Gauge covariant neural network for 4 dimensional non-abelian gauge theory

Y. Nagai
原子力機構

RBRC
RIKEN BNL Research Center

富谷昭夫 (RIKEN-BNL)

akio.tomiya_AT_riken.jp

https://arxiv.org/abs/2103.11965

ゲージ共変なニューラルネットと4次元非可換ゲージ理論への応用

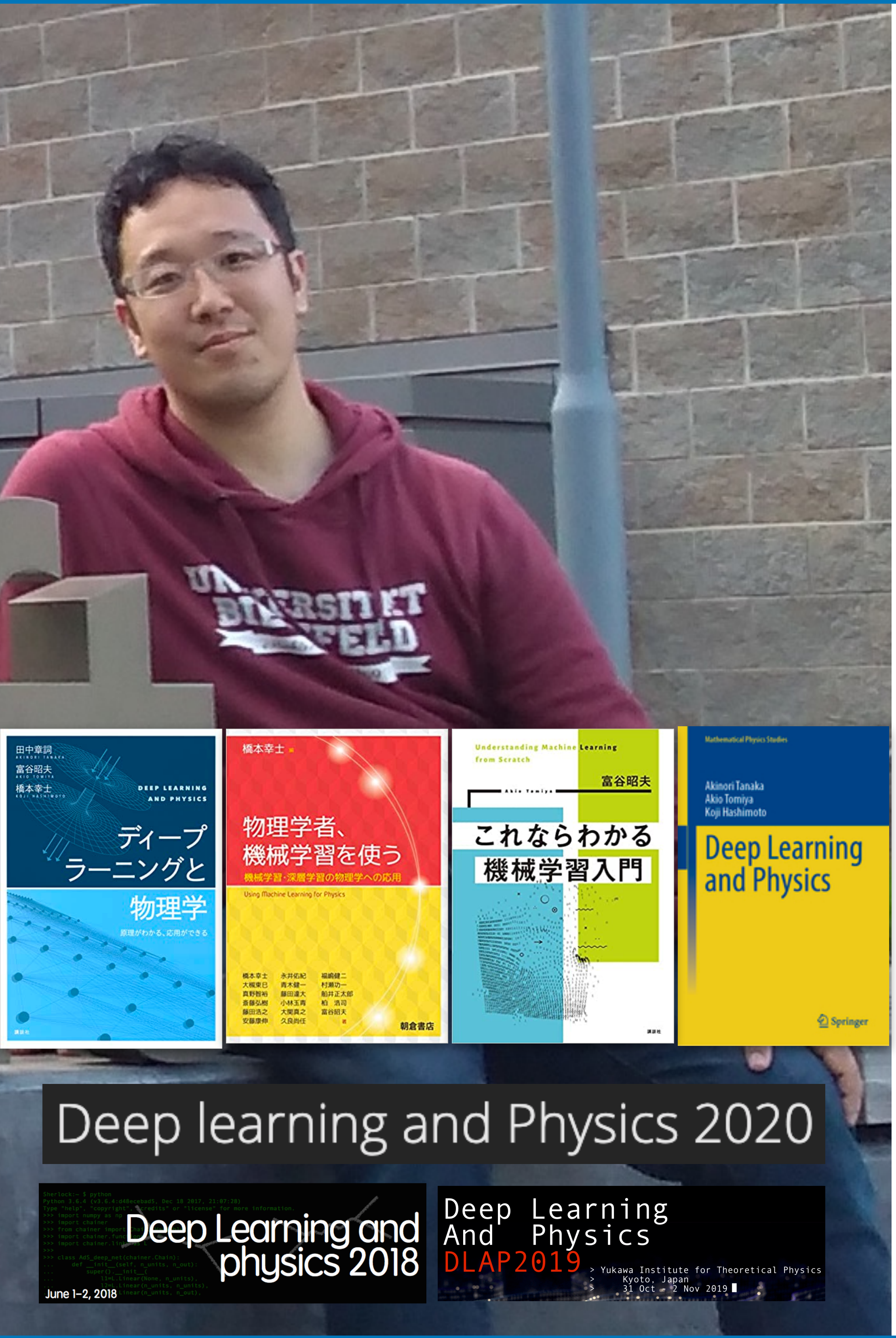
富谷昭夫 (RIKEN-BNL → 大阪国際工科専門職大学, テニュア助教)
RBRC
RIKEN BNL Research Center
akio.tomiya_AT_riken.jp

https://arxiv.org/abs/2103.11965

Y. Nagai
原子力機構

自己紹介

理研-ブルックヘブン国立研究所, 素粒子物理、機械学習



なにをしてる人？

素粒子物理の理論的研究をしています。

機械学習を理論計算の効率化に使いたいです。

主な論文

Detection of phase transition via convolutional neural networks
A Tanaka, A Tomiya
Journal of the Physical Society of Japan 86 (6), 063001
ニューラルネットを使った相検出

Evidence of effective axial $U(1)$ symmetry restoration at high temperature QCD
A Tomiya, G Cossu, S Aoki, H Fukaya, S Hashimoto, T Kaneko, J Noaki, ...
Physical Review D 96 (3), 034509
格子QCDを用いたU(1)量子異常の消失の証拠

Digital quantum simulation of the schwinger model with topological term via adiabatic state preparation
B Chakraborty, M Honda, T Izubuchi, Y Kikuchi, A Tomiya
arXiv preprint arXiv:2001.00485
量子コンピュータ

略歴

2010 : 兵庫県立大学理学部物質科学科卒、超伝導

2015 : 大阪大学で博士号取得。素粒子論。

2015 - 2018 : 華中師範大学でポスドク (中国、武漢)

2018 - 2021 : 理研/BNLでポスドク (米国、NY)

