Machine learning action parameters in lattice quantum chromodynamics

Reference

PRD**97**, 094506 (2018) [arXiv:1801.05784]

大野浩史

文献紹介　2018年6月8日

概要

- 機械学習を用いて、ゲージ配位から作用のパラメータを予想する。

$$U_\mu(x) \xrightarrow{\text{ML}} \beta,\ m_0$$

- 入力するデータを工夫するとある程度うまくいく。

シミュレーションセットアップ

- N_c = 2 Wilson gauge
- N_f = 2 dynamical Wilson quarks
- 最初の 500 trajectory を除き 10000 trajectory まで 10 trajectory ごとの保存 = 950 配位

- Grid A: 12^3x36, 20パラメータ
- Grid B: 12^3x36, 25パラメータ
- Grid C: 16^3x36, Grid A と同じパラメータ
- Set D & E: 同じプラケット値を持つパラメータセット
- Set F: パラメータの同じ10個の独立なアンサンブル

そもそもゲージ配位の見分けがつくのか？

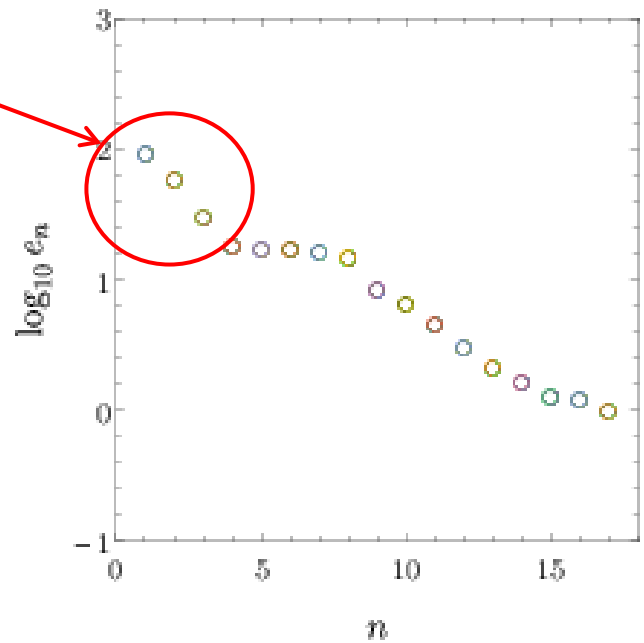
- 主成分分析で確認
- ウィルソンループの相関関数

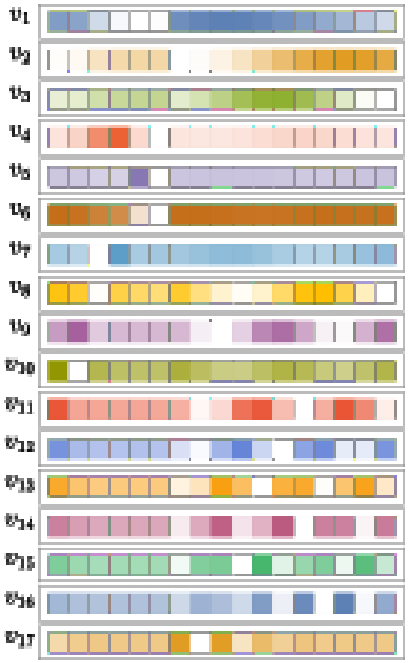
l_i = 1x1, 1x2…: ループの種類
e: アンサンブル
c: 配位

$$\mathcal{M}_{\ell_i,\ell_j} = \sum_e \sum_c \frac{\left[W_{\ell_i}(e,c) - \overline{W}_{\ell_i}(e)\right]\left[W_{\ell_j}(e,c) - \overline{W}_{\ell_j}(e)\right]}{\sigma(W_{\ell_i}(e))\sigma(W_{\ell_j}(e))}$$

