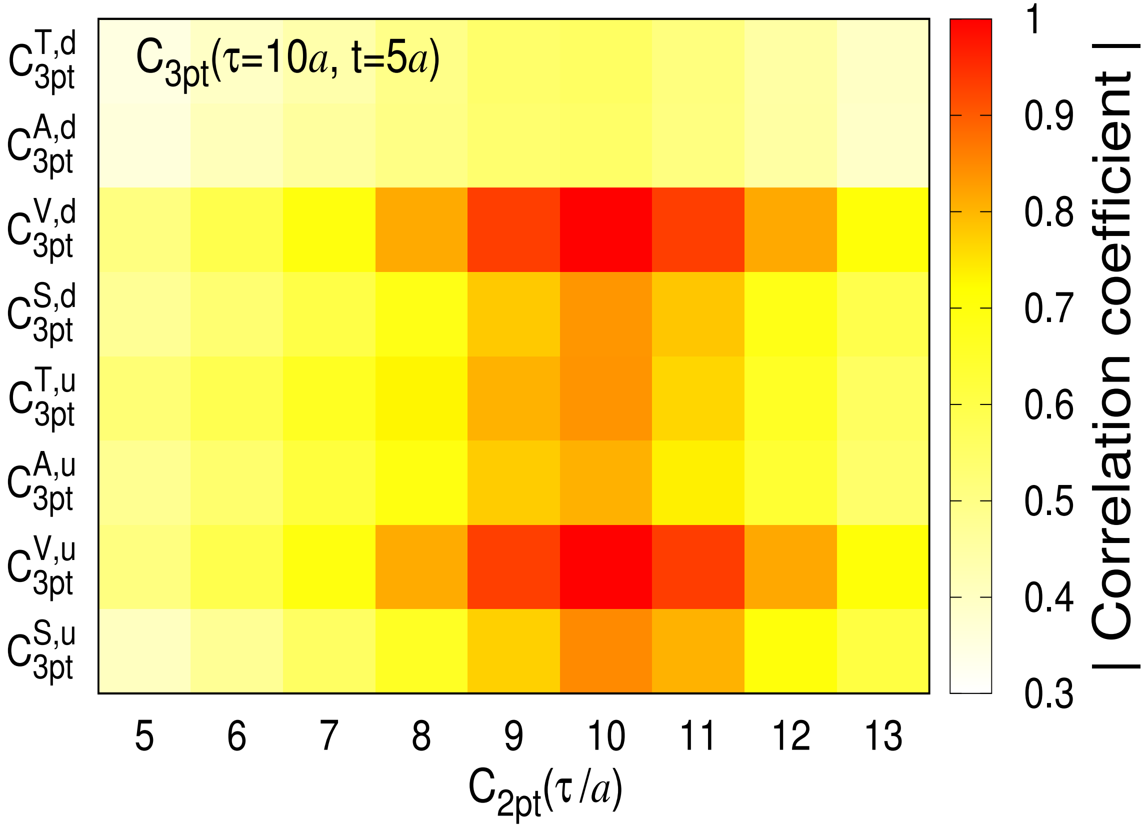
Calculation of Lattice QCD Observables using Machine Learning

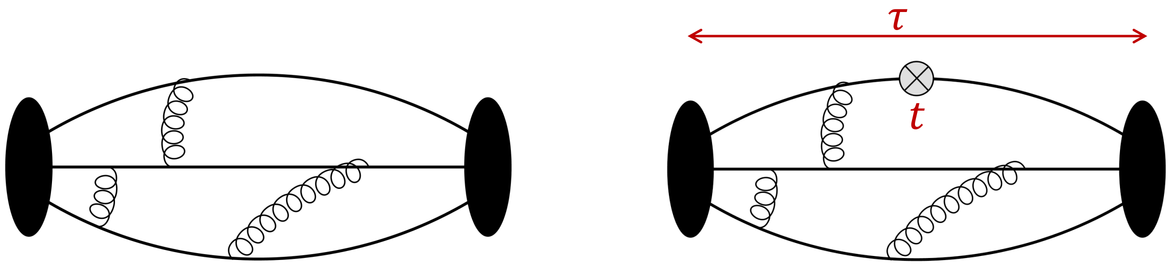
Boram Yoon
in collaboration with
Tanmoy Bhattacharya and Rajan Gupta

Los Alamos National Laboratory

Correlation Map of Nucleon Observables



- Correlation between proton(uud) 3-pt and 2-pt correlation functions

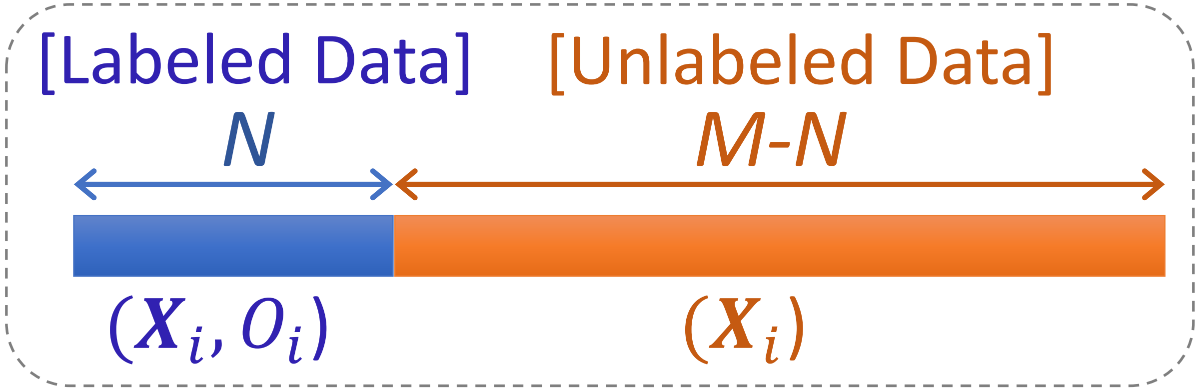


$C_{2pt} \sim \langle N(\tau)N^{\dagger}(0)\rangle$ $C_{3pt}^{A,S,T,V} \sim \langle N(\tau)O(t)N^{\dagger}(0)\rangle$

- Using these correlations, C_{3pt} can be estimated from C_{2pt} on each configuration

