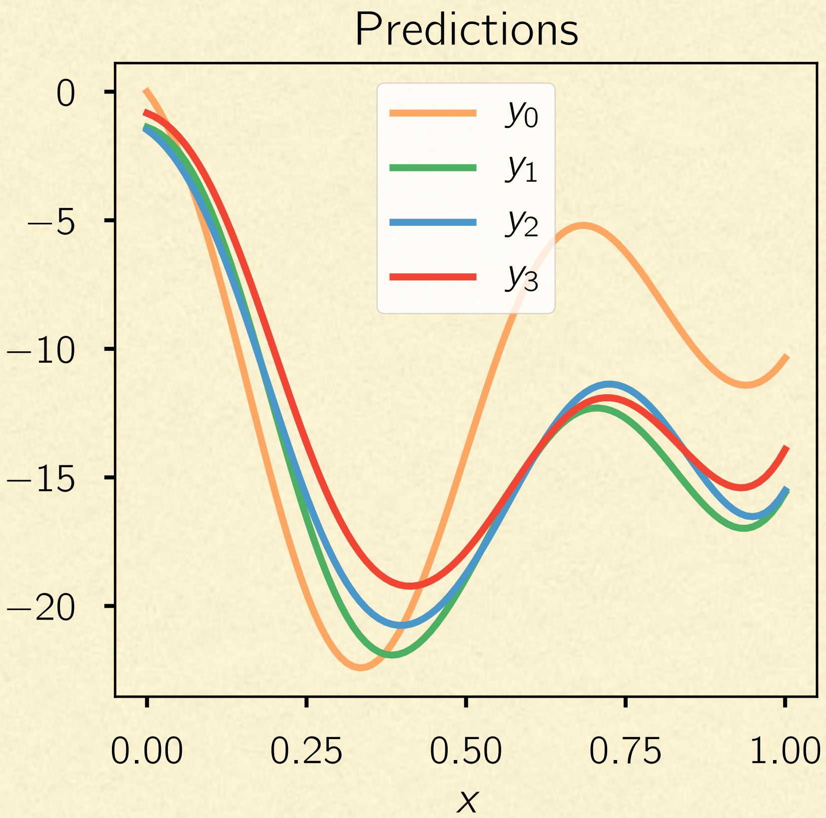
S. Wesolowski, R. Furnstahl, J. Melendez, DP, J. Phys. G **46**, 045102 (2019)

An EFT expansion in pictures

- General EFT series for observable to order k: $y = y_{\rm ref} \sum_{n=0}^{k} c_n Q^n$
- In ChiEFT $Q = \dfrac{(p, m_\pi)}{\Lambda_b}; \quad \Lambda_b \approx 600 \text{ MeV}$

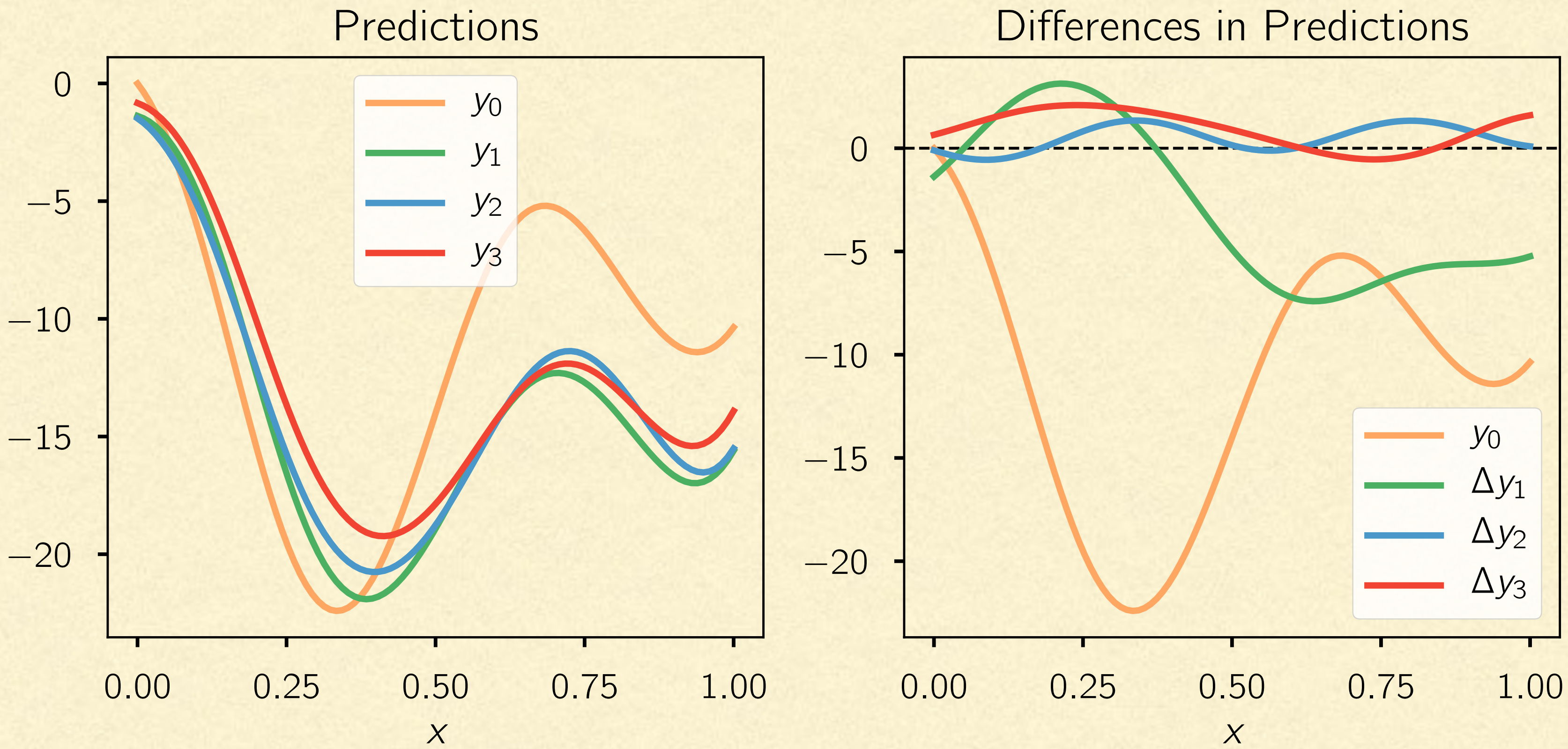
An EFT expansion in pictures

- General EFT series for observable to order k: $y = y_{\text{ref}} \sum_{n=0}^{k} c_n Q^n$
- In ChiEFT $Q = \dfrac{(p, m_\pi)}{\Lambda_b}; \quad \Lambda_b \approx 600 \text{ MeV}$



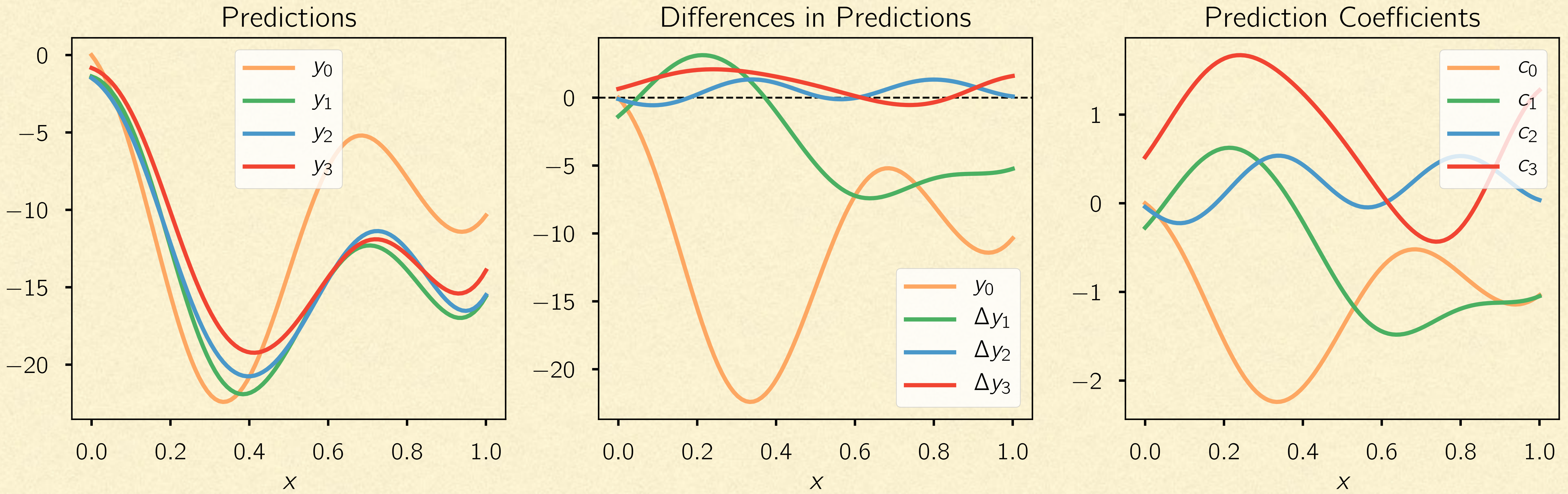
An EFT expansion in pictures

- General EFT series for observable to order k: $y = y_{\mathrm{ref}} \sum_{n=0}^{k} c_n Q^n$
- In ChiEFT $Q = \frac{(p, m_\pi)}{\Lambda_b}; \quad \Lambda_b \approx 600 \text{ MeV}$



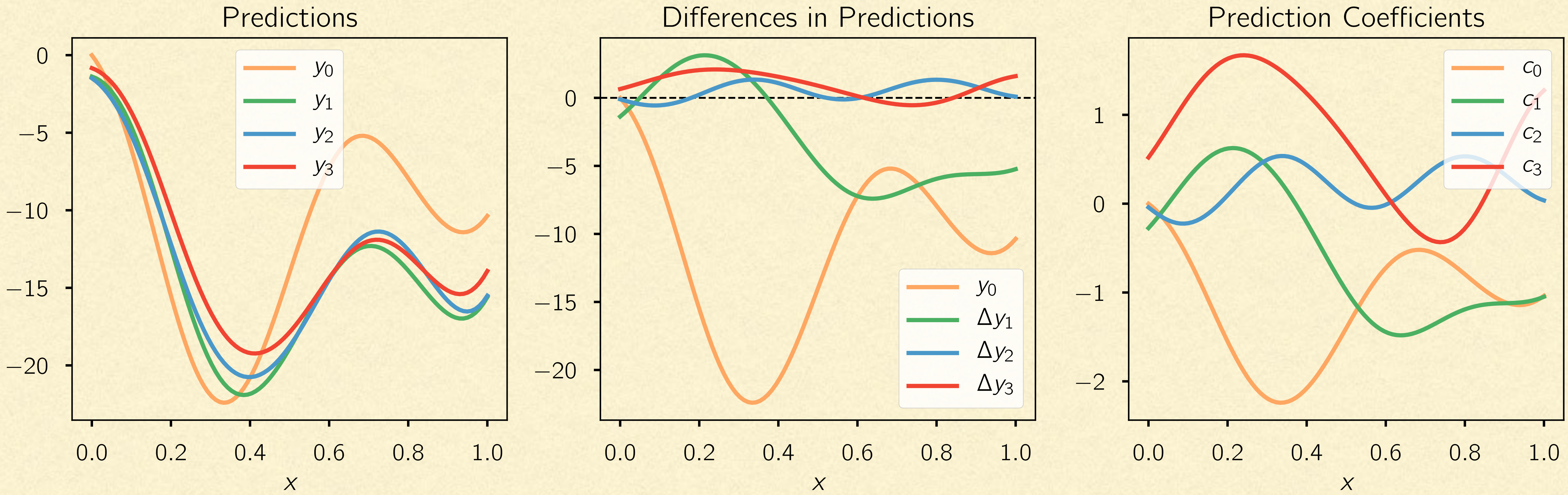
An EFT expansion in pictures

- General EFT series for observable to order k: $y = y_{\mathrm{ref}} \sum_{n=0}^{k} c_n Q^n$
- In ChiEFT $Q = \frac{(p, m_\pi)}{\Lambda_b}; \quad \Lambda_b \approx 600 \text{ MeV}$



An EFT expansion in pictures

- General EFT series for observable to order k: $y = y_{\rm ref} \sum_{n=0}^{k} c_n Q^n$
- In ChiEFT $Q = \frac{(p, m_\pi)}{\Lambda_b}; \quad \Lambda_b \approx 600 \text{ MeV}$



This is what a healthy observable expansion looks like: bounded coefficients, that do not grow or shrink with order.

Probability for EFT coefficients

Furnstahl, Klco, DP, Wesolowski, PRC, 2015 after Cacciari and Houdeau, JHEP, 2011

Probability for EFT coefficients

Furnstahl, Klco, DP, Wesolowski, PRC, 2015 after Cacciari and Houdeau, JHEP, 2011

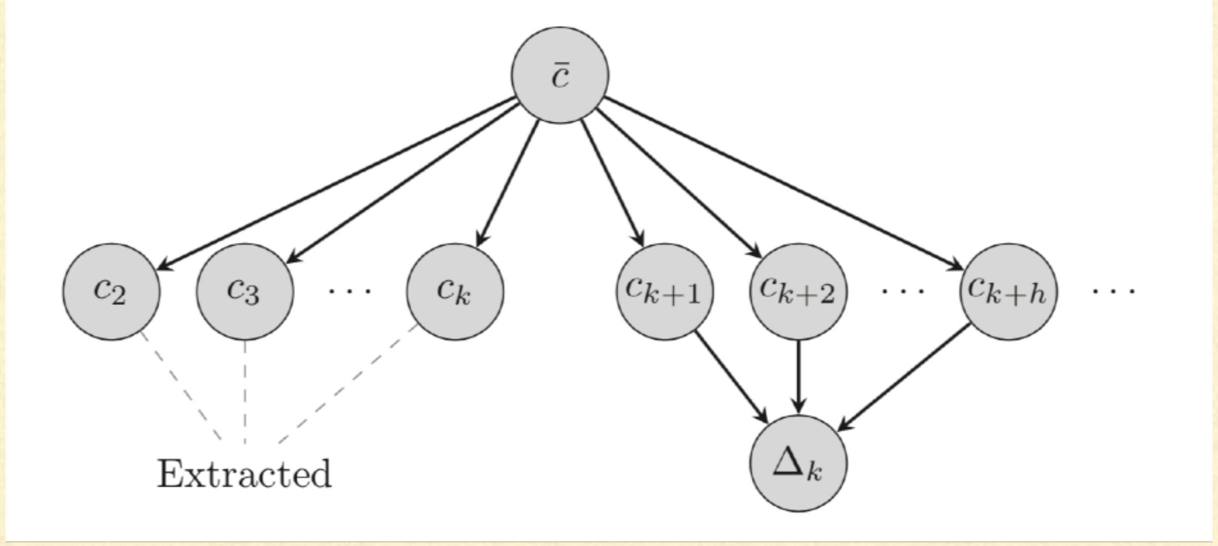
- So can we use extracted c_0, c_1, c_2, …, c_k to estimate (in a probabilistic way) c_{k+1}? From there construct $\Delta_k = y_{ref}\, c_{k+1}\, Q^{k+1}$:truncation error

Probability for EFT coefficients

Furnstahl, Klco, DP, Wesolowski, PRC, 2015 after Cacciari and Houdeau, JHEP, 2011

- So can we use extracted c_0, c_1, c_2, …, c_k to estimate (in a probabilistic way) c_{k+1}? From there construct $\Delta_k = y_{ref}\, c_{k+1}\, Q^{k+1}$:truncation error

- Bayesian model:



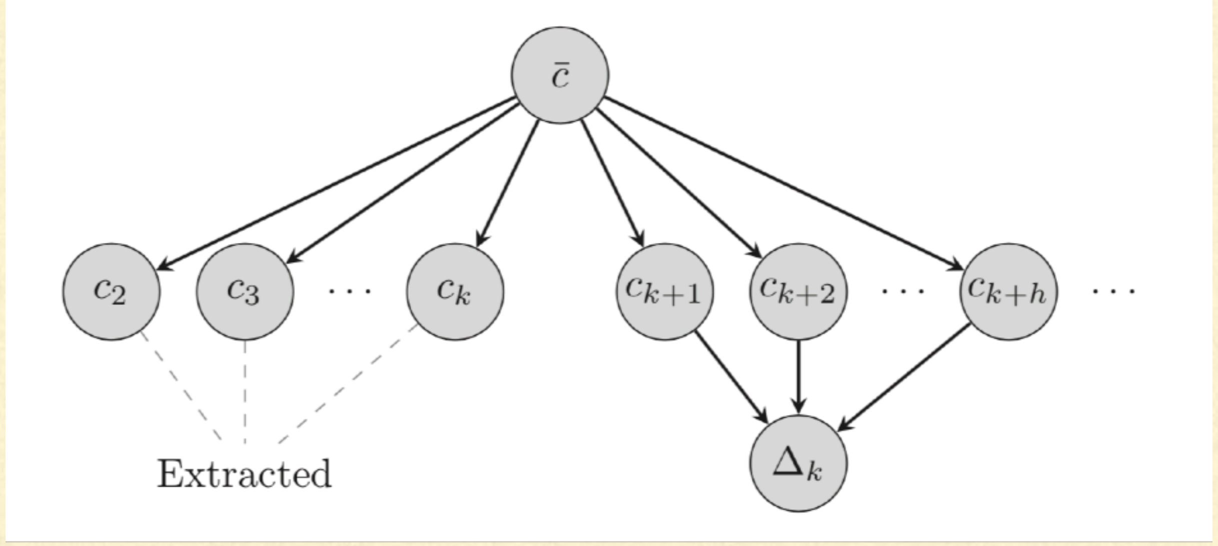
Parameter cbar sets size of all dimensionless coefficients

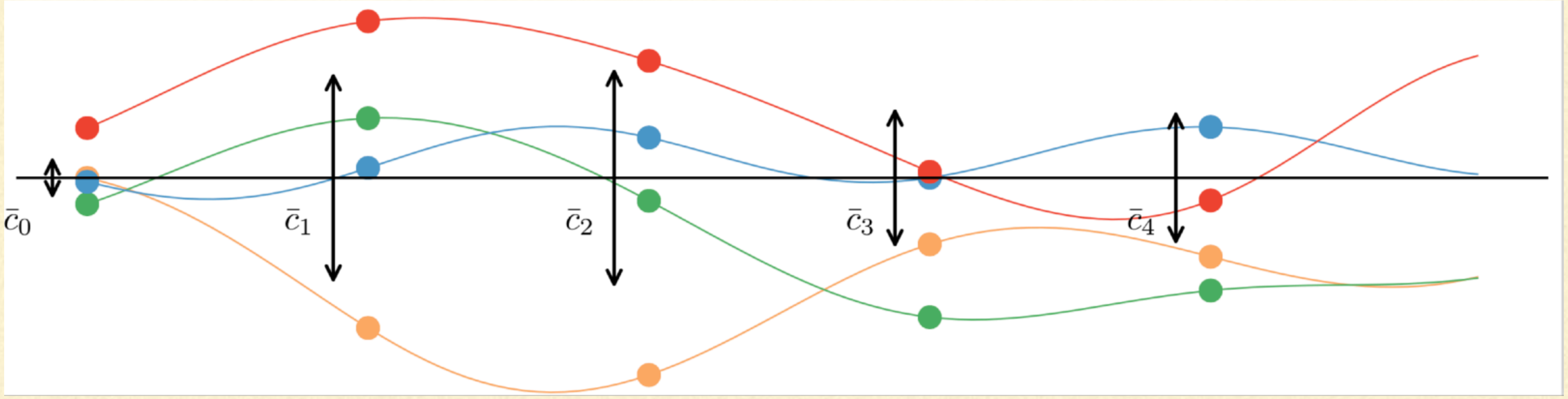
Probability for EFT coefficients

Furnstahl, Klco, DP, Wesolowski, PRC, 2015 after Cacciari and Houdeau, JHEP, 2011

- So can we use extracted $c_0, c_1, c_2, \ldots, c_k$ to estimate (in a probabilistic way) c_{k+1}? From there construct $\Delta_k = y_{ref}\, c_{k+1}\, Q^{k+1}$:truncation error
- Bayesian model:

Parameter cbar sets size of all dimensionless coefficients





First shot: cbar can be different at different kinematic points: "uncorrelated model"

Normal naturalness

Furnstahl, Klco, DP, Wesolowski, PRC, 2015; Melendez, Furnstahl, Wesolowski, PRC, 2017

- c_n's are normally distributed, with mean 0 and standard deviation cbar. that is a) fixed or b) distributed uniformly in its logarithm

$$\mathrm{pr}(c_n|\bar{c}) = \frac{1}{\sqrt{2\pi}\bar{c}} e^{-c_n^2/2\bar{c}^2};\ \mathrm{pr}(\bar{c}) \propto \frac{1}{\bar{c}}\theta(\bar{c}-\bar{c}_<)\theta(\bar{c}_> - \bar{c})$$

- Marginalization:

$$\mathrm{pr}(c_{k+1}|c_0, c_1, \ldots, c_k) = \int_0^\infty d\bar{c}\, \mathrm{pr}(c_{k+1}|\bar{c})\mathrm{pr}(\bar{c}|c_0, c_1, \ldots c_k)$$

$$= \int_0^\infty \frac{d\bar{c}}{\bar{c}^{k+3}} \exp\left(-\frac{c_{k+1}^2}{2\bar{c}^2}\right) \exp\left(-\frac{(k+1)\langle c^2\rangle}{2\bar{c}^2}\right)$$

- Student's t-distribution results:

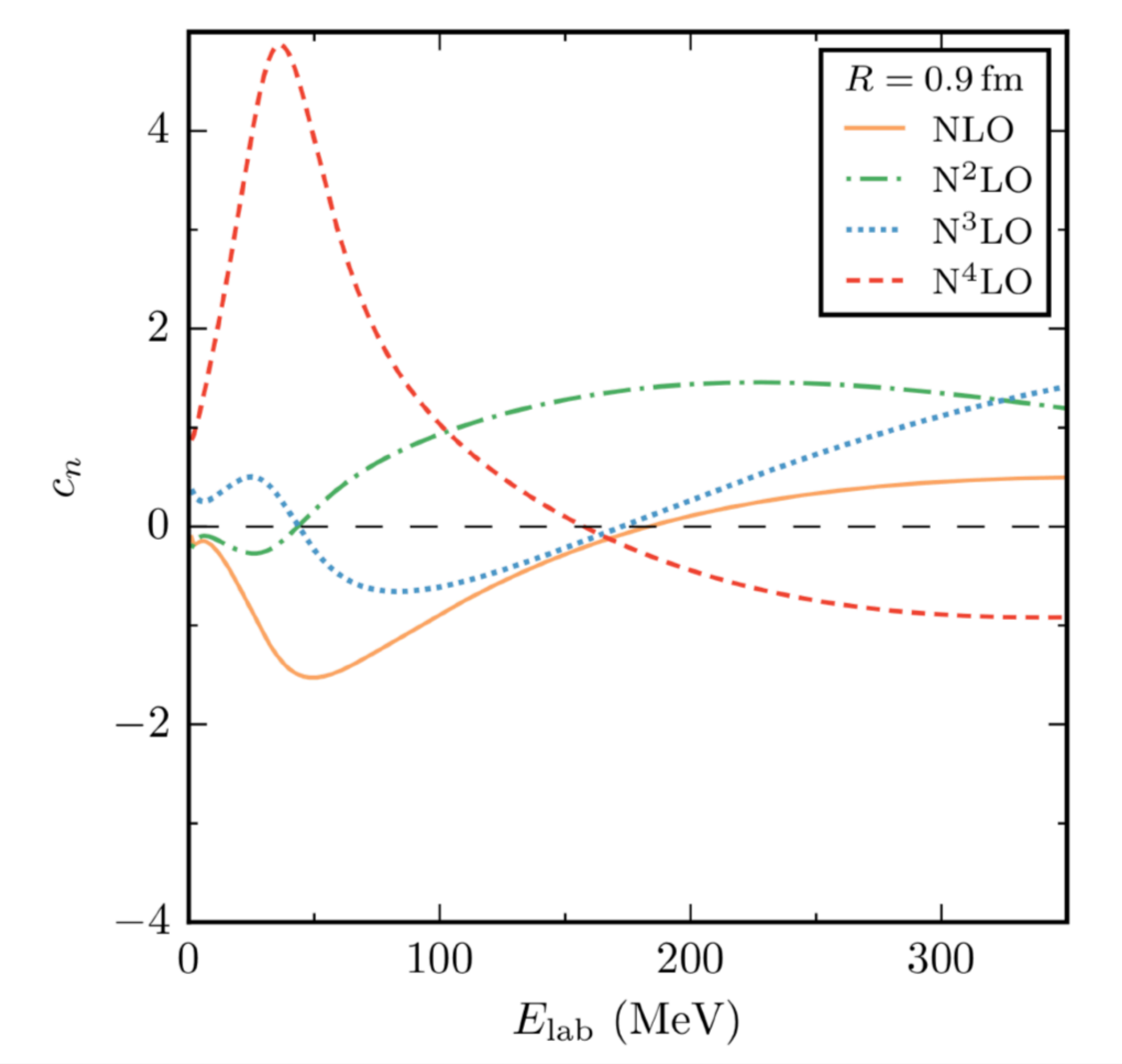
$$\mathrm{pr}(c_{k+1}|c_0, c_1, \ldots, c_k) \propto \frac{\Gamma\left(\frac{k+2}{2}\right)}{\Gamma\left(\frac{k+1}{2}\right)} \left(\frac{(k+1)\langle c^2\rangle}{(k+1)\langle c^2\rangle + c_{k+1}^2}\right)^{(k+2)/2}$$

- DoB intervals computed using known results for this distribution. Size of error bar set by $\langle c^2 \rangle$, k, Q^{k+1}, and y_{ref}.