2021 - : 大阪国際工科専門職大学、助教

もくじ

2つの話

1. ニューラルネットと畳込み、フィルタ
2. 格子QCD
3. スメアリング
4. ゲージ共変ニューラルネット

ニューラルネットと対称性の話
ゲージ対称性を保ったニューラルネットの話

---

5. 自己学習HMCでの応用
6. まとめ

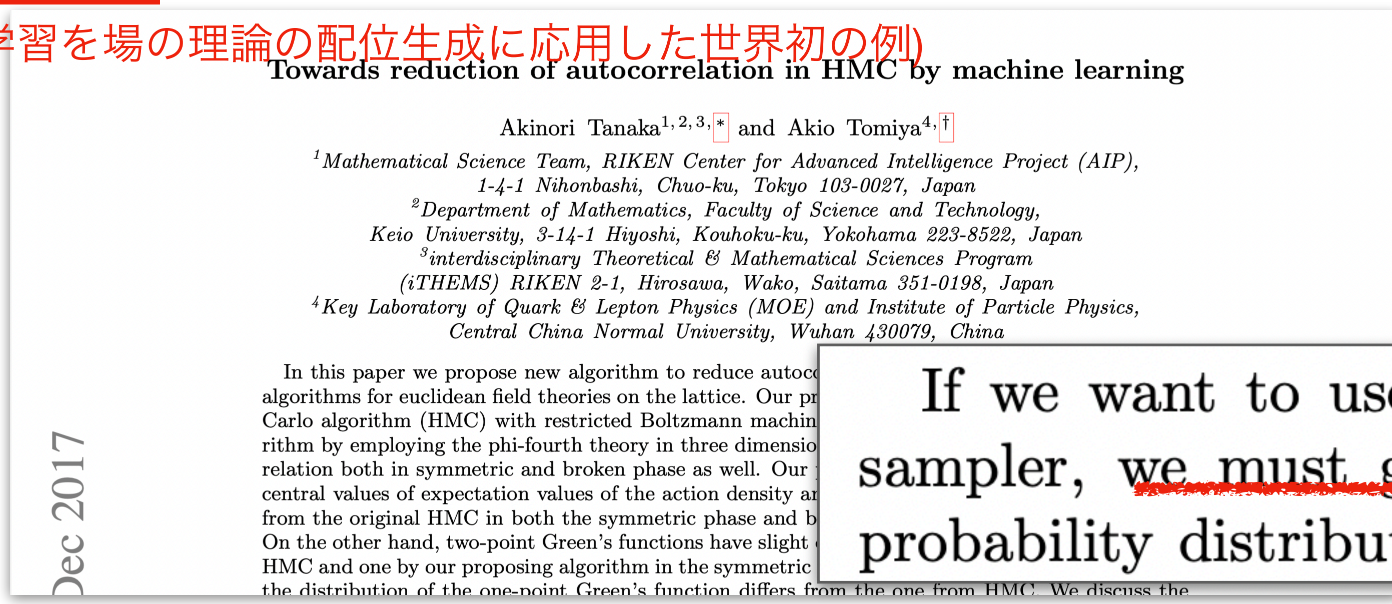
応用編

動機と結果
ゲージ対称性とニューラルネット

(個人的な)動機 2017年からゲージ対称性を尊重するニューラルネットを作りたかった

(機械学習を場の理論の配位生成に応用した世界初の例)

arXiv: 1712.03893

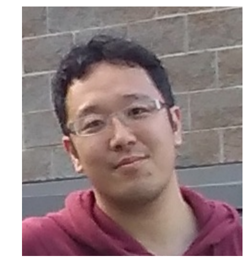
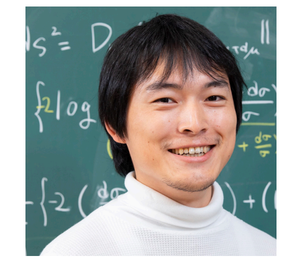

Towards reduction of autocorrelation in HMC by machine learning

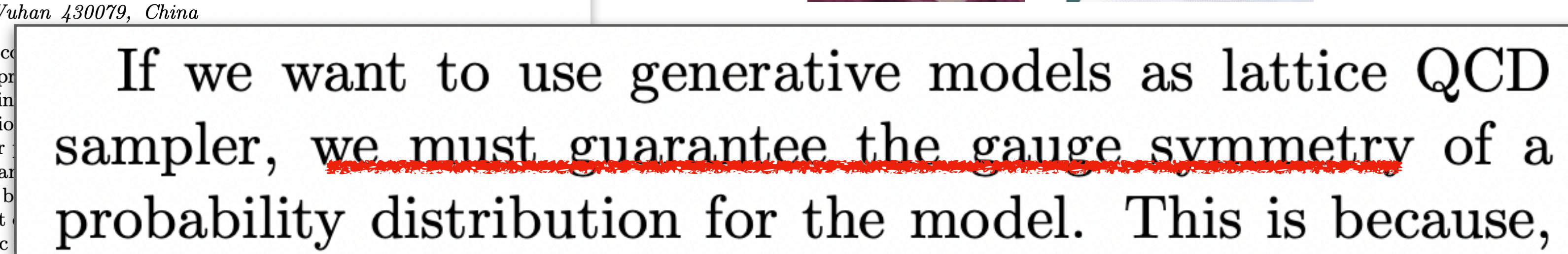
Akinori Tanaka[1,2,3,*] and Akio Tomiya[4,†]

[1] Mathematical Science Team, RIKEN Center for Advanced Intelligence Project (AIP), 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan
[2] Department of Mathematics, Faculty of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kouhoku-ku, Yokohama 223-8522, Japan
[3] interdisciplinary Theoretical & Mathematical Sciences Program (iTHEMS) RIKEN 2-1, Hirosawa, Wako, Saitama 351-0198, Japan
[4] Key Laboratory of Quark & Lepton Physics (MOE) and Institute of Particle Physics, Central China Normal University, Wuhan 430079, China

In this paper we propose new algorithm to reduce autoc... algorithms for euclidean field theories on the lattice. Our pr... Carlo algorithm (HMC) with restricted Boltzmann machin... rithm by employing the phi-fourth theory in three dimensio... relation both in symmetric and broken phase as well. Our ... central values of expectation values of the action density ar... from the original HMC in both the symmetric phase and b... On the other hand, two-point Green's functions have slight ... HMC and one by our proposing algorithm in the symmetric ... the distribution of the one-point Green's function differs from the one from HMC. We discuss the

Dec 2017

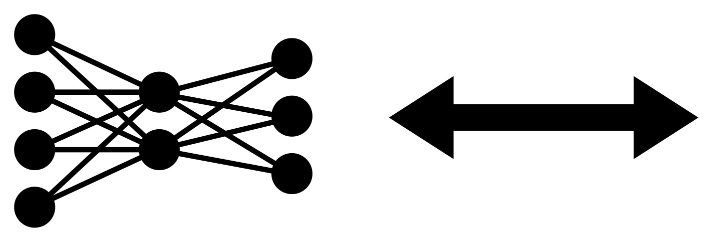




If we want to use generative models as lattice QCD sampler, we must guarantee the gauge symmetry of a probability distribution for the model. This is because,