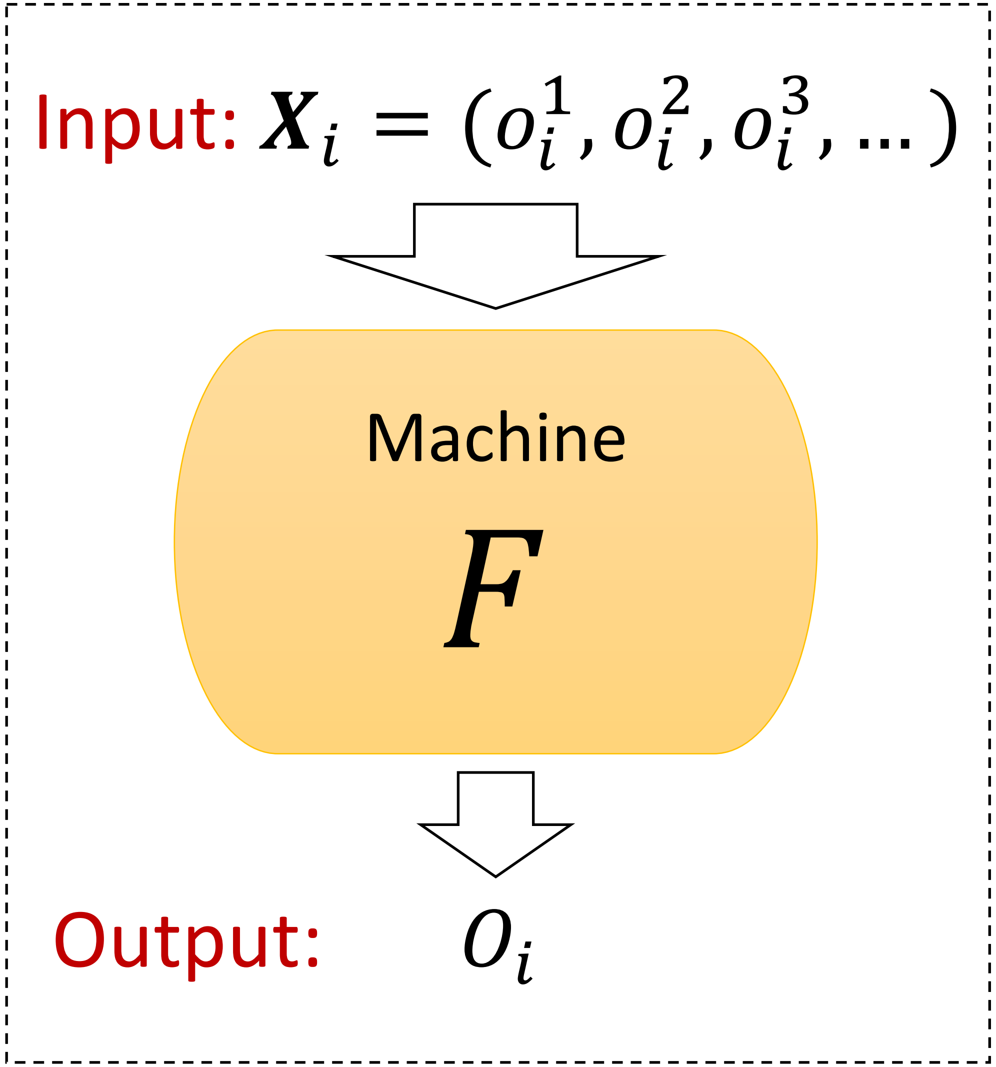
Machine Learning on Lattice QCD Observables

- Assume *M* indep. measurements
- Common observables $\boldsymbol{X}_i$ on all *M*
  Target observable O_i on first *N*



1) **Train** machine ***F*** to yield O_i from $\boldsymbol{X}_i$ on the Labeled Data
2) **Predict** O_i of the Unlabeled data from $\boldsymbol{X}_i$

$$\boldsymbol{F}(\boldsymbol{X}_i) = O_i^P \approx O_i$$



Simple ML algorithms (1): Linear Regression

- Consider prediction of $C_{3pt}^{A}(10,5)$ from $\{C_{2pt}(0 \leq \tau/a \leq 20)\}$

  Input: $\mathbf{X} = \{C_{2pt}(0),\ C_{2pt}(1),\ C_{2pt}(2),\ \cdots,\ C_{2pt}(20)\}$

  Output: $O = C_{3pt}^{A}(10,5)$

- **Linear regression**:

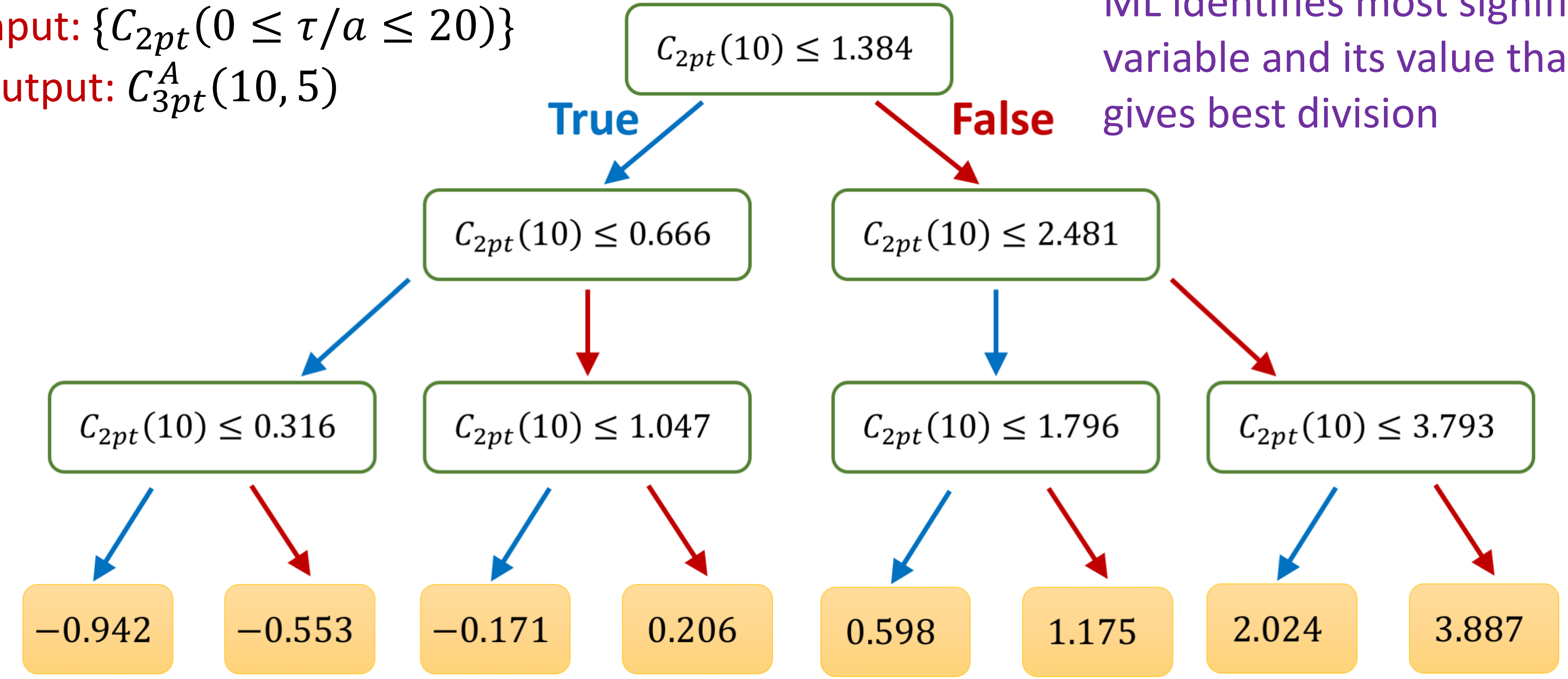
$$C_{3pt}^{A}(10,5)^{\text{prediction}} = b + \sum_{\tau=0}^{20} w_\tau\, C_{2pt}(\tau)$$