Error bands for NN observables

Melendez, Furnstahl, Wesolowski, PRC, 2017

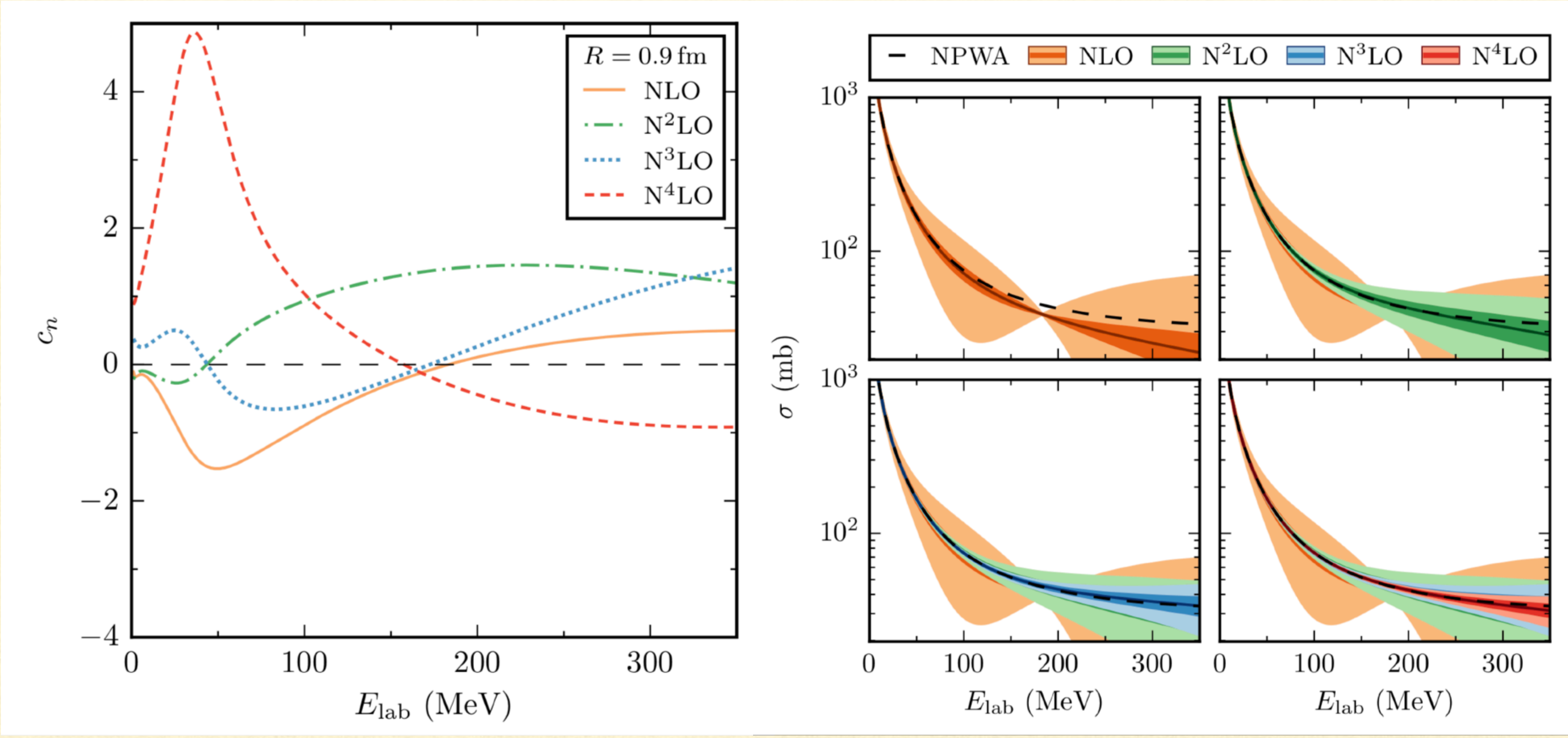
EKM R=0.9 fm potential



Error bands for NN observables

Melendez, Furnstahl, Wesolowski, PRC, 2017

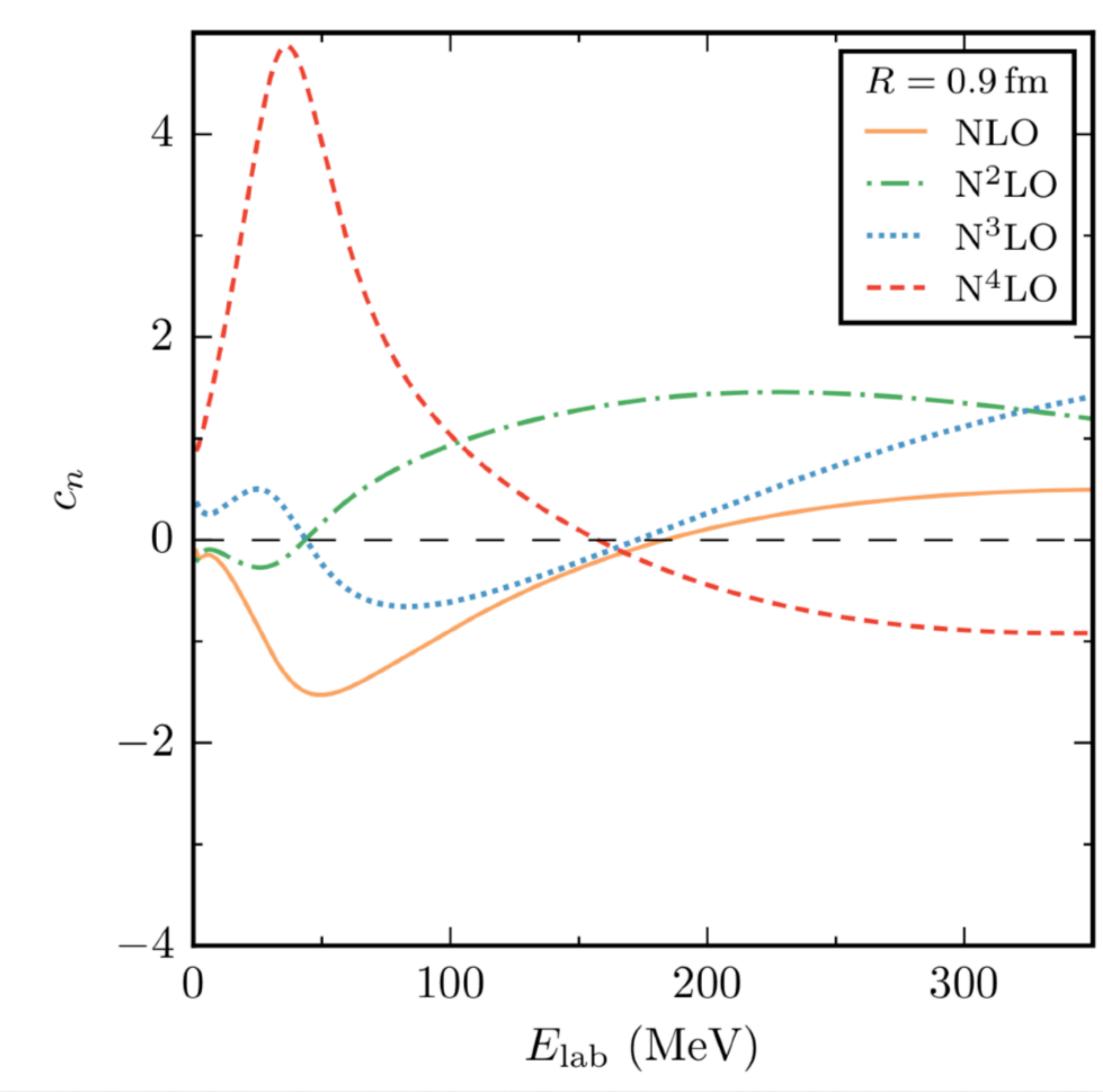
EKM R=0.9 fm potential

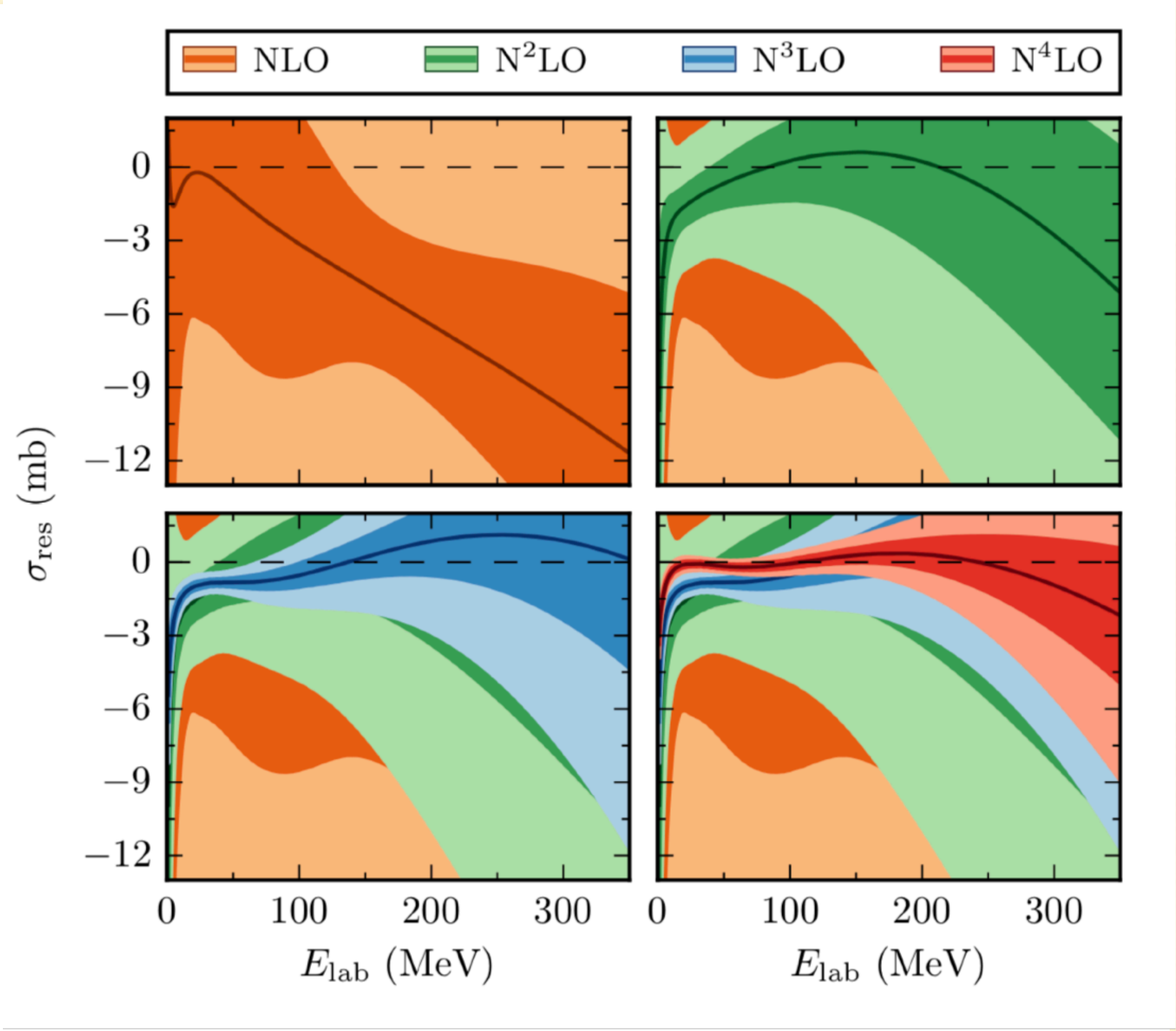


Error bands for NN observables

Melendez, Furnstahl, Wesolowski, PRC, 2017

EKM R=0.9 fm potential

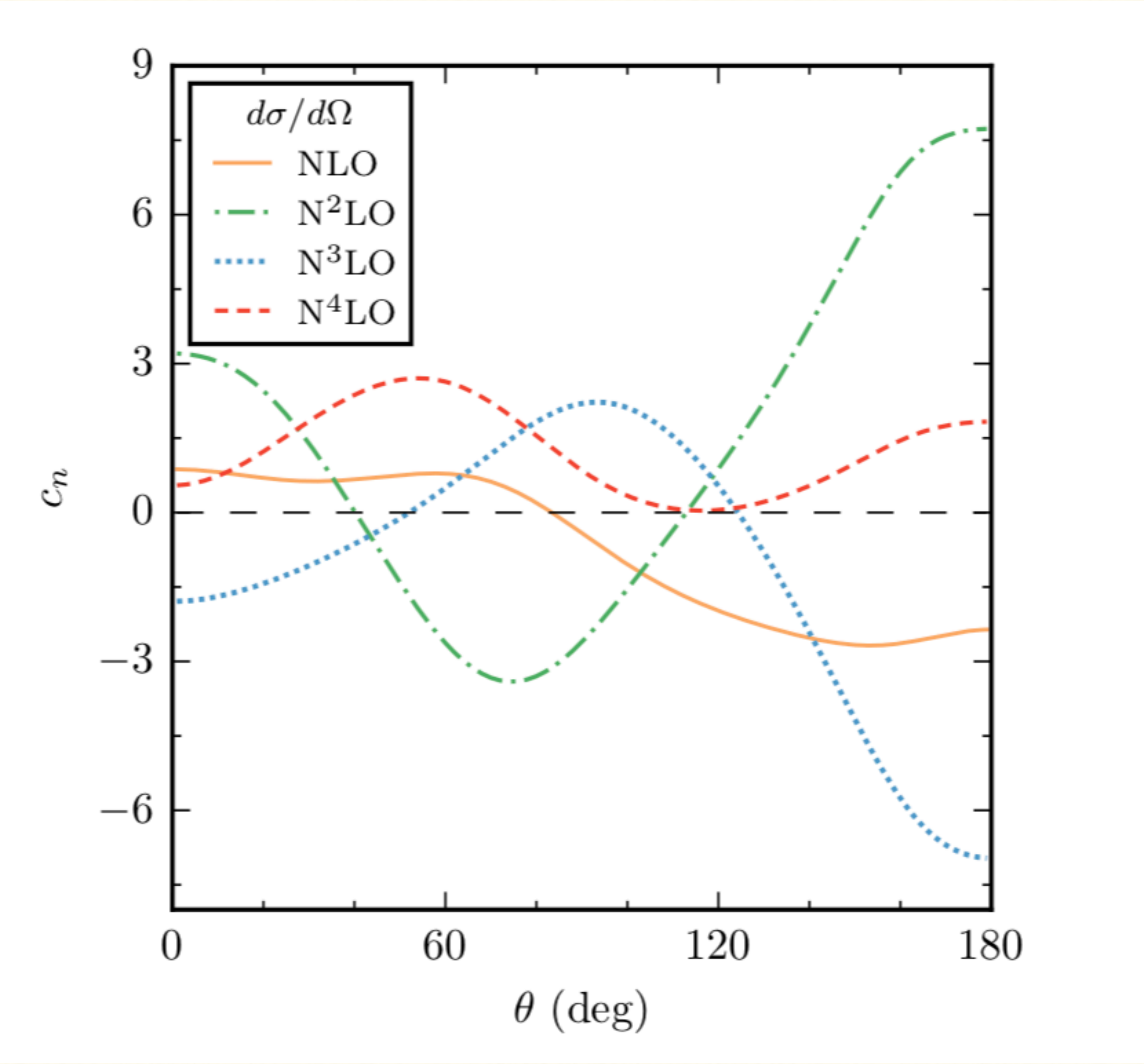




Error bands for NN observables

Melendez, Furnstahl, Wesolowski, PRC, 2017

EKM R=0.9 fm potential

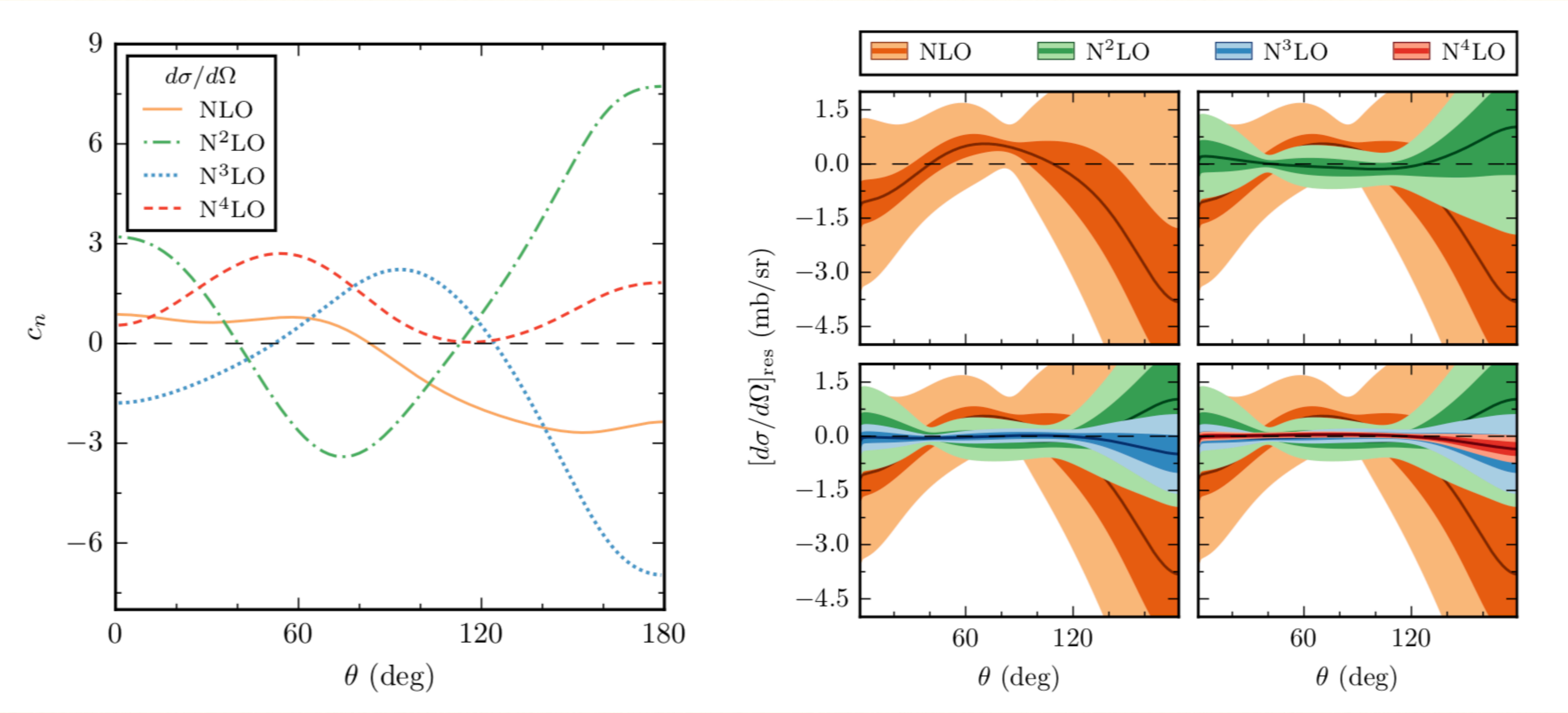


E_{lab}=96 MeV

Error bands for NN observables

Melendez, Furnstahl, Wesolowski, PRC, 2017

EKM R=0.9 fm potential

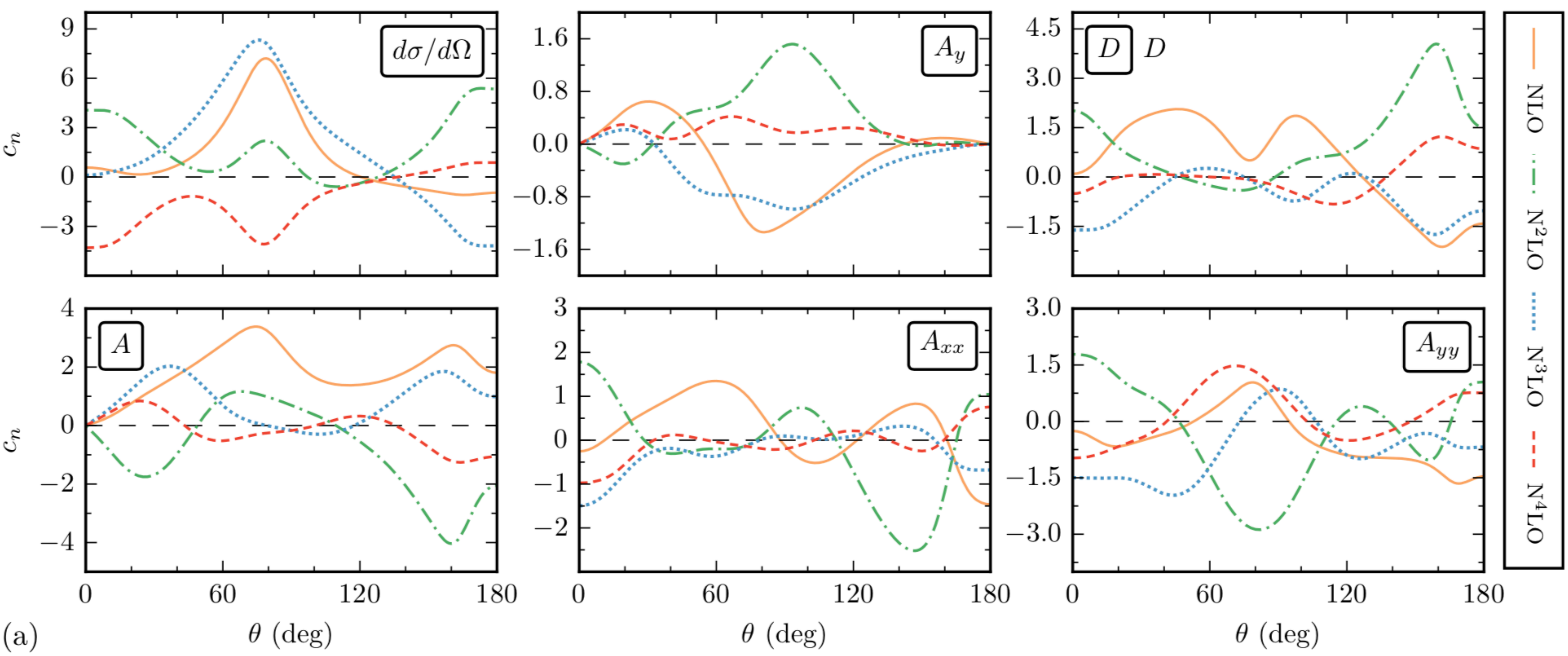


E_{lab}=96 MeV

Error bands for NN observables

Melendez, Furnstahl, Wesolowski, PRC, 2017

EKM R=0.9 fm potential

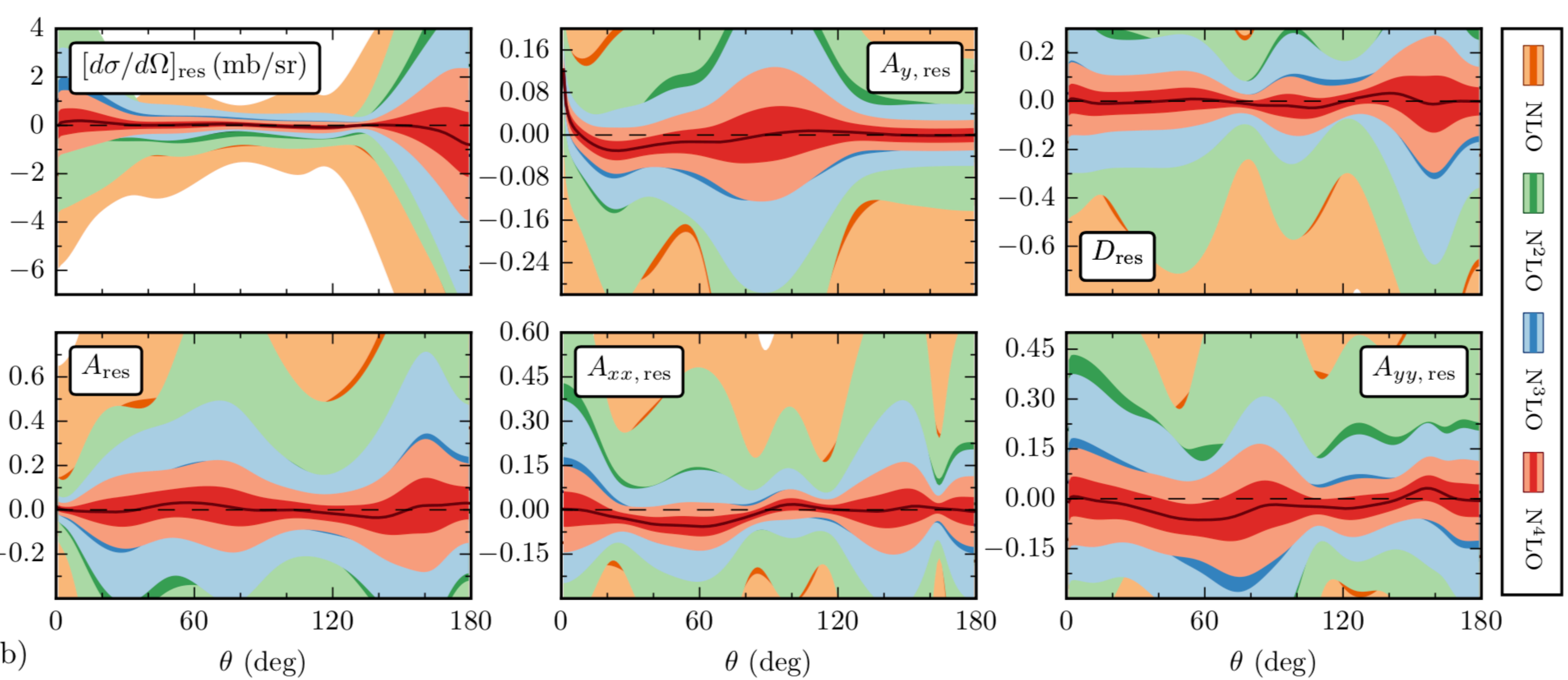


E_{lab}=250 MeV

Error bands for NN observables

Melendez, Furnstahl, Wesolowski, PRC, 2017

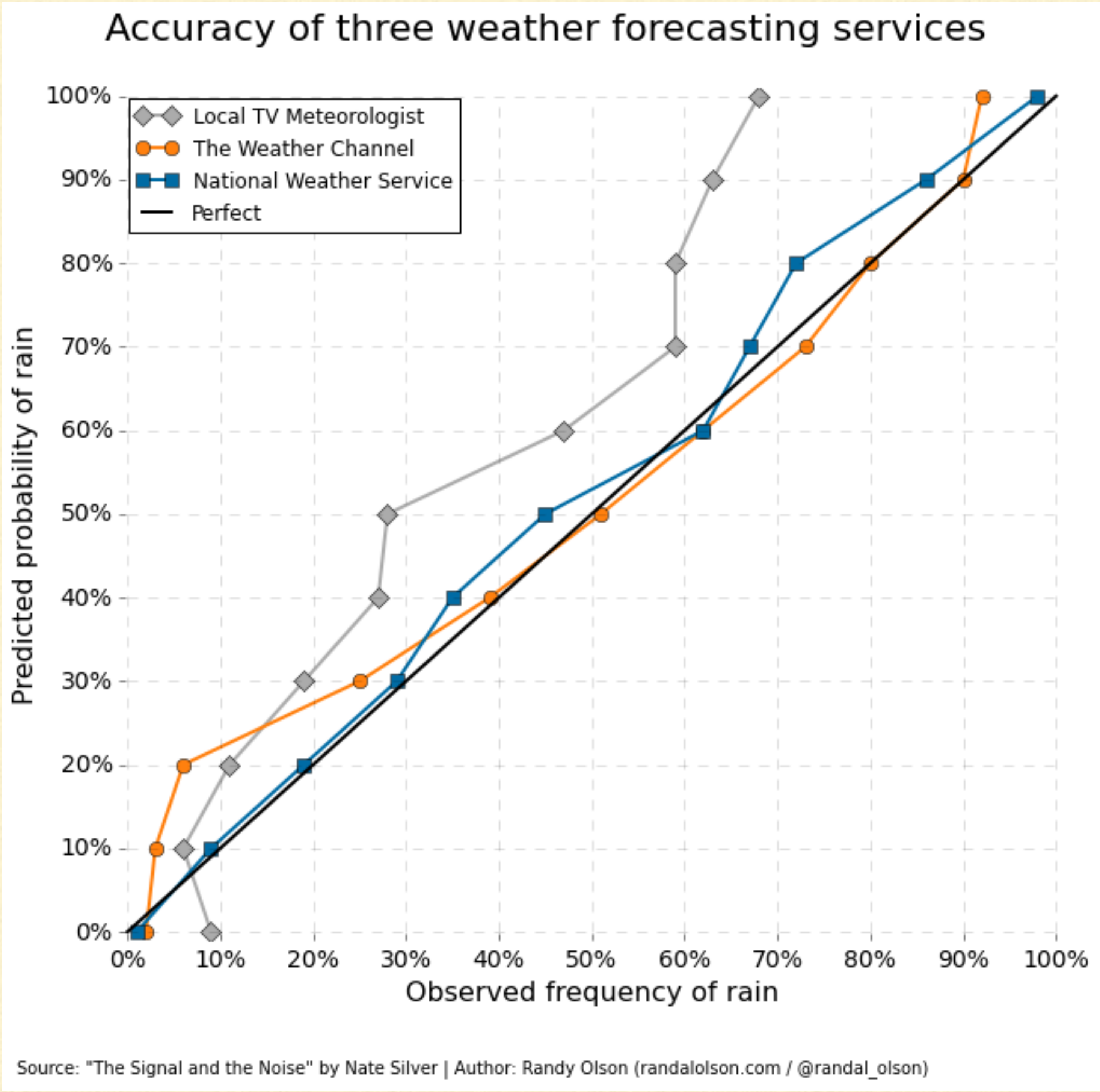
EKM R=0.9 fm potential



$E_{\rm lab}$=250 MeV

The well-calibrated EFTer

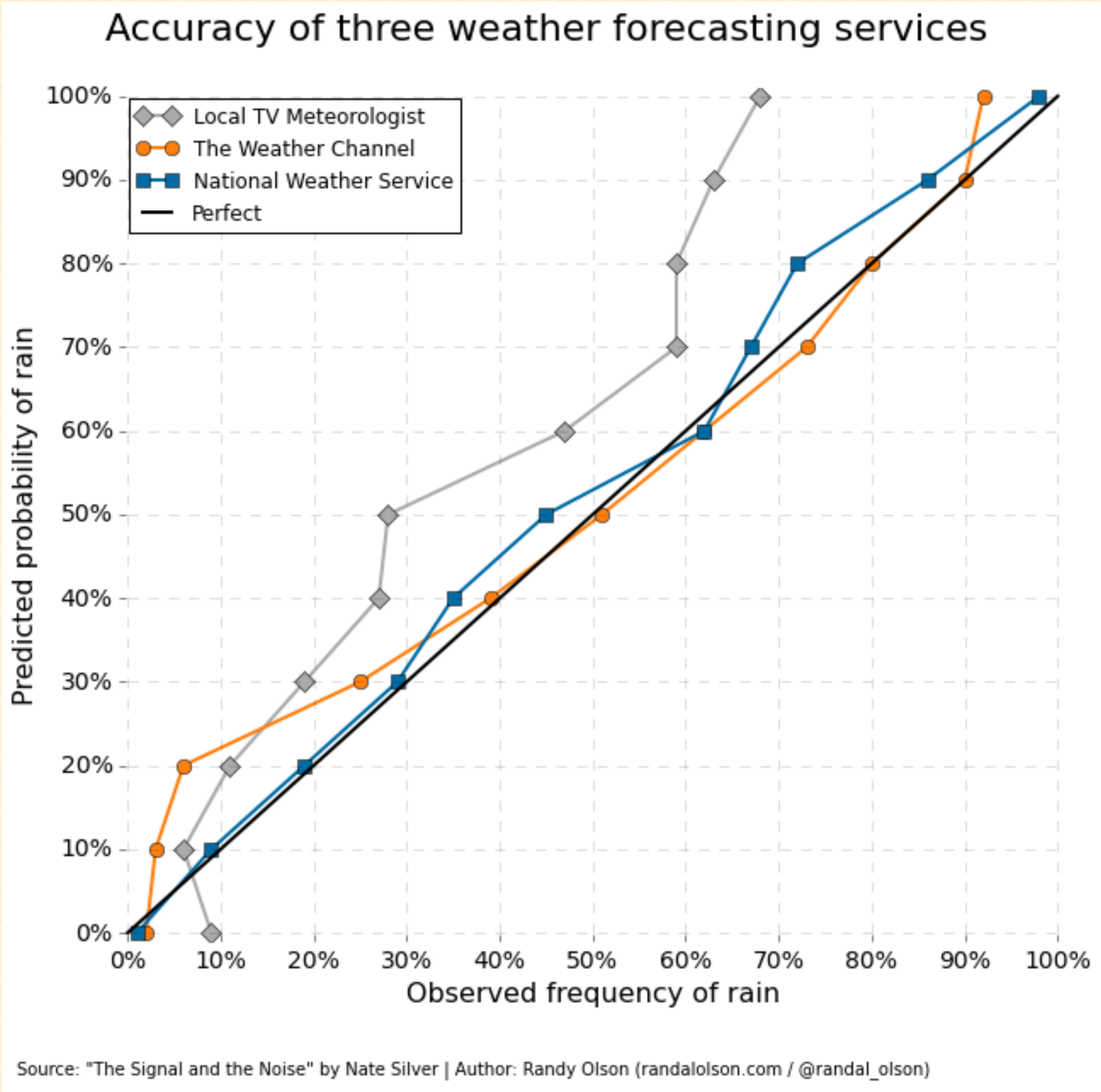
The well-calibrated EFTer



Melendez, Furnstahl, Wesolowski, PRC, 2017
after Furnstahl, Klco, DP, Wesolowski, PRC, 2015
after: Bagnaschi, Cacciari, Guffanti, Jenniches, 2015

- Consider predictions at each order, with their error bars, as data and test them to see if the procedure is consistent
- Fix a given DOB interval: compute success ratio, compare

The well-calibrated EFTer

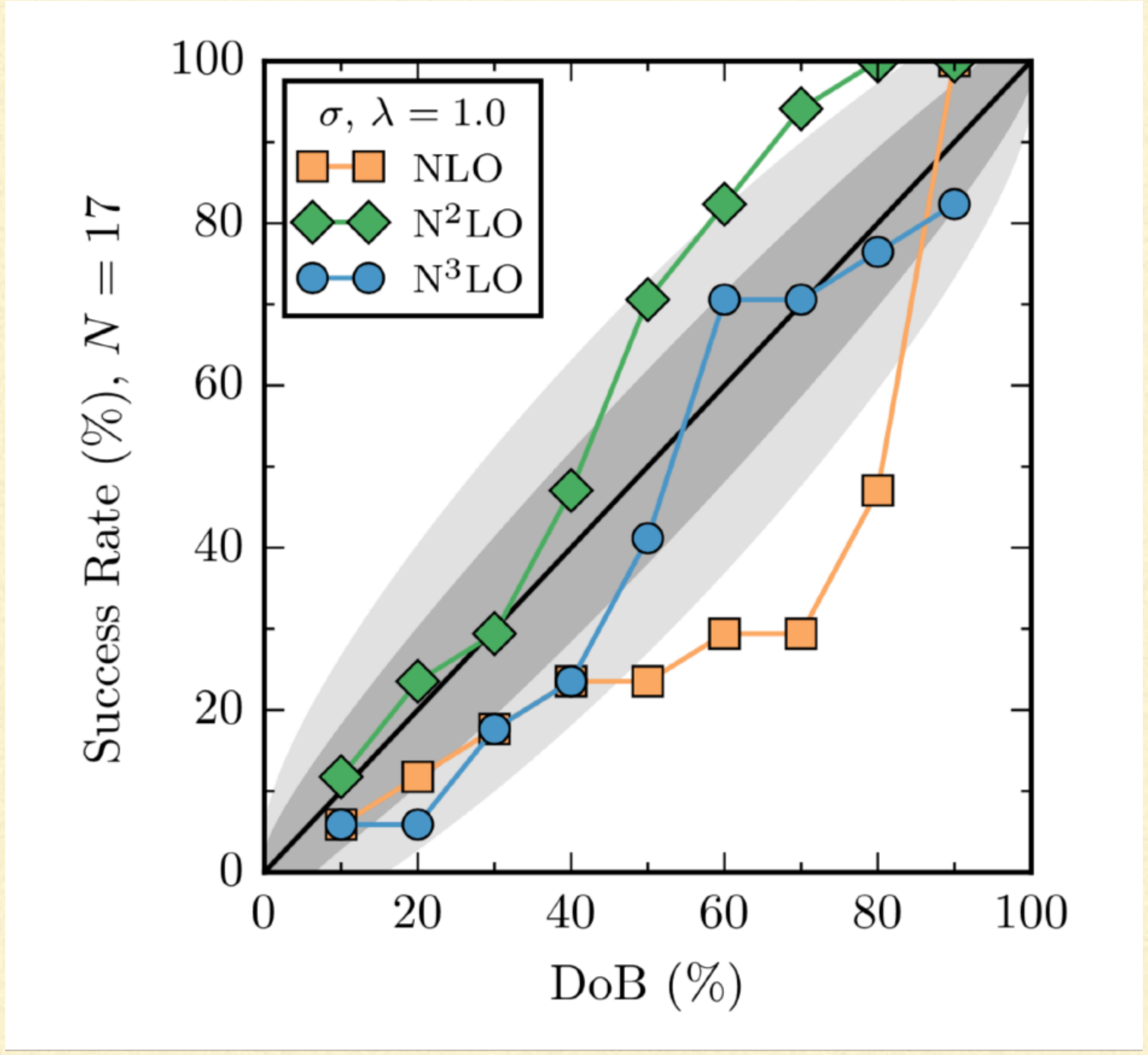


Melendez, Furnstahl, Wesolowski, PRC, 2017
after Furnstahl, Klco, DP, Wesolowski, PRC, 2015
after: Bagnaschi, Cacciari, Guffanti, Jenniches, 2015

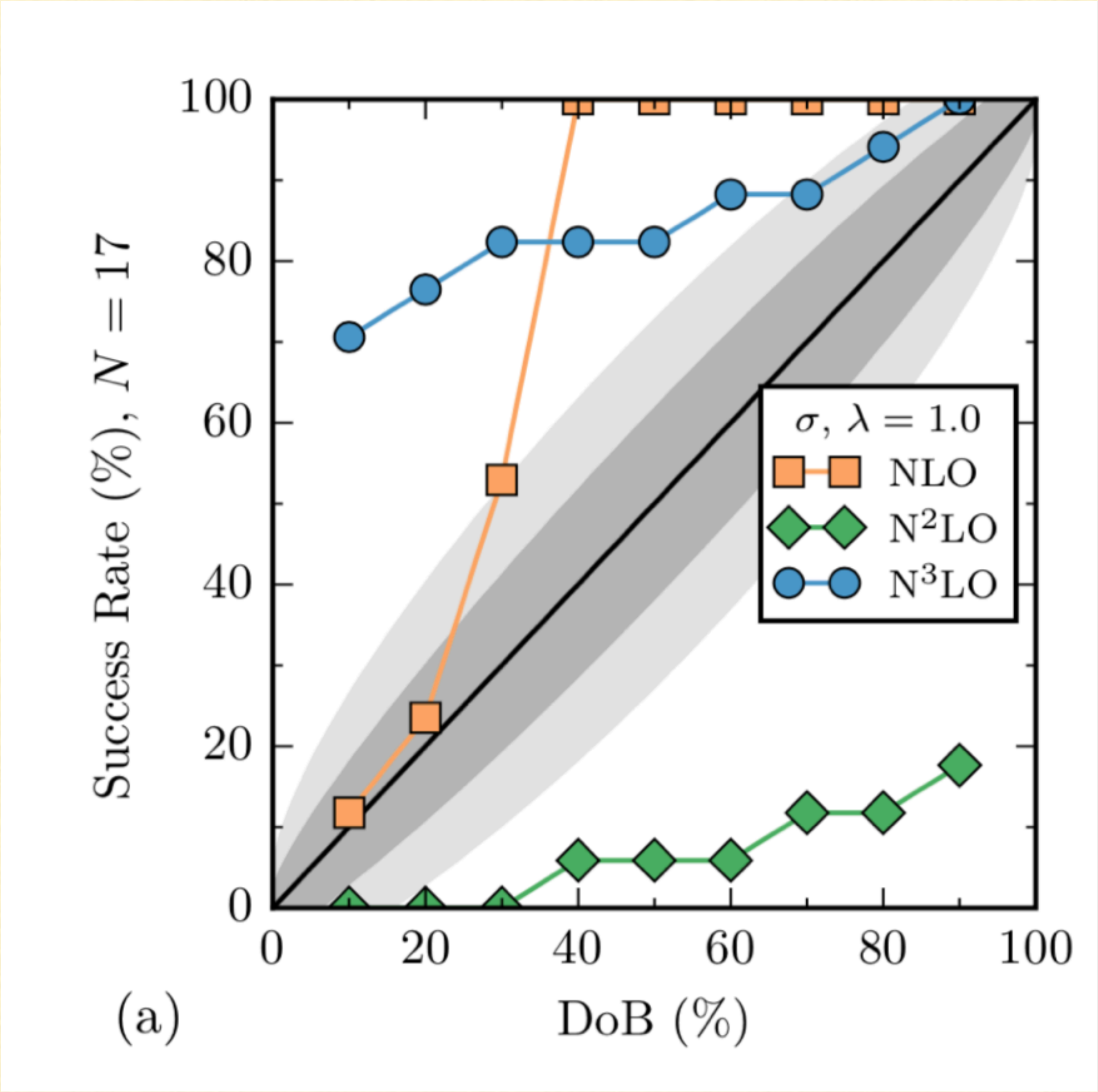
- Consider predictions at each order, with their error bars, as data and test them to see if the procedure is consistent
- Fix a given DOB interval: compute success ratio, compare
- Look at this for EKM predictions
- Should be near diagonal, with fluctuations described by Poisson statistics

Physics from calibration plots

R=0.9 fm



R=1.2 fm



- Allows assessment of order-by-order convergence
- Can look at differential cross section and spin observables too

Breakdown-scale Inference

- Λ_b determines the size of the c_n's. Choose it too big, and they'll be too big. Choose it too small, they'll be too small. And progressively so as one moves to higher and higher order.