Grid A

支配的な3固有値





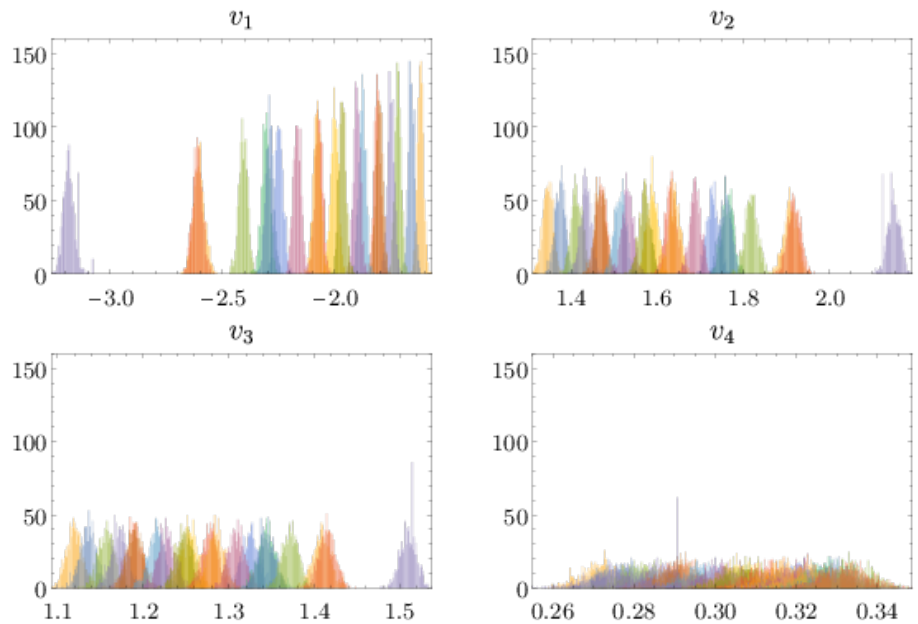
そもそもゲージ配位の見分けがつくのか？

- 主成分分析で確認
- ウィルソンループの相関関数

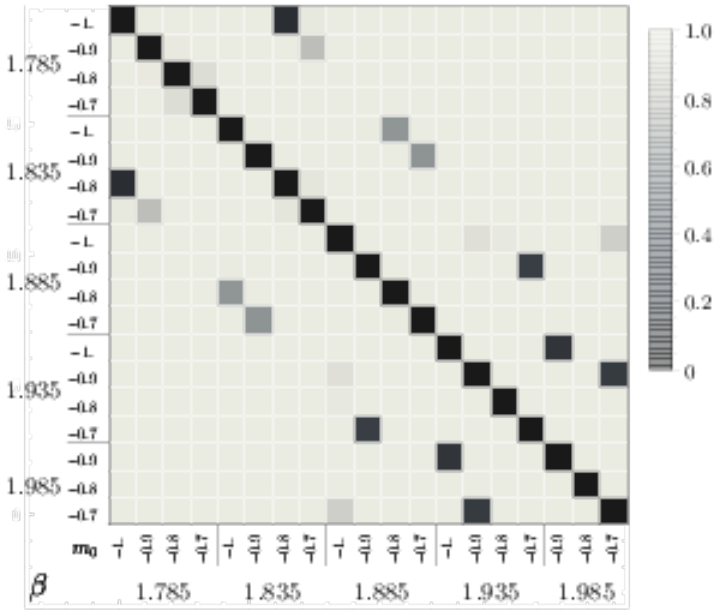
l_i = 1x1, 1x2…: ループの種類
e: アンサンブル
c: 配位

$$\mathcal{M}_{\ell_i,\ell_j} = \sum_e \sum_c \frac{\left[W_{\ell_i}(e,c) - \overline{W}_{\ell_i}(e)\right]\left[W_{\ell_j}(e,c) - \overline{W}_{\ell_j}(e)\right]}{\sigma(W_{\ell_i}(e))\sigma(W_{\ell_j}(e))}$$

Grid A



Jensen-Shannon divergence
= 確率分布の重なり具合の指標



そもそもゲージ配位の見分けがつくのか？

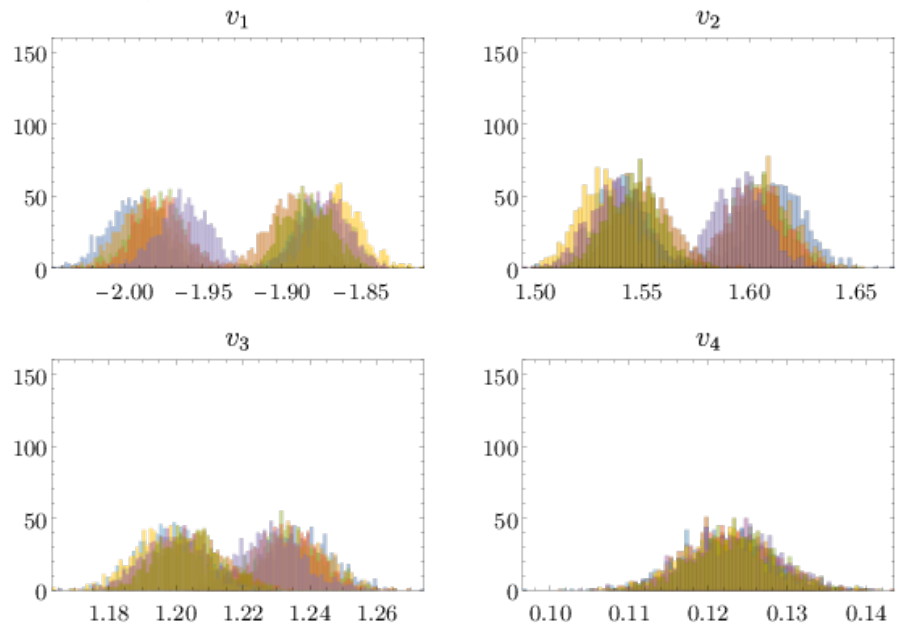
- 主成分分析で確認
- ウィルソンループの相関関数

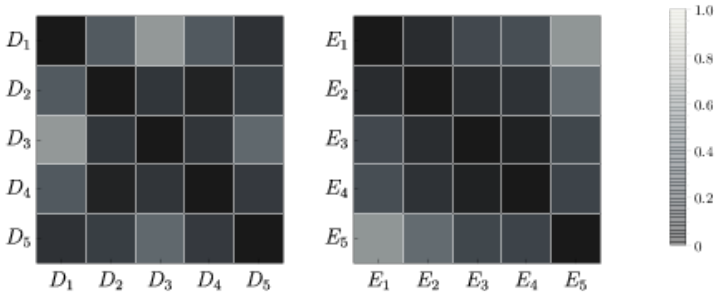
l_i = 1x1, 1x2…: ループの種類
e: アンサンブル
c: 配位

$$\mathcal{M}_{\ell_i,\ell_j} = \sum_e \sum_c \frac{\left[W_{\ell_i}(e,c) - \overline{W}_{\ell_i}(e)\right]\left[W_{\ell_j}(e,c) - \overline{W}_{\ell_j}(e)\right]}{\sigma(W_{\ell_i}(e))\sigma(W_{\ell_j}(e))}$$

Set D&E

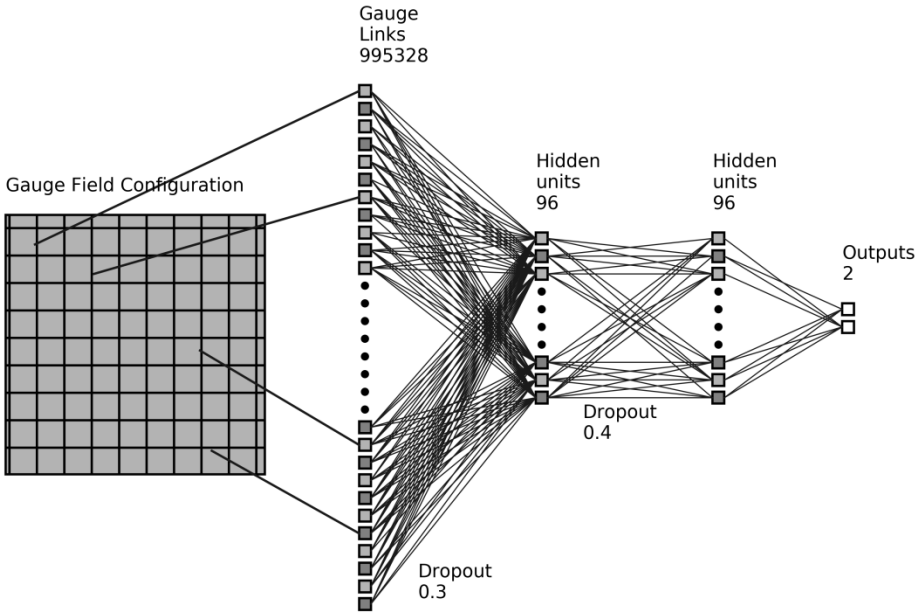
Jensen-Shannon divergence
= 確率分布の重なり具合の指標