ML finds (least-square fit) optimal values of intercept b and coefficients w_τ using the training data

Simple ML algorithms (2): Decision Tree

Input: $\{C_{2pt}(0 \le \tau/a \le 20)\}$
Output: $C^{A}_{3pt}(10, 5)$

ML identifies most significant variable and its value that gives best division



$$C^{A}_{3pt}(\tau/a = 10, t/a = 5)$$

Prediction Bias

- $\boldsymbol{F}(\boldsymbol{X}_i) = O_i^P \approx O_i$
- Simple average

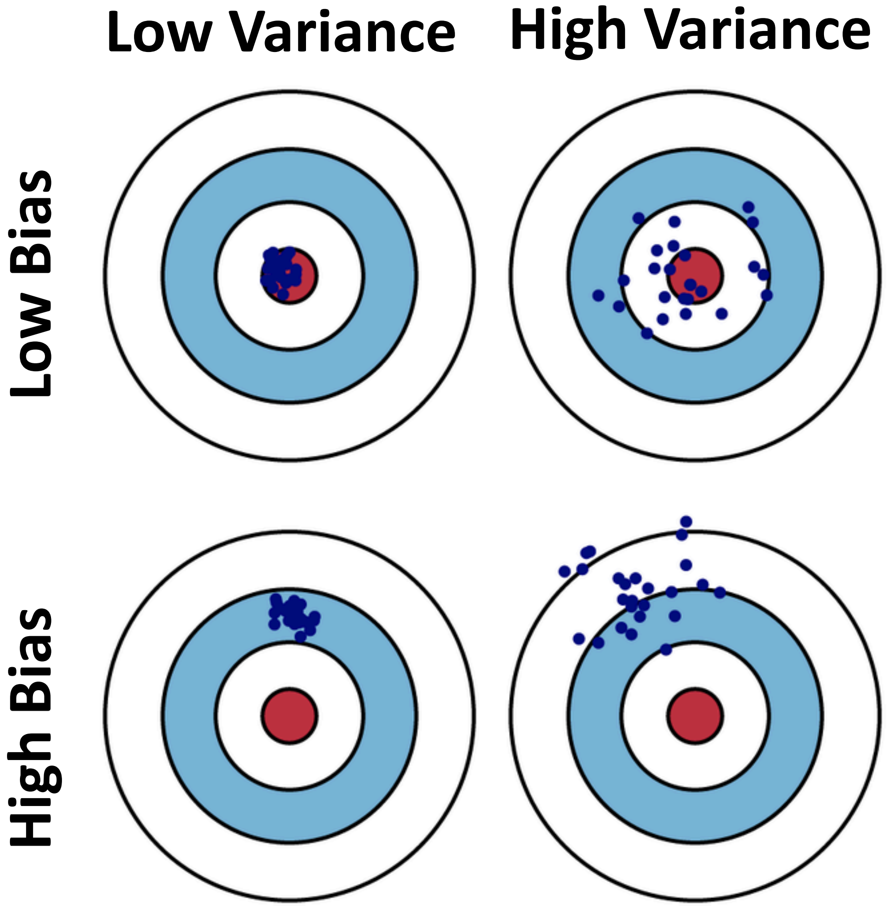
$$\overline{O} = \frac{1}{M-N} \sum_{i \in \text{Unlabeled}} O_i^P$$

is not correct due to prediction bias

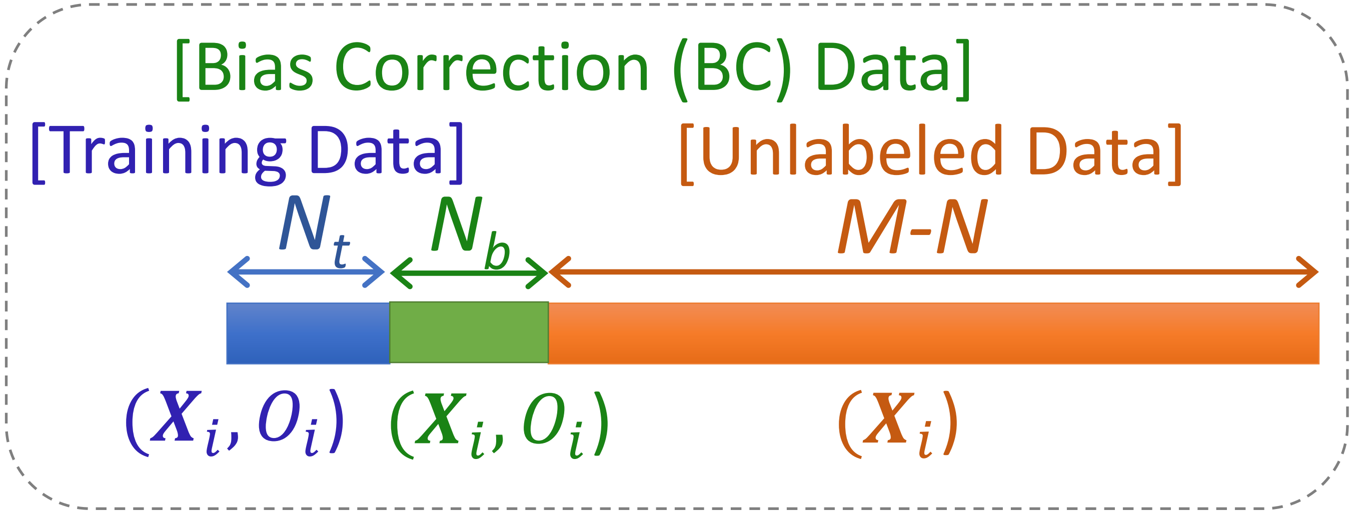
- Prediction = TrueAnswer + Noise + Bias
- ML prediction may have bias