- We have a theory for $pr(c_n|c_0, c_1, \ldots, c_k)$: now use Bayes' theorem to see how (im)probable are the c_n's that dimensionful EFT coefficients (b_n's) produce for a given Λ_b.

Breakdown-scale Inference

- Λ_b determines the size of the c_n's. Choose it too big, and they'll be too big. Choose it too small, they'll be too small. And progressively so as one moves to higher and higher order.

- We have a theory for pr($c_n|c_0, c_1, \ldots, c_k$): now use Bayes' theorem to see how (im)probable are the c_n's that dimensionful EFT coefficients (b_n's) produce for a given Λ_b.

At one energy:

$$\mathrm{pr}(\Lambda_b|b_2,\ldots,b_k) \propto \frac{1}{\Lambda_b}\left(\frac{\Lambda_b^{k+2}}{(k+1)\langle b^2\rangle}\right)^{\frac{k-1}{2}}$$

(NLO: k=2, NNLO: k=3, N^3LO: k=4, etc.)

Breakdown-scale Inference

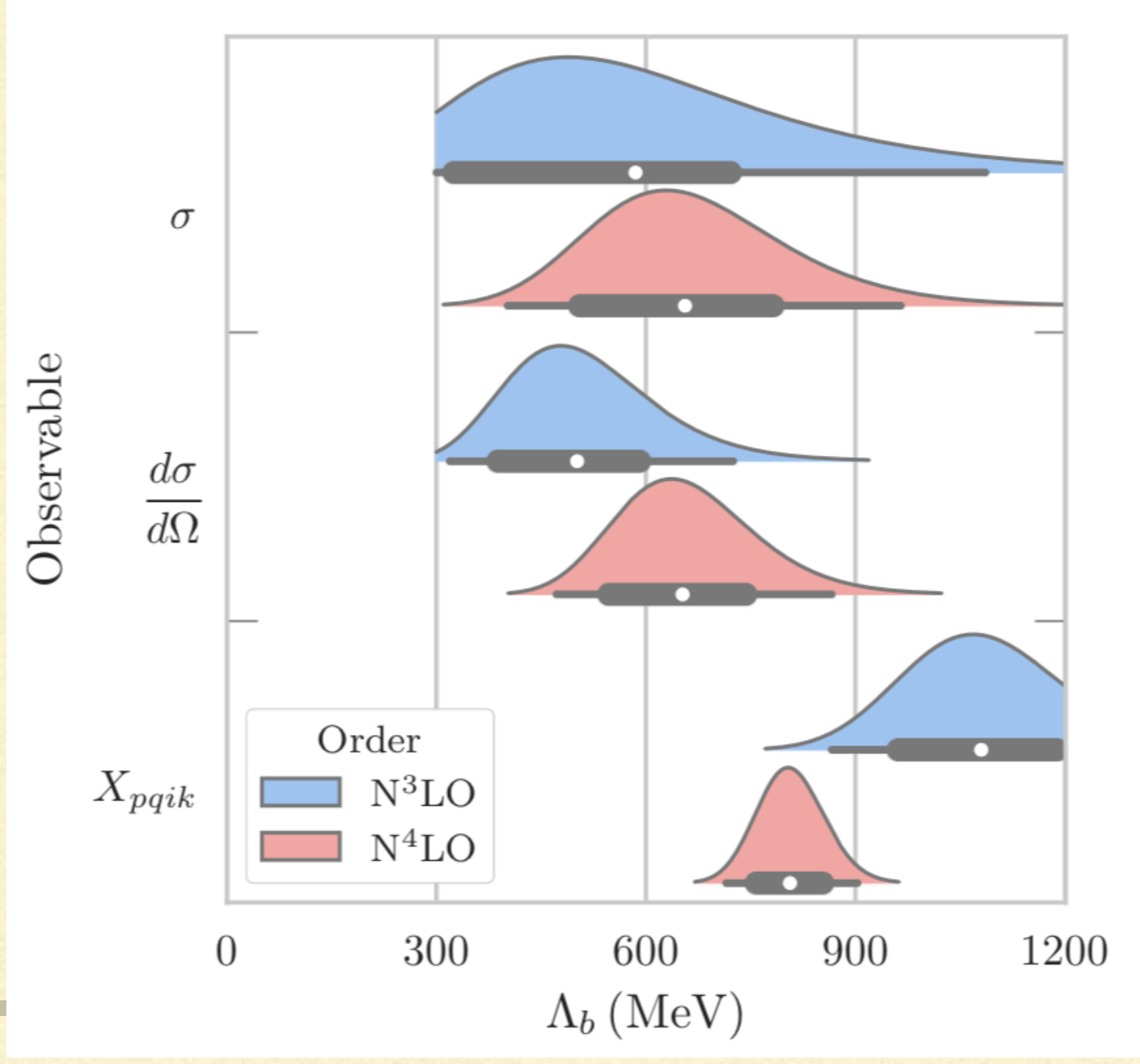
- Λ_b determines the size of the c_n's. Choose it too big, and they'll be too big. Choose it too small, they'll be too small. And progressively so as one moves to higher and higher order.
- We have a theory for pr($c_n|c_0, c_1, \ldots, c_k$): now use Bayes' theorem to see how (im)probable are the c_n's that dimensionful EFT coefficients (b_n's) produce for a given Λ_b.

At one energy:

$$\mathrm{pr}(\Lambda_b|b_2,\ldots,b_k) \propto \frac{1}{\Lambda_b}\left(\frac{\Lambda_b^{k+2}}{(k+1)\langle b^2\rangle}\right)^{\frac{k-1}{2}}$$

(NLO: k=2, NNLO: k=3, N^3LO: k=4, etc.)

Using 5 energies (and 2 angles):



Breakdown-scale Inference

- Λ_b determines the size of the c_n's. Choose it too big, and they'll be too big. Choose it too small, they'll be too small. And progressively so as one moves to higher and higher order.

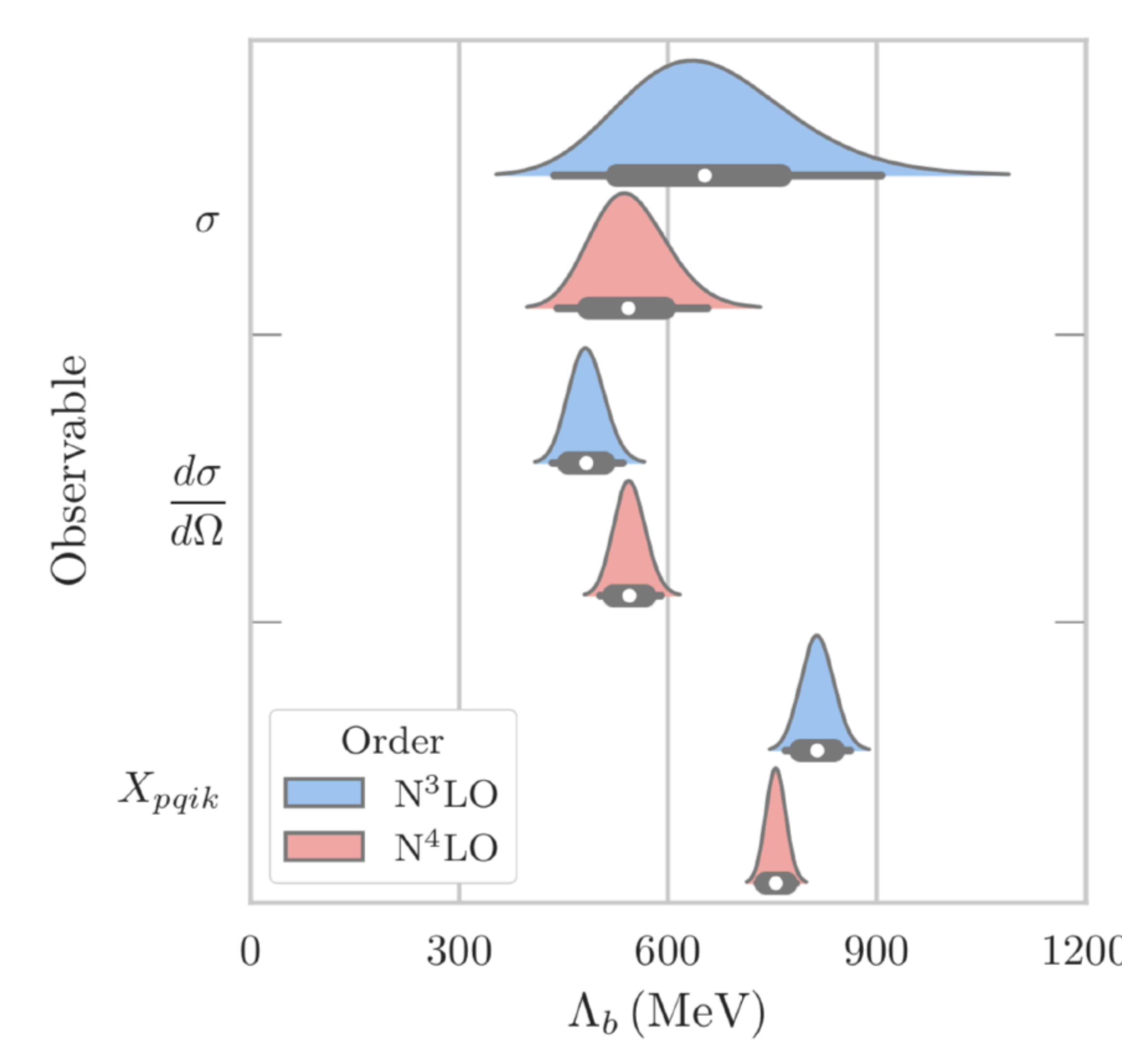
- We have a theory for pr($c_n|c_0, c_1, \ldots, c_k$): now use Bayes' theorem to see how (im)probable are the c_n's that dimensionful EFT coefficients (b_n's) produce for a given Λ_b.

At one energy:

$$\mathrm{pr}(\Lambda_b|b_2,\ldots,b_k) \propto \frac{1}{\Lambda_b}\left(\frac{\Lambda_b^{k+2}}{(k+1)\langle b^2\rangle}\right)^{\frac{k-1}{2}}$$

(NLO: k=2, NNLO: k=3, N^3LO: k=4, etc.)

Using 17 energies (and 7 angles):



Breakdown-scale Inference

- Λ_b determines the size of the c_n's. Choose it too big, and they'll be too big. Choose it too small, they'll be too small. And progressively so as one moves to higher and higher order.

- We have a theory for $\mathrm{pr}(c_n|c_0, c_1, \ldots, c_k)$: now use Bayes' theorem to see how (im)probable are the c_n's that dimensionful EFT coefficients (b_n's) produce for a given Λ_b.