明らかにしたこと

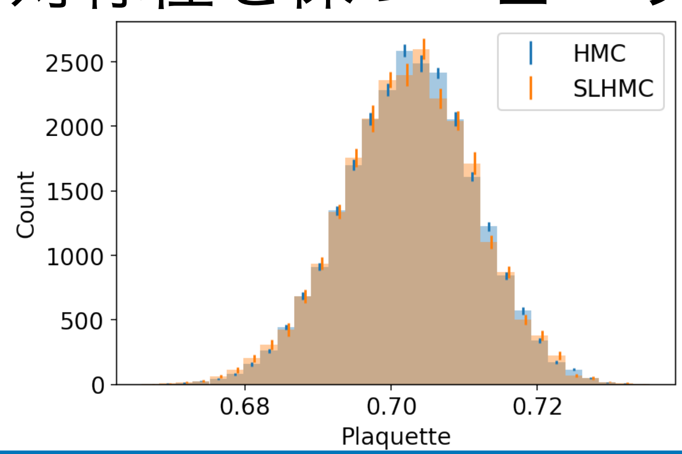
arXiv: 2103.11965

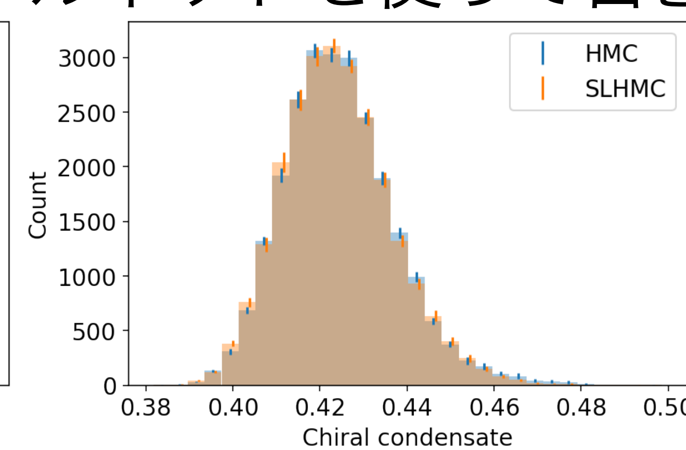
1. “非可換ゲージ対称性を保つニューラルネット” = (調整できる)スメアリング

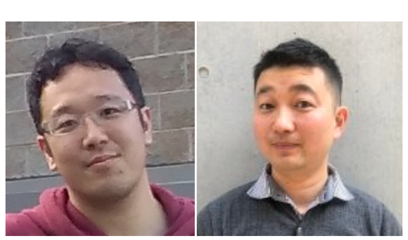


$$U_\mu^{\mathrm{NN}}(n)[U] = U_\mu^{(3)}(n)\left[U_\mu^{(2)}(n)\left[U_\mu^{(1)}(n)\left[U_\mu(n)\right]\right]\right]$$

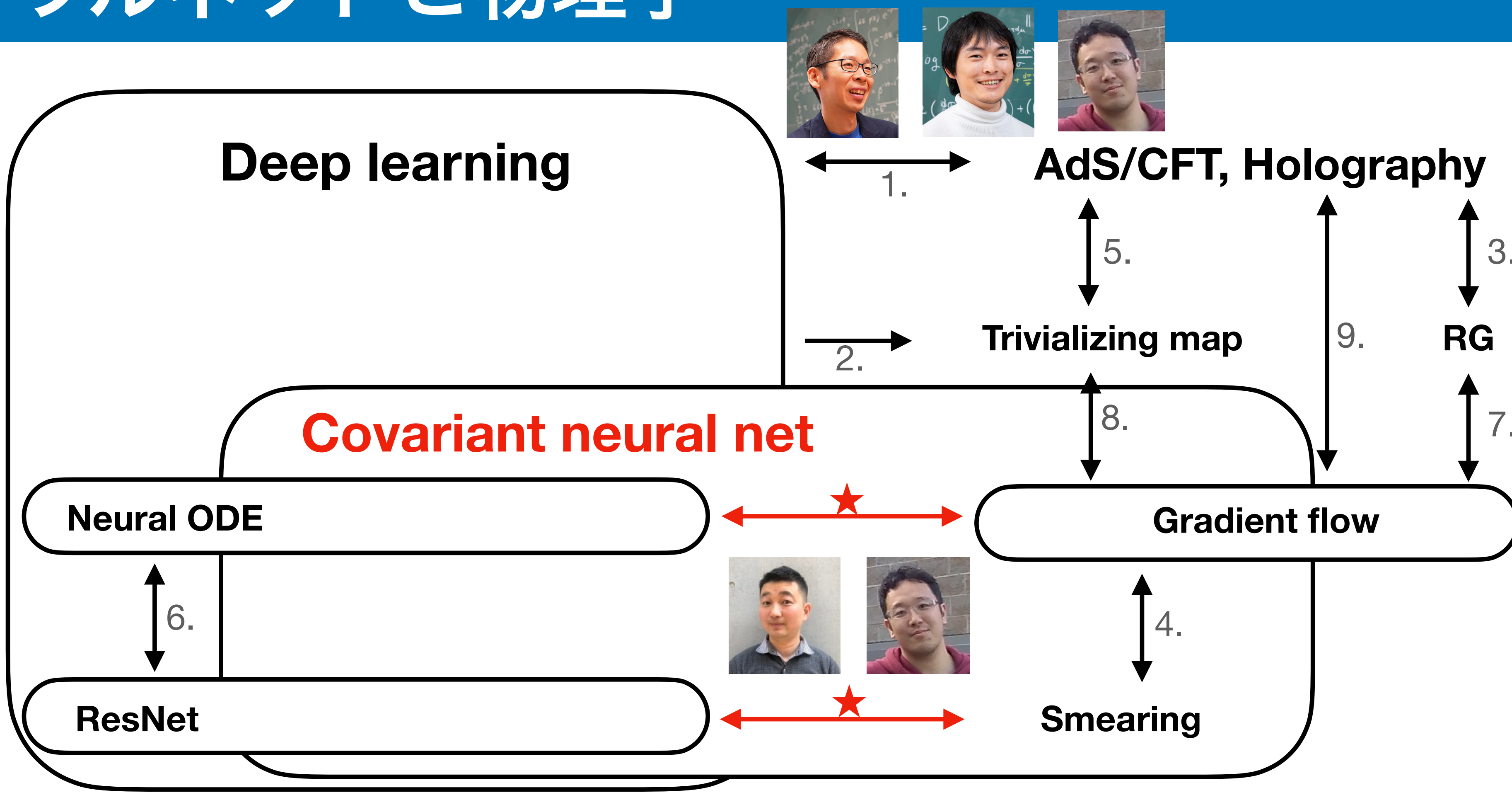
2. 非可換ゲージ対称性を保つニューラルネットを使って自己学習HMC。成功。







背景
ニューラルネットと物理学



黒: 先行研究
1.arXiv: 1802.08313
2.arXiv: 1904.12072
3.arXiv:hep-th/0212314
4.arXiv: 1006.4518
5.arXiv:1903.00804
6.arXiv: 1806.07366
7.arXiv: 1802.07897
8.arXiv: 0907.5491
9. arXiv: 1707.03982
赤: この仕事
★. arXiv: 2103.11965