$$\langle O_i^P \rangle \neq \langle O_i \rangle$$

$$\text{Bias} = \langle O_i^P \rangle - \langle O_i \rangle$$



Bias Correction



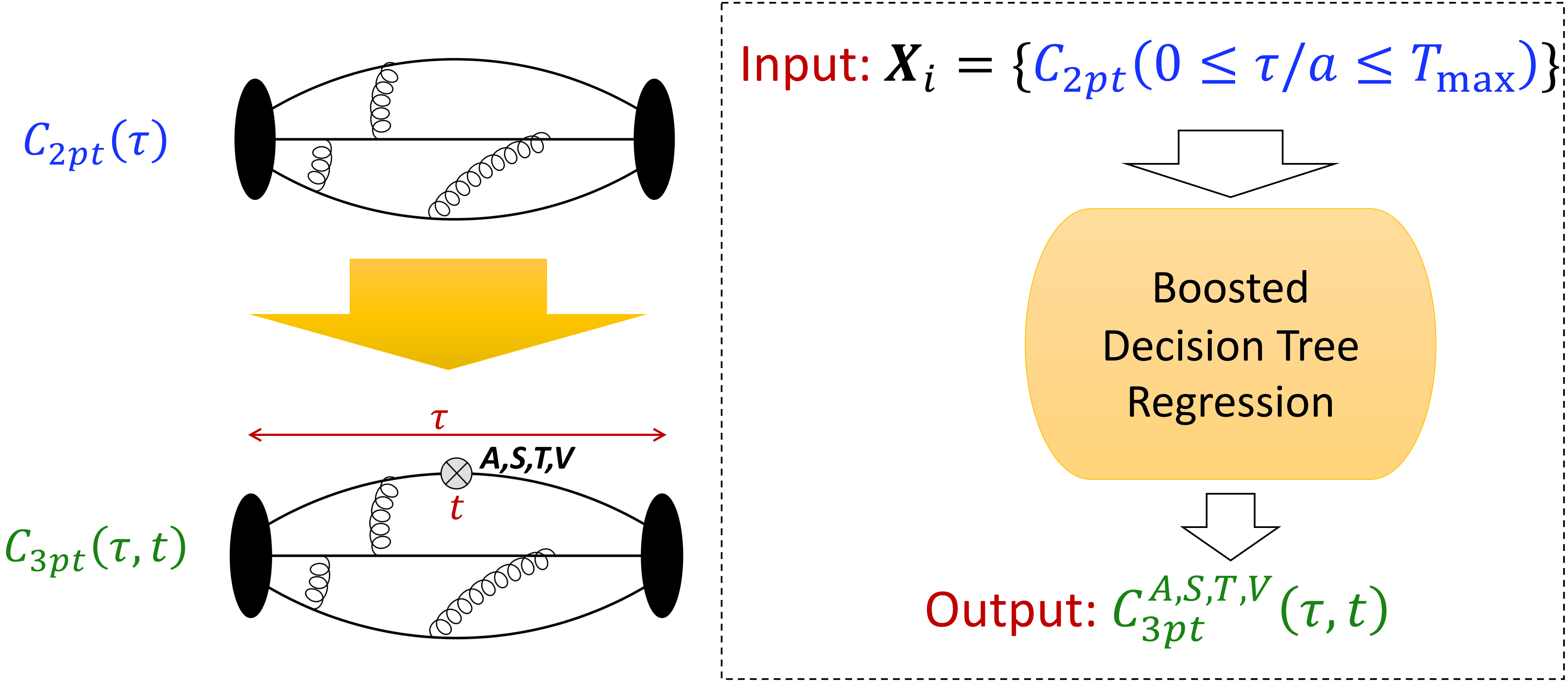
- Split labeled data $N = N_t + N_b$
- Average of predictions on test data with bias correction

$$\overline{O} = \frac{1}{M-N} \sum_{i \in \text{Unlabeled}} O_i^P + \frac{1}{N_b} \sum_{i \in BC} \left(O_i - O_i^P\right)$$

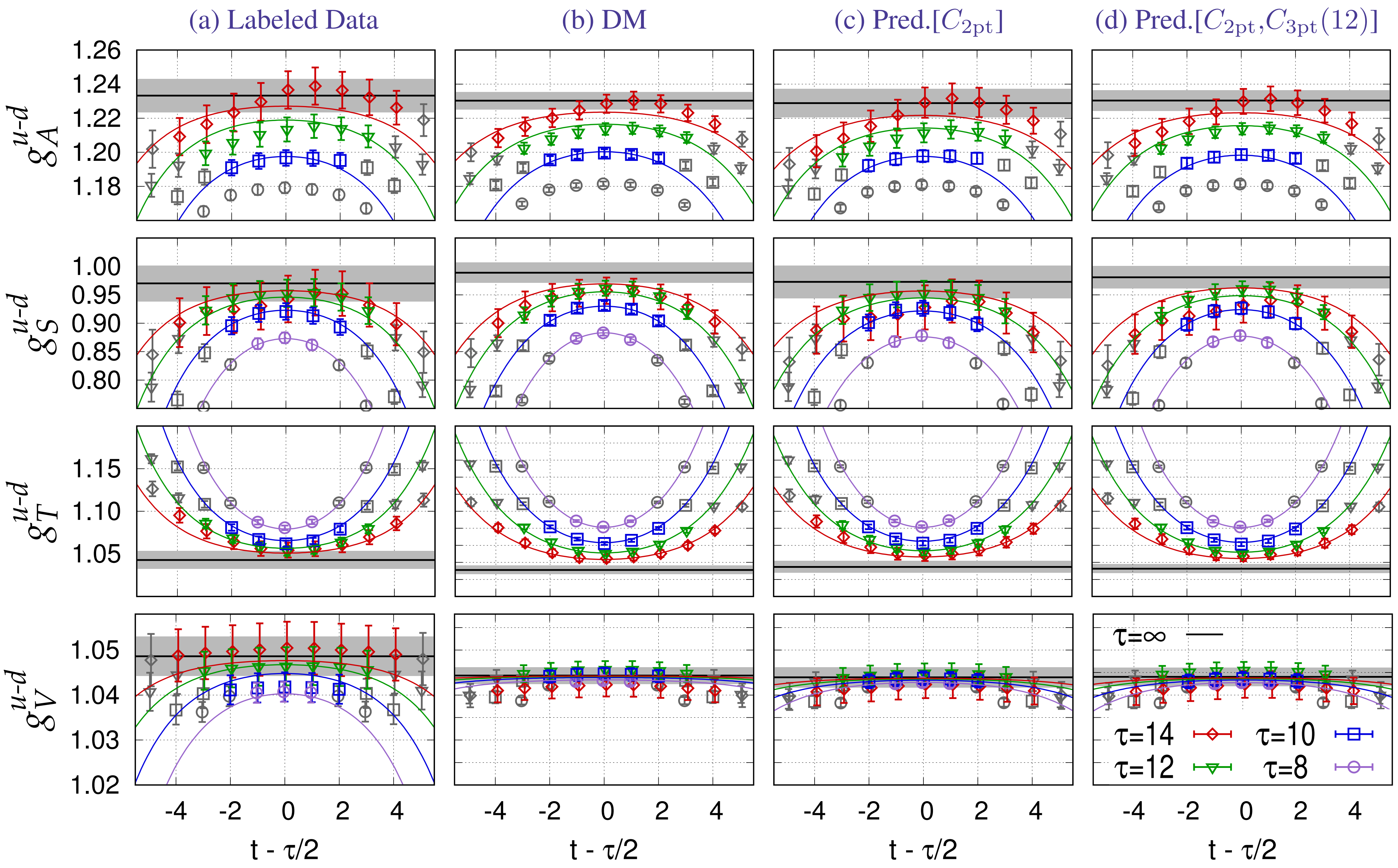
- Expectation value, $\langle \overline{O} \rangle = \langle O_i^P \rangle + \langle O_i - O_i^P \rangle = \langle O_i \rangle$
- BC term converts systematic error of prediction to statistical uncertainty