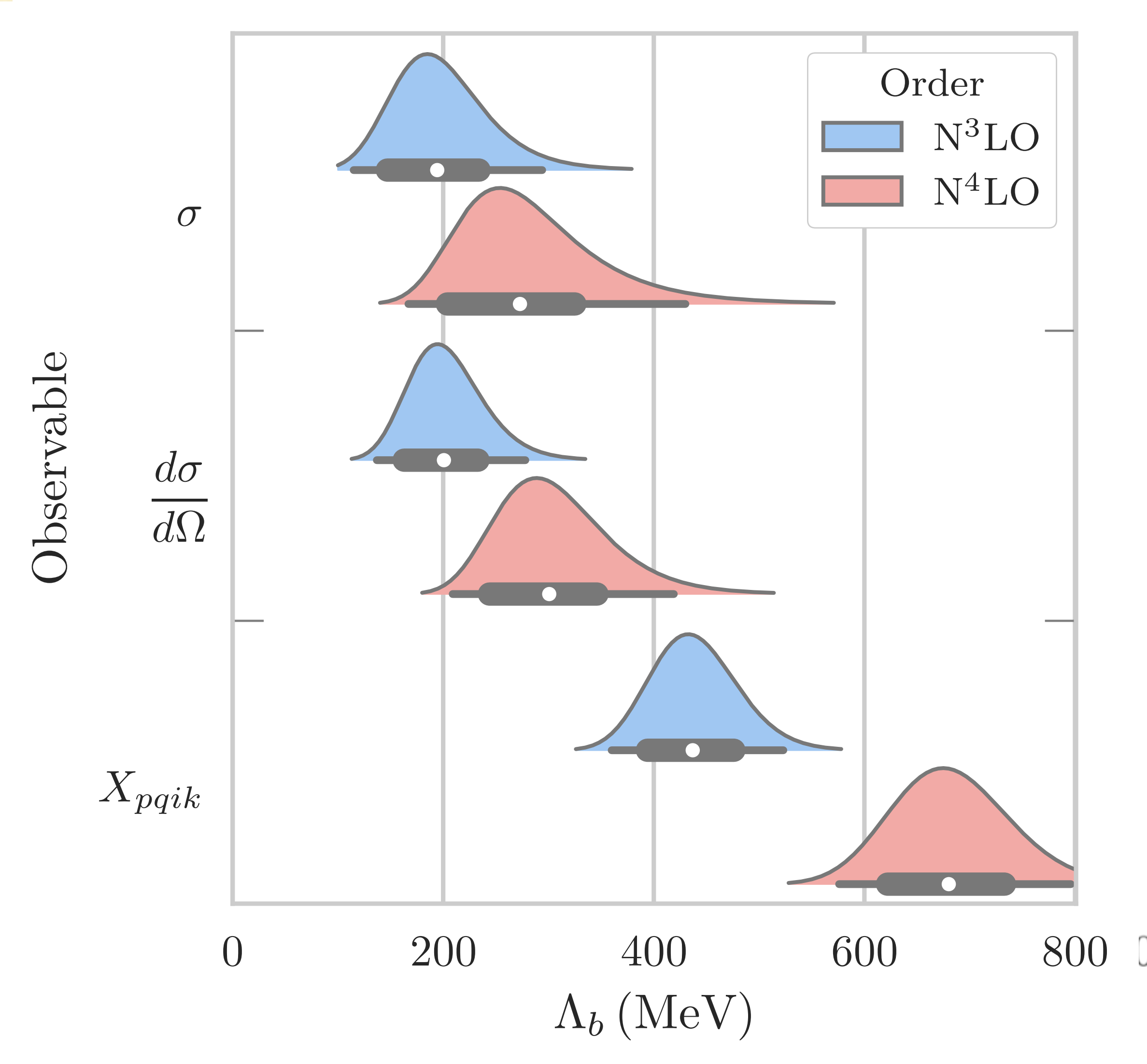
At one energy:

$$\mathrm{pr}(\Lambda_b|b_2,\ldots,b_k) \propto \frac{1}{\Lambda_b}\left(\frac{\Lambda_b^{k+2}}{(k+1)\langle b^2\rangle}\right)^{\frac{k-1}{2}}$$

(NLO: k=2, NNLO: k=3, N^3LO: k=4, etc.)

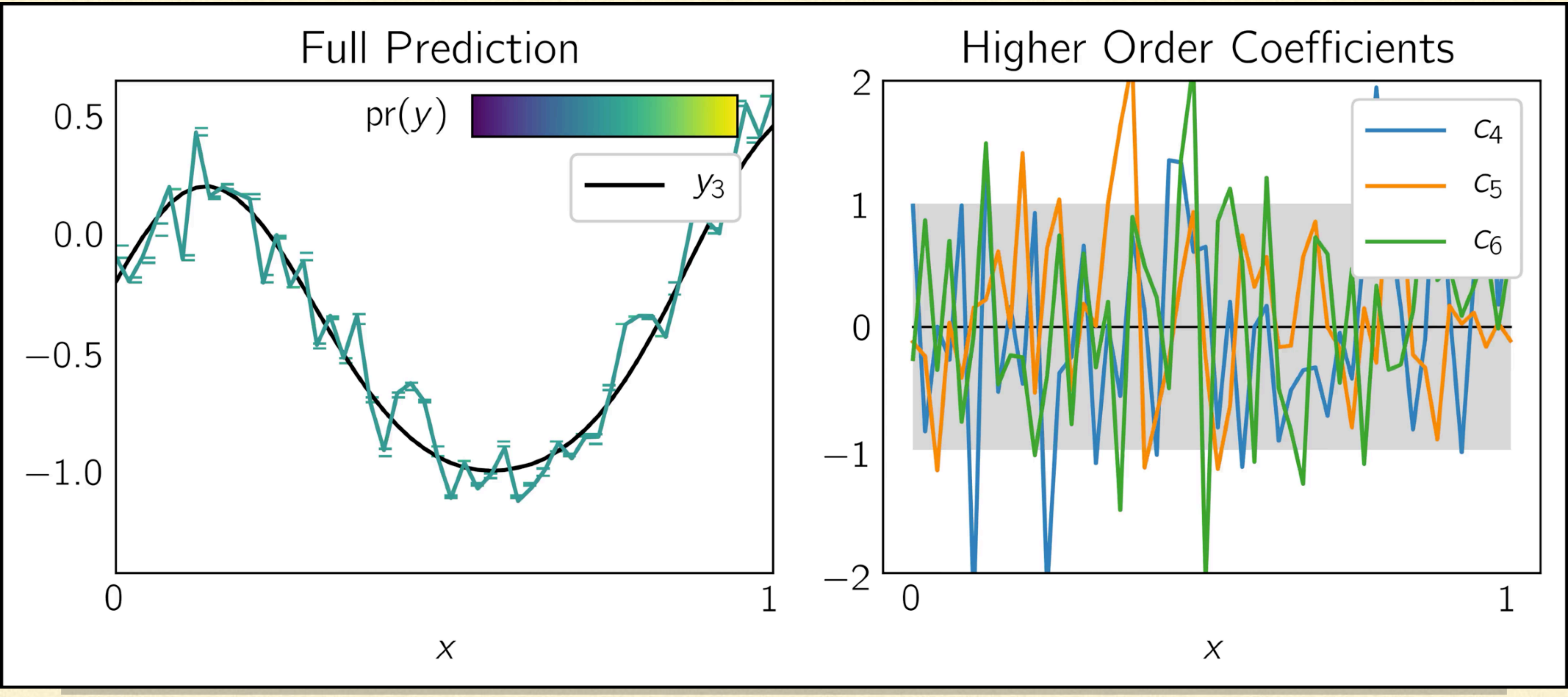
Using 17 energies (and 7 angles):

R=1.2 fm



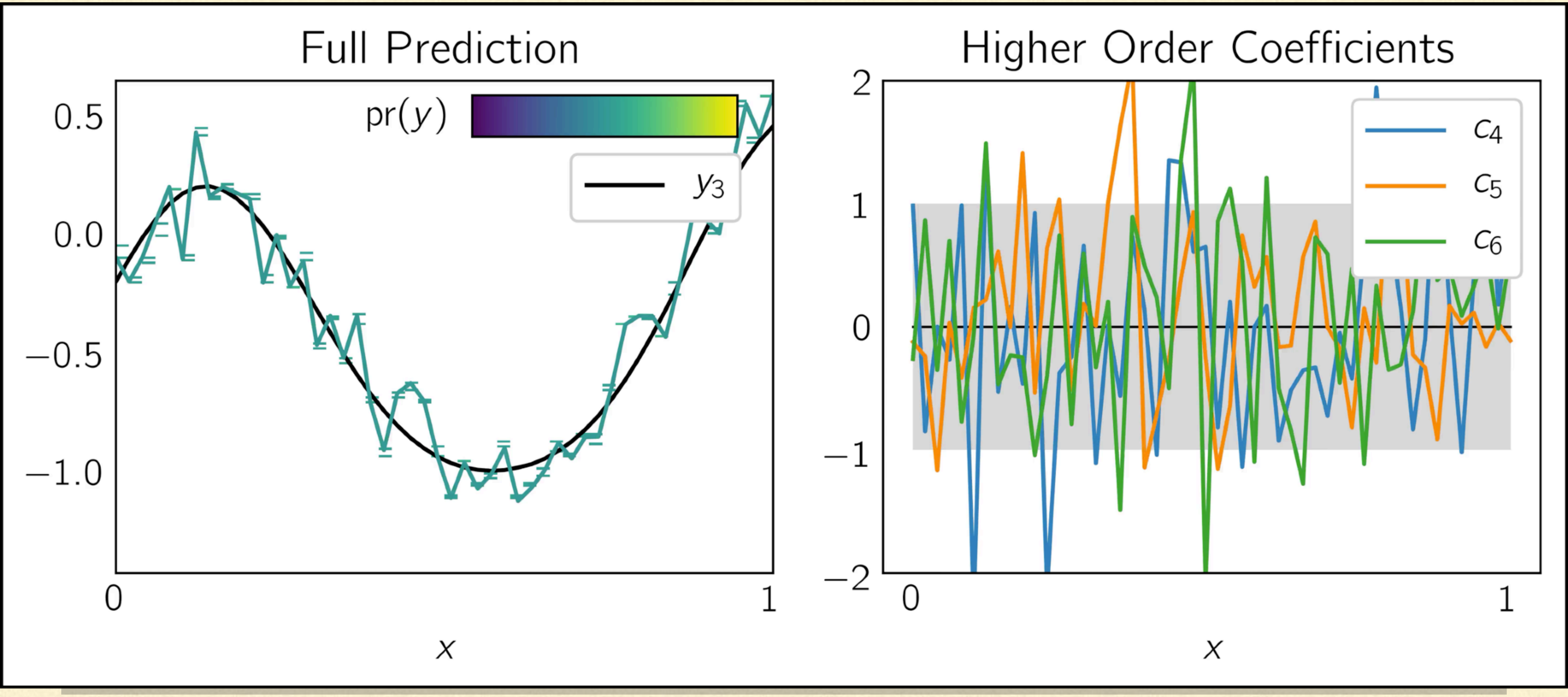
Why GPs?

Assumptions about correlation structure have significant impact on parameter estimation, NN model assessments, physics extraction, ….



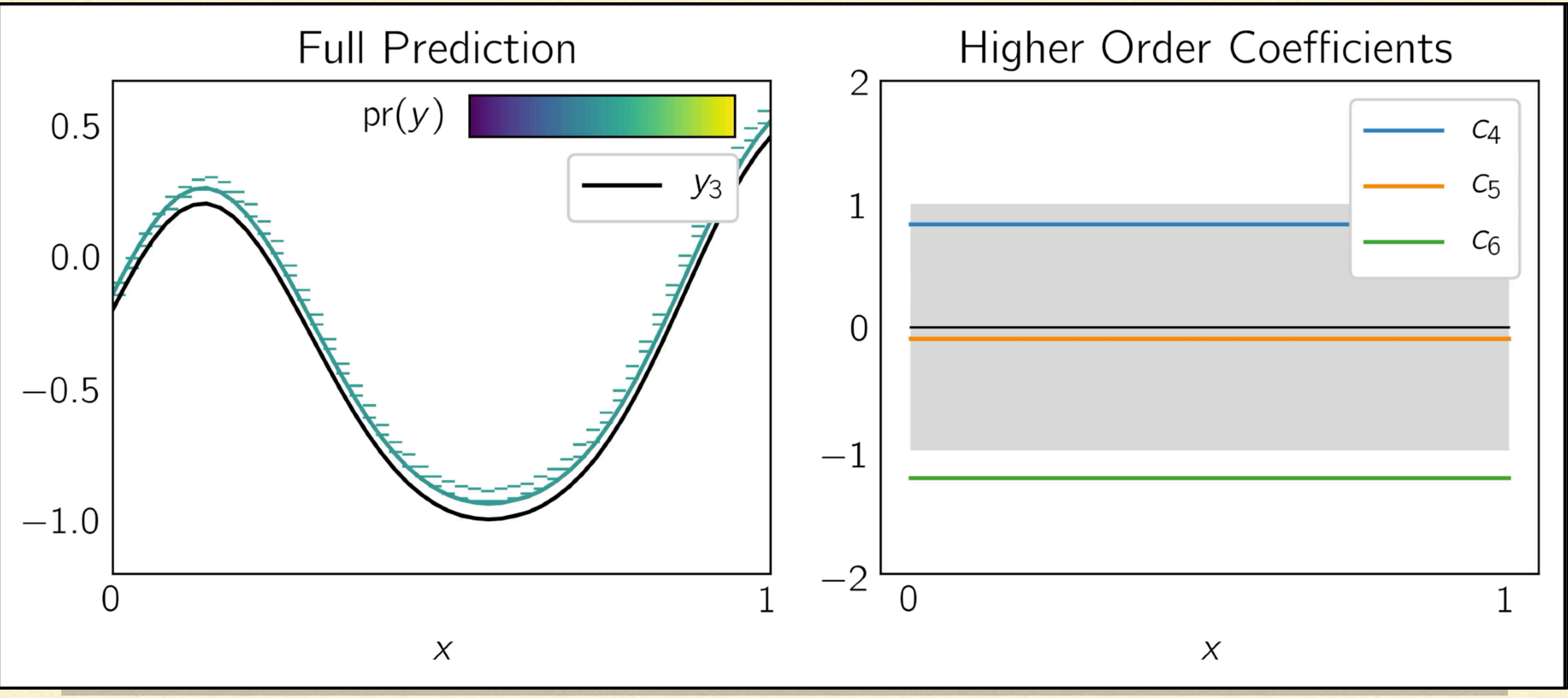
Why GPs?

Assumptions about correlation structure have significant impact on parameter estimation, NN model assessments, physics extraction, ….



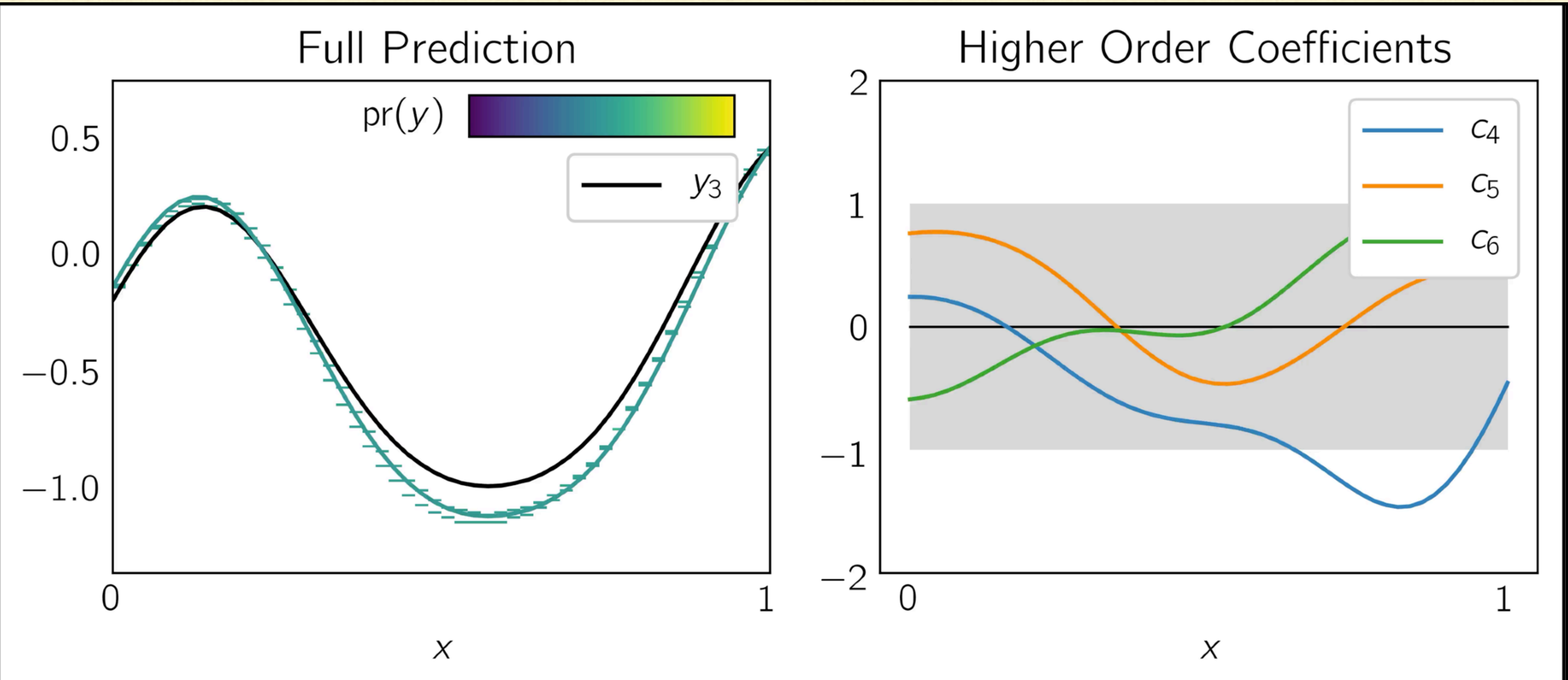
Why GPs?

Assumptions about correlation structure have significant impact on parameter estimation, NN model assessments, physics extraction, ….

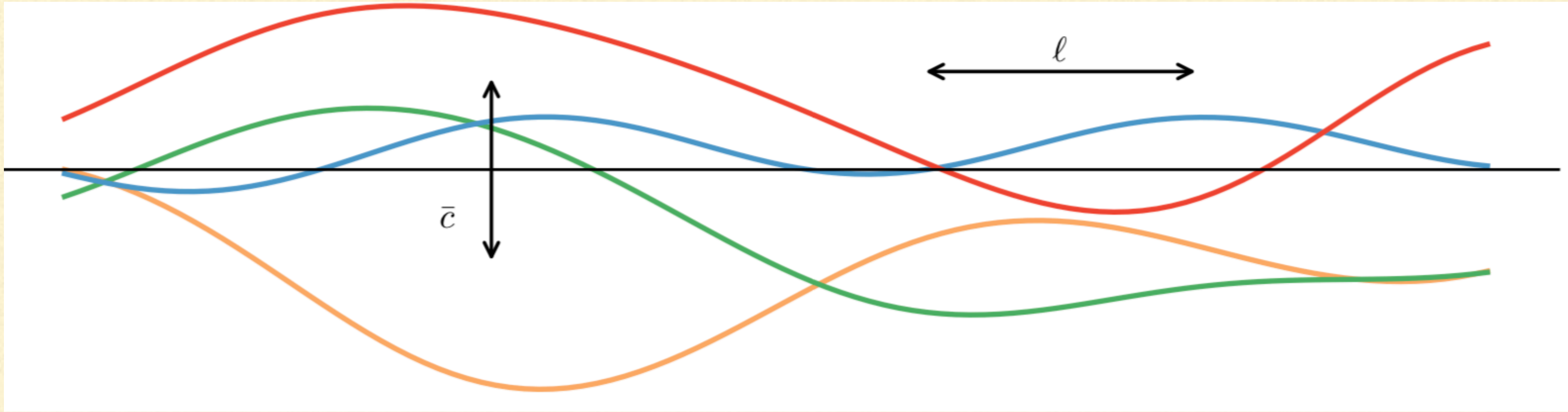


Why GPs?

Assumptions about correlation structure have significant impact on parameter estimation, NN model assessments, physics extraction, ….



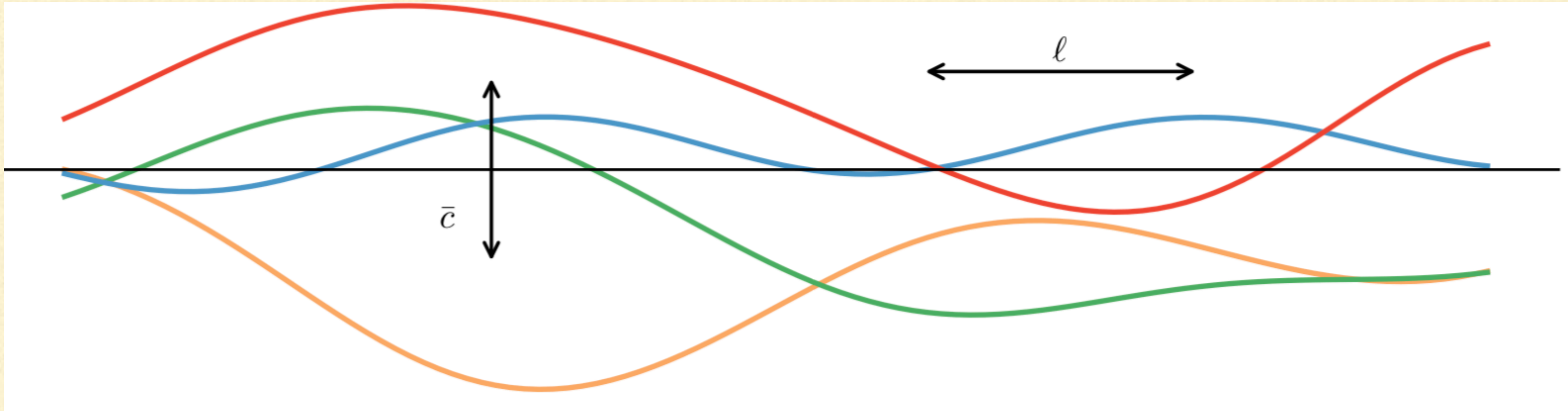
Hypothesis: the c_n's are a GP



GP describes "functional data" as

- Gaussian distribution at each point.
- With specified correlation structure.

Hypothesis: the c_n's are a GP



GP describes "functional data" as

- Gaussian distribution at each point.
- With specified correlation structure.

Our hypothesis:

EFT coefficients at different orders can be modeled as:

- independent but identical realizations of one Gaussian Process;
- with a correlation structure; here we use a "squared exponential" (Gaussian) kernel, but we test it

A bit more on Gaussian Processes

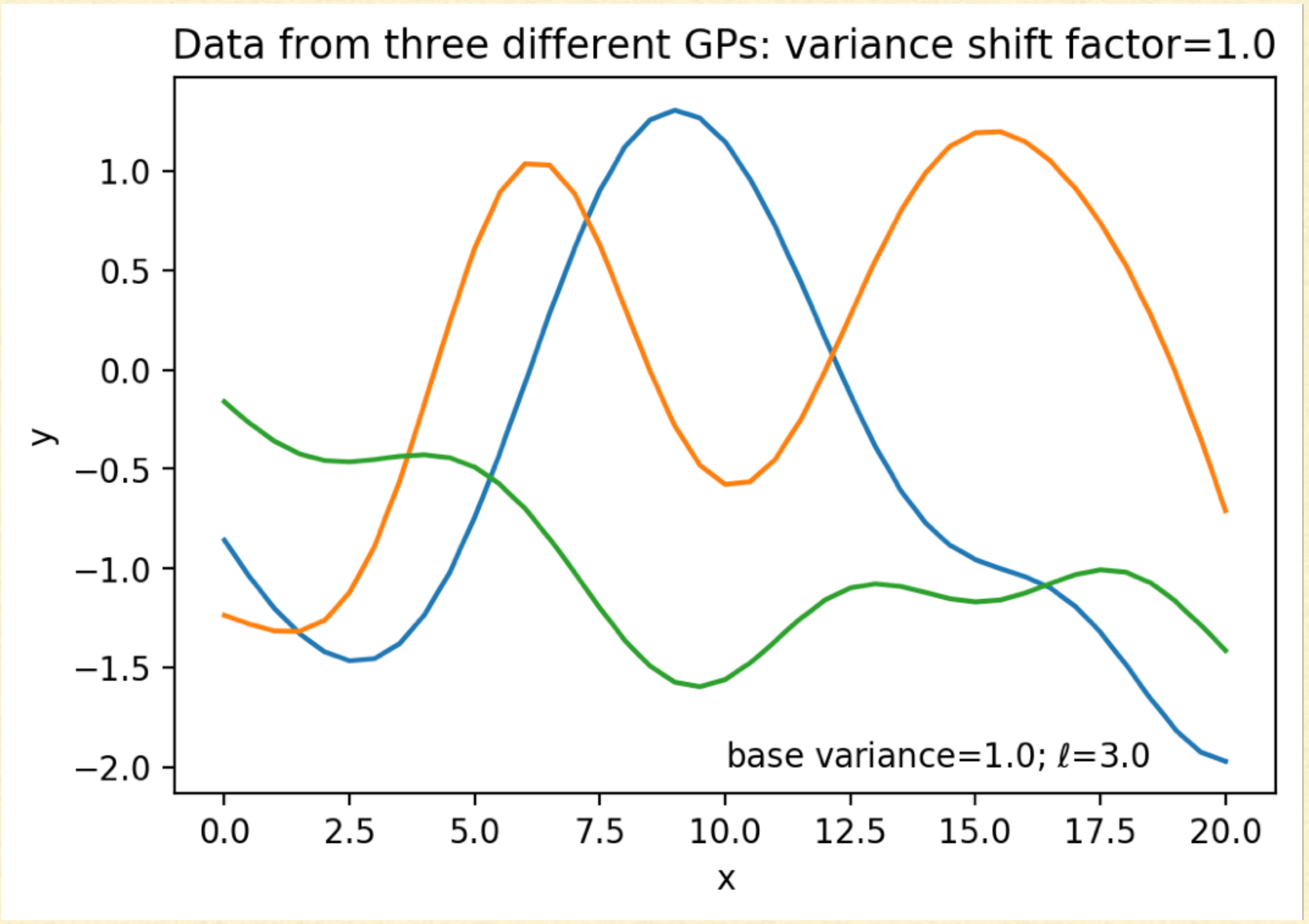
- Non-parametric, probabilistic model for a function
- Suppose we already know f at $x_1, x_2, x_3, \ldots, x_n$.
- Specify how f(y) is correlated with $f(x_1)$, $f(x_2)$, …..; don't specify underlying functional form.
- But value of f(y) is not deterministic: it's given by a (Gaussian) probability distribution.
- Correlation decreases as points get further away from each other.
- Specify correlation matrix of f at x and y, e.g.:

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

- Two parameters cbar and ℓ

Learning cbar and ℓ: Toy Example

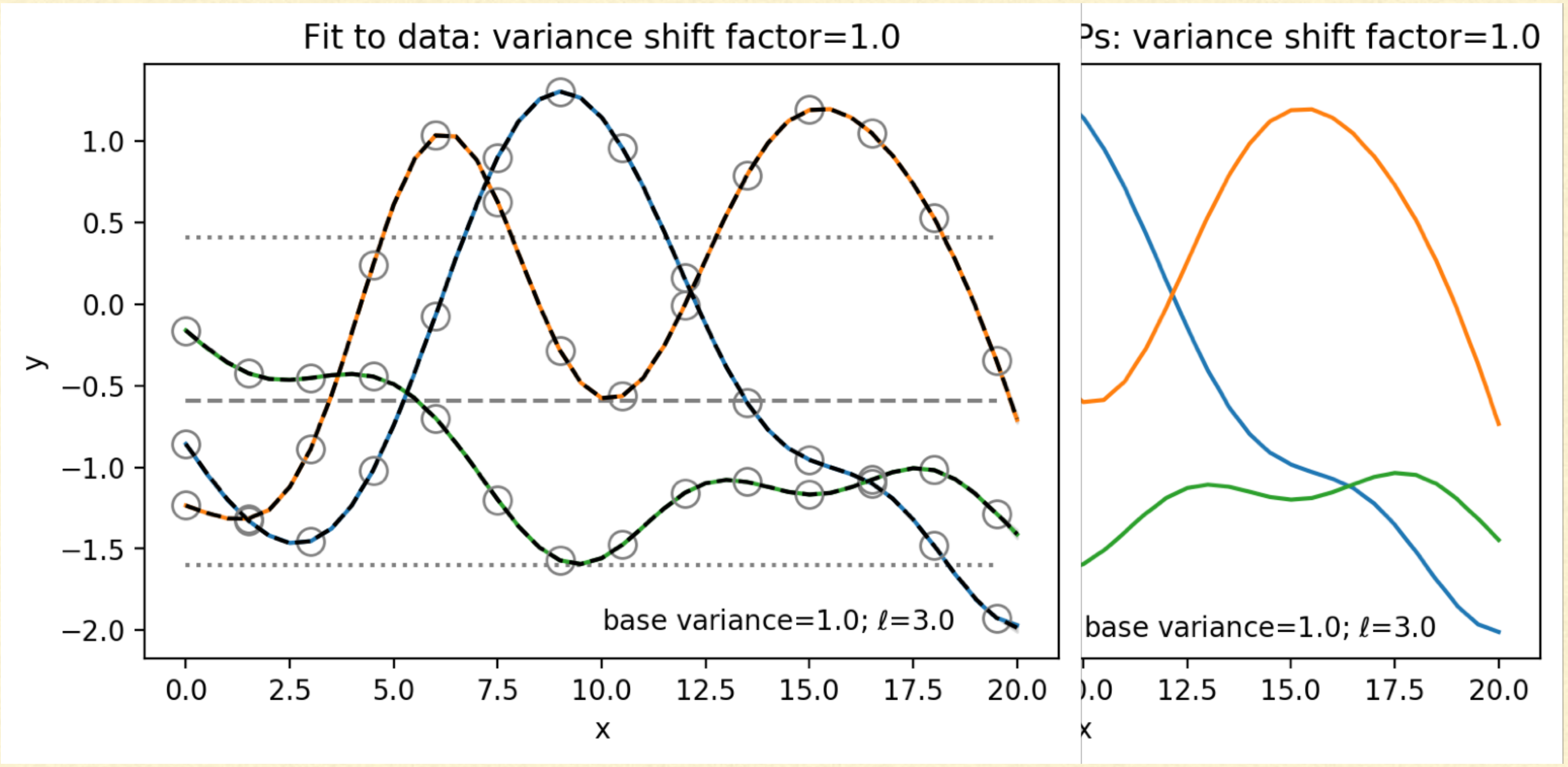
- Generate three data sets from an underlying GP with mean 0, variance 1 and length scale 3.