今までの研究で知られていたこと

1.深層ニューラルネット　〜　AdS　〜　gradient flow

2.ニューラルネット　→　ニューラルODE

3. (ゲージ)対称性は、物理でもニューラルネットでも重要

本研究では、 上記の3つの文脈に沿ったニューラルネットの構造を発見

ゲージ共変ニューラルネットのニューラルODE　〜　“gradient flow” ということも見つけた

ニューラルネットと畳込み、フィルタ

(ここに書いてある説明です)

ニューラルネットとは？

線形変換と要素ごとの非線形変換の合成関数

行列

$$[W\vec{x}]_i = \sum_j w_{ij} x_j$$

任意の線形変換を表現できる

ニューラルネットの構成

$$u_i(x_j) = \begin{cases} z_i^{(l)} = \sum_j w_{ij}^{(l)} x_j + b_i^{(l)} & \text{アフィン変換} \\ u_i = \sigma^{(l)}(z_i^{(l)}) & \text{要素ごと (局所的)の関数} \end{cases}$$

全結合ニューラルネット

$$f_\theta(\vec{x}) = \sigma^{(l=2)}(W^{(l=2)}\sigma^{(l=1)}(W^{(l=1)}\vec{x}))$$

θ はパラメータの集合, $w_{ij}^{(l)}, b_i^{(l)}, \cdots$

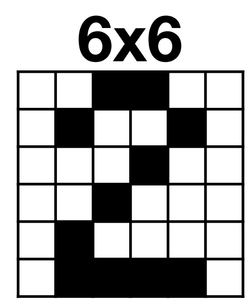
ニューラルネットはベクトルを入出力する複合関数

ニューラルネットとは？

Akio Tomiya

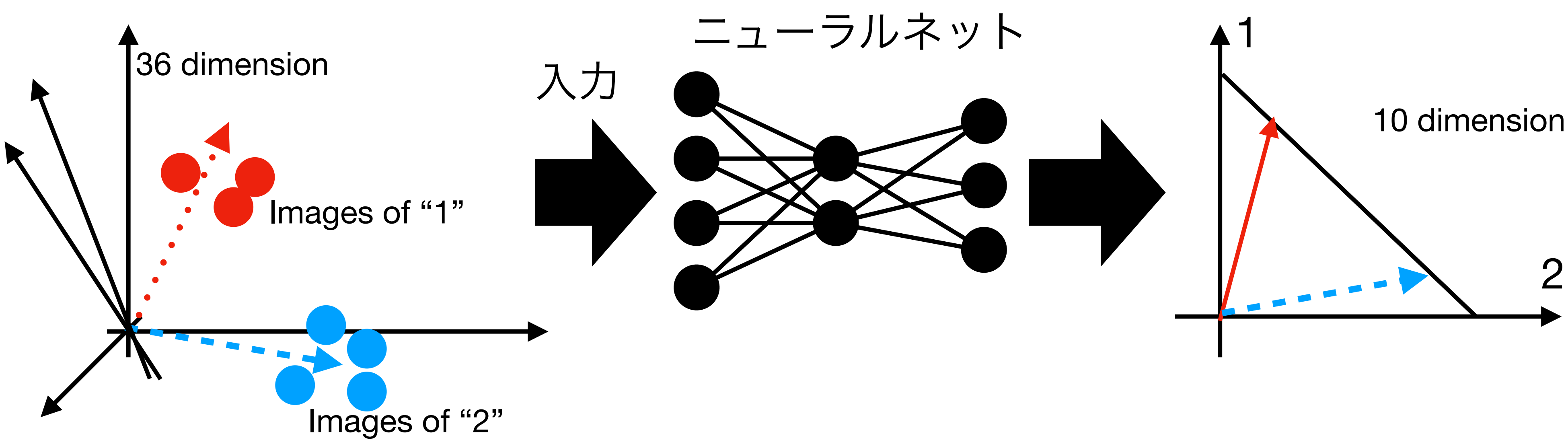
ニューラルネット = 万能関数近似器

~



$$= \begin{pmatrix} 0.000 \\ 0.000 \\ 0.8434 \\ 0.756 \\ 0.3456 \\ 0.64 \\ 0.251 \\ 0.11 \\ 0.23 \\ \vdots \end{pmatrix}$$

画像はベクトル



事実: ニューラルネットは、”どんな関数”でも近似できる！

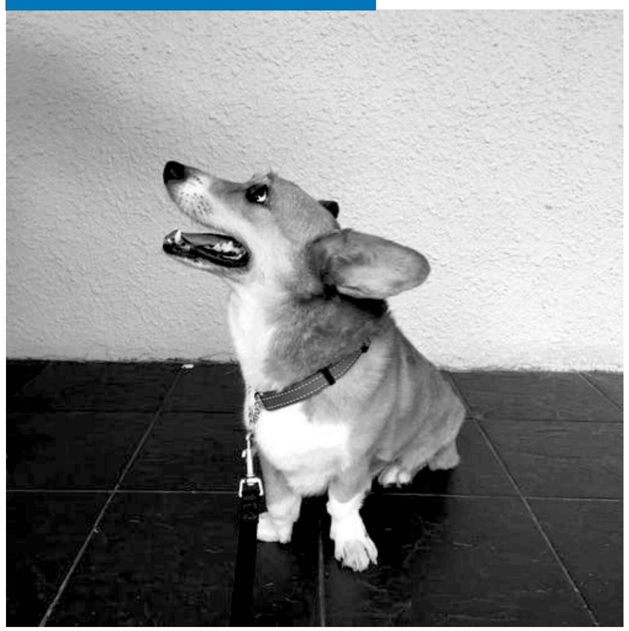
(直感的には、内部のユニット数 ~ フーリエ展開の項の数)

この例では、画像(36次元ベクトル)と数字のラベル(10次元ベクトル)を結ぶ写像を”真似た”