Applications

Prediction of C_{3pt} from C_{2pt}



Prediction of C_{3pt} from C_{2pt} and more



- *a12m310*
- 2263 confs
  - 680 labeled
  - 1158 unlab.

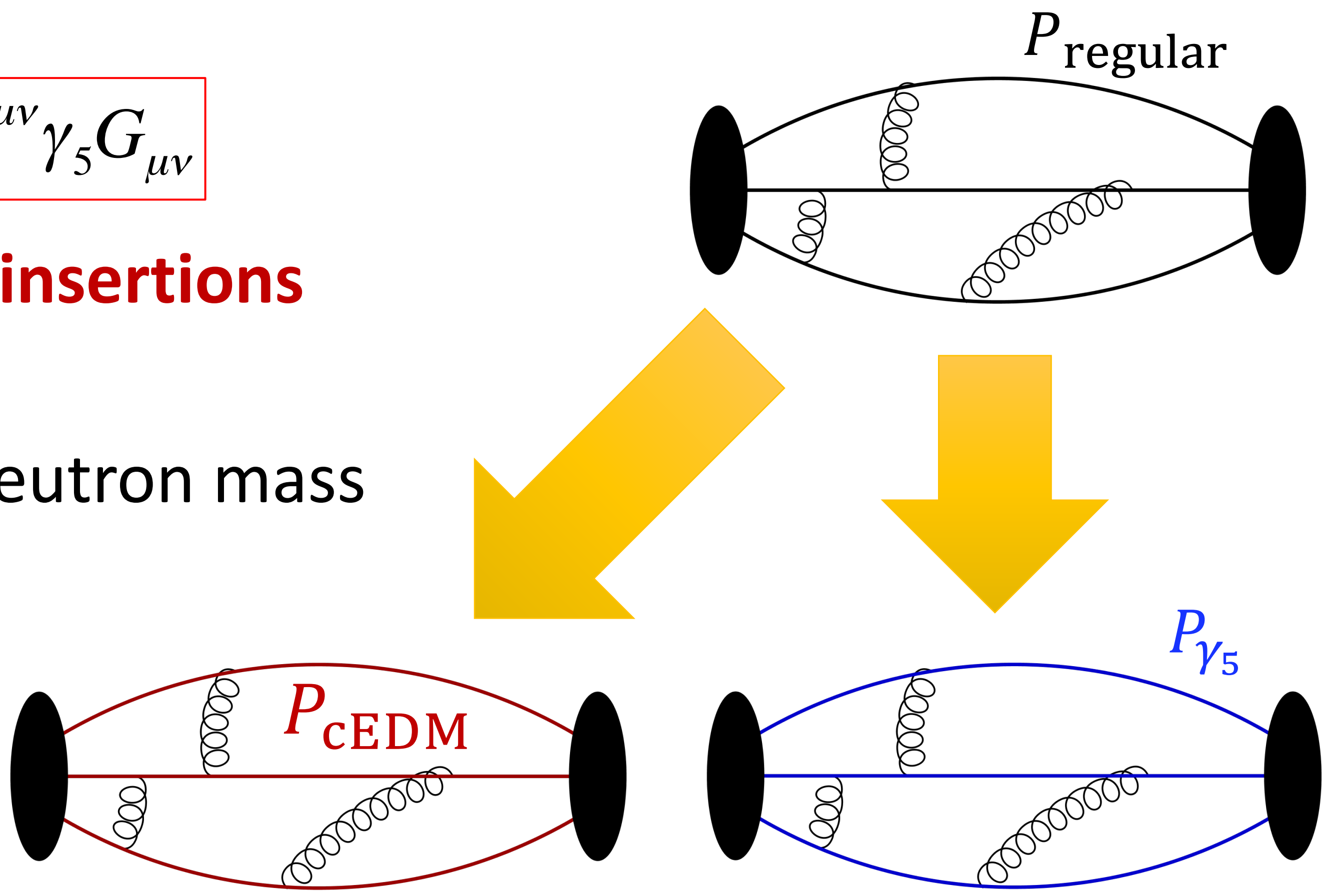
Quark Chromo EDM (cEDM)

- **CP violating interactions** using Schwinger source method: Include cEDM & γ_5 terms in valence quark propagators by modifying Dirac operator

