Machine Learning Estimators for Lattice QCD Observables

arXiv:1807.05971v2 [hep-lat]

B. Yoon, T. Bhattacharya, R. Gupta

目的:

Machine Learning を使って２点関数から３点関数を予測する

結果:

概ね良好な結果が得られた

Machine Learning in Lattice QCD

Machine Learning:

Training data を使って適当(適切)な予測モデル(F)を作る方法

$$\{O_i, X_i\} \text{の組} N \text{個} \xrightarrow[F(X_i)=O_i^P \approx O_i]{\text{Machine Learning}} O_j^P = F(X_j)$$

X_j:Training data N 個と重ならない M 個

Machine Learning を使った Lattice QCD 関係の研究

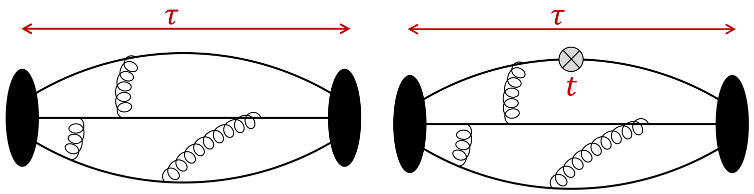
1. Lefschetz Thimble
   Alexandru, Bedaque, Lamm, Lawrence, PRD96:9:094505(2017)
   Mori, Kashiwa, Ohnishi, PTEP2018:2:023B04(2018)
   Kashiwa, Mori, Ohnishi, PRD99:1:014033(2019)

2. Order parameter or action paramters from configuration
   Wetzel, Scherzer, PRB96:18:184410(2017)
   Tanaka, Tomiya, JPSJ86:6:063001(2017)
   Shanahan, Trewartha, Detmold, PRD97:9:094506(2018) 大野さん文献紹介(2018.6.8)

3. AdS/CFT
   Hashimoto, Sugishita, Tanaka, Tomiya, PRD98:4:046019(2018)

今回の話はこれまでよりもシンプルな応用

Nucleon chargeの計算



核子２点相関関数 $t \gg 1$

$$C_2(t) = \langle 0|N(t)\overline{N}(0)|0\rangle = Z_N^2 e^{-M_N t}, \quad Z_N = \langle 0|N(0)|N\rangle$$

核子３点相関関数 $\tau - t, t \gg 1 \rightarrow$ 励起状態を落とすため

$$C_3^\Gamma(\tau, t) = \langle 0|N(\tau)J_\Gamma(t)\overline{N}(0)|0\rangle = Z_N^2 e^{-M_N \tau} g_\Gamma$$