- Fit one GP to "training data" from all three data sets ("three realizations"): estimate μ, cbar, and ℓ.

Learning cbar and ℓ: Toy Example

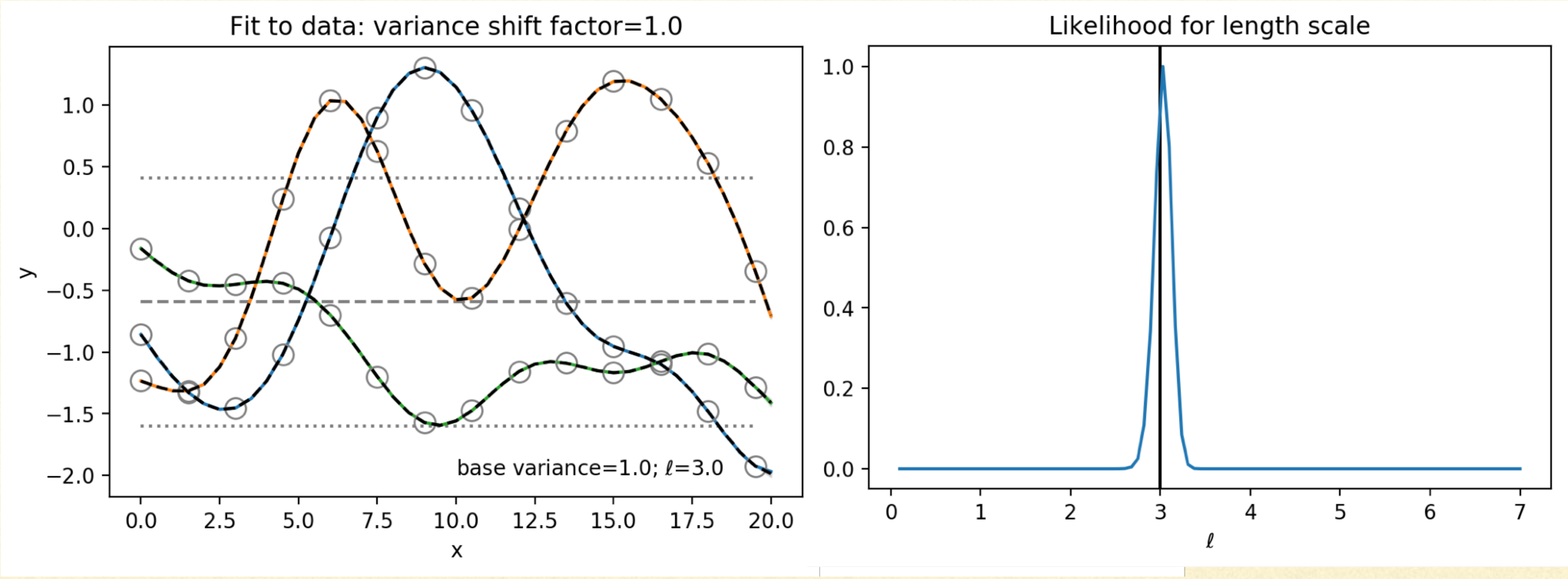
- Generate three data sets from an underlying GP with mean 0, variance 1 and length scale 3.



- Fit one GP to "training data" from all three data sets ("three realizations"): estimate μ, cbar, and ℓ.

Learning cbar and ℓ: Toy Example

- Generate three data sets from an underlying GP with mean 0, variance 1 and length scale 3.



- Fit one GP to “training data” from all three data sets (“three realizations”): estimate μ, cbar, and ℓ.

Diagnostics

Melendez, Furnstahl, DP, Pratola, Wesolowski, in preparation
after Bastos & O'Hagan, Technometrics, 2009

Diagnostics

Melendez, Furnstahl, DP, Pratola, Wesolowski, in preparation
after Bastos & O'Hagan, Technometrics, 2009

- Assess performance of fitted GP on "testing data" set.

Diagnostics

Melendez, Furnstahl, DP, Pratola, Wesolowski, in preparation
after Bastos & O'Hagan, Technometrics, 2009

- Assess performance of fitted GP on "testing data" set.
- Errors are correlated, so can't just add up number of sigmas. "Consistency plot" does not account for correlations.

Diagnostics

Melendez, Furnstahl, DP, Pratola, Wesolowski, in preparation
after Bastos & O'Hagan, Technometrics, 2009

- Assess performance of fitted GP on "testing data" set.
- Errors are correlated, so can't just add up number of sigmas. "Consistency plot" does not account for correlations.
- Define Mahalanobis distance, which does account for correlations

$$\mathrm{MD} \equiv (y(x_i) - \mu)^T [k(y(x_i), y(x_j))]^{-1} (y(x_j) - \mu)$$

Diagnostics

Melendez, Furnstahl, DP, Pratola, Wesolowski, in preparation
after Bastos & O'Hagan, Technometrics, 2009

- Assess performance of fitted GP on "testing data" set.

- Errors are correlated, so can't just add up number of sigmas. "Consistency plot" does not account for correlations.

- Define Mahalanobis distance, which does account for correlations

$$\mathrm{MD} \equiv (y(x_i) - \mu)^T [k(y(x_i), y(x_j))]^{-1} (y(x_j) - \mu)$$

- Write $k(y(x_i),y(x_j))=G\, G^T$ with G from PC or SV decomposition then form "Pivoted Cholesky" or "Eigen" errors from $G(y(x_i)-\mu)$

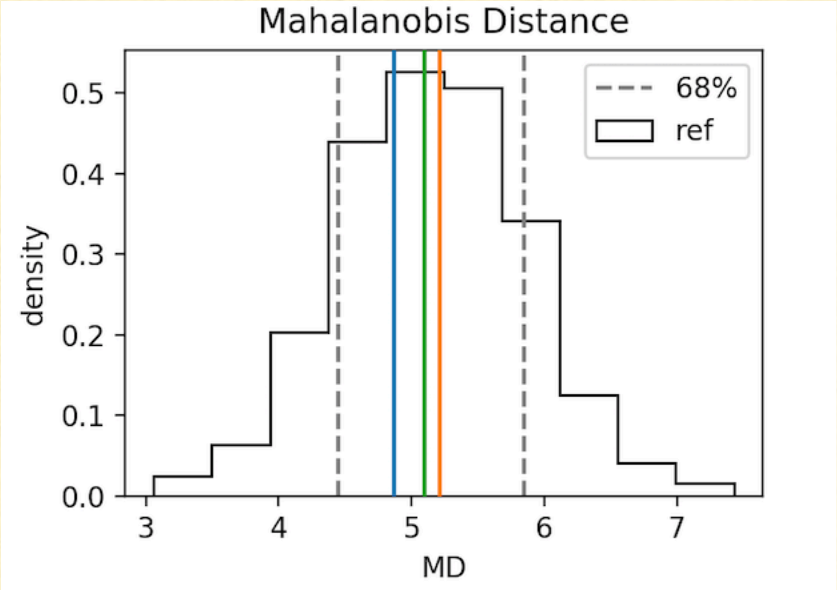
Diagnostics

Melendez, Furnstahl, DP, Pratola, Wesolowski, in preparation
after Bastos & O'Hagan, Technometrics, 2009

- Assess performance of fitted GP on "testing data" set.
- Errors are correlated, so can't just add up number of sigmas. "Consistency plot" does not account for correlations.
- Define Mahalanobis distance, which does account for correlations

$$\mathrm{MD} \equiv (y(x_i) - \mu)^T [k(y(x_i), y(x_j))]^{-1} (y(x_j) - \mu)$$

- Write $k(y(x_i),y(x_j))=G\,G^T$ with G from PC or SV decomposition then form "Pivoted Cholesky" or "Eigen" errors from $G(y(x_i)-\mu)$



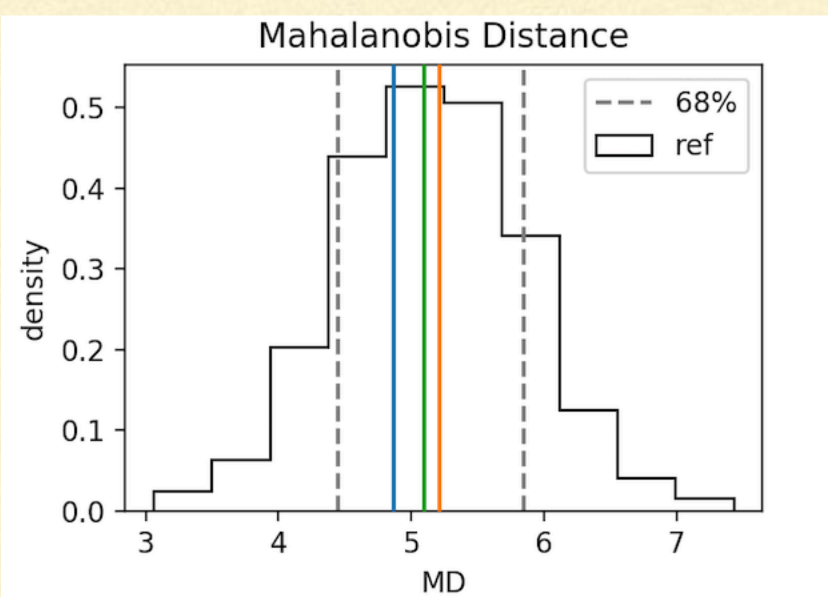
Diagnostics

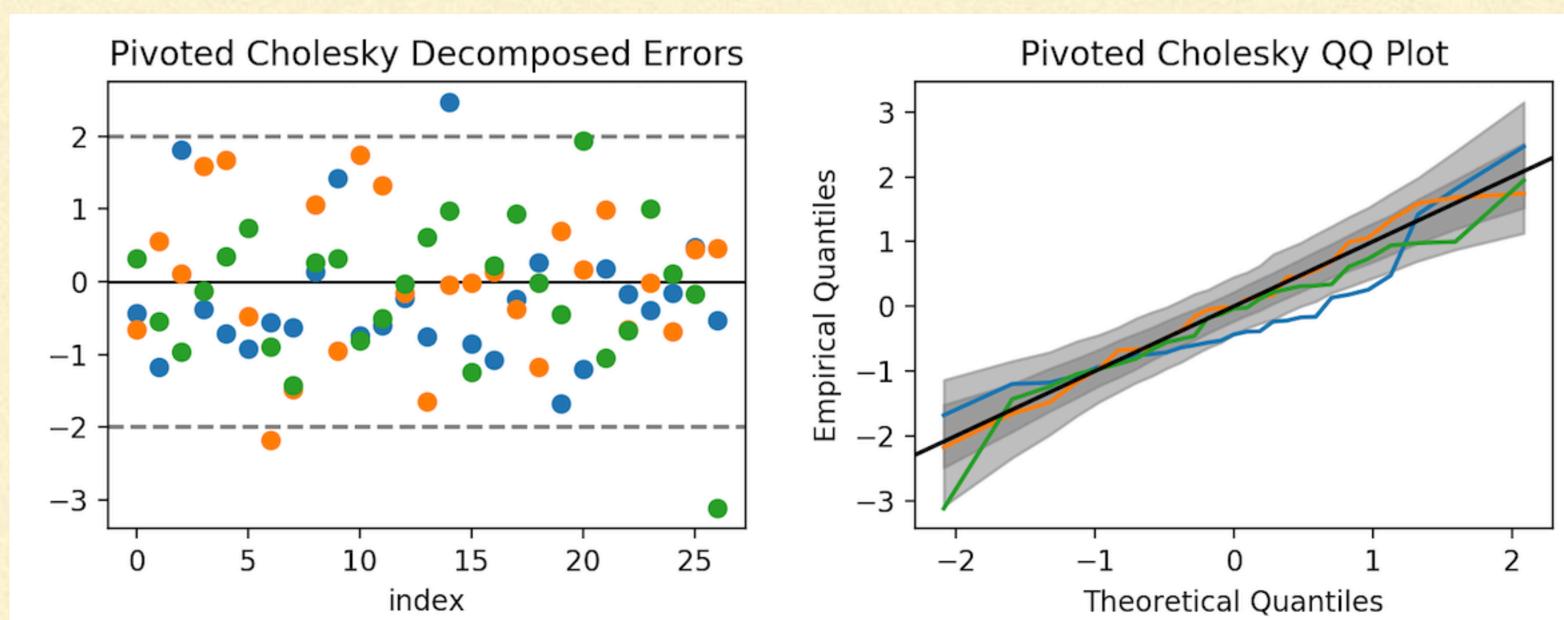
Melendez, Furnstahl, DP, Pratola, Wesolowski, in preparation
after Bastos & O'Hagan, Technometrics, 2009

- Assess performance of fitted GP on "testing data" set.
- Errors are correlated, so can't just add up number of sigmas. "Consistency plot" does not account for correlations.
- Define Mahalanobis distance, which does account for correlations

$$\mathrm{MD} \equiv (y(x_i) - \mu)^T [k(y(x_i), y(x_j))]^{-1} (y(x_j) - \mu)$$

- Write $k(y(x_i),y(x_j))=G\,G^T$ with G from PC or SV decomposition then form "Pivoted Cholesky" or "Eigen" errors from $G(y(x_i)-\mu)$



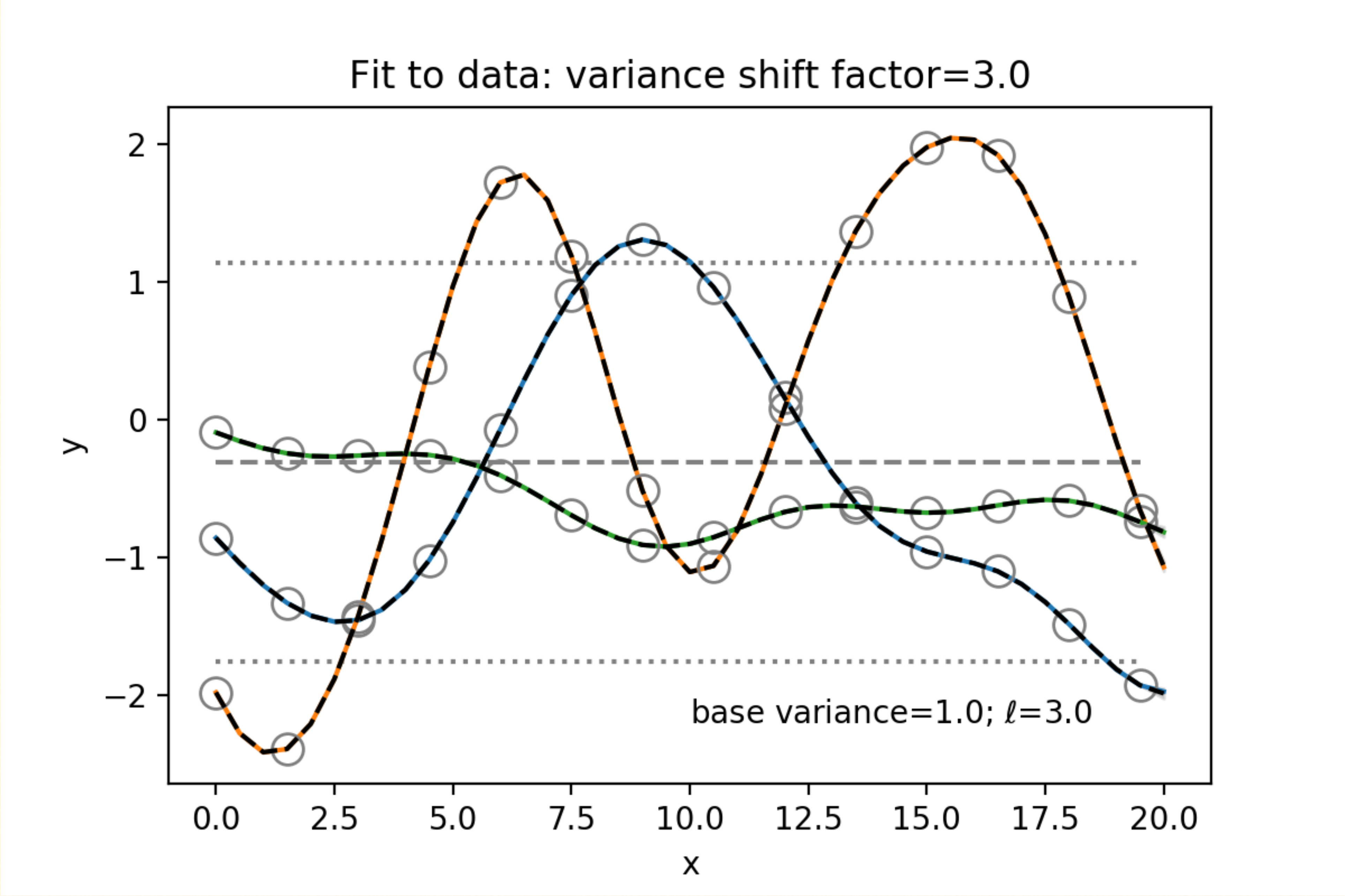


What can go wrong 1: different cbar's

- Try to fit a single GP to data generated using different variances

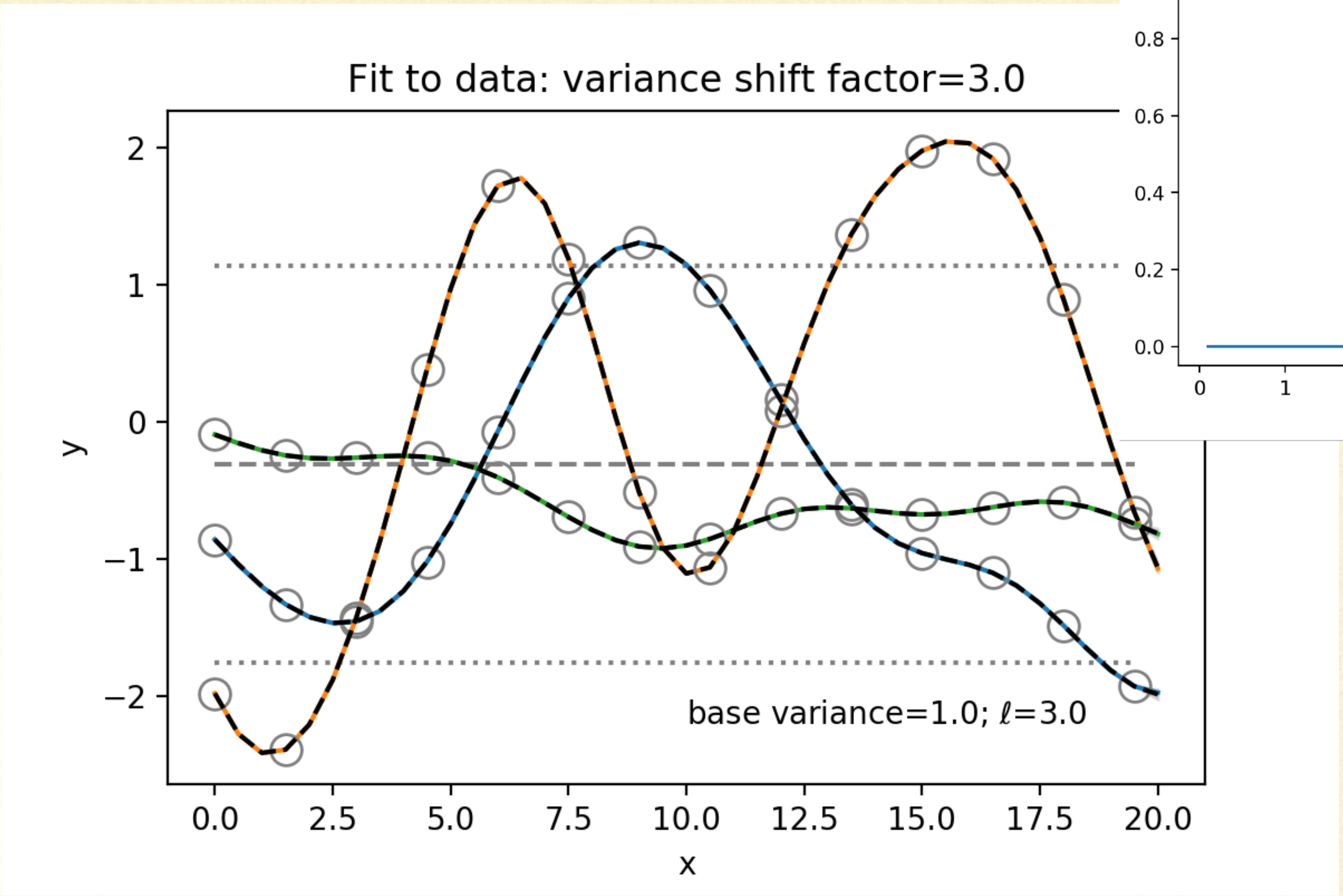
What can go wrong 1: different cbar's

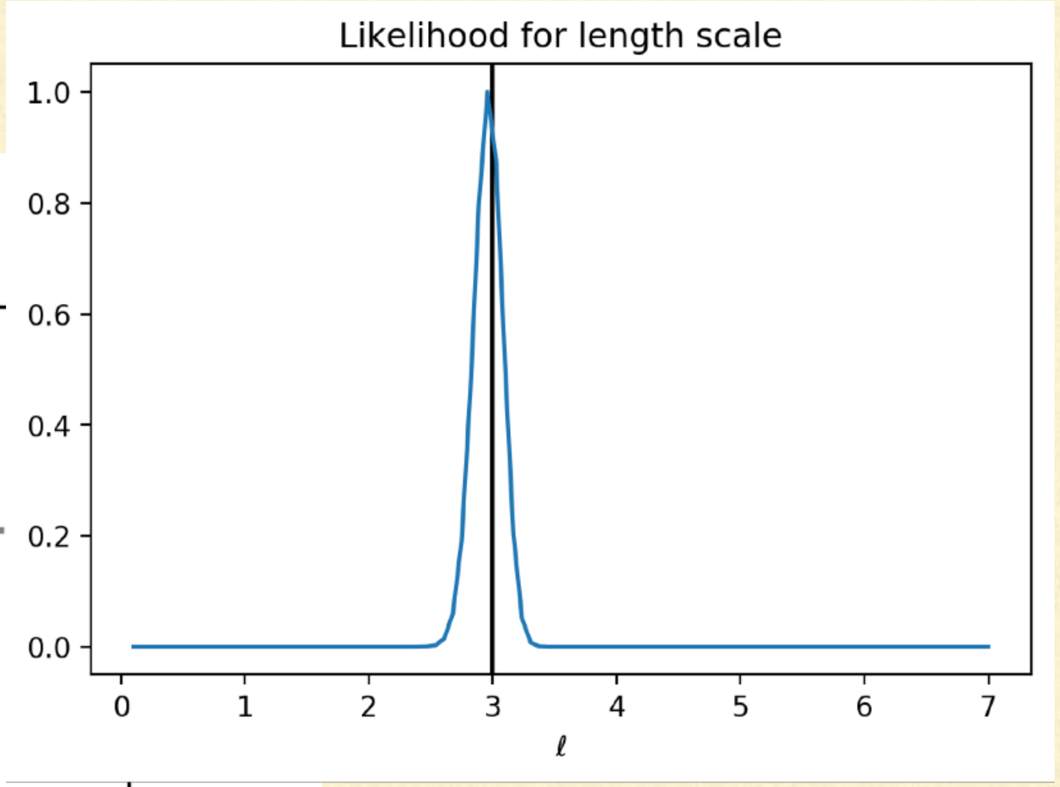
- Try to fit a single GP to data generated using different variances