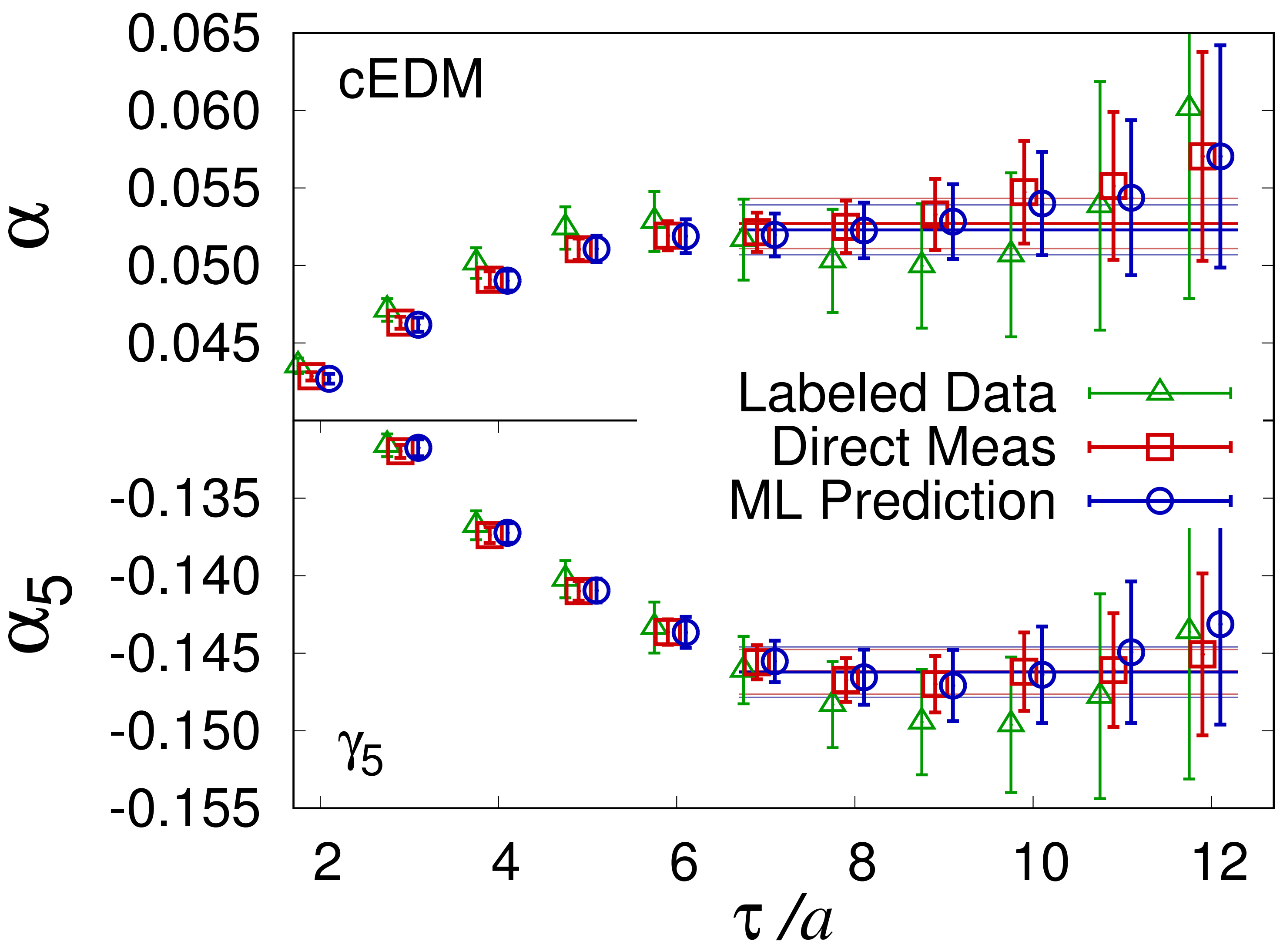
$$D_{\mathrm{clov}} \to D_{\mathrm{clov}} + i\varepsilon\sigma^{\mu\nu}\gamma_5 G_{\mu\nu}$$

- **Predict C_{2pt} for cEDM and γ_5 insertions from C_{2pt} without CPV**
- CPV interactions ➔ phase in neutron mass

$$\left(ip_\mu\gamma_\mu + me^{-2i\alpha\gamma_5}\right)u_N = 0$$



Prediction of C_{2pt}^{CPV} from C_{2pt}



- *a12m310*
  400 confs × 64 srcs

➢ DM:
  DM on 400 confs

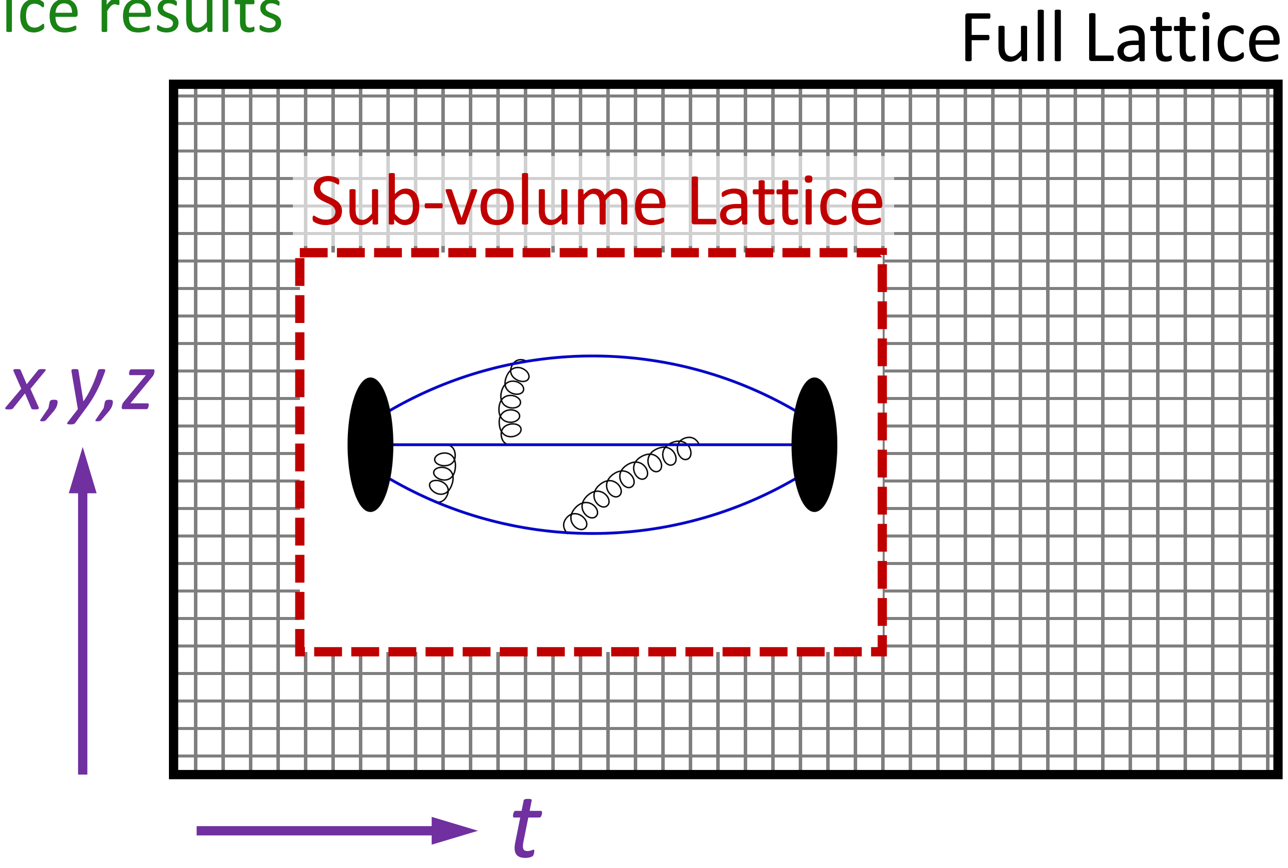
➢ Prediction:
  DM on 120 confs
  + ML pred. on 280 confs

