# Obtaining quark self-energy from chiral perturbation theory

蒋绍周

广西大学 物理科学与工程技术学院

2024.10.21 湖南长沙

# Outline

- Background & motivations
- Bayesian statistics & MCMC
- Tests & experience conclusions
- Results
- Summary

# Background & motivation

- Theory: SDE + $\alpha_s \to$ quark self-energy $\Sigma \to$ LECs

  Yang, H. et al., PRD **66**, 014019 (2002).
  Jiang, S.-Z. et al., PRD **81**, 014001 (2010).
  Jiang, S.-Z. et al., PRD, **81**, 094037, (2010).
  Jiang, S.-Z. et al., PRD, **87**, 094014, (2013).
  Jiang, S.-Z. et al., PRD, **92**, 025014, (2015).

- $\alpha_s \leftarrow$ gluon propagator $\leftarrow$ BSE

  Aoki, K.-I. et al., Prog. Theor. Phys., **84**, 683, (1990).
  Maris, P. et al., PRC, **60**, 055214, (1999).

- Experiment: Global fit $\to$ LECs

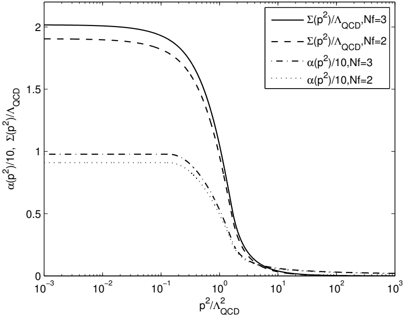
  Amorós, G. et al., NPB, **585**, 293, (2000).
  Amoros, G. et al., NPB, **602**, 87, (2001).
  Bijnens, J. et al., Nucl. Phys. B, **854**, 631, (2012).
  Bijnens, J. et al., Ann. Rev. Nucl. Part. Sci., **64**, 149, (2014),
  Yang, Q.-H. et al., PRD, **102**, 094009, (2020),
  Pan, H.-X. et al., Front. Phys. **19**, 64203, (2024).

- AI in physics

  Experimental data analysis, theoretical simulation, ...

# Background & motivation

- Direct problem: quark self-energy $\Sigma(p^2) \rightarrow$ LECs relatively simple




$$\xrightarrow{\int} \begin{cases} L_1 = 0.92^{+0.03}_{-0.04} \\ L_2 = 1.84^{+0.05}_{-0.08} \\ \ldots \\ L_{10} = -5.43^{-0.29}_{+0.44} \end{cases}$$

$$L_i = \int f_i(q^2, \Sigma(q^2), \Sigma'(q^2)) d^4q$$

- Inverse problem: LECs $\rightarrow$ low-energy $\Sigma(p^2)$ and its error

  solvability? uniqueness? efficiency?

- Fit a curve $\rightarrow$ test out theories

# Background & motivation

- Bayesian statistics has already achieved a great success in artificial intelligence AlphaGo, ChatGPT, ...

- MCMC Algorithm:

  Metropolis-Hastings algorithm J. Chem. Phys., **21**, 1087, (1953). Biometrika, **57**, 97, (1970)

  Hamiltonian Monte Carlo PLB **195**, 216-222 (1987)

  Adaptive Metropolis Algorithm Bernoulli, 7, (2), 223, (2001)

  No-U-turn Sampler J. Mach. Learn. Res., **15**, 1593, (2011)

- Bayesian statistics has been adopted in physics
  - EFT: Annals Phys. **324**, 682-708 (2009)
  - EFT with truncation errors: PRC **96**, 024003 (2017), PRC **105**, 014004 (2022) JPG **46**, 095101 (2019), JPG **46**, 045102 (2019)
  - Nuclear physics: JPG **43**, 074001 (2016)
  - MCMC algorithm: Hamiltonian Monte Carlo PLB **195**, 216-222 (1987)