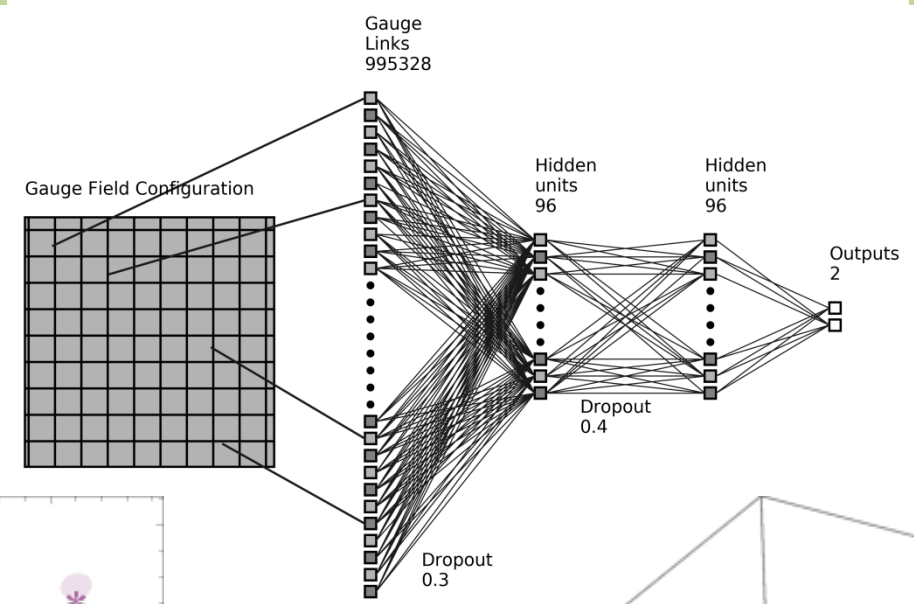
機械学習

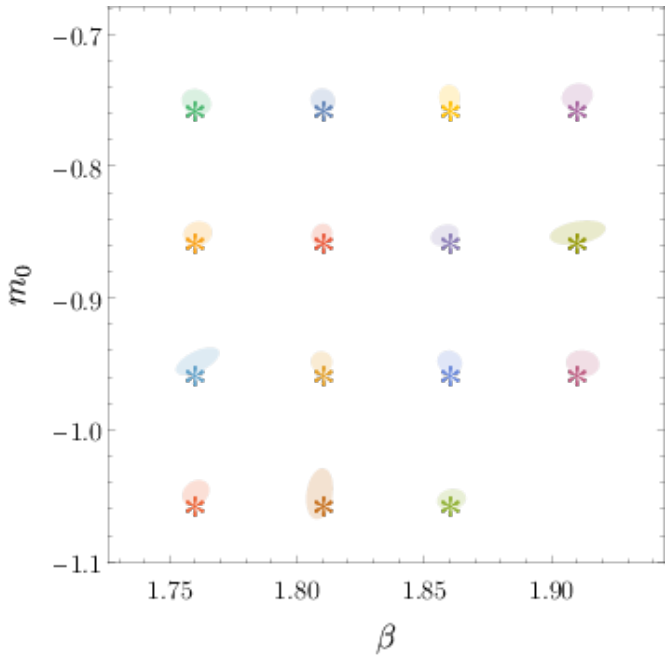
- 多層パーセプトロン
  - 3層以上、全結合
  - 非線形活性化関数（tanh等）
- ランダムに選んだ850配位で学習、残りの100配位で予想の精度を確認

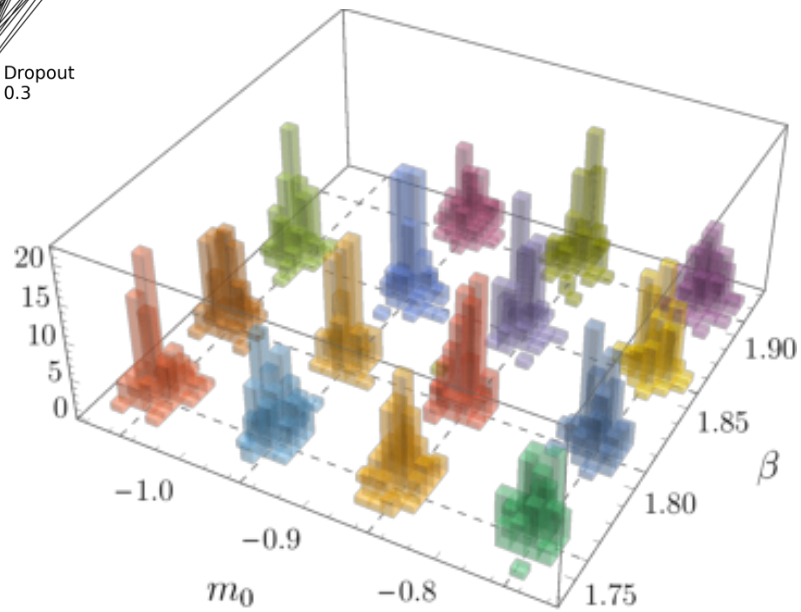


ゲージ配位そのものを入力にした場合

＊:実際のパラメータ
楕円:予想の1σ信頼区間







ゲージ配位そのものを入力にした場合

- 一見うまくいっているように見える
- 実際にはうまくいっていない（予想が一般化できない）
  - Set F からは学習に用いたすべてのパラメータの平均が得られる
  - 学習に用いた配位から連続的に生成された配位を入力した時、trajectory数が増えるほど正しい値からずれて、平均値へ近づく
- 原因
  - ゲージ配位の自由度に対して、学習データ（配位数）が少なすぎる
  - 物理とは関係のない長相関の特徴にとらわれている