What can go wrong 1: different cbar's

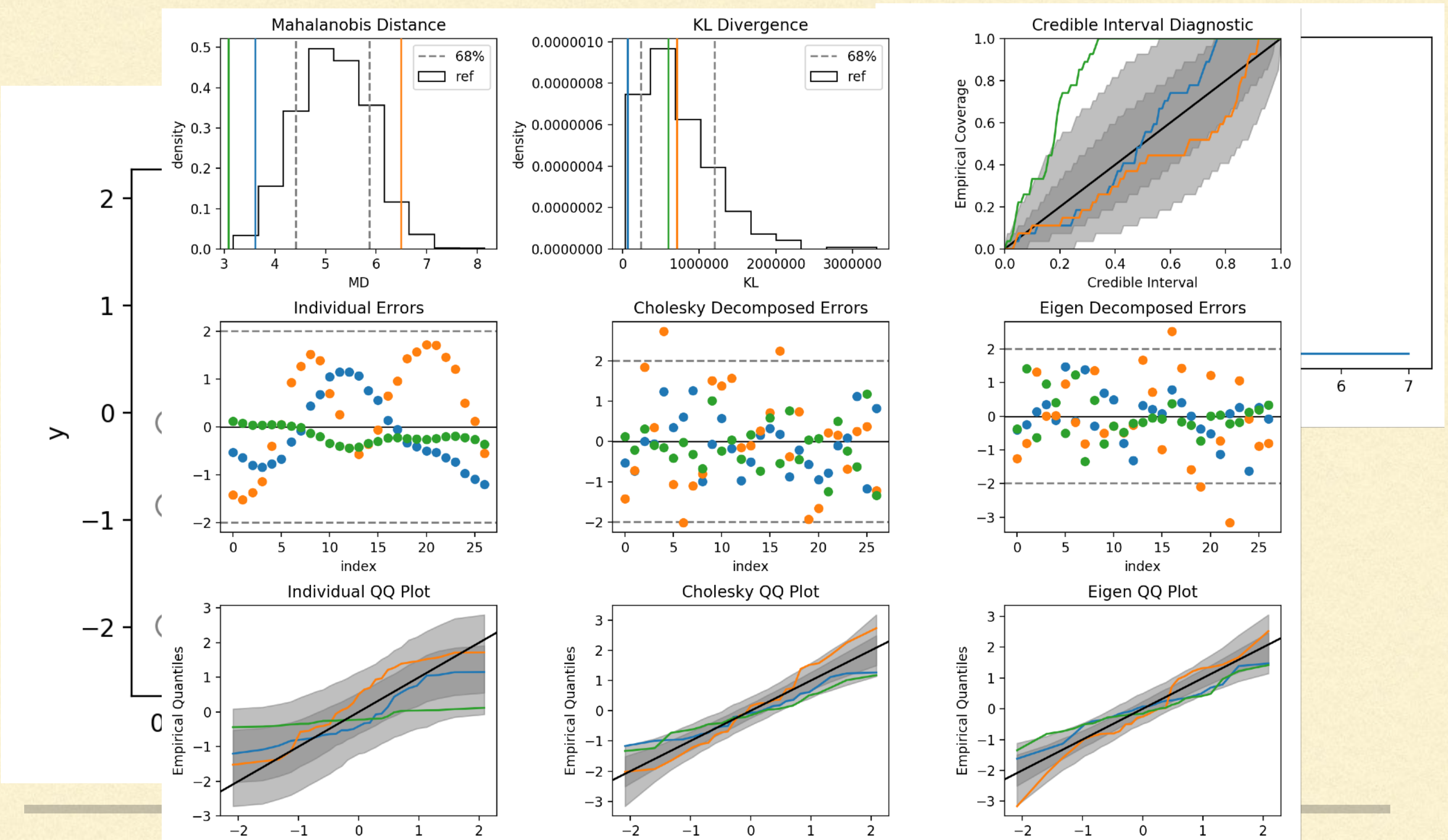
- Try to fit a single GP to data generated using different variances





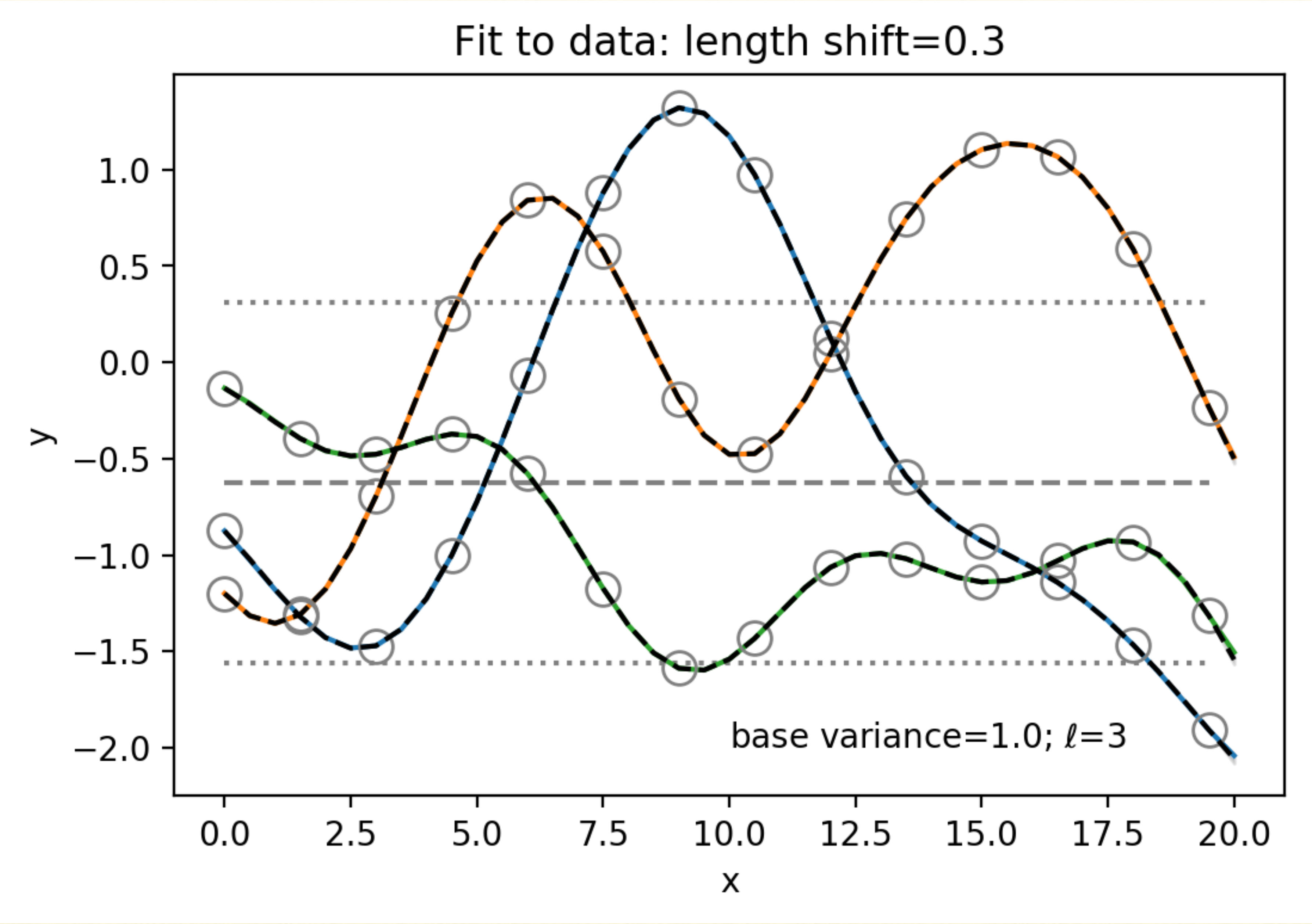
What can go wrong I: different cbar's

- Try to fit a single GP to data generated using different variances



What can go wrong II: different ℓs

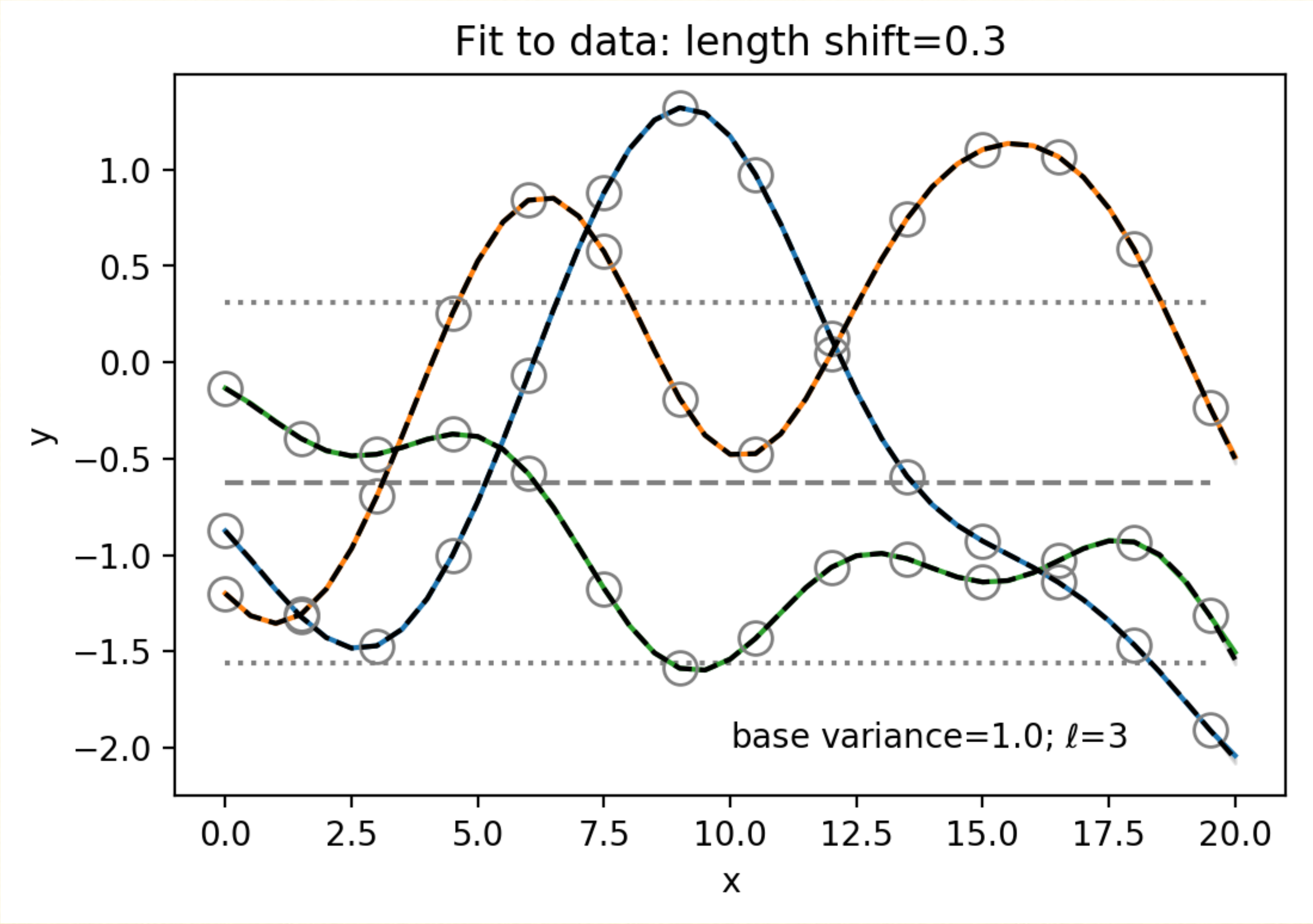
- Try to fit single GP to data with different ℓs



What can go wrong II: different ℓs

- Try to fit single GP to data with different ℓs

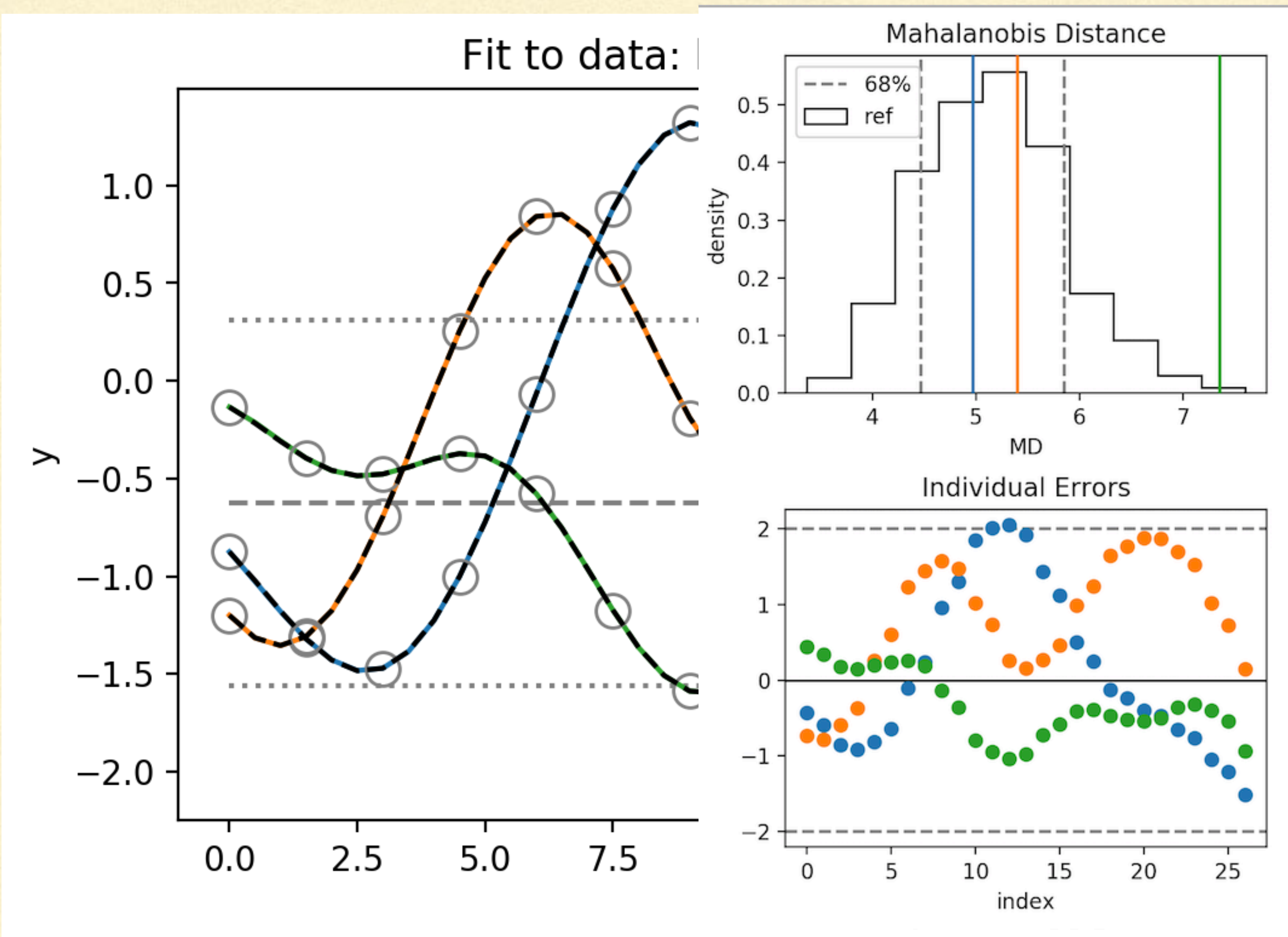
ℓ=3.1

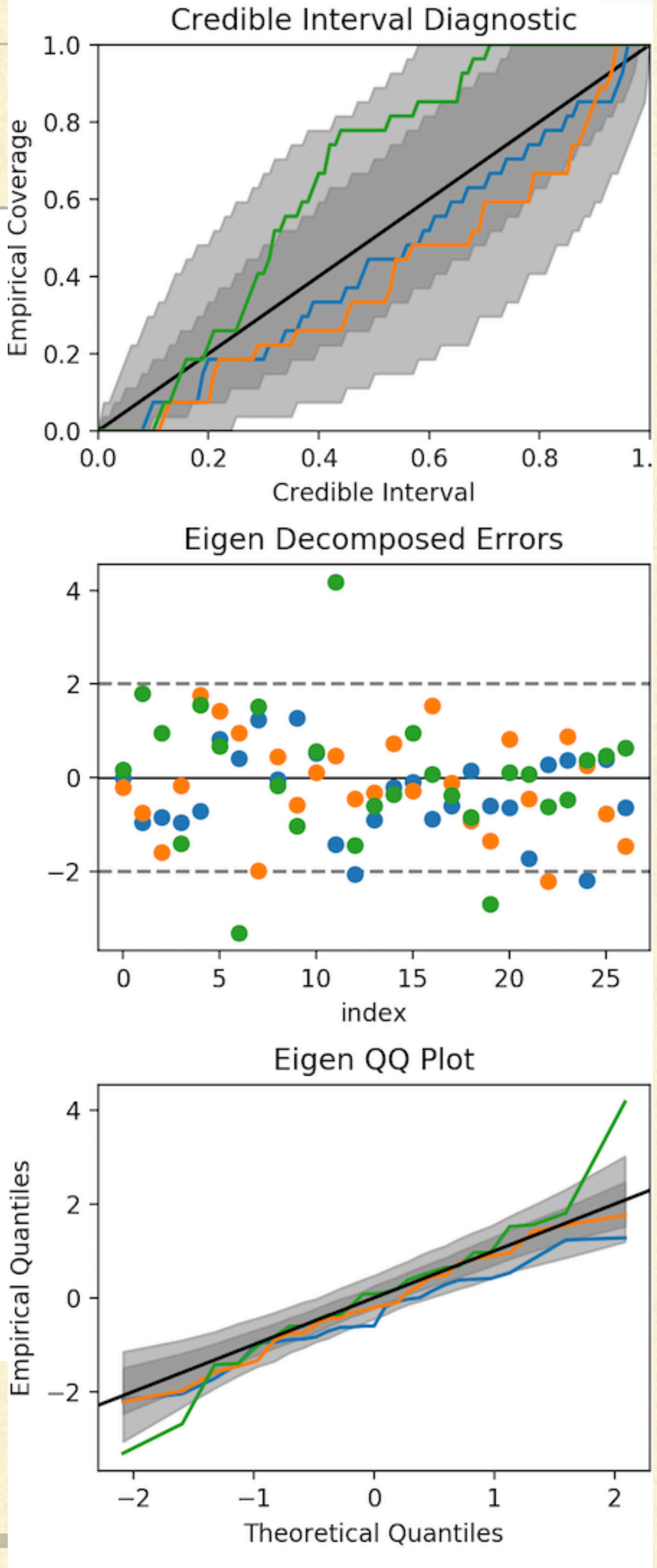


What can go wrong II: different ℓs

- Try to fit single GP to data with different ℓs

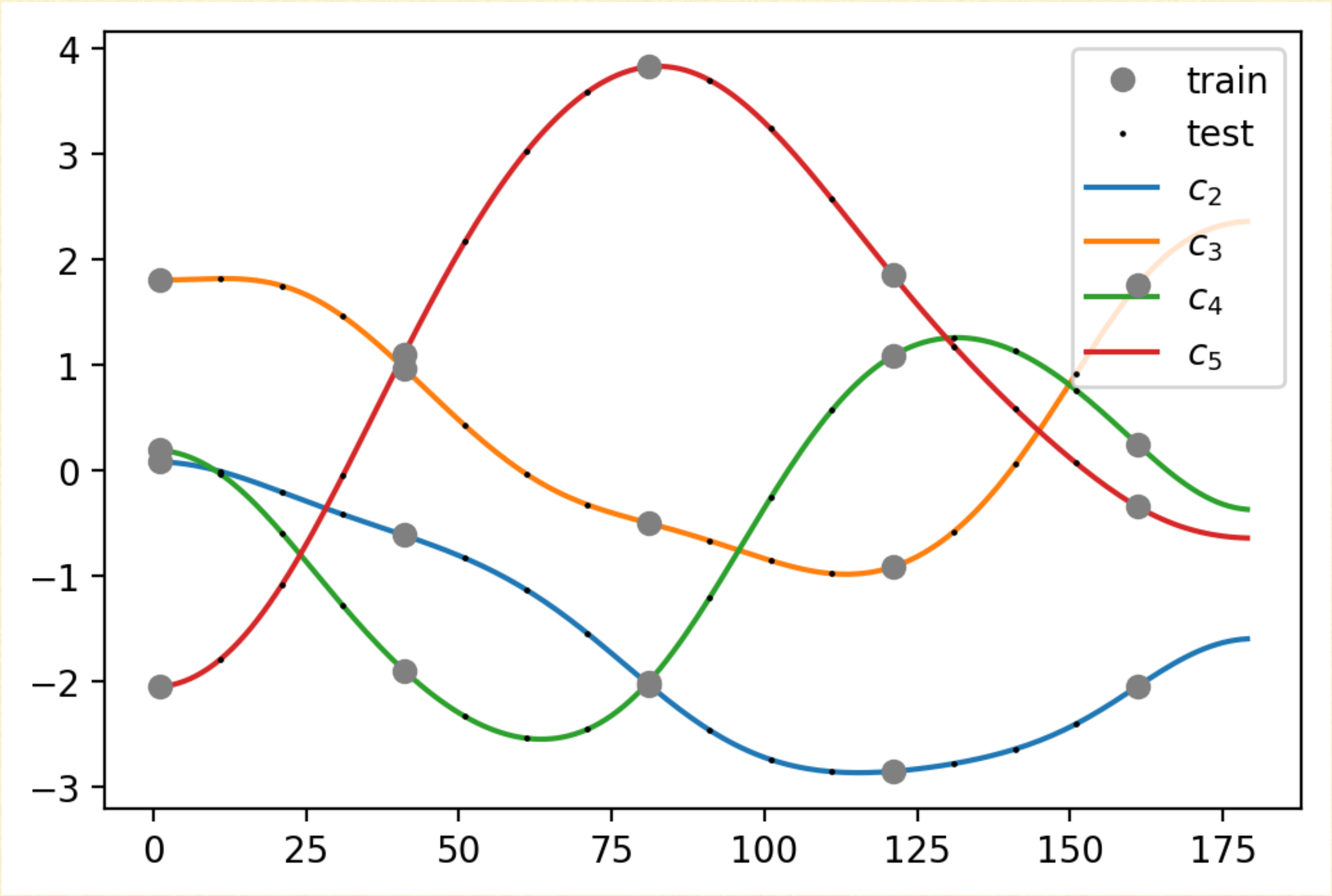
ℓ=3.1





Differential cross section vs. angle

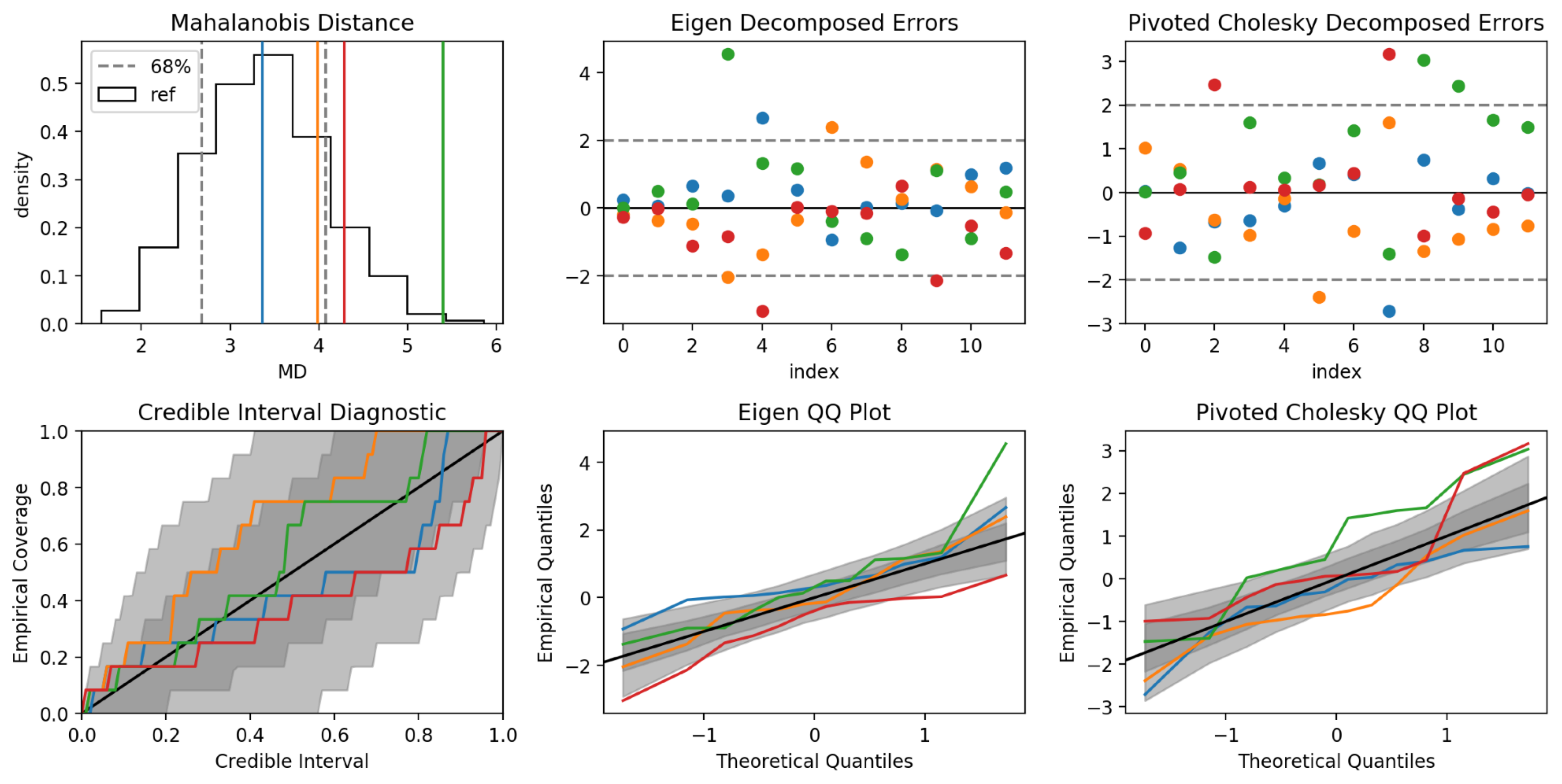
$\frac{d\sigma}{d\Omega}$ E_{lab}=96 MeV **PRELIMINARY**