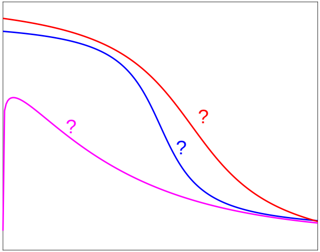
# Bayesian statistics

- Bayes formula

$$\mathrm{pr}(\Sigma|L_i) = \frac{\mathrm{pr}(L_i|\Sigma)\,\mathrm{pr}(\Sigma)}{\mathrm{pr}(L_i)} \propto \mathrm{pr}(L_i|\Sigma)\,\mathrm{pr}(\Sigma)$$

| | | | |
|---|---|---|---|
| $\Sigma$: | quark self-energy | $L_i$: | LECs |
| $\mathrm{pr}(\Sigma)$: | prior PDF | $\mathrm{pr}(L_i\|\Sigma)$: | likelihood PDF |
| $\mathrm{pr}(\Sigma\|L_i)$: | posterior PDF | $\mathrm{pr}(L_i)$: | Bayesian evidence |

- How to solve $\mathrm{pr}(\Sigma)$? Solutions may not be uniqueness



# MCMC with Metropolis-Hasting algorithm

- Discretizate $\Sigma(p^2) \to \Sigma(p_j^2) = \Sigma^{(j)}, \; j = 1, 2, \cdots, d$

- Markov Chain $\Sigma_i = \Sigma_i^{(j)}$ all $j$: $(\Sigma_1, \Sigma_2, \Sigma_3, ..., \Sigma_N)$, $g(\Sigma_{i+1}|\Sigma_i)$

  $i$: chain number

- Stationary distribution $\Rightarrow$ Detailed balance condition

$$Q(\Sigma_m|\Sigma_n)\,\mathrm{pr}(\Sigma_n) = Q(\Sigma_n|\Sigma_m)\,\mathrm{pr}(\Sigma_m)$$

  $Q(\Sigma_m|\Sigma_n)$: transition matrix, $m, n$: position

- How to determine $Q(\Sigma_m|\Sigma_n)$?

- Generating $\Sigma_i$, not determining $Q(\Sigma_m|\Sigma_n)$ directly

  Metropolis-Hasting, Hamiltonian Monte Carlo,

  Adaptive Metropolis, No-U-turn Sampler, ...

# MCMC with Metropolis-Hasting algorithm

1. Generating $\Sigma_1$ any or special function
2. Generating $\Sigma'_{i+1}$ by $\Sigma_i$, $\mathrm{pr}(\Sigma'_{i+1}|\Sigma_i) = N(\Sigma_i, \sigma^2_{\Sigma_i})$
3. Receiving probability $\alpha$, $\Sigma_{i+1} = \Sigma'_{i+1}$ $(\alpha)$, $\Sigma_{i+1} = \Sigma_i$ $(1-\alpha)$

$$\alpha = \min\left(\frac{Q'(\Sigma_i|\Sigma'_{i+1})}{Q'(\Sigma'_{i+1}|\Sigma_i)}\frac{\mathrm{pr}(\Sigma'_{i+1}|L_i)}{\mathrm{pr}(\Sigma_i|L_i)}, 1\right) = \min\left(\frac{\mathrm{pr}(L_i|\Sigma'_{i+1})\,\mathrm{pr}(\Sigma'_{i+1})}{\mathrm{pr}(L_i|\Sigma_i)\,\mathrm{pr}(\Sigma_i)}, 1\right)$$

$Q'$: guessing $Q$, $Q'(\Sigma_i|\Sigma'_{i+1}) = Q'(\Sigma'_{i+1}|\Sigma_i)$

$$\mathrm{pr}(L_i|\Sigma) = \exp\left\{-\frac{1}{2}\left[L_i^{\mathrm{th}} - L_i\right]^T (\Sigma^2_{L_i})^{-1}\left[L_i^{\mathrm{th}} - L_i\right]\right\}$$

4. Going to Step 2, until $N$ times to get Markov Chain $\Sigma_i$
5. Result: mean value $\pm$ 68% highest posterior density

Different algorithms obtain similar results

# Adaptive Metropolis algorithm

- $\mathrm{pr}(\Sigma'_{i+1}|\Sigma_i) = N(\Sigma_i, \sigma^2_{\Sigma_i})$ may not be smooth, introduce covariance matrix

$$C_{ij}^{(0)} = \sigma^2 \exp\left(-\frac{p_i^2 - p_j^2}{2l}\right) + \epsilon\delta_{ij}$$

$\sigma$ controls accept, $l$ controls vibration. $\epsilon$ makes $C_{ij}$ positive, smaller is better