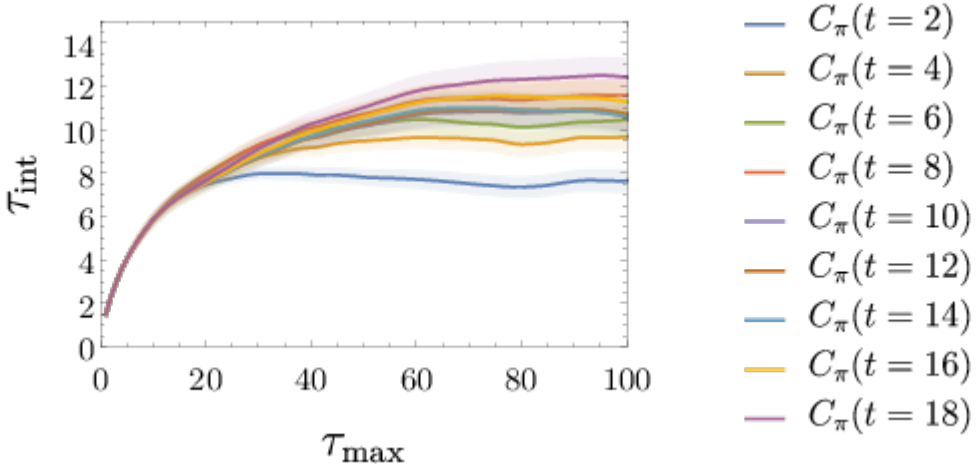
ゲージ配位そのものを入力にした場合

- Set F を用いて長相関を検証
- 2つのアンサンブルペアを見分けられる確率を推定

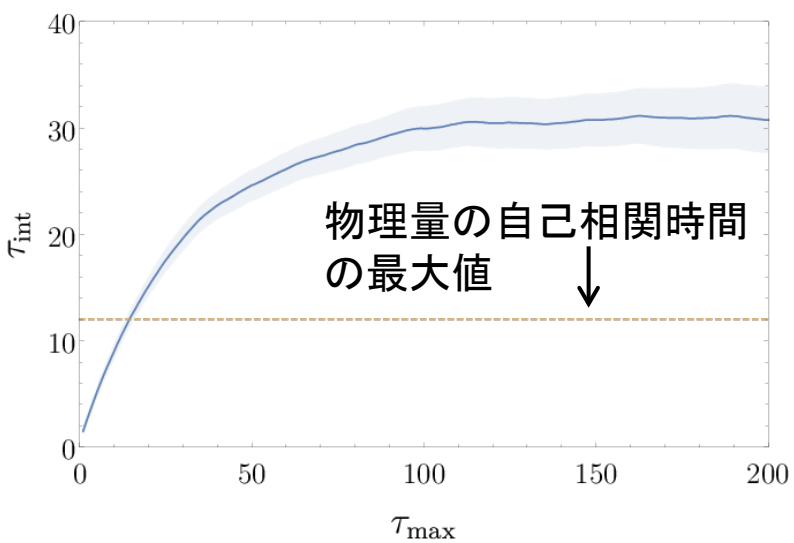
→ 自己相関関数

$$\rho(\tau) = 2\left[P_\alpha\left(c^\alpha(\tau)\right) + P_\beta\left(c^\beta(\tau)\right)\right] - 1$$

π中間子相関関数
の自己相関時間

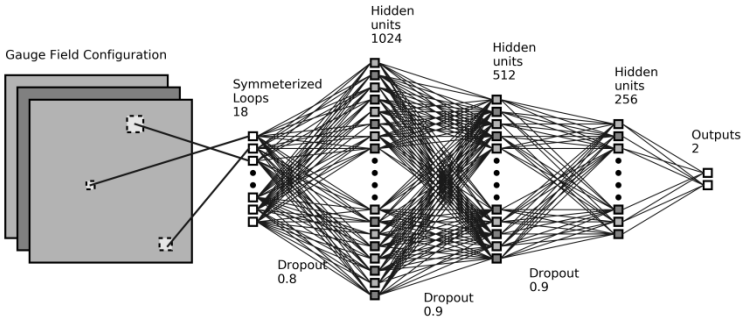


機械学習に現れる
自己相関時間



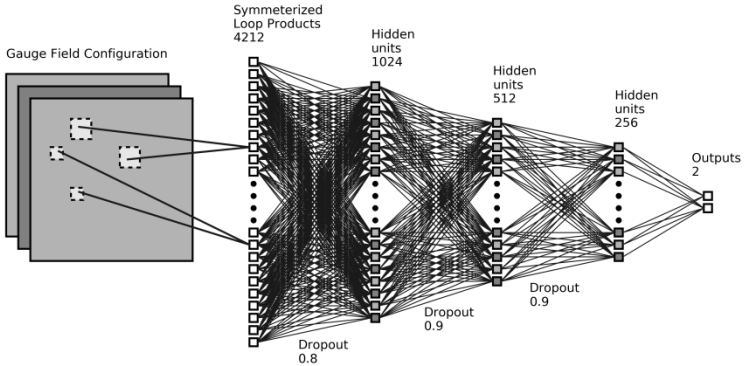
ウィルソンループを入力にした場合

SL



ウイルソンループそのもの
（nxn, 1xn のみ）

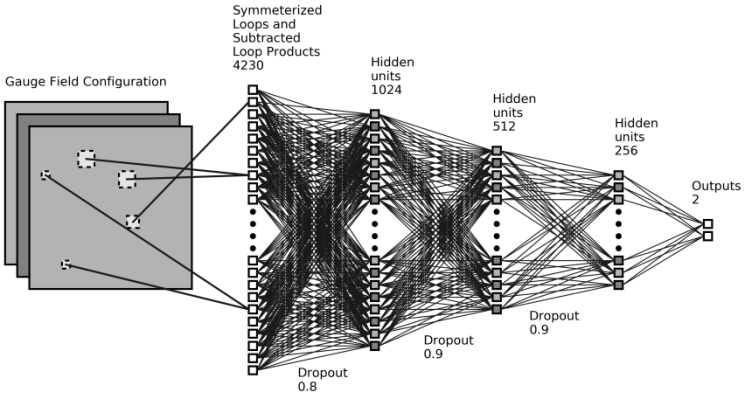
CP



ウイルソンループ同士の相関

$$\mathcal{W}_{j\times k,l\times m}(R) = \sum_{|r|=R} \sum_{\ell\in\mathcal{O}(j\times k)} \sum_{\ell'\in\mathcal{O}(l\times m)} \sum_{x} \mathcal{W}_{\ell}(x)\mathcal{W}_{\ell'}(x+r)$$