Calculation on Sub-volume Lattice

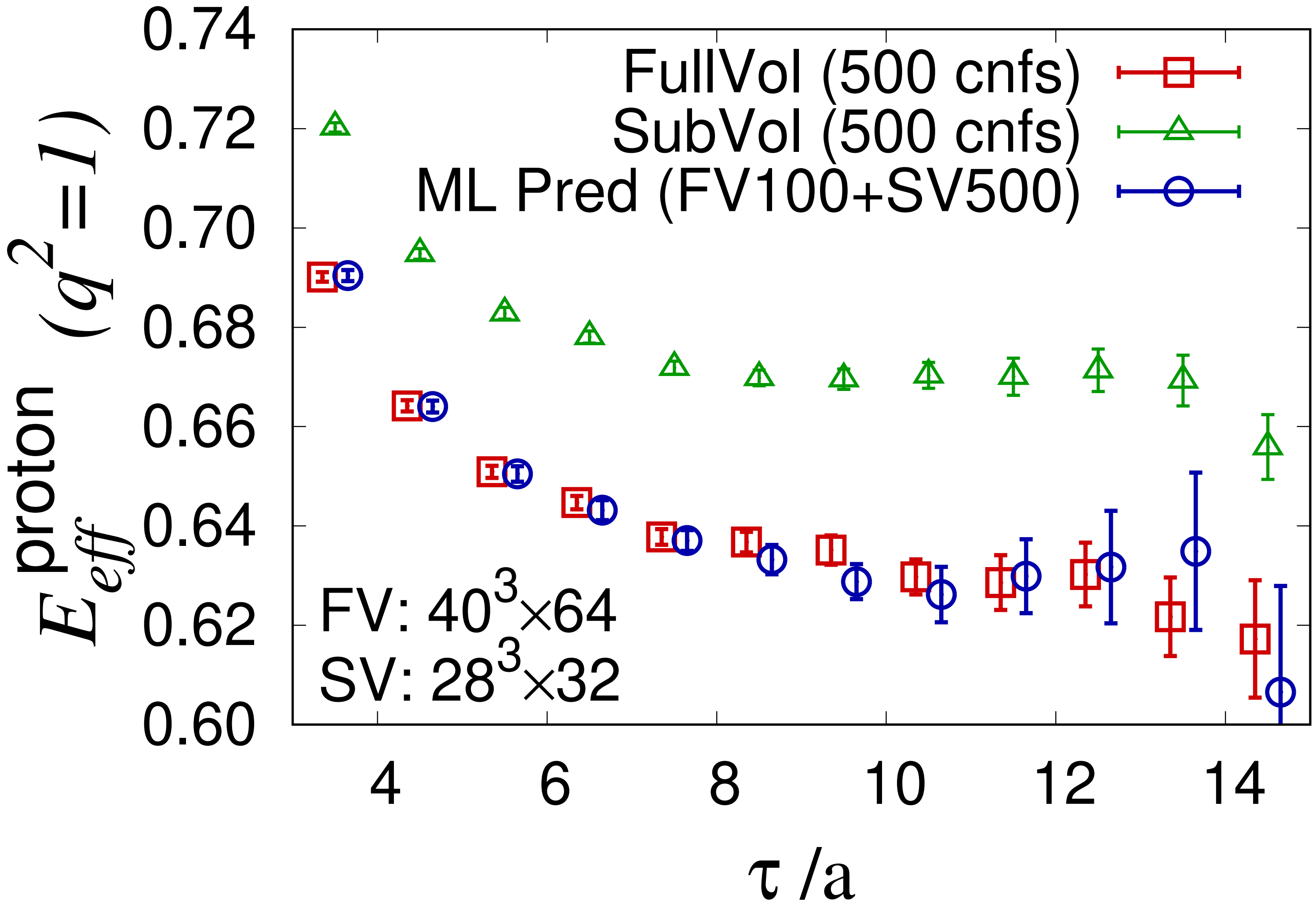
1) Take sub-volume (SV) lattice around C_{2pt} source and sink
2) Carry out all calculations (src/snk smearing, inversion, …) on SV
3) Convert *SV* results to full-lattice results using ML (Linear Regression)

INPUT: C_{2pt} *of Sub-volume*

OUTPUT: C_{2pt} *of Full-Lattice*



Calculation on Sub-volume Lattice



- **a12m220L**
  - $a = 0.12$ fm
  - $M_\pi = 228$ MeV
  - $M_\pi L = 5.49$
  - Vol = $40^3 \times 64$
- **Reduced volume**

  $40^3 \times 64 \rightarrow 28^3 \times 32$

  (reduced to $1/6$)
- **Better condition number**

  BiCG iterations for LP (10^{-4})

  $613 \rightarrow 159$ (reduced to $1/4$)

Data Compression

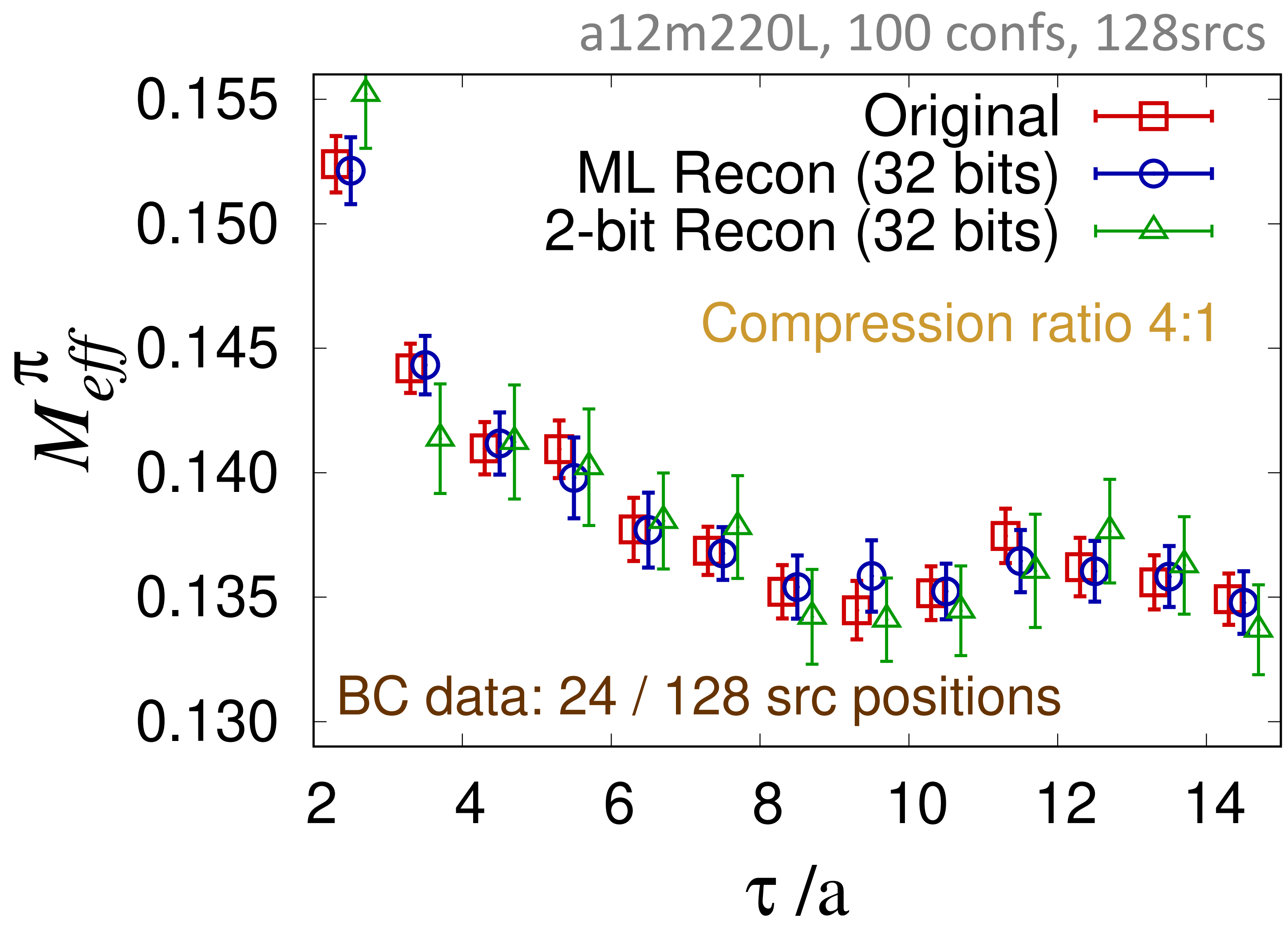
- **Correlated data can be represented by the coefficients of basis vectors**