ニューラルネットとは？
畳み込みレイヤー = 学習可能フィルタ

Akio Tomiya

画像フィルタ



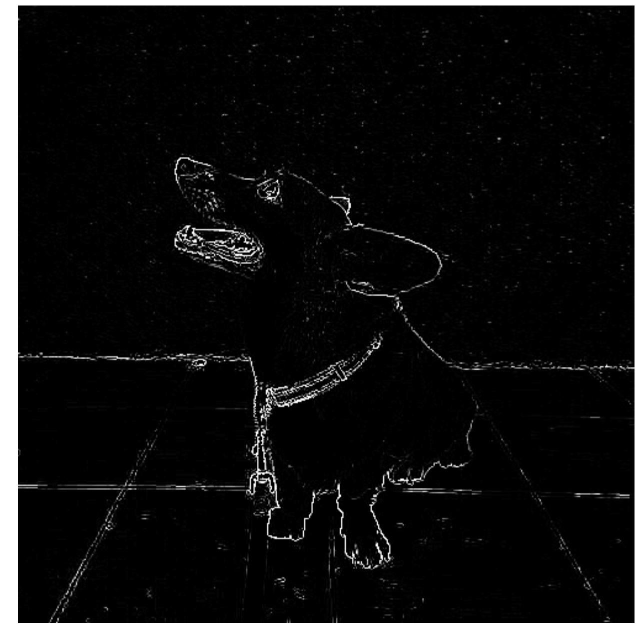
ラプラシアンフィルタ

$*$

0	1	0
1	-2	1
0	1	0

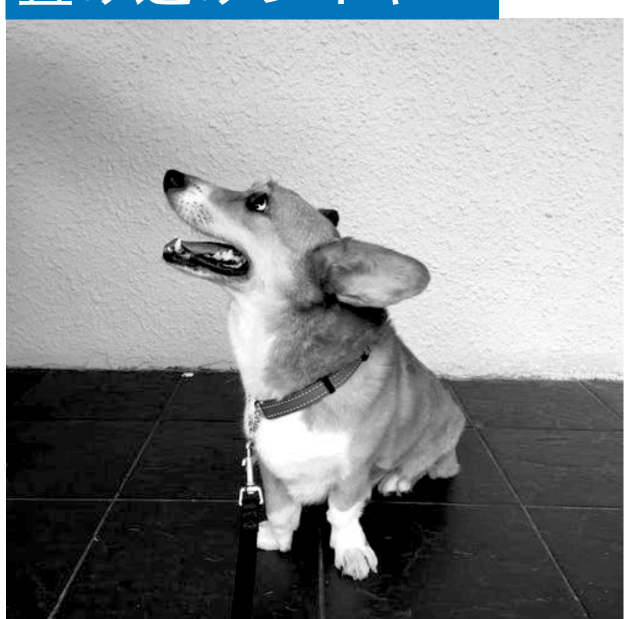
(∂^2 の離散化
= 色の変曲点を検出)

=



エッジ検出

畳み込みレイヤー



調整可能なフィルタ

$*$

w_{11}	w_{12}	w_{13}
w_{21}	w_{22}	w_{23}
w_{31}	w_{32}	w_{33}

=

エッジ検出

平滑化(ぼかし)

...

Fukushima, Kunihiko (1980)
Zhang, Wei (1988) + a lot!

平滑化: ガウシアンフィルタ

$\frac{1}{16}$

1	2	1
2	4	2
1	2	1

(どういうフィルタになるかはトレーニングによる)
うまく特徴を抽出する

ニューラルネットとは？
畳み込みも合成できる

もちろん畳み込みレイヤーも合成関数としてつかえる

Fukushima, Kunihiko (1980)
Zhang, Wei (1988) + a lot!

$$\left(\left(\quad * \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix}\right) * \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix}\right) * \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix}$$

1. 畳み込みが画像認識タスクの向上に役立ち、近年のブームの1つのきっかけになった

2. フィルターは画素の絶対的な座標には依存しない = 入力データの並進対称性を尊重する

下記の画像はどっちも「犬」である

現代的な視点:

対称性をつかうとニューラルネットの性能をあげられる (T. Cohen+, group equivariant NN)
(回転対称性: Spherical convolution T. Cohen+), 模倣する関数が対称性をもつことを保証している。

機械学習の文脈でも、「ゲージ対称性を持ったデータ」を
取り扱うことがある。専用のニューラルネットを作れる。

(T. Cohen+, gauge equivariant NN)

https://towardsdatascience.com/translational-invariance-vs-translational-equivariance-f9fbc8fca63a

ニューラルネットとは？
畳み込みレイヤー + 全結合ニューラルネットもOK

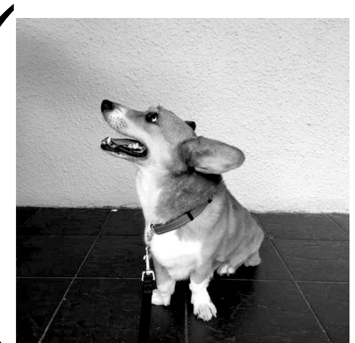
Akio Tomiya

Fukushima, Kunihiko (1980)
Zhang, Wei (1988) + a lot!

例:

$$f_\theta(\vec{x}) = \sigma^{(l=2)}(W^{(l=2)}\sigma^{(l=1)}(W^{(l=1)}\vec{v}))$$

$$\vec{v} = \mathrm{Flatten}\Bigg(\Bigg(\Bigg(\Bigg($$



$$* \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix}\Bigg) * \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix}\Bigg) * \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix}\Bigg)$$

畳み込みレイヤーのパラメータも全結合ネットと同じようにトレーニングできる

(現代的にはmulti-channel にして、Global average pooling とかにすべきだが、置いておく)

物理系に使うと重みから系の情報を抽出できたりする (AT+ 2016)

格子QCD

格子QCD

電磁気学 = U(1) ゲージ対称性を持つ理論

(相対論的記法の)電磁気学

$$S = \int d^4x \left[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi} \left(i\partial\!\!\!/ + eA\!\!\!/ - m \right) \psi \right]$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \qquad A_\mu(x) \in \mathbb{R} \quad \mu = 0,1,2,3$$

- 身の回りの電気・磁気現象を記述: F_{0i}が電場、F_{ij} が磁場 ($i, j = 1,2,3$)
- U(1)ゲージ対称性が鍵: Sは、以下の変換で不変

$$A_\mu(x) \to A_\mu(x) - \partial_\mu \Omega(x) \qquad \Omega(x) \in \mathbb{R}$$