RKE potential; Λ=500 MeV

Differential cross section vs. angle

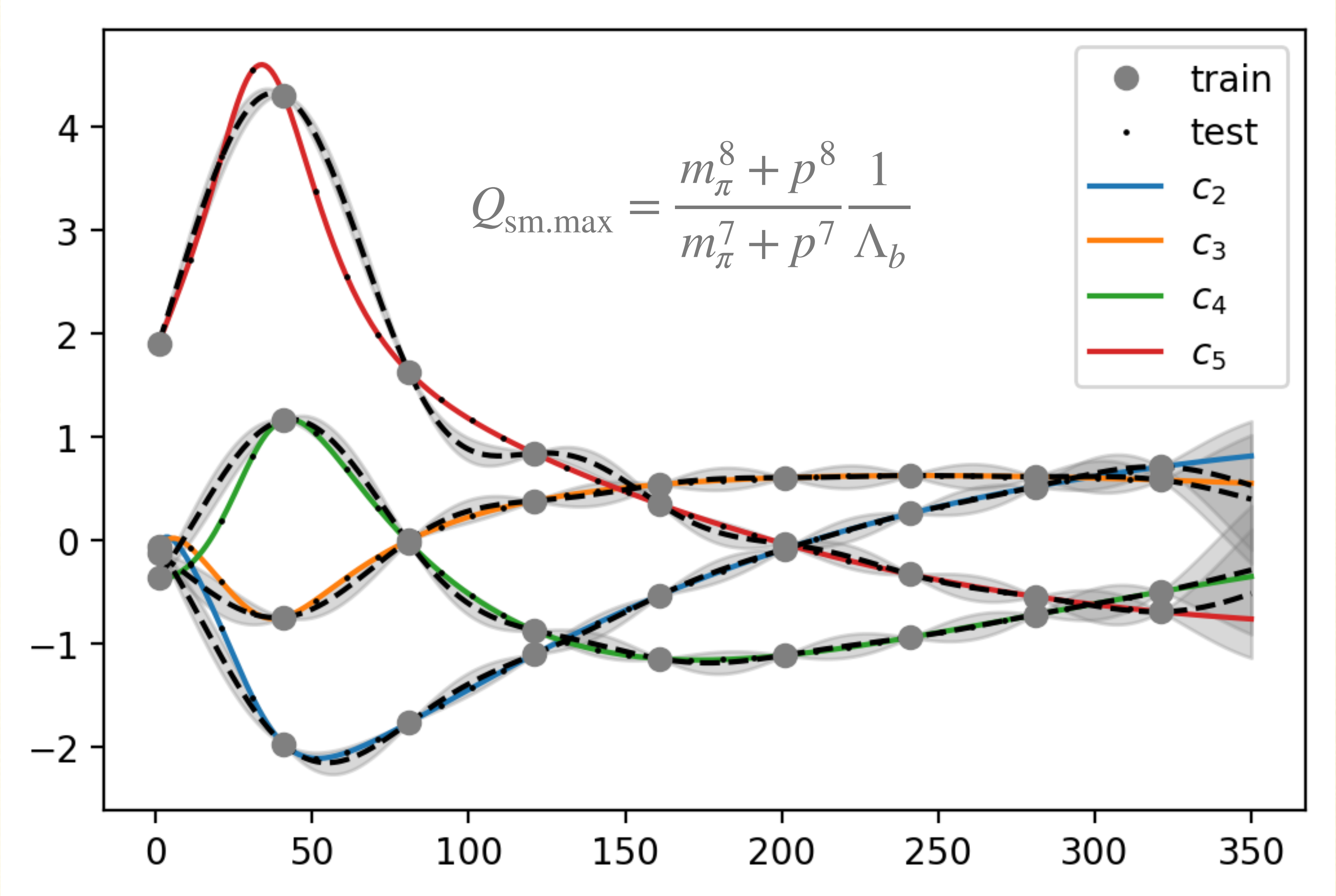
$\frac{d\sigma}{d\Omega}$ E_{lab}=96 MeV **PRELIMINARY**



RKE potential; Λ=500 MeV

Total cross section

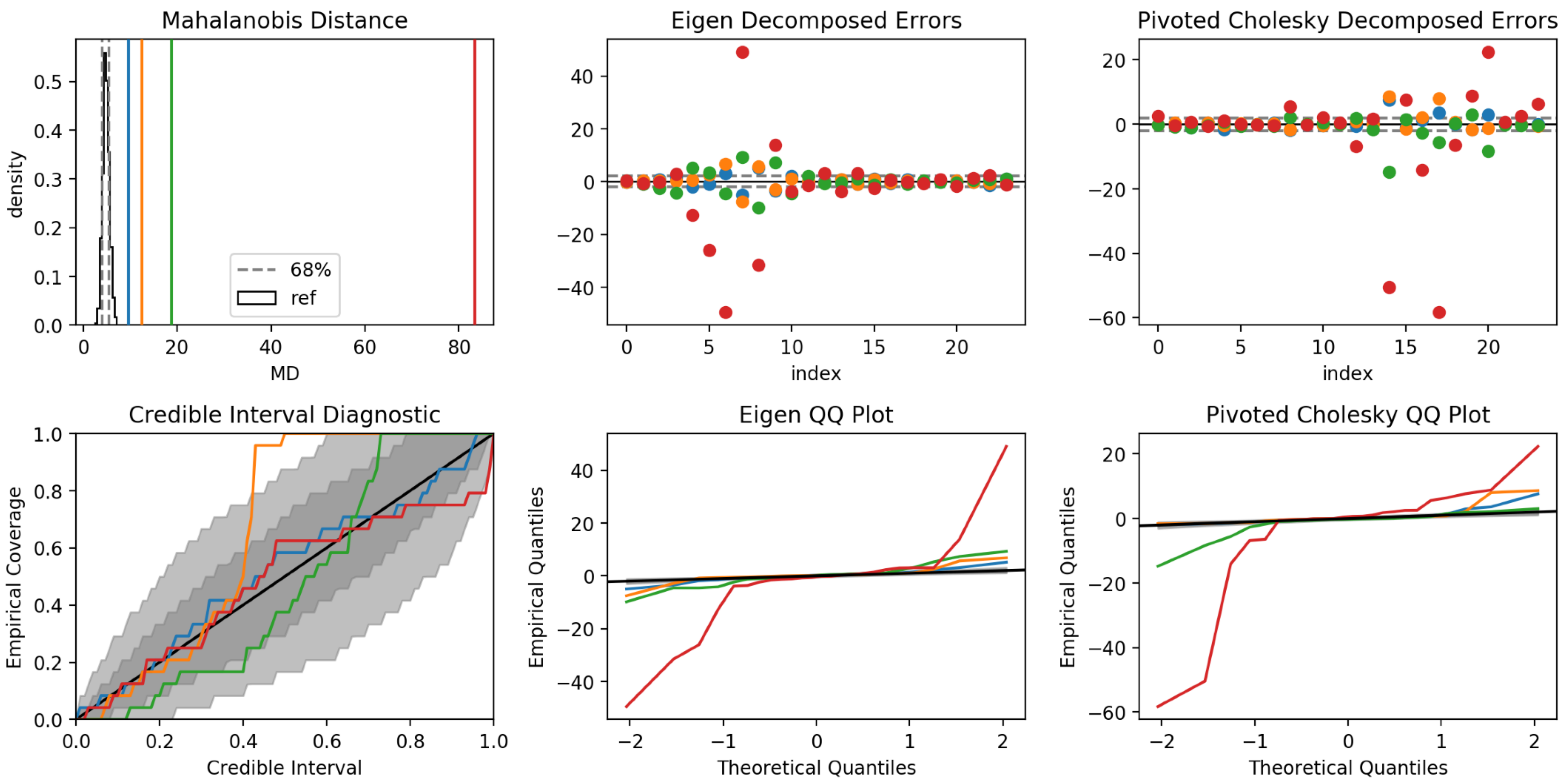
PRELIMINARY



RKE potential;
Λ=500 MeV

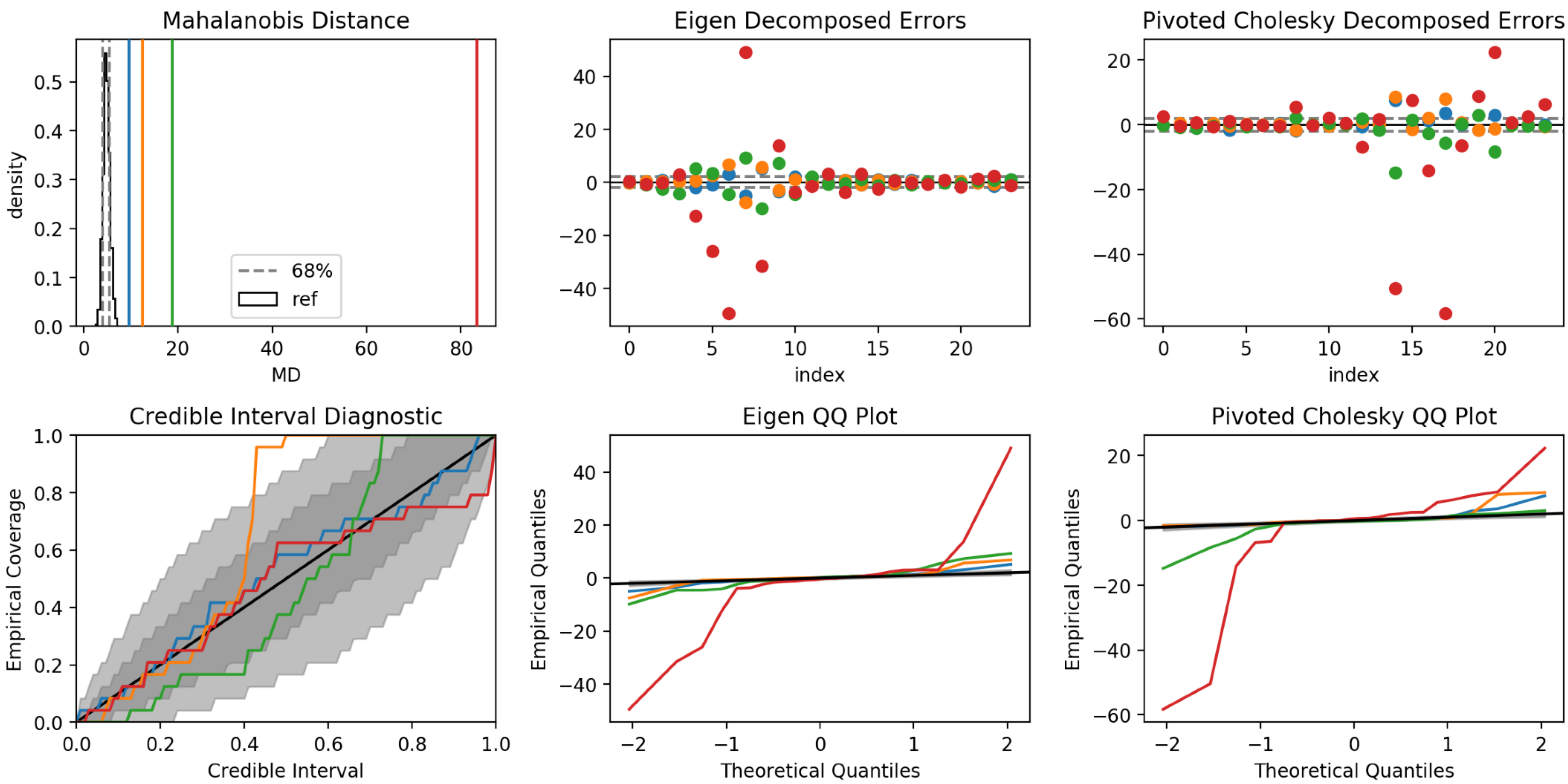
Total cross section

PRELIMINARY



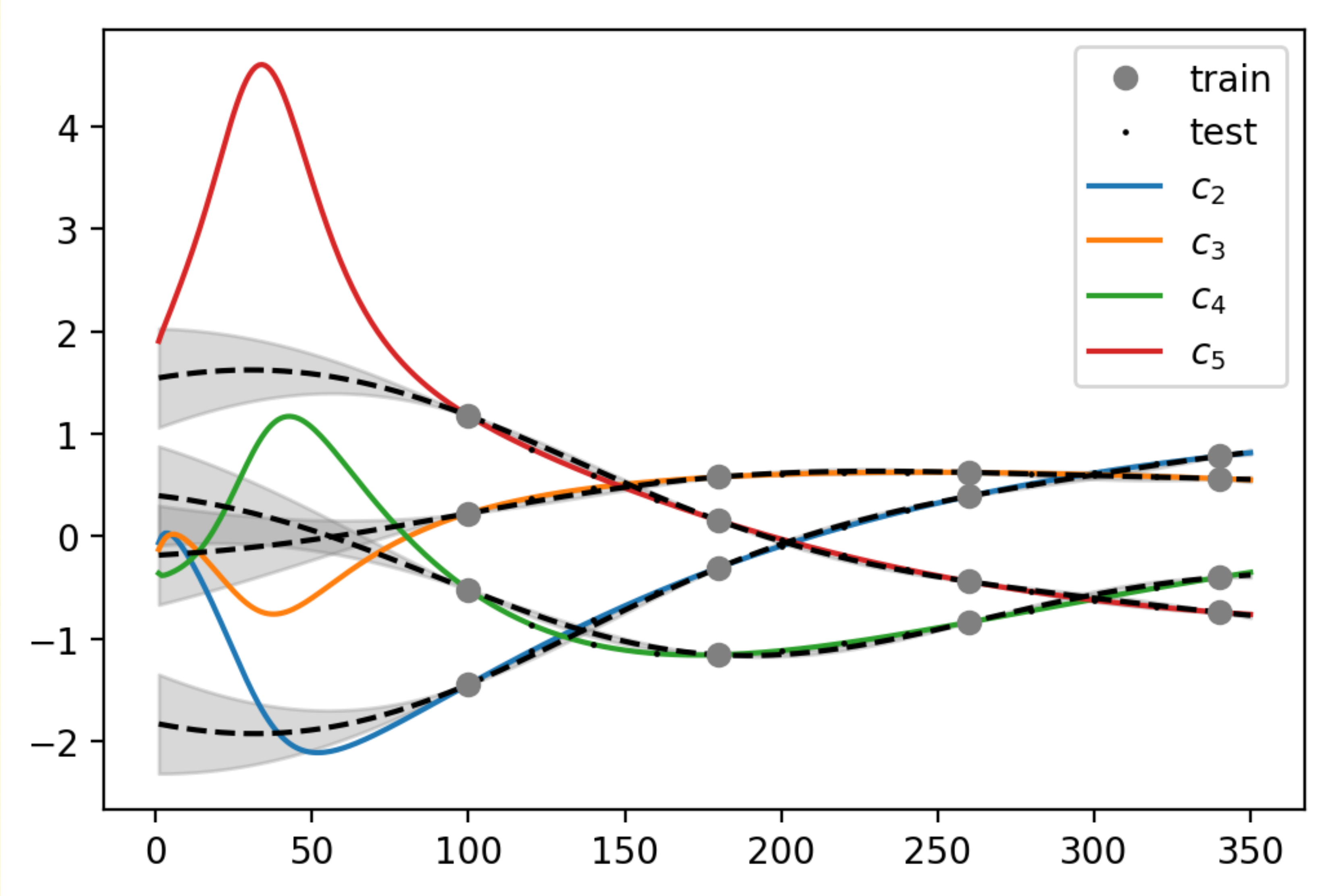
Total cross section

PRELIMINARY

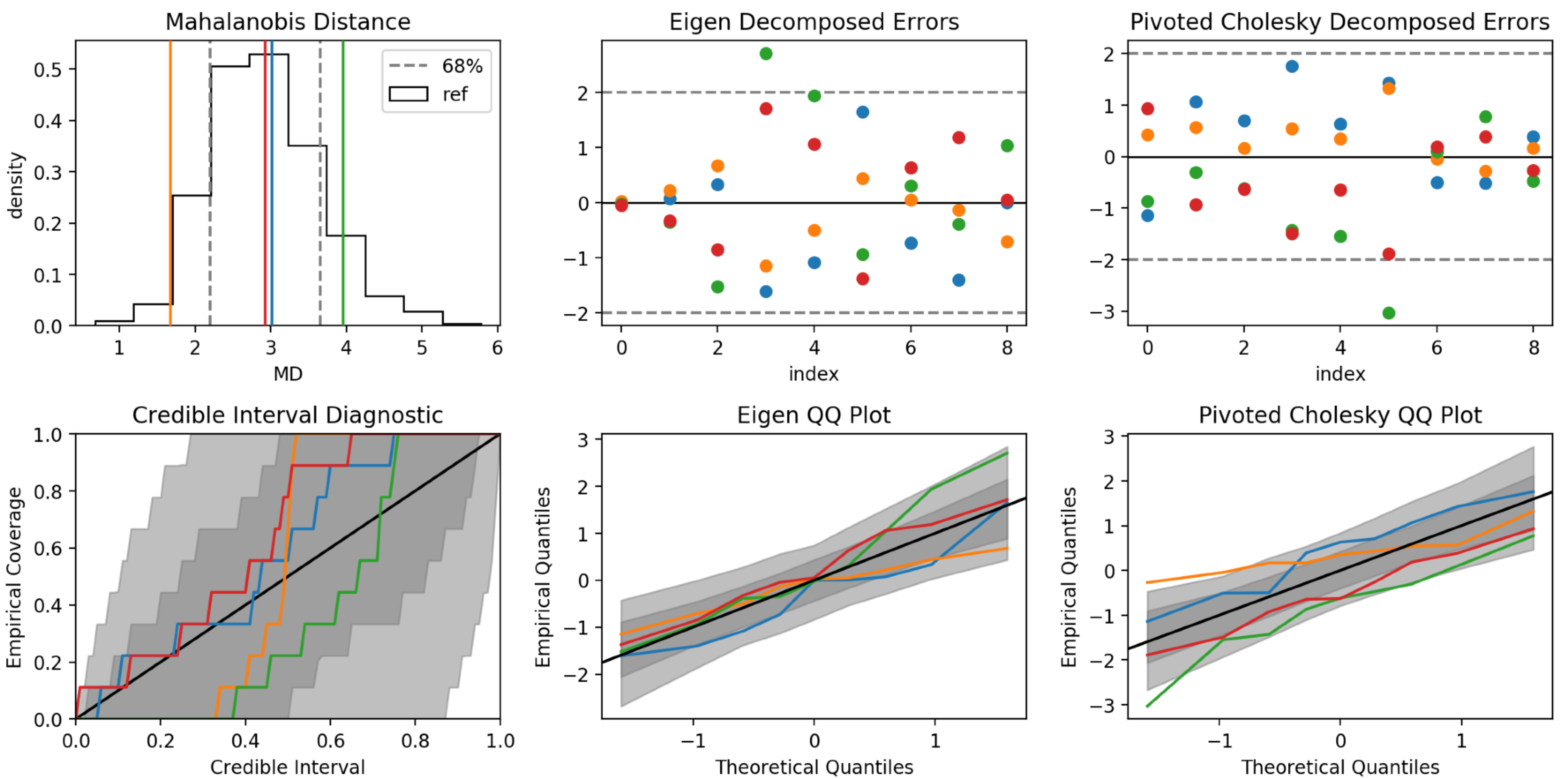


- Even with change of expansion parameter built in, change of length scale messes up GP estimation: "non-stationarity"

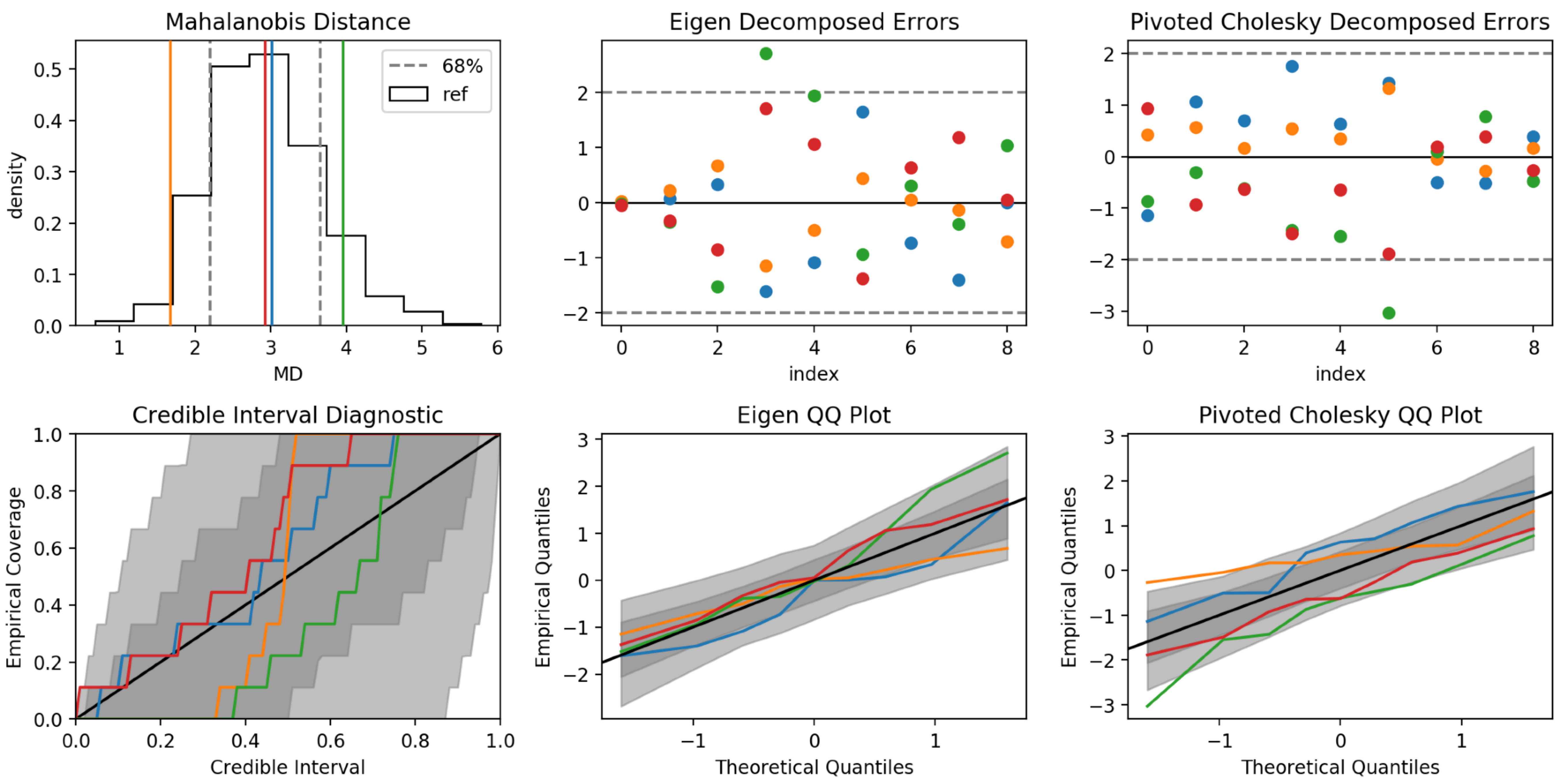
One way to deal with low energies



One way to deal with low energies



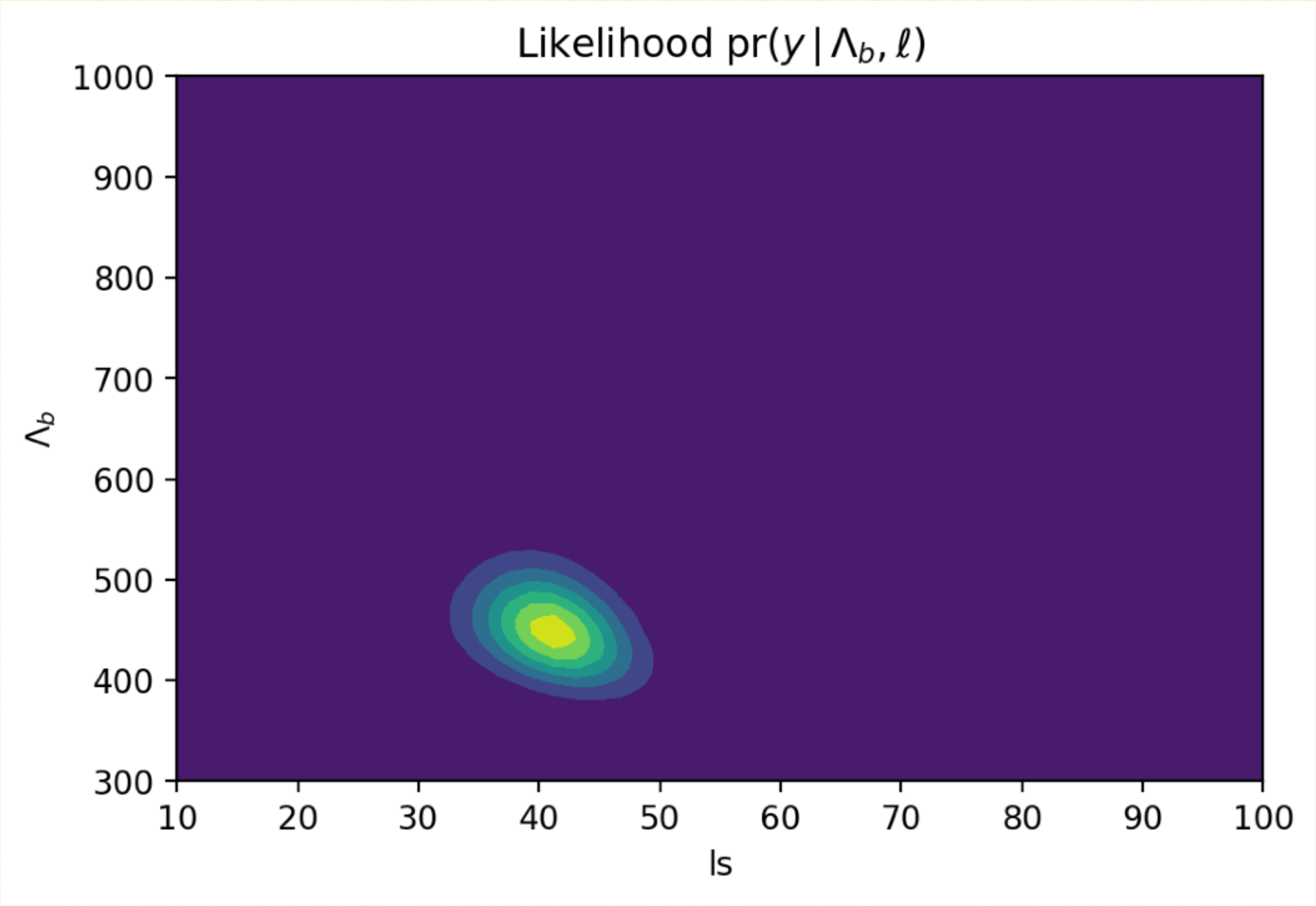
One way to deal with low energies



- Hypothesis of EFT coefficients as draws from a GP looks very healthy for cross section at E_{lab}>100 MeV: next we need to model change of behavior