- MH algorithm often has low acceptance rate
- Adaptive Metropolis (AM) algorithm is better

$$\boldsymbol{C}_k = \begin{cases} C^{(0)} & k \leqslant m \\ s_d \operatorname{cov}(\Sigma_0, \cdots, \Sigma_{k-1}) + s_d \varepsilon I_d & k > m \end{cases}$$

$$\operatorname{cov}(\Sigma_0, \cdots, \Sigma_{k-1}) = \frac{1}{k}\left(\sum_{i=0}^{k} \Sigma_i^\top \Sigma_i - (k+1)\bar{\Sigma}_i^\top \bar{\Sigma}_i\right)$$

$\bar{\Sigma}_i = \sum_{i=0}^{k} \Sigma_i/(k+1)$, $s_d = 2.4^2/d$, $\varepsilon$ small, $I_d$ $d \times d$ unit matrix

- In most cases, acceptance rate, AM $> \times 10$ MH

# Tests

# Tests functions

$$y_1 = \int_0^1 g(x) f(x) dx \qquad y_6 = \int_0^1 g(x) f(x) f'(x) x dx$$

$$y_2 = \int_0^1 g(x) f^2(x) dx \qquad y_7 = \int_0^1 g(x) f'(x) f''(x) x dx$$

$$y_3 = \int_0^1 g(x) f(x) f''(x) dx \qquad y_8 = \int_0^1 g(x) \frac{f(x) x}{(x + f(x))^2} dx$$

$$y_4 = \int_0^1 g(x) \frac{f^2(x)}{x + f(x)} dx \qquad y_9 = \int_0^1 g(x) \frac{f(x) f'(x) x}{x + f(x)} dx$$

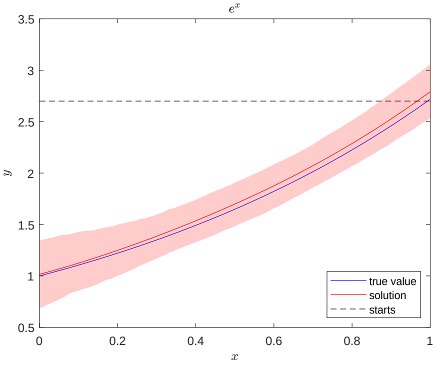
$$y_5 = \int_0^1 g(x) \frac{f(x)^2}{(x + f(x))^2} dx \qquad y_{10} = \int_0^1 g(x) \frac{f^3(x)}{(x + f(x))^2} dx$$

$$g(x) = \frac{x}{x + f(x)} e^{-\frac{x + f(x)}{\Lambda}}$$

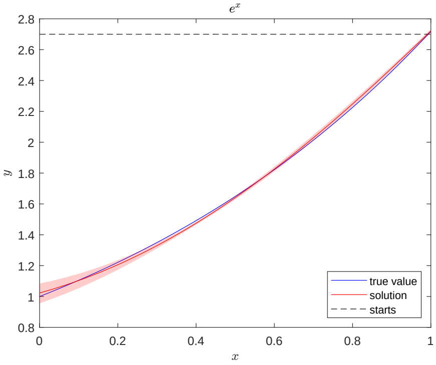
similar to the actual situation Yang, H. et al., PRD **66**, 014019 (2002).