相対論的量子電磁気学

$$[E(x), A(y)] \sim \delta(x - y) \qquad \text{量子化}$$

$$|\psi(t)\rangle = \mathrm{e}^{-iHt} |\psi(0)\rangle$$

H: 上記のSから作ったハミルトニアン

- 量子化した電磁気学
- 世界一精密な理論(実験と12桁あう)

格子QCD

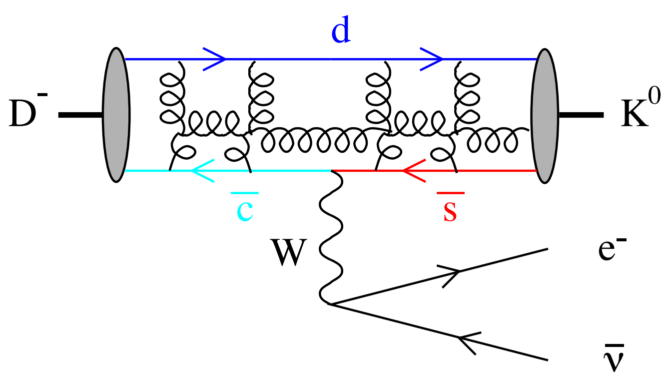
QCD = 量子色力学、電磁場を行列にした。クォークとグルーオンの理論

QCD (Quantum Chromo-dynamics) in 3 + 1 dimension

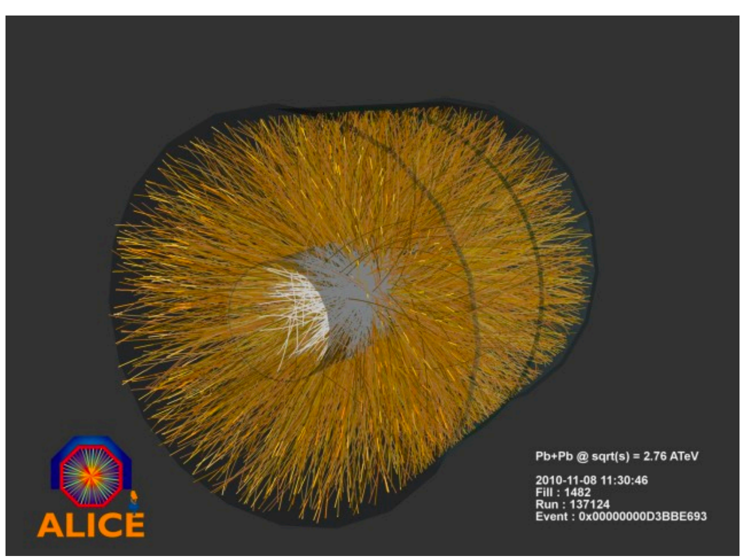
$$S = \int d^4x \left[-\frac{1}{2}\mathrm{tr}\, F_{\mu\nu}F^{\mu\nu} + \bar{\psi}\left(\mathrm{i}\partial\!\!\!/ + gA\!\!\!/ - m\right)\psi \right]$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - \mathrm{i}g[A_\mu, A_\nu]$$ $A_\mu(x) \in su(3)$, 3x3 トレースレス、エルミート行列

$$|\psi(t)\rangle = \mathrm{e}^{-\mathrm{i}Ht}|\psi(0)\rangle$$ H: 上記のSから作ったハミルトニアン



- 量子電磁気学の一般化 $A_\mu(x)$が行列 (Yang-Mills-内山)
- これ1つで以下が(原理的には)計算できる
  - 中性子星などの状態方程式、相転移温度、原子核間の力
  - 原子核内部のクォークとグルーオンの様子、散乱
  - 陽子と中性子の仲間(ハドロン)の質量, などなど
  - カイラル対称性の自発的破れ(南部)、とじこめ力(南部)
- g >> 1 なので電磁気の計算法は使えない
- 基底状態の性質を解き明かせば1億円もらえる (ミレニアム懸賞金)



格子QCD
ゲージ対称性が理論を決める

QCD (Quantum Chromo-dynamics) in 3 + 1 dimension

$$S = \int d^4x \left[-\frac{1}{2} \mathrm{tr}\, F_{\mu\nu}F^{\mu\nu} + \bar{\psi}\left(\mathrm{i}\partial\!\!\!/ + gA\!\!\!/ - m\right)\psi \right]$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - \mathrm{i}g[A_\mu, A_\nu]$$ $A_\mu(x) \in su(3)$, 3x3 トレースレス、エルミート行列

$$|\psi(t)\rangle = \mathrm{e}^{-\mathrm{i}Ht}|\psi(0)\rangle$$ H: 上記のSから作ったハミルトニアン

U(1)のゲージ変換: $\alpha(x) \in \mathbb{R}$ $\quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$

$$\left\{ \begin{array}{l} A_\mu(x) \to A_\mu(x) - \partial_\mu \alpha(x) \\ \psi(x) \to \mathrm{e}^{\mathrm{i}\alpha(x)}\psi(x) \end{array} \right\}$$ ゲージ共変 covariant

$$F_{\mu\nu}F^{\mu\nu} \to F_{\mu\nu}F^{\mu\nu}$$ ゲージ不変 invariant

$$\left(\mathrm{i}\partial\!\!\!/ + gA\!\!\!/ - m\right)\psi \to \mathrm{e}^{\mathrm{i}\alpha(x)}\left(\mathrm{i}\partial\!\!\!/ + gA\!\!\!/ - m\right)\psi$$ ゲージ共変

$$\bar{\psi}\left(\mathrm{i}\partial\!\!\!/ + gA\!\!\!/ - m\right)\psi \to \bar{\psi}\left(\mathrm{i}\partial\!\!\!/ + gA\!\!\!/ - m\right)\psi$$ ゲージ不変

ゲージ対称性が作用の形を決定する(ゲージ原理)

格子QCD

実時間でも虚時間でも同じハミルトニアン、得意な計算が違う

4次元ユークリッド時空の経路積分表示のQCD

$$S = \int d^4x \left[+ \frac{1}{2} \mathrm{tr}\, F_{\mu\nu} F_{\mu\nu} + \bar{\psi}\left(\partial\!\!\!/ - \mathrm{i}g A\!\!\!/ + m\right)\psi \right]$$

$$Z = \int \mathscr{D}A \mathscr{D}\bar{\psi} \mathscr{D}\psi e^{-S} \longleftrightarrow |\psi(\tau)\rangle = \mathrm{e}^{-H\tau} |\psi(0)\rangle$$