SLCP

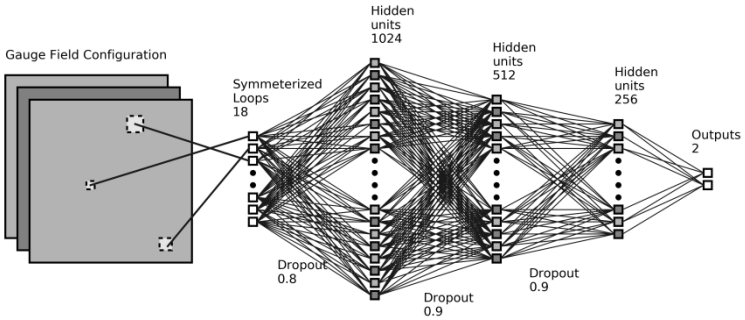


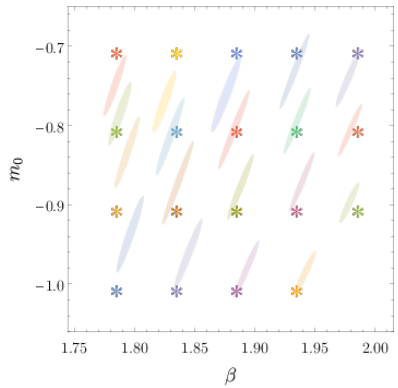
両方

$$\mathcal{W}^{(\mathrm{sub})}_{j\times k,l\times m}(R) = \sum_{|r|=R} \sum_{\ell\in\mathcal{O}(j\times k)} \sum_{\ell'\in\mathcal{O}(l\times m)} \left[\sum_{x} \mathcal{W}_{\ell}(x)\mathcal{W}_{\ell'}(x+r) - \sum_{x} \mathcal{W}_{\ell}(x) \sum_{x} \mathcal{W}_{\ell'}(x)\right]$$