Λ_b and ℓ inference

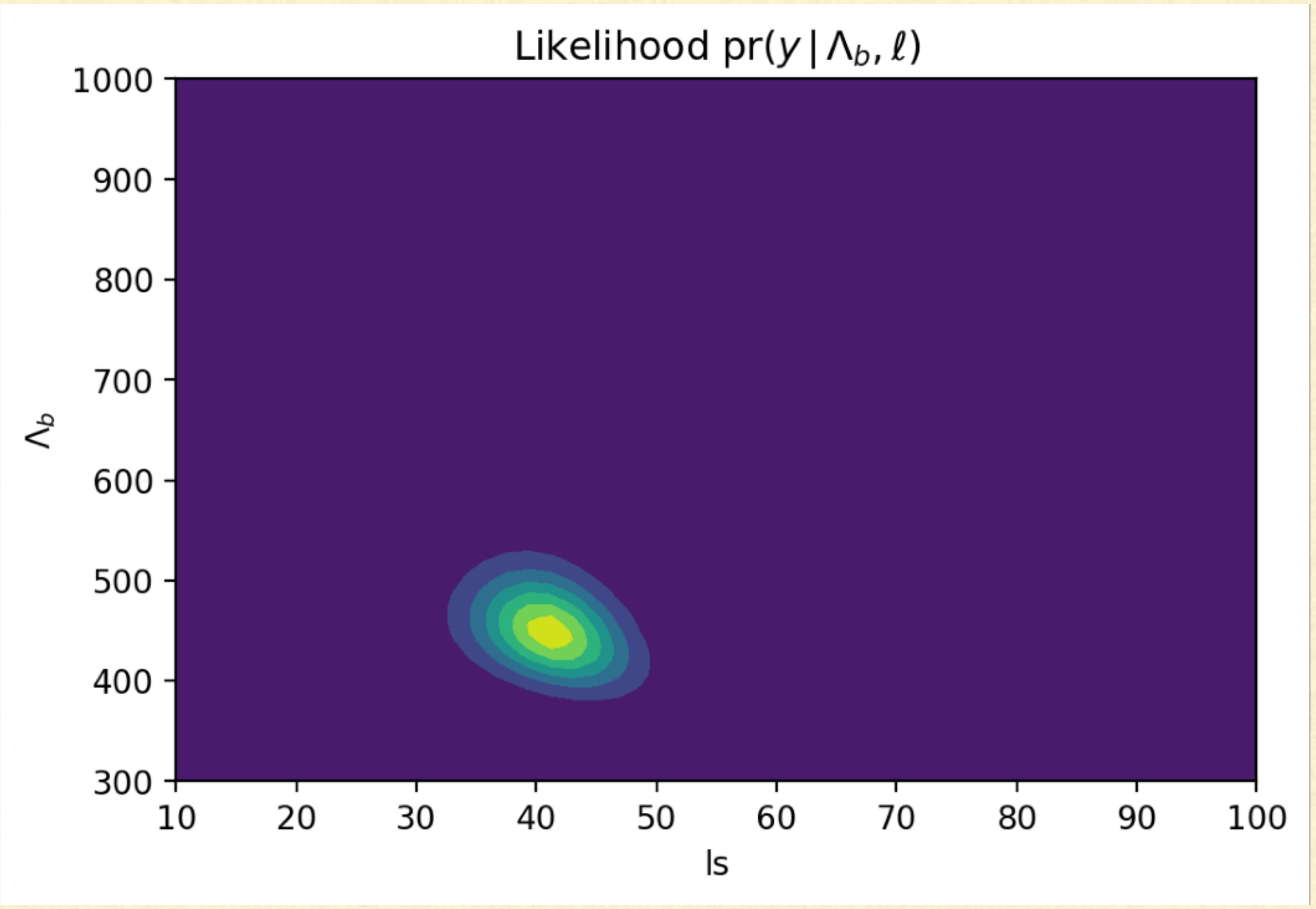
PRELIMINARY



RKE potential; Λ=500 MeV;
Total cross section only

Λ_b and ℓ inference

PRELIMINARY

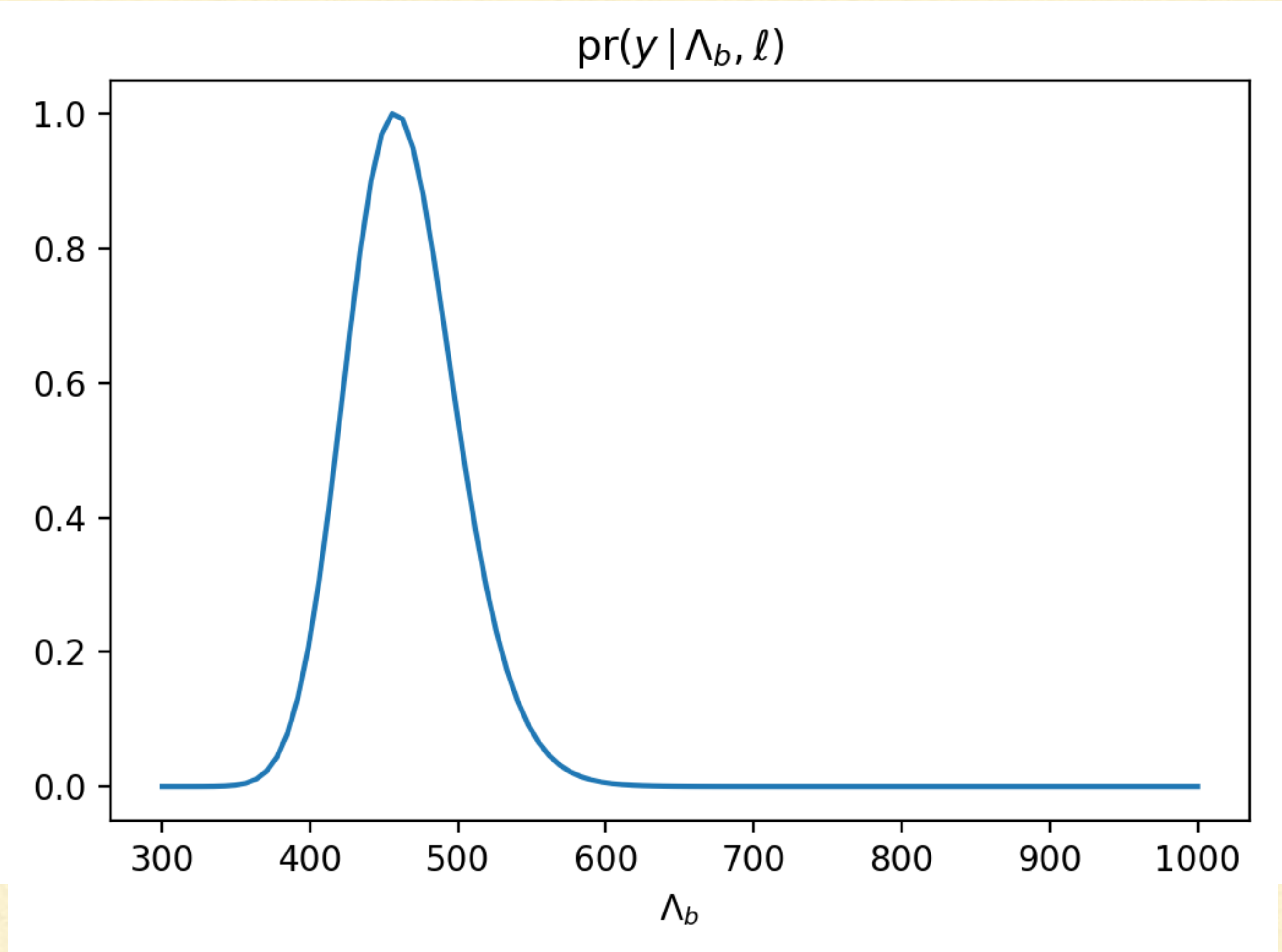


RKE potential; Λ=500 MeV;
Total cross section only

$\ell \approx 40$ MeV

Λ_b and ℓ inference

PRELIMINARY

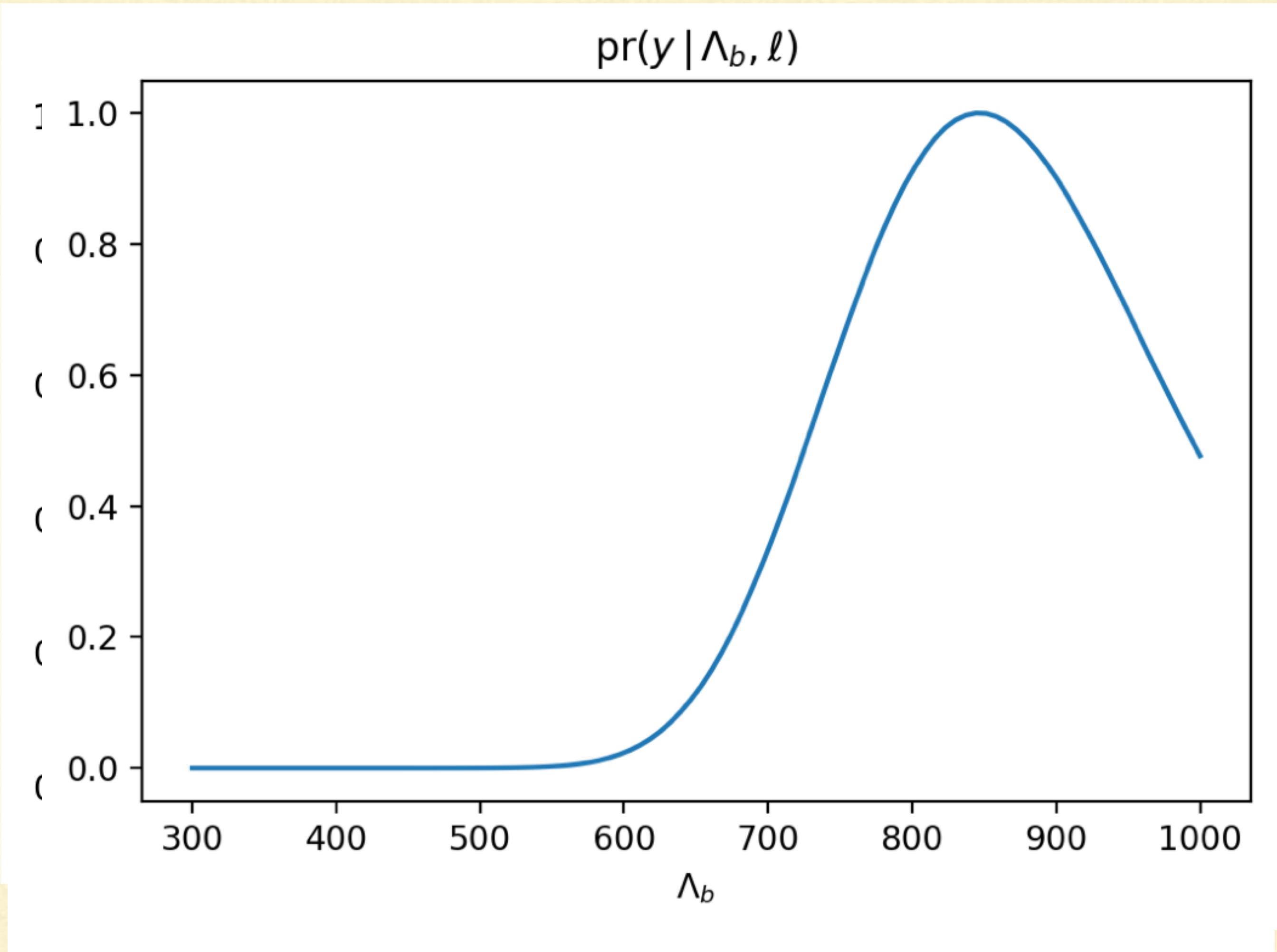


RKE potential; Λ=500 MeV;
Total cross section only

$\ell \approx$ 40 MeV

Λ_b and ℓ inference

PRELIMINARY



RKE potential; Λ=500 MeV;

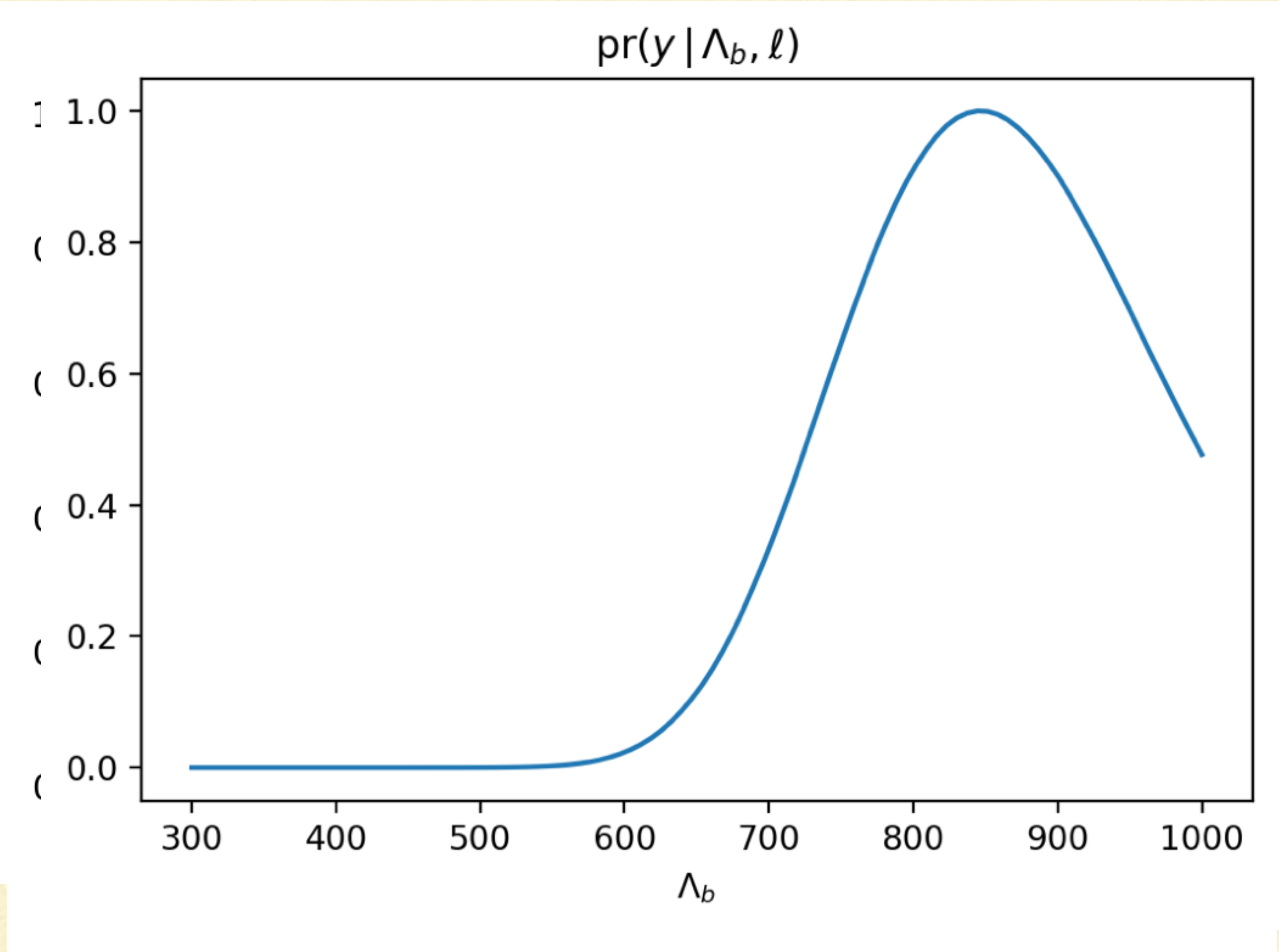
Total cross section only

Restricted domain

$\ell \approx 116$ MeV

Λ_b and ℓ inference

PRELIMINARY



RKE potential; Λ=500 MeV;

Total cross section only

Restricted domain

$\ell \approx 116$ MeV

To do:

- Model non-stationarity
- Play with "switch over" of expansion parameter at $p \approx m_\pi$
- Include other observables in analysis

Applications/future work

- Analyze different ChiEFT NN potentials for breakdown-scale, order-by-order calibration, etc.
- Gaussian Process models for ChiEFT expansion of E/N in nuclear and neutron matter

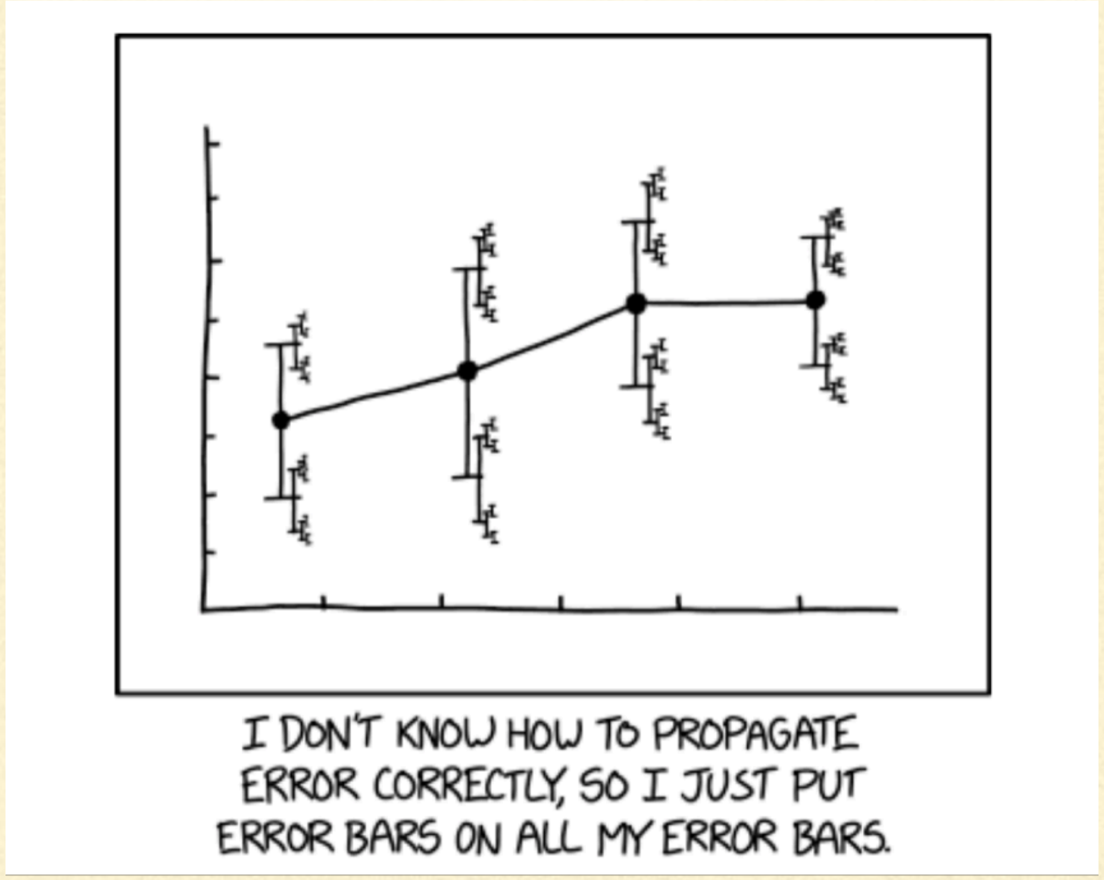
  With C. Drischler
- Gaussian Process models of Compton scattering observables

Applications/future work

- Analyze different ChiEFT NN potentials for breakdown-scale, order-by-order calibration, etc.
- Gaussian Process models for ChiEFT expansion of E/N in nuclear and neutron matter

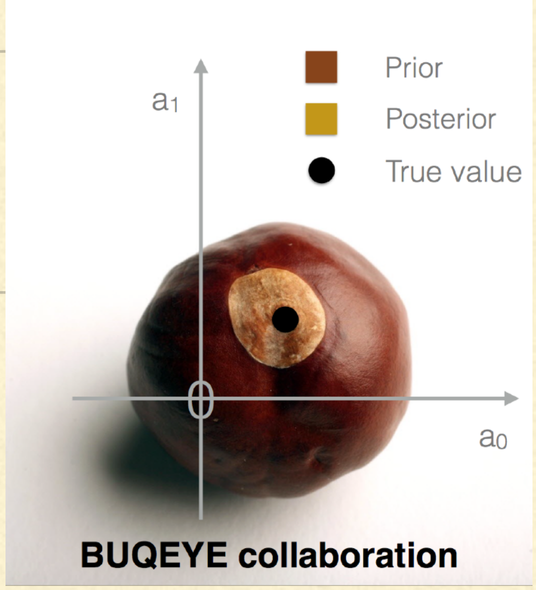
With C. Drischler

- Gaussian Process models of Compton scattering observables



http://xkcd.com

Summary



- Bayesian analysis makes explicit what the assumptions about the EFT convergence pattern are
- A rigorous treatment of truncation errors can validate (or falsify) claims that potentials respect the ChiEFT power counting
- And improve parameter estimation

Sarah's talk

- But to do this properly we need a good model for the way that truncation errors are correlated across different observables
- Gaussian Process models of EFT truncation errors provide understanding of those correlations

gsum package

- Also provide an inexpensive way to simulate expensive observables
- Honest parameter estimation, calibration plots, breakdown-scale inference…

A Generic EFT

$$g(x) = \sum_{i=0}^{k} \mathcal{A}_i(x) x^i \qquad x = \frac{p}{\Lambda_b}$$

A Generic EFT

- Suppose we are interested in a quantity as a function of a momentum, p, that is small compared to some high scale, Λ_b.
- EFT expansion for quantity is $g(x) = \sum_{i=0}^{k} \mathcal{A}_i(x) x^i$

$$x = \frac{p}{\Lambda_b}$$

A Generic EFT

- Suppose we are interested in a quantity as a function of a momentum, p, that is small compared to some high scale, Λ_b.

- EFT expansion for quantity is $g(x) = \sum_{i=0}^{k} \mathcal{A}_i(x) x^i$ $\qquad x = \frac{p}{\Lambda_b}$

$$\mathcal{A}_i(x) = c_i(\mu) + f_i(x, \mu) \qquad c_i, f_i = \mathcal{O}(1) \text{ for } \mu \sim \Lambda_b, x \sim 1$$

A Generic EFT

- Suppose we are interested in a quantity as a function of a momentum, p, that is small compared to some high scale, $\Lambda_{b.}$
- EFT expansion for quantity is $g(x) = \sum_{i=0}^{k} \mathcal{A}_i(x) x^i$ $\quad x = \frac{p}{\Lambda_b}$

$$\mathcal{A}_i(x) = c_i(\mu) + f_i(x, \mu) \qquad c_i, f_i = \mathcal{O}(1) \text{ for } \mu \sim \Lambda_b, x \sim 1$$

- $f_i(x, \mu)$ is a calculable function, that encodes IR physics at order i
- c_i is a low-energy constant (LEC): encodes UV physics at order i. Must be fit to data
- Complications: multiple light scales, multiple functions at a given order, skipped orders, ….

Bayesian tools

Thomas Bayes (1701?-1761)

http://www.bayesian-inference.com

$$\mathrm{pr}(A|B,I) = \frac{\mathrm{pr}(B|A,I)\mathrm{pr}(A|I)}{\mathrm{pr}(B|I)}$$

Bayesian tools

Thomas Bayes (1701?-1761)

http://www.bayesian-inference.com

$$\mathrm{pr}(A|B,I) = \frac{\mathrm{pr}(B|A,I)\mathrm{pr}(A|I)}{\mathrm{pr}(B|I)}$$

$$\mathrm{pr}(x|\mathrm{data},I) = \frac{\mathrm{pr}(\mathrm{data}|x,I)\mathrm{pr}(x|I)}{\mathrm{pr}(\mathrm{data}|I)}$$

Bayesian tools

Thomas Bayes (1701?-1761)

http://www.bayesian-inference.com

Probability as degree of belief

$$\mathrm{pr}(A|B,I) = \frac{\mathrm{pr}(B|A,I)\mathrm{pr}(A|I)}{\mathrm{pr}(B|I)}$$

$$\mathrm{pr}(x|\mathrm{data},I) = \frac{\mathrm{pr}(\mathrm{data}|x,I)\mathrm{pr}(x|I)}{\mathrm{pr}(\mathrm{data}|I)}$$

Bayesian tools

Thomas Bayes (1701?-1761)

http://www.bayesian-inference.com

Probability as degree of belief

$$\mathrm{pr}(A|B,I) = \frac{\mathrm{pr}(B|A,I)\mathrm{pr}(A|I)}{\mathrm{pr}(B|I)}$$

Likelihood

↓

$$\mathrm{pr}(x|\mathrm{data},I) = \frac{\mathrm{pr}(\mathrm{data}|x,I)\mathrm{pr}(x|I)}{\mathrm{pr}(\mathrm{data}|I)}$$

Bayesian tools

Thomas Bayes (1701?-1761)

http://www.bayesian-inference.com

Probability as degree of belief

$$\mathrm{pr}(A|B,I) = \frac{\mathrm{pr}(B|A,I)\mathrm{pr}(A|I)}{\mathrm{pr}(B|I)}$$

Likelihood

Prior

$$\mathrm{pr}(x|\mathrm{data},I) = \frac{\mathrm{pr}(\mathrm{data}|x,I)\mathrm{pr}(x|I)}{\mathrm{pr}(\mathrm{data}|I)}$$

Bayesian tools

Thomas Bayes (1701?-1761)

http://www.bayesian-inference.com

Probability as degree of belief

$$\mathrm{pr}(A|B,I) = \frac{\mathrm{pr}(B|A,I)\mathrm{pr}(A|I)}{\mathrm{pr}(B|I)}$$

Likelihood   Prior

$$\mathrm{pr}(x|\mathrm{data},I) = \frac{\mathrm{pr}(\mathrm{data}|x,I)\mathrm{pr}(x|I)}{\mathrm{pr}(\mathrm{data}|I)}$$

Posterior

Bayesian tools

Thomas Bayes (1701?-1761)

http://www.bayesian-inference.com

Probability as degree of belief

$$\mathrm{pr}(A|B,I) = \frac{\mathrm{pr}(B|A,I)\mathrm{pr}(A|I)}{\mathrm{pr}(B|I)}$$

Likelihood

Prior

$$\mathrm{pr}(x|\mathrm{data},I) = \frac{\mathrm{pr}(\mathrm{data}|x,I)\mathrm{pr}(x|I)}{\mathrm{pr}(\mathrm{data}|I)}$$

Posterior

Normalization

Bayesian tools

Thomas Bayes (1701?-1761)

http://www.bayesian-inference.com

Probability as degree of belief

$$\mathrm{pr}(A|B,I) = \frac{\mathrm{pr}(B|A,I)\mathrm{pr}(A|I)}{\mathrm{pr}(B|I)}$$

Likelihood

Prior

$$\mathrm{pr}(x|\mathrm{data},I) = \frac{\mathrm{pr}(\mathrm{data}|x,I)\mathrm{pr}(x|I)}{\mathrm{pr}(\mathrm{data}|I)}$$

Posterior

Normalization

Marginalization: $$\mathrm{pr}(x|\mathrm{data},I) = \int dy\,\mathrm{pr}(x,y|\mathrm{data},I)$$

Bayesian tools

Thomas Bayes (1701?-1761)

http://www.bayesian-inference.com

Probability as degree of belief

$$\mathrm{pr}(A|B,I) = \frac{\mathrm{pr}(B|A,I)\mathrm{pr}(A|I)}{\mathrm{pr}(B|I)}$$

Likelihood

Prior

$$\mathrm{pr}(x|\mathrm{data},I) = \frac{\mathrm{pr}(\mathrm{data}|x,I)\mathrm{pr}(x|I)}{\mathrm{pr}(\mathrm{data}|I)}$$

Posterior

Normalization

Marginalization: $$\mathrm{pr}(x|\mathrm{data},I) = \int dy\, \mathrm{pr}(x,y|\mathrm{data},I)$$

Allows us to integrate out "nuisance" (e.g. higher-order) parameters