quark bilinear current $J_\Gamma =$ Vector, Axialvector, Tensor, Scalar

$$g_\Gamma = \frac{C_3^\Gamma(\tau, t)}{C_2(\tau)} =$$

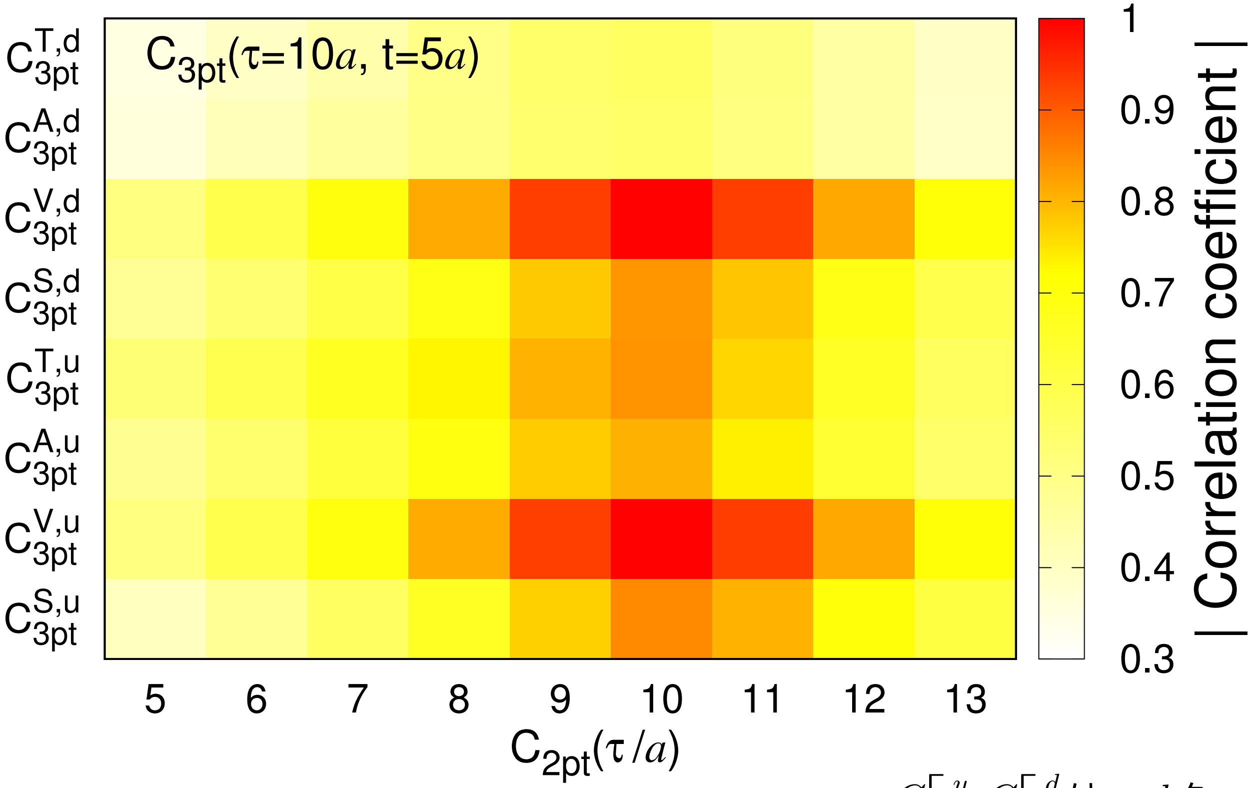
g_V	電荷
g_A	axial charge (最も基本的な量)
g_T	標準模型を超えた物理探索で重要
g_S	標準模型を超えた物理探索で重要

1. 励起状態の寄与を調べる $\rightarrow C_3^\Gamma(\tau, t)$ の τ 依存性を調べる必要
2. τ ごとに $C_3^\Gamma(\tau, t)$ の計算が必要 (異なる Γ, t は同時に計算できる)
3. でも計算コストは抑えたい

$\Rightarrow$ 既知のデータから $C_3^\Gamma(\tau, t)$ を正しく予測できれば計算コストを抑えることが可能

２点関数と３点関数の相関

$C_2(\tau)$ と $C_3^\Gamma(\tau, \tau/2)$ には強い相関がある (特に C_3^V)