# Some tests: $e^x$

## MH algorithm



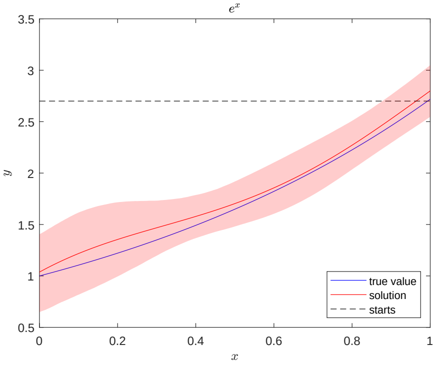

Large errors



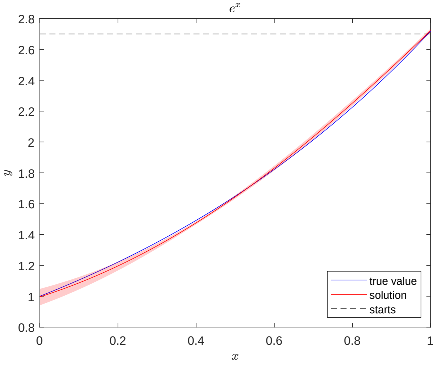

Small errors

# Some tests: $e^x$

## AM algorithm



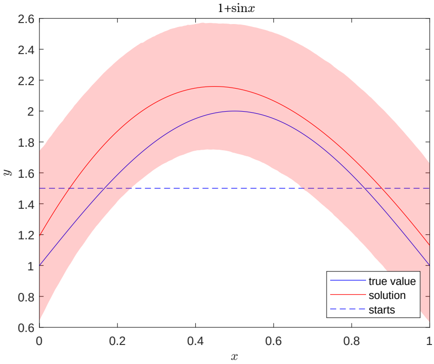

Large errors



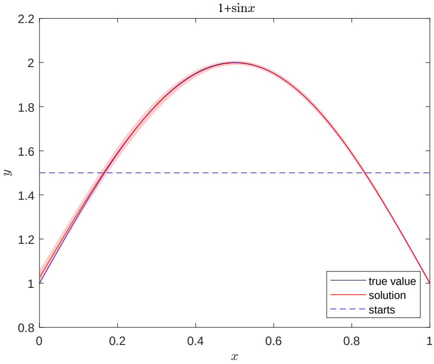

Small errors

# Some tests: $1 + \sin x$

## MH algorithm, semiperiod



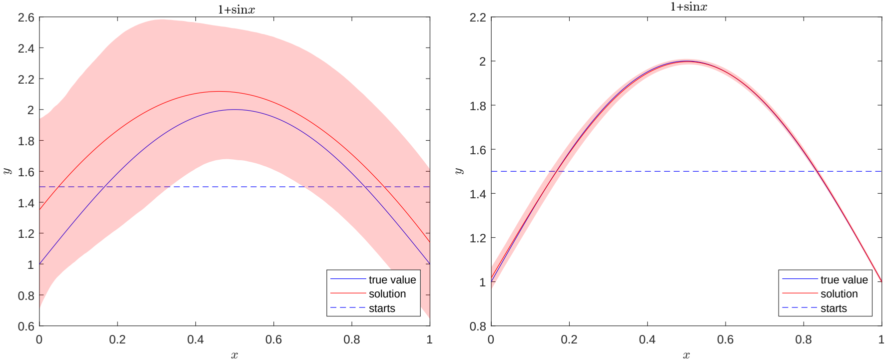

Large errors




Small errors

# Some tests: $1 + \sin x$

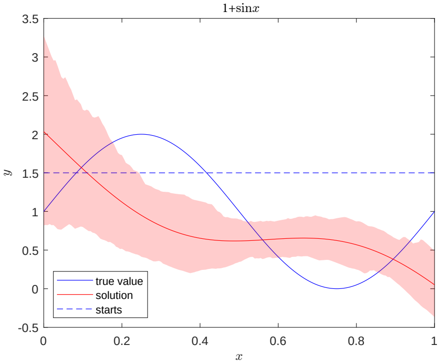
## AH algorithm, semiperiod



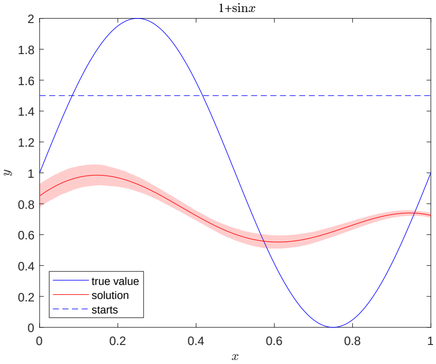

Large errors　　　　Small errors

# Some tests: $1 + \sin x$

## MH algorithm, one period



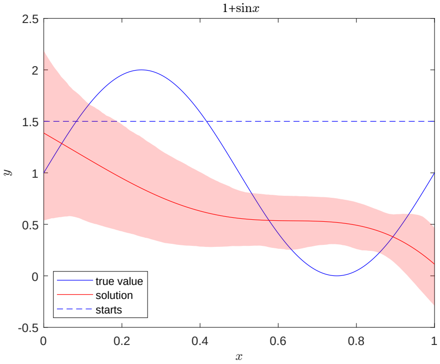

Large errors



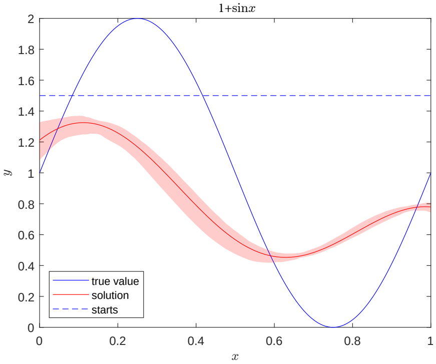

Small errors

# Some tests: $1 + \sin x$

## AH algorithm, one period




Large errors




Small errors