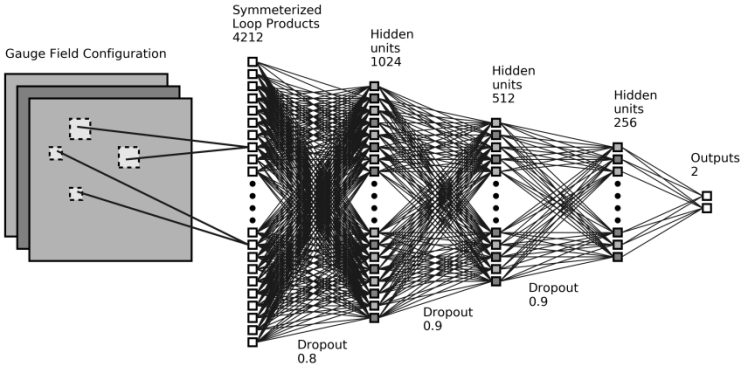
ウィルソンループを入力にした場合

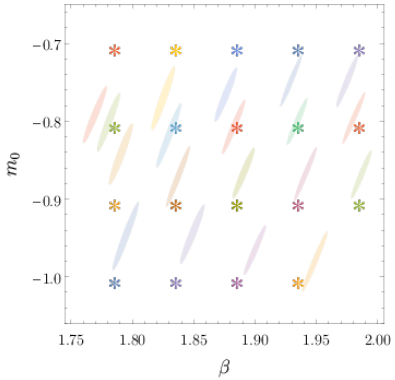
SL



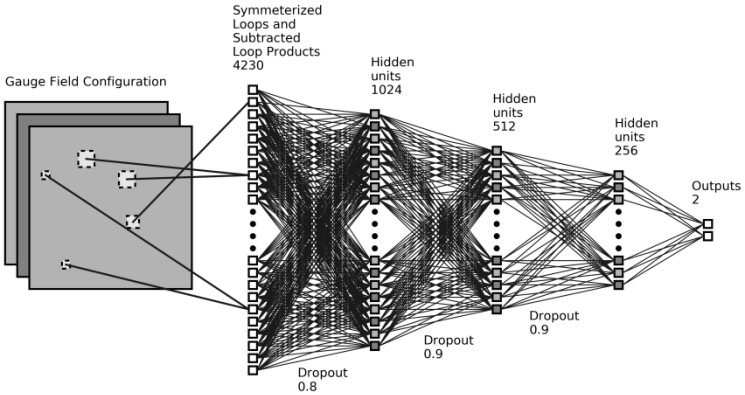


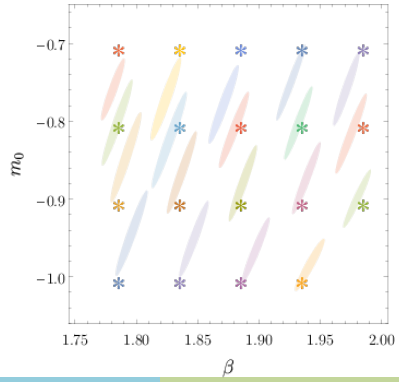
CP





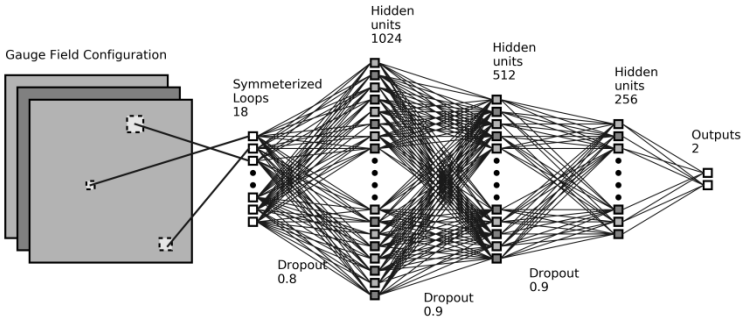
SLCP



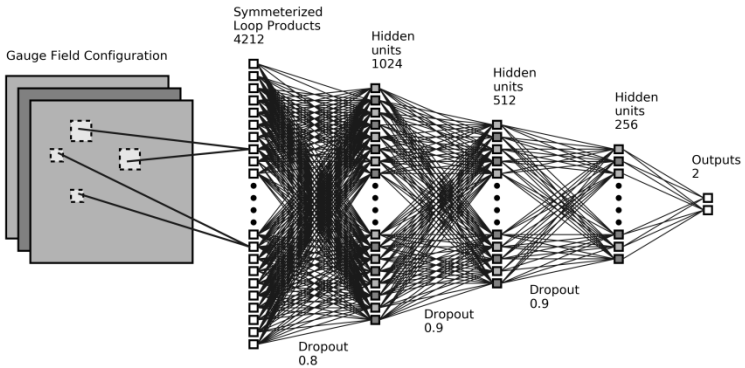


ウィルソンループを入力にした場合