$C_3^{\Gamma,u}, C_3^{\Gamma,d}$ は u, d クォークカレントの計算

この相関を使って $C_2(t)$ から $C_3^\Gamma(\tau, t)$ を予測するモデルを作りたい

$\Rightarrow$ Machine Learning

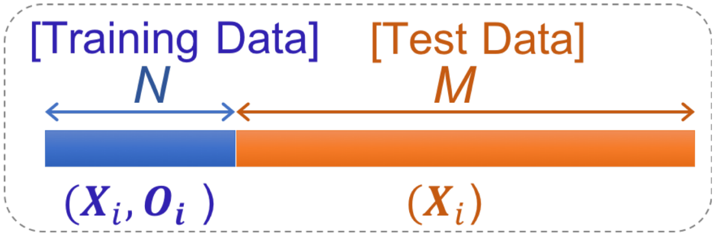
Bias Correction

- Simple average

例) X_i を $N+M$ 配位、O_i を N 配位で計算

Traning Data N 個を使って予測モデル $F(X_i) = O_i^P$ を作る

$$\overline{O} = \frac{1}{M}\sum_{i\in M} O_i^P(X_i)$$



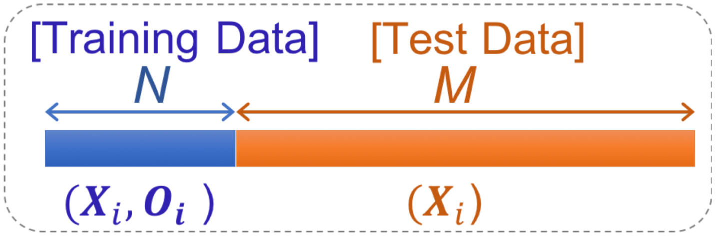
Bias Correction

- Simple average

例) X_i を $N+M$ 配位、O_i を N 配位で計算

Traning Data N 個を使って予測モデル $F(X_i) = O_i^P$ を作る

$$\overline{O} = \frac{1}{M}\sum_{i\in M} O_i^P(X_i)$$



- Bias corrected average

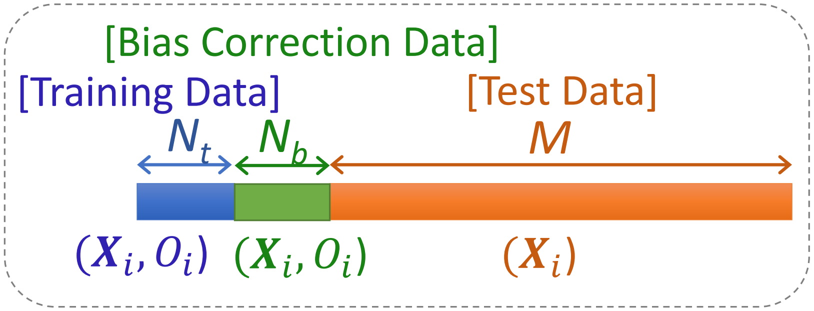
O_i^P は Bias($\langle O_i\rangle - \langle O_i^P\rangle \neq 0$) を持つかもしれないので補正項をいれる

Traning Data N_t 個を使って予測モデル $F(X_i) = O_i^P$ を作る

$$\overline{O} = \frac{1}{M}\sum_{i\in M} O_i^P(X_i) + \frac{1}{N_b}\sum_{i\in N_b}(O_i - O_i^P(X_i))$$

第２項は $F(X_i)$ が適切なら小さいはず (効果は後で見る)

N_t と N_b は重ならないように選ぶ



シミュレーションパラメータ

- $M_\pi \approx 310$ MeV, $a \approx 0.089$ fm, Clover on HISQ
- $C_3^\Gamma(\tau, t)$ with $\tau = 8, 10, 12, 14$
- 2263配位 (64測定/配位)

  3 配位飛ばし 680 配位を選択

  Traning Data $N_t = 60$, Bias Correction Data $N_b = 620$

  Test Data $M = 1583$
- 統計誤差評価はBootstrap

- Machine Learning

  Boosted Decision Tree Regression algorithmというものを使った

  調整パラメータが少なく、overfittingの危険性が低い

  数値安定性のため $C_3^\Gamma(\tau, t)/\mathcal{N}, C_2(\tau)/\mathcal{N}$ と規格化 ($\mathcal{N} = \langle C_2(\tau) \rangle$)

  モデル生成+予測 = ノートパソコンで数分

直接計算した結果と Machine Learning で予測した結果を比較し再現性を調べる

結果1 $\tau = 10$

Training Data 60 配位 + Bias Correction Data 620 配位

$$X = C_2(\tau)(0 \le \tau \le 20),\ O = C_3^\Gamma(\tau, t)(\tau = 8, 10, 12, 14, 0 \le t \le \tau, \Gamma = S, A, T, V)$$

Test Data 1583 配位

イ ン プ ッ ト: $X = C_2(\tau)(0 \le \tau \le 20)$, 予測: $O^P = C_3^\Gamma(\tau, t)(\tau = 8, 10, 12, 14, 0 \le t \le \tau)$

$g_\Gamma = C_3^\Gamma(\tau = 10, t = 5)/C_2(\tau = 10)$ の Test Data 結果

Γ	DM	BC-Prediction	Raw-Prediction	Bias
S	0.936(10)	0.933(15)	0.931(45)	+0.002(46)
A	1.2011(41)	1.1997(48)	1.1999(109)	−0.0003(105)
T	1.0627(34)	1.0638(39)	1.0642(79)	−0.0004(78)
V	1.0462(36)	1.0455(36)	1.0453(39)	+0.0002(20)