- 3+1次元ミンコフスキー時空の理論でも4次元ユークリッド時空の理論でも同じハミルトニアン、時間発展が違う
- ハミルトニアンの固有値(ハドロンの質量！)などは同じ。
- (時間依存するもの, 例えば n点グリーン関数は解析接続の関係)

$$Z = \int \mathscr{D}A \mathscr{D}\bar{\psi} \mathscr{D}\psi e^{-S}$$ この積分ができればいい

格子QCD

格子化=時空の離散化。巨視的な物理量の期待値は同じ

K. Wilson 1974

$$S = \int d^4x \left[+ \frac{1}{2} \mathrm{tr}\, F_{\mu\nu} F_{\mu\nu} + \bar{\psi}\left(\partial\!\!\!/ - ig A\!\!\!/ + m \right) \psi \right]$$

格子正則化

$U_\mu = \mathrm{e}^{aigA_\mu}$

$$S[U, \psi, \bar{\psi}] = a^4 \sum_n \left[-\frac{1}{g^2} \mathrm{Re}\ \mathrm{tr}\ U_{\mu\nu} + \bar{\psi}\left(D\!\!\!\!/ + m \right) \psi \right]$$

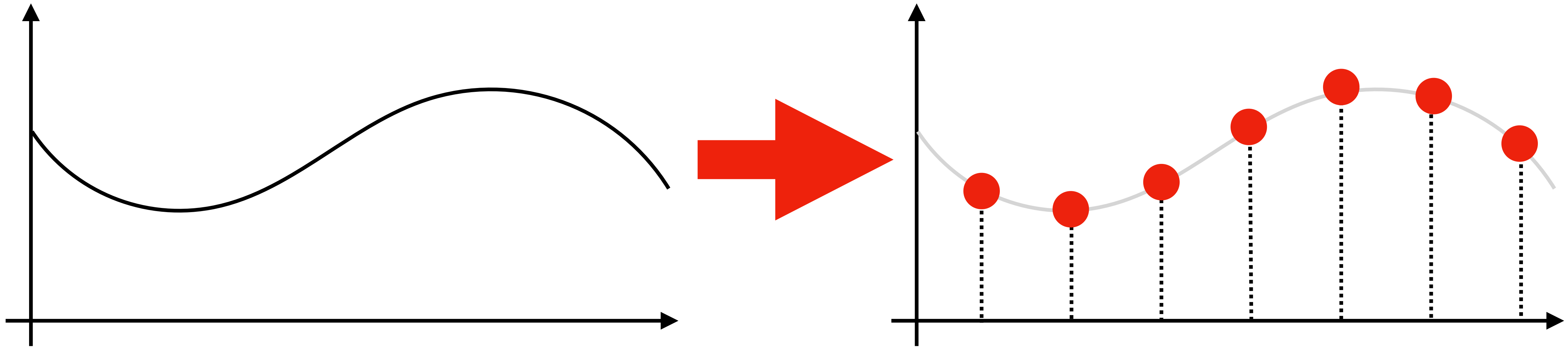
a は格子間隔 (カットオフa^{-1})

$$\mathrm{Re}\ U_{\mu\nu} \sim \frac{-1}{2} g^2 a^4 F_{\mu\nu}^2 + O(a^6)$$

どちらも(低エネルギーでは)同じ期待値を返す

(無限個のIrrelevant operatorsしか違わない)

例



積分値は(離散化が十分細ければ)精度の範囲で同じにできる

格子QCD
(2次元時空の)場 ~ 画像

Akio Tomiya

K. Wilson 1974

$$S = \int d^4x \left[+ \frac{1}{2} \mathrm{tr}\, F_{\mu\nu}F_{\mu\nu} + \bar{\psi}\left(\partial\!\!\!/ - \mathrm{i}gA\!\!\!/ + m\right)\psi \right]$$

格子正則化

$U_\mu = \mathrm{e}^{aigA_\mu}$

$$S[U, \psi, \bar{\psi}] = a^4 \sum_n \left[- \frac{1}{g^2} \mathrm{Re\ tr}\ U_{\mu\nu} + \bar{\psi}\left(D\!\!\!\!/ + m\right)\psi \right]$$

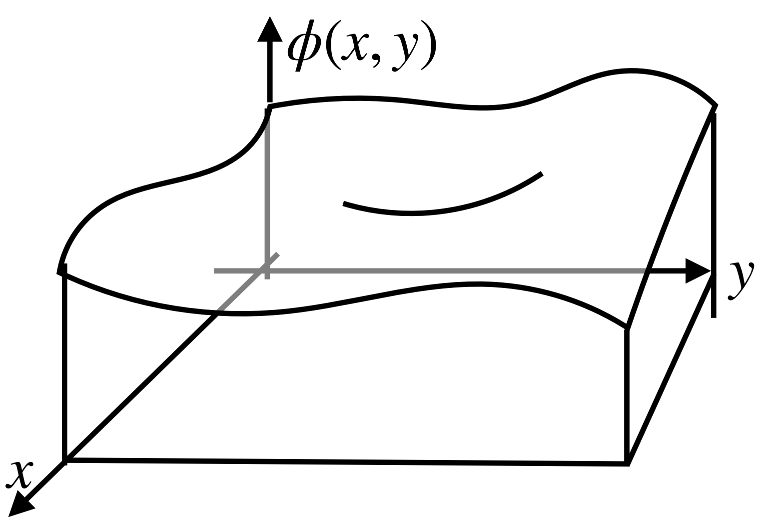
a は格子間隔 (カットオフa^{-1})

$$\mathrm{Re}\ U_{\mu\nu} \sim \frac{-1}{2} g^2 a^4 F_{\mu\nu}^2 + O(a^6)$$

どちらも(低エネルギーでは)同じ期待値を返す

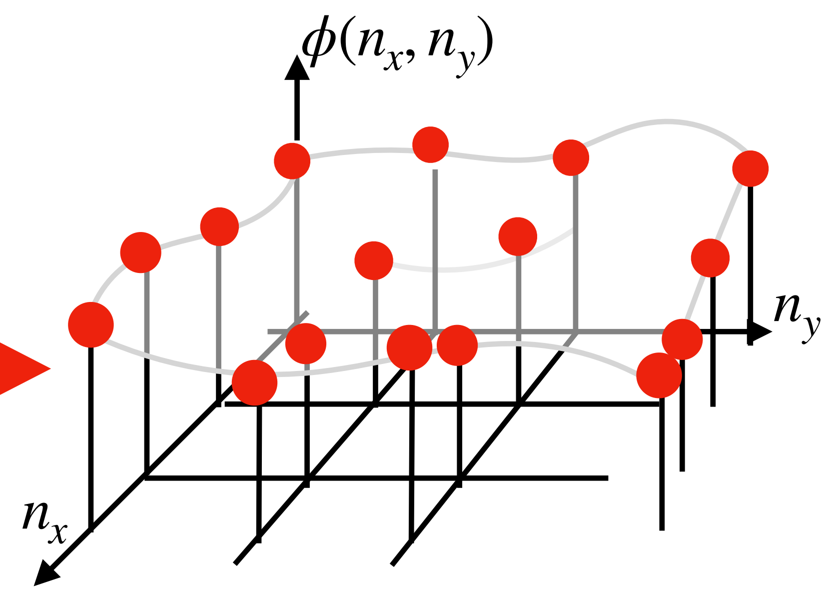
(無限個のIrrelevant operatorsしか違わない)