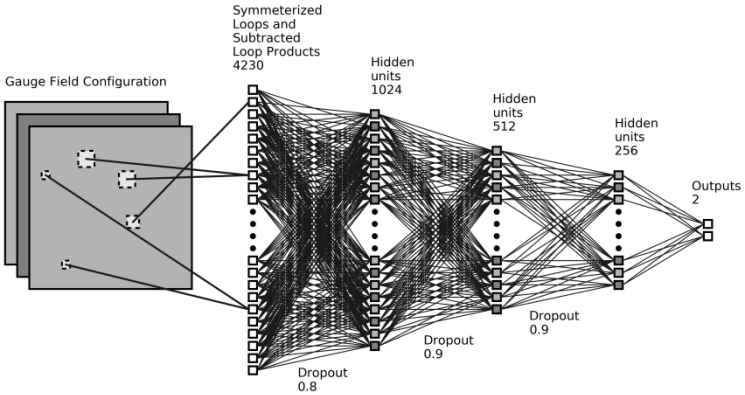
SL



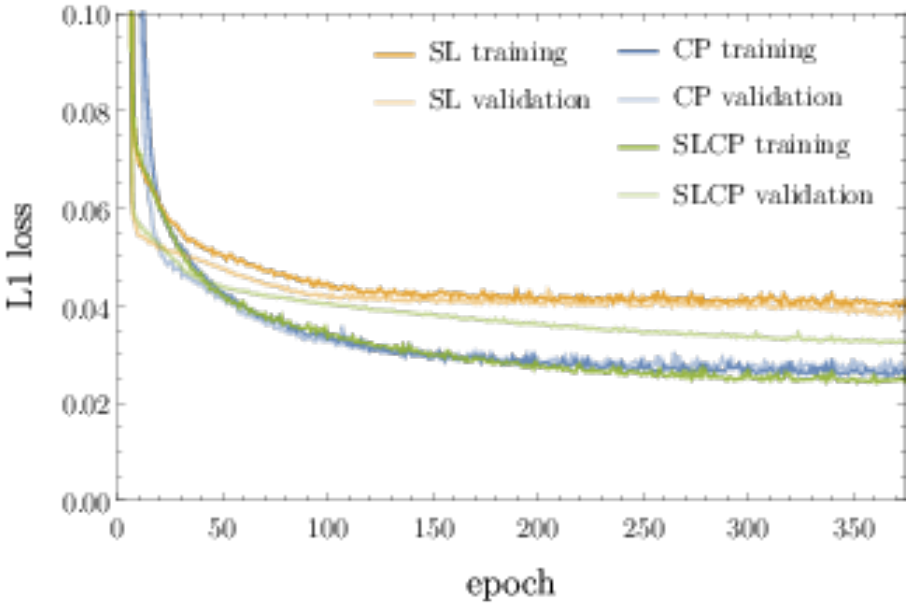
CP



SLCP



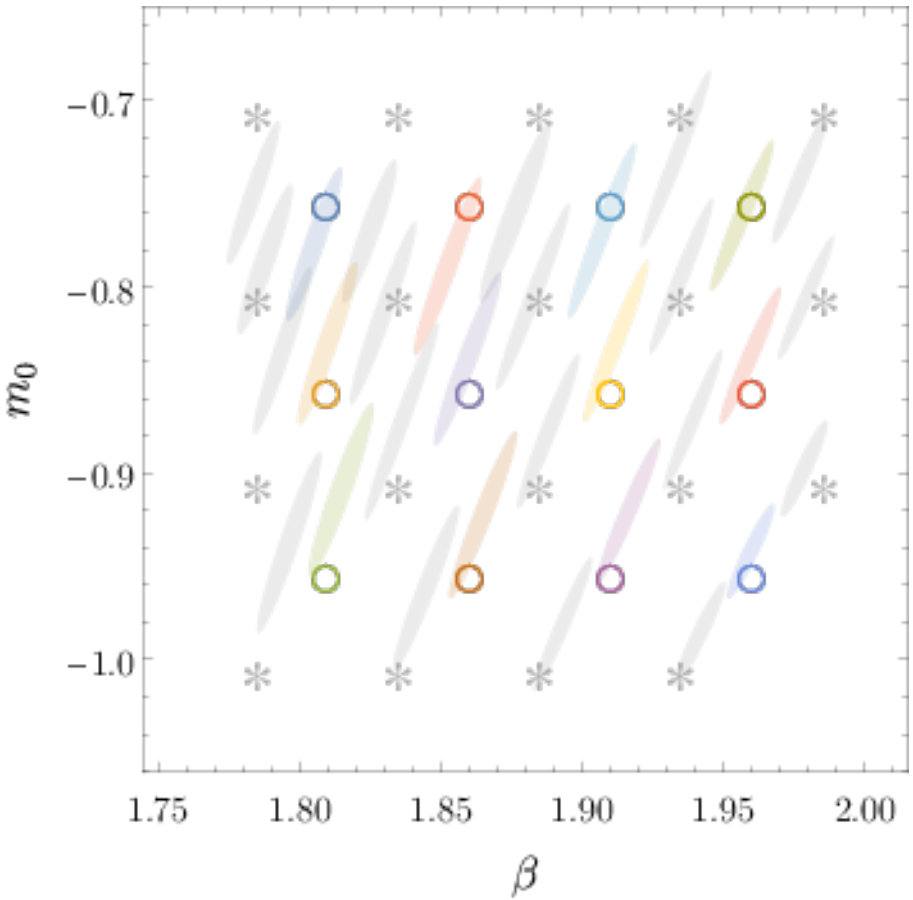
L1 loss = 差の絶対値の和



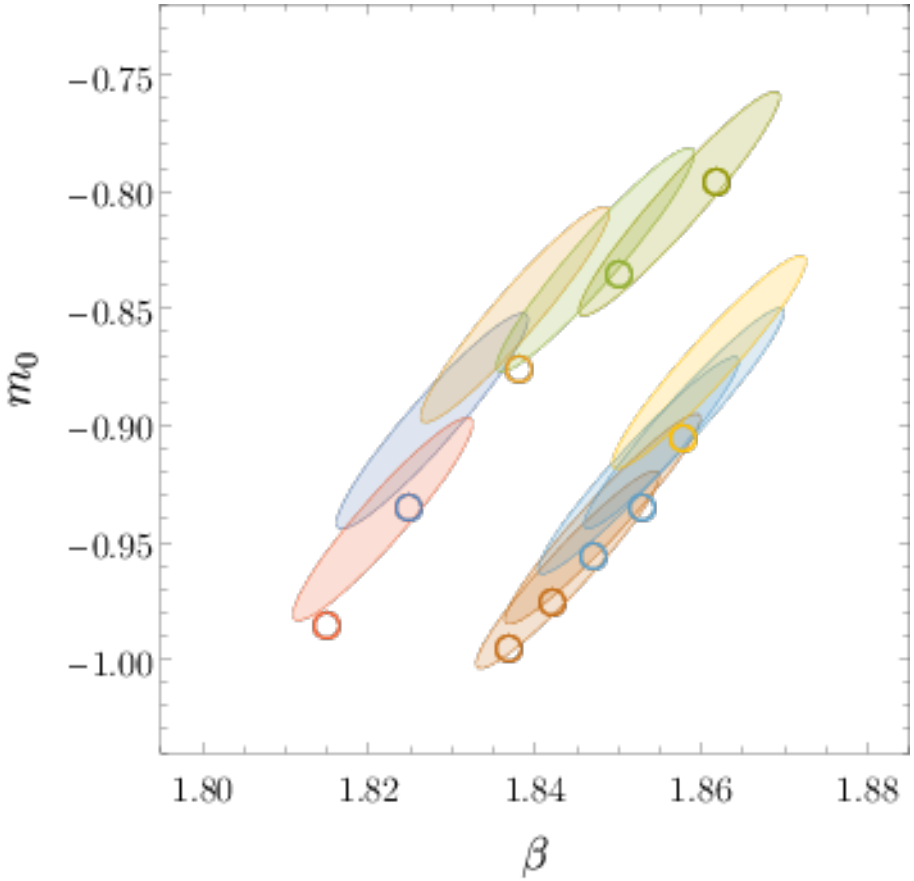
CPが一番よさそう

CPを用いた予想

Grid A で学習 → Grid Bを予想

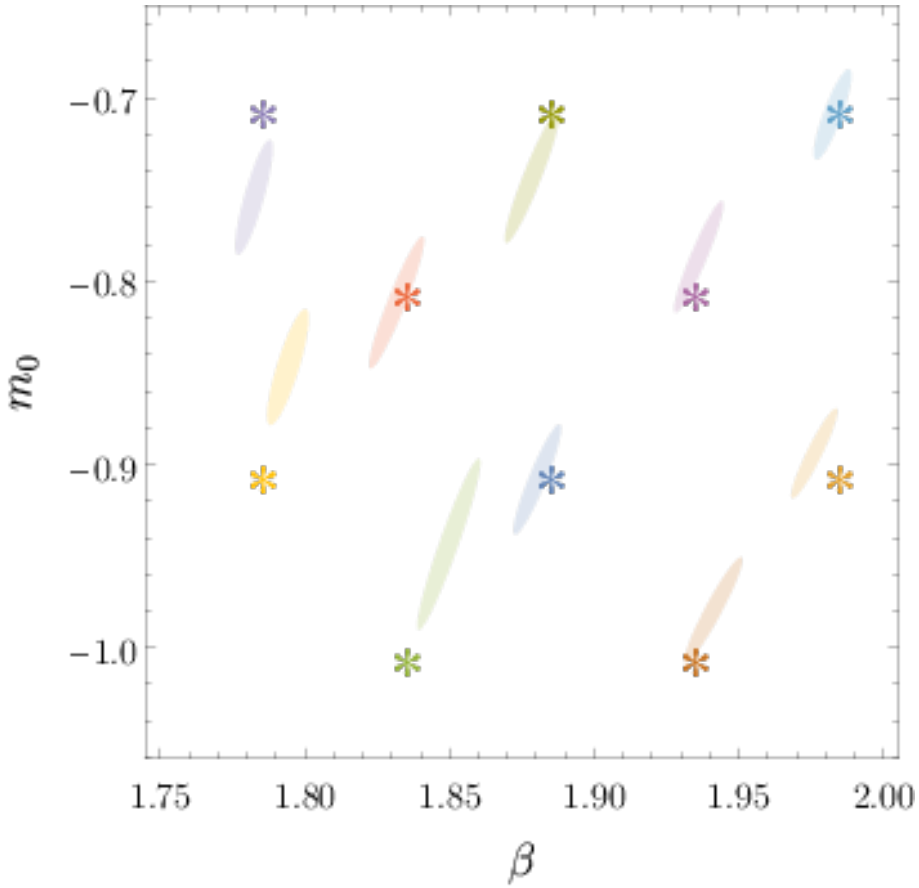