DM: Direct Measurement[目標とする結果], BC-Prediction: Bias Corrected average, Raw-Prediction: Simple average, Bias: Bias Correction

Bias は小さいが誤差が大きい。Raw-Prediction も DM より誤差が大きい

BC-Prediction = Raw-Prediction + Bias は誤差が DM と同程度

→ DM を再現できた (Bias は入れる必要がある)

結果1 $\tau = 10$

Training Data 60 配位 + Bias Correction Data 620 配位

$$X = C_2(\tau)(0 \le \tau \le 20),\ O = C_3^\Gamma(\tau, t)(\tau = 8, 10, 12, 14, 0 \le t \le \tau, \Gamma = S, A, T, V)$$

Test Data 1583 配位

インプット: $X = C_2(\tau)(0 \le \tau \le 20)$, 予測: $O^P = C_3^\Gamma(\tau, t)(\tau = 8, 10, 12, 14, 0 \le t \le \tau)$

$g_\Gamma = C_3^\Gamma(\tau = 10, t = 5)/C_2(\tau = 10)$ の Test Data 結果

Γ	DM	BC-Prediction	Raw-Prediction	Bias
S	0.936(10)	0.933(15)	0.931(45)	+0.002(46)
A	1.2011(41)	1.1997(48)	1.1999(109)	−0.0003(105)
T	1.0627(34)	1.0638(39)	1.0642(79)	−0.0004(78)
V	1.0462(36)	1.0455(36)	1.0453(39)	+0.0002(20)

DM: Direct Measurement[目標とする結果], BC-Prediction: Bias Corrected average,
Raw-Prediction: Simple average, Bias: Bias Correction