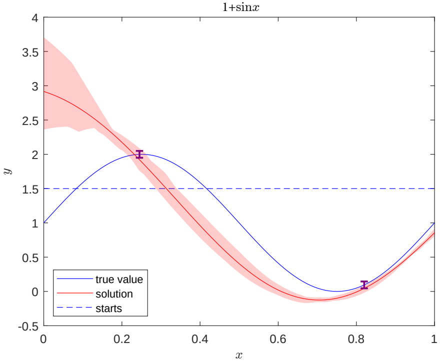
# Experience conclusions

$$C_{ij}^{(0)} = \sigma^2 \exp\left(-\frac{p_i^2 - p_j^2}{2l}\right) + \epsilon\delta_{ij}$$

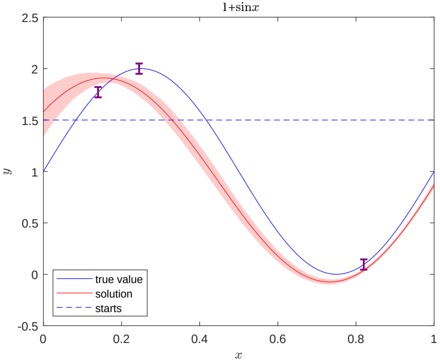
- Large $l \rightarrow$ small oscillation, small $l \rightarrow$ large oscillation
- $\sigma$ effects convergence, smaller $\sigma$ lead higher accept ratio
- One-oscillation curve can be fitted, two-oscillation curve is hard to be fitted
- In actual case, some properties of curve are known, i.e. prior

# Tests with prior

## MH algorithm, one period



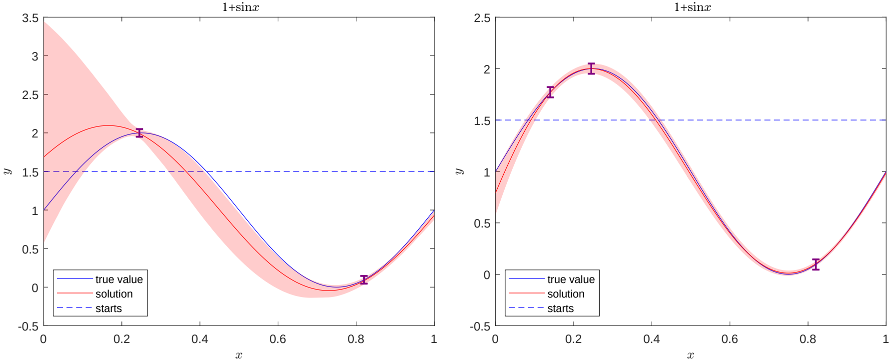

Two points




Three points

# Tests with prior

## AH algorithm, one period




Two points

Three points

# Tests with prior

- Bad Fit come from the local extreme point
- Existing bad Fit is not a disadvantage
- Removing most possible shapes
- Choosing a special shape
- Which is the best? Not very large $\chi^2$