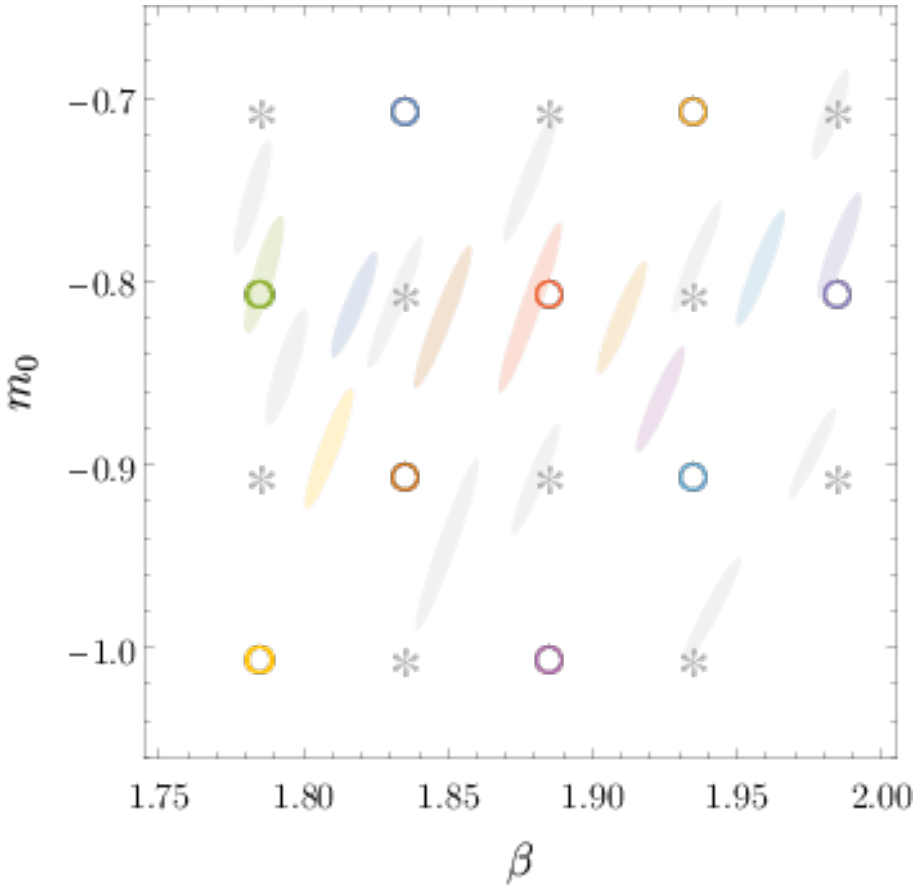
Grid A で学習 → Set D&Eを予想



CPを用いた予想

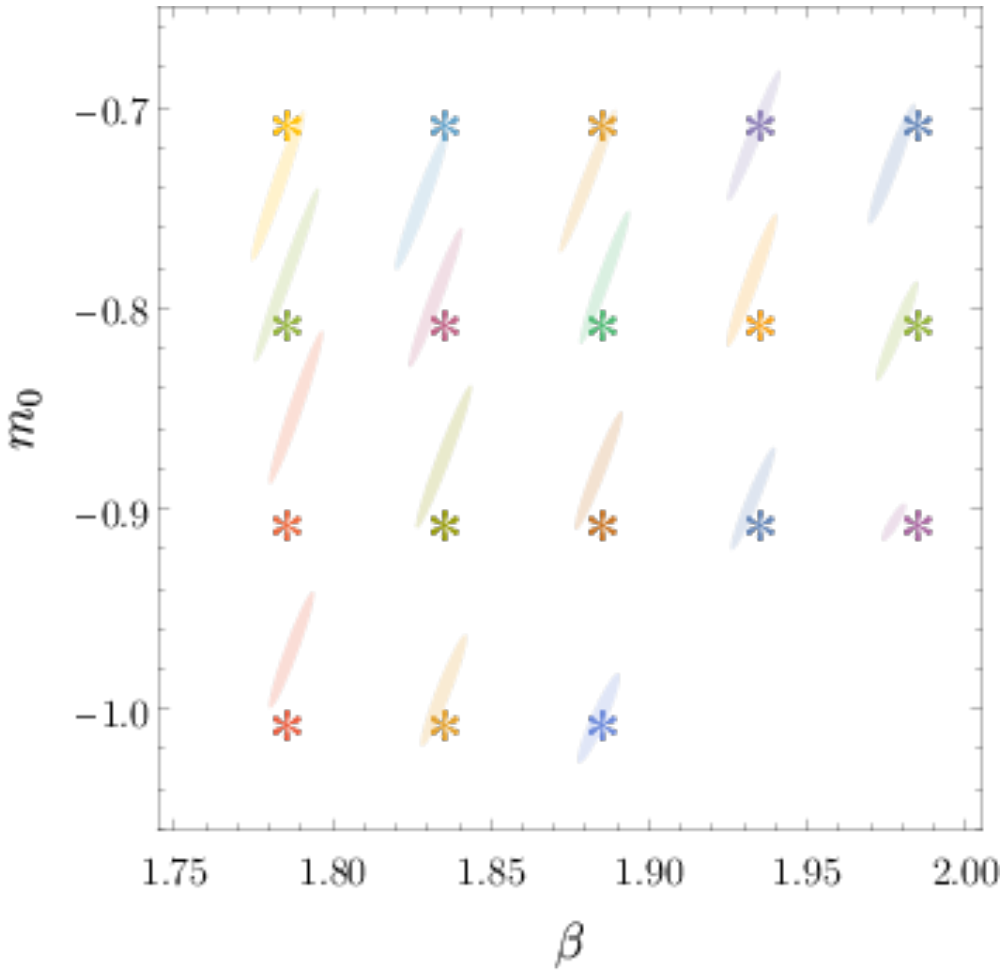
Grid Aで学習 → Grid B を予想
（パラメータの間隔が広い場合）





CPを用いた予想

Grid C（より体積が大きい場合）



まとめ

- 入力を工夫すると、機械学習を用いてゲージ配位から作用のパラメータをある程度予想することができる

- 疑問
  - 学習データを用意する手間と、得られる結果が本当につりあうのか？
  - 既存の方法より、優れているのか？

End