Biasは小さいが誤差が大きい。Raw-PredictionもDMより誤差が大きい

BC-Prediction = Raw-Prediction + Bias は誤差がDMと同程度

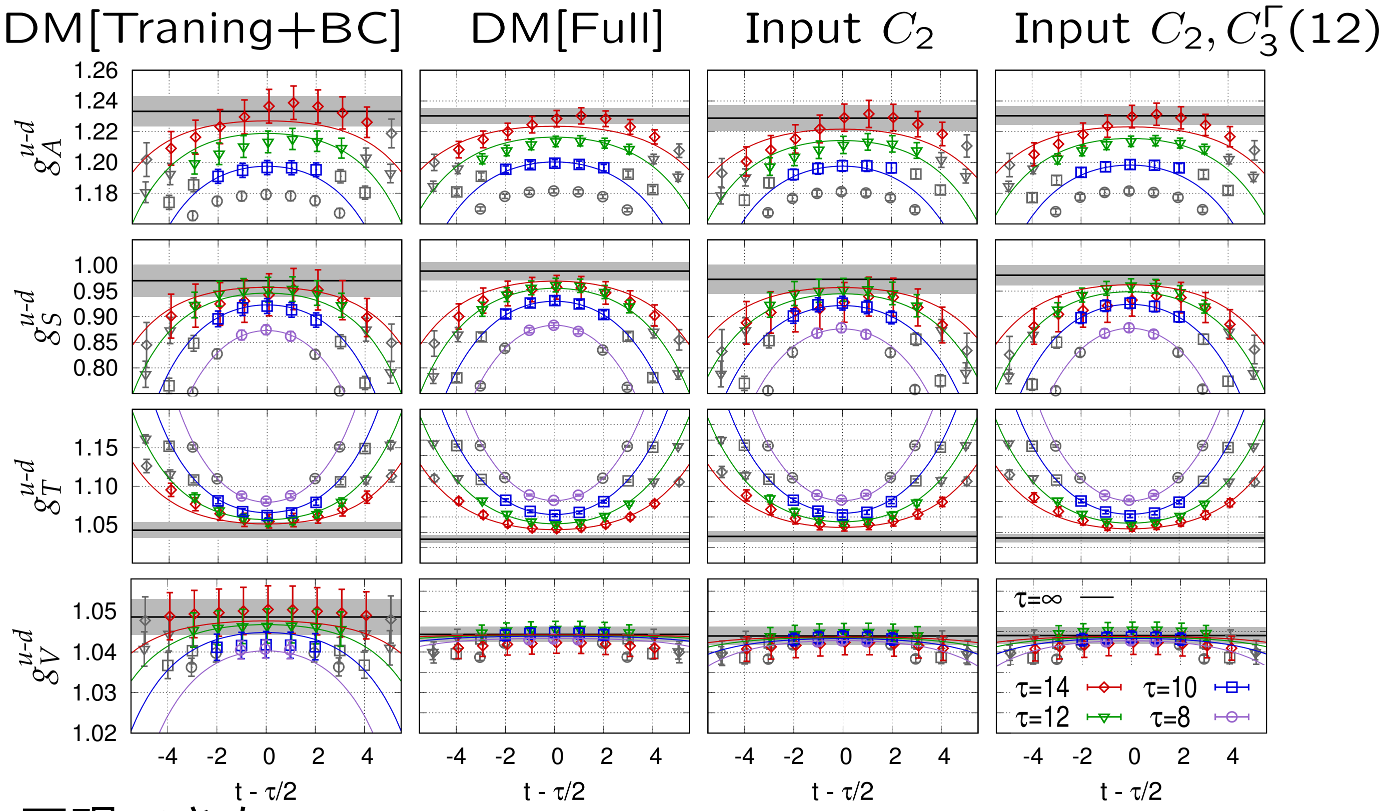
→ Raw-Prediction と Bias は強く相関している

相関の原因はモデル $F(X)$ と配位のauto correlation?

結果2 $\tau = 8, 10, 12, 14$

インプット:$X = C_2(\tau)(0 \le \tau \le 20)$, 予測:$O^P = C_3^\Gamma(\tau, t)(\tau = 8, 10, 12, 14, 0 \le t \le \tau)$
インプット:$X = C_2(\tau)(0 \le \tau \le 20) + C_3^\Gamma(12, t)$, 予測:$O^P = C_3^\Gamma(\tau, t)(\tau = 8, 10, 14, 0 \le t \le \tau)$
統計誤差はTraning+BCのO_iも含めたもの

$g_\Gamma = C_3^\Gamma(\tau, t)/C_2(\tau)$ の結果 (線は$\tau \to \infty$フィット)



DMを概ね再現できた
$C_3^\Gamma(12)$をインプットに加えると再現性は若干改善する

結果2 $\tau = 8, 10, 12, 14$

インプット:$X = C_2(\tau)(0 \le \tau \le 20)$, 予測:$O^P = C_3^\Gamma(\tau, t)(\tau = 8, 10, 12, 14, 0 \le t \le \tau)$

インプット:$X = C_2(\tau)(0 \le \tau \le 20) + C_3^\Gamma(12, t)$, 予測:$O^P = C_3^\Gamma(\tau, t)(\tau = 8, 10, 14, 0 \le t \le \tau)$

統計誤差は Traning+BC の O_i も含めたもの

$g_\Gamma(\tau \to \infty)$ の結果

	DM	Pred.[$C_{\rm 2pt}$]	Pred.[$C_{\rm 2pt}$,$C_{\rm 3pt}(12)$]
g_S	0.989(18)	0.973(29)	0.981(20)
g_A	1.2303(51)	1.2289(83)	1.2304(61)
g_T	1.0311(51)	1.0347(68)	1.0326(54)
g_V	1.0443(19)	1.0439(22)	1.0440(21)

DM を概ね再現できた (統計誤差は若干大きくなる)

$C_2 + C_3^\Gamma(12)$ では DM より誤差が 10–20%大きくなる

$\to$ 計算コストは 35% 抑制できた (統計誤差増加分を考慮)

論文では CP violating $C_2(t)$ の結果も議論しているが同じような結論なので省略

まとめ

- Machine Learning を使って、C_2 から C_3^Γ を概ね再現できた

  統計誤差を再現するためには Bias Correction が必要

  O_i^P と Bias には強い相関がある (理由はよくわからない)

  統計誤差は若干大きくなる

- インプットを増やすと再現性は若干改善する

- $C_2 + C_3^\Gamma(12)$ では直接計算より誤差が 10%大きくなる

  $\rightarrow$ 計算コストを 35% 抑制 (統計誤差増加分を考慮)