$$\chi^2 = \sum_i \left( \frac{y_i^{\mathrm{th}} - y_i}{\Delta y_i} \right)^2$$

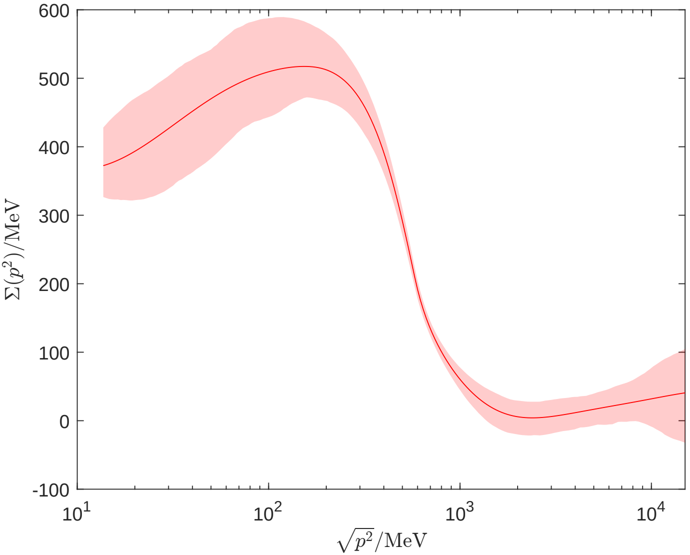
# For $\Sigma(p^2)$

- Analytical formulas are not exact, but also not very bad
- Approximate formulas + approximate LECs
  $\rightarrow$ approximate $\Sigma(p^2)$
- Formula errors + LECs errors $\rightarrow$ $\Sigma(p^2)$ errors
  a main motivation
- Prior: high-energy behavior is known theory or experiment

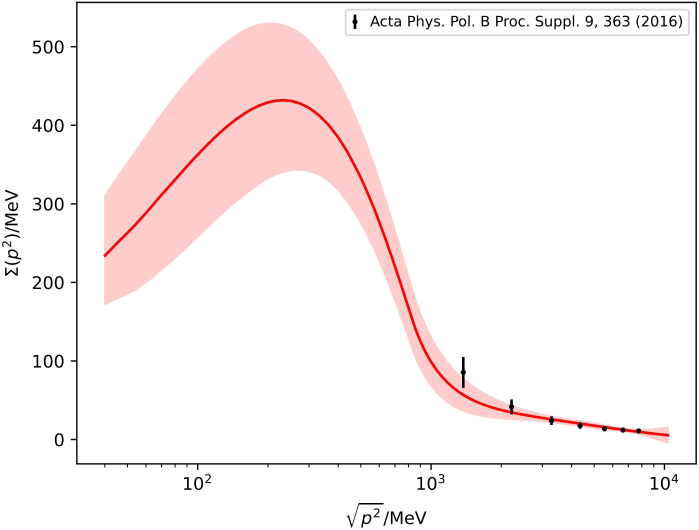
$$\Sigma(p^2) \xrightarrow{p^2\to\infty} C\frac{\ln^{\gamma-1}(p^2/\Lambda^2_{\text{QCD}})}{p^2}, \quad \gamma = \frac{9N_C}{33-2N_f}$$

  No violent oscillation
- Unknown shape needs be explored
- LECs come from Pan, H.-X. et al., Front. Phys. 19, 64203, (2024)

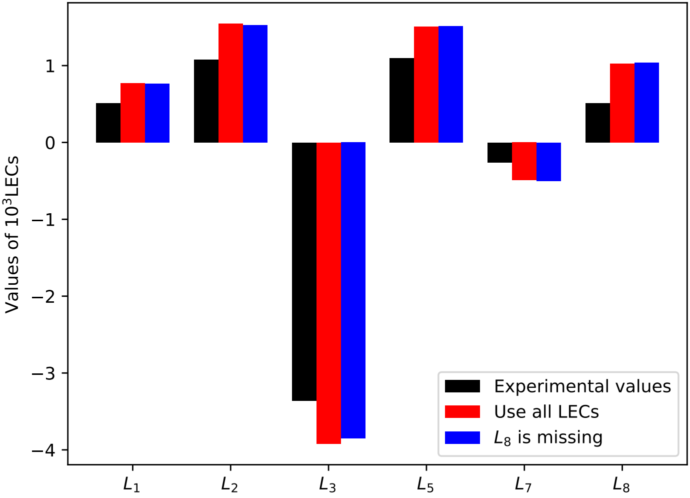
# Results of $\Sigma(p^2)$: free fitting



# Results of $\Sigma(p^2)$: with prior



# Prediction



Values of $10^3$LECs
1
0
−1
−2
−3
−4
$L_1$
$L_2$
$L_3$
$L_5$
$L_7$
$L_8$
Experimental values
Use all LECs
$L_8$ is missing

# Discussion & Summary

- Free fitting: error range at the low and high energy is large
- A peak at the low energy
- Providing an approach to fit curves
- Obtaining the quark self-energy with errors
- More information, more precise

# Thank you!