$$\vec{I}^{(k)} = \sum_i a_i^{(k)} \vec{v}_i \equiv \boldsymbol{\Phi} \vec{a}^{(k)}$$

$$\begin{pmatrix}3\\3\\2\end{pmatrix} = 2\begin{pmatrix}1\\1\\1\end{pmatrix} + 1\begin{pmatrix}1\\1\\-1\end{pmatrix}, \quad \begin{pmatrix}3\\3\\3\end{pmatrix} = 3\begin{pmatrix}1\\1\\1\end{pmatrix} + 0\begin{pmatrix}1\\1\\-1\end{pmatrix}$$

  - $\boldsymbol{\Phi}$: Dictionary (basis vectors) that is common for all data ($\forall k$)
  - $\vec{a}^{(k)}$: Representation of a data vector $\vec{I}^{(k)}$
- **Restrict** $a_i^{(k)}$ = 0 or 1 so that it can be stored in a bit (COMPRESSION)
  - Eg) floating point numbers at 16 timeslices (32×16 bits)
    → $a_{i=1,2,\ldots,32}$ with 32 dictionary vectors (32 bits) + Bias Correction (BC) data
  - Arithmetic and ML can be done on $\vec{a}^{(k)}$ instead of $\vec{I}^{(k)}$
- **Machine learning** finds optimal $\boldsymbol{\Phi}$ and $\vec{a}^{(k)}$ that reconstruct $\vec{I}^{(k)} \approx \boldsymbol{\Phi}\vec{a}^{(k)}$
  - **D-Wave quantum annealer** is used to find $\vec{a}^{(k)}$ efficiently
- **Anomaly detector**: data vectors with large reconstruction error are anomalous

Data Compression

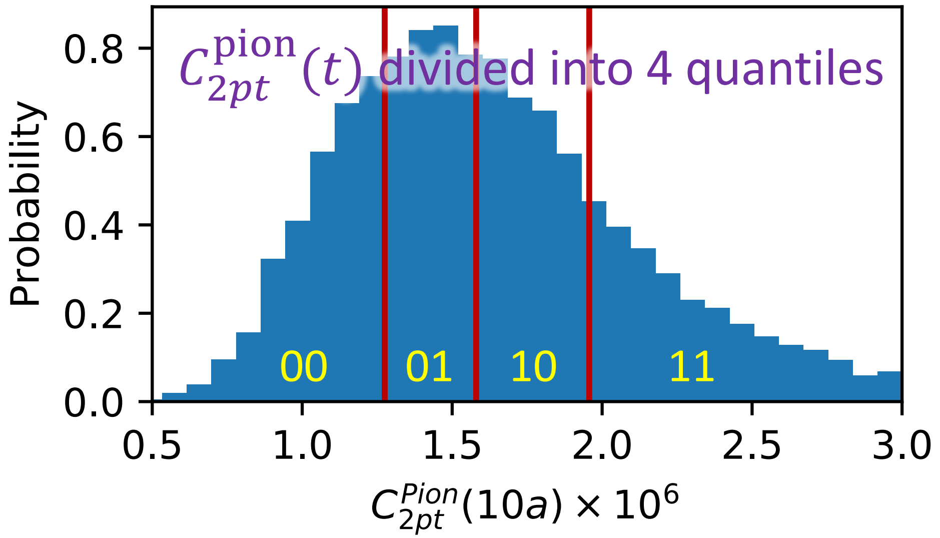


- $C_{2pt}^{\text{pion}}(t)$ at 16 timeslices are reconstructed from a $\mathbf{\Phi}$ and 32 coefficients $a_i^{(k)} = 0\ or\ 1$

$$\left[\vec{C}_{2pt}^{(k)}\right]_{16} \approx [\mathbf{\Phi}]_{16\times32}\left[\vec{a}^{(k)}\right]_{32}$$

$k(=0,\cdots,100\times128)$: confs&meas index

- Compared with a 2-bit (4 bins) representation (16×2=32 bits)



Summary

- Machine learning (ML) is employed to predict unmeasured observables from measured observables
- Bias correction and bootstrap error estimation are used
- Demonstrated for four lattice QCD calculations
  1) Prediction of C_{3pt} from C_{2pt}
  2) Prediction of C_{2pt}^{CPV} from C_{2pt}
  3) Prediction of full-lattice results from sub-vol. calculation
  4) Data compression (using D-Wave quantum annealer)

Acknowledgement

- This work was supported by the DOE Office of Science, Office of High Energy Physics, by the DOE Office of Science , Office of Advanced Scientific Computing Research and Office of Nuclear Physics, Scientific Discovery through Advanced Computing (SciDAC) program under Contract No. 89233218CNA000001, and by the LANL LDRD program.
- Computations were carried out using NERSC, OLCF, USQCD Collaboration, LANL IC and D-Wave 2000Q.