２次元時空、スカラー場での図示



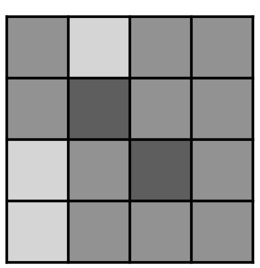
連続時空上の場

格子正則化



格子上の場

=



上から見た

(色の濃さ = 高さ)

画像みたいなもの

格子QCD

格子化上でもゲージ変換は定義できる。作用はゲージ不変

Akio Tomiya

K. Wilson 1974

$$S = \int d^4x \left[+ \frac{1}{2} \mathrm{tr}\, F_{\mu\nu} F_{\mu\nu} + \bar{\psi} \left(\partial\!\!\!/ - \mathrm{i} g A\!\!\!/ + m \right) \psi \right]$$

格子正則化

$U_\mu = \mathrm{e}^{aigA_\mu}$

$$S[U, \psi, \bar{\psi}] = a^4 \sum_n \left[- \frac{1}{g^2} \mathrm{Re}\ \mathrm{tr}\ U_{\mu\nu} + \bar{\psi} \left(D\!\!\!\!/ + m \right) \psi \right]$$

a は格子間隔 (カットオフa^{-1})

$$\mathrm{Re}\ U_{\mu\nu} \sim \frac{-1}{2} g^2 a^4 F_{\mu\nu}^2 + O(a^6)$$

どちらも(低エネルギーでは)同じ期待値を返す

(無限個のIrrelevant operatorsしか違わない)

ゲージ変換は格子上のほうが単純

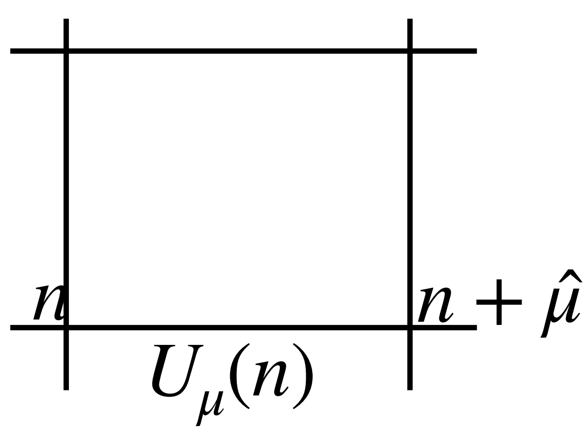
$$\boxed{A_\mu(x) \to G(x) A_\mu(x) G^{-1}(x) - G(x) \partial_\mu G^{-1}(x)}$$

格子正則化

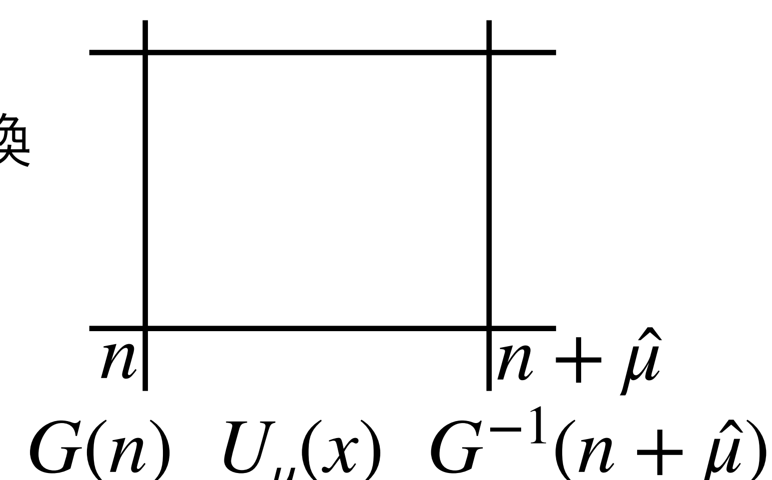
$$\boxed{U_\mu(x) \to G(n) U_\mu(x) G^{-1}(n + \hat{\mu})}$$

ゲージ場が辺に乗ってて

ゲージ変換は点の上の関数



ゲージ変換 $\longrightarrow$



格子QCD
格子化した経路積分

K. Wilson 1974

$$S=\int d^4x\left[+\frac{1}{2}\mathrm{tr}\,F_{\mu\nu}F_{\mu\nu}+\bar{\psi}\left(\partial\!\!\!/ -\mathrm{i}gA\!\!\!/ +m\right)\psi\right]$$

格子正則化

$U_\mu = \mathrm{e}^{aigA_\mu}$

$$S[U,\psi,\bar{\psi}]=a^4\sum_n\left[-\frac{1}{g^2}\mathrm{Re\ tr}\ U_{\mu\nu}+\bar{\psi}\left(D\!\!\!\!/ +m\right)\psi\right]$$

a は格子間隔 (カットオフa^{-1})

どちらも(低エネルギーでは)同じ期待値を返す

(無限個のIrrelevant operatorsしか違わない)

$$\mathrm{Re}\ U_{\mu\nu}\sim\frac{-1}{2}g^2a^4F^2_{\mu\nu}+O(a^6)$$

$$\langle\mathcal{O}\rangle=\frac{1}{Z}\int\mathcal{D}U\mathcal{D}\bar{\psi}\mathcal{D}\psi e^{-S}\mathcal{O}(U)=\frac{1}{Z}\int\mathcal{D}Ue^{-S_{\mathrm{gauge}}[U]}\det(D+m)\mathcal{O}(U)$$

$$=\frac{1}{Z}\int\mathcal{D}Ue^{-S_{\mathrm{eff}}[U]}\mathcal{O}(U)$$

$$S_{\mathrm{eff}}[U]=S_{\mathrm{gauge}}[U]-\log\det(D\!\!\!\!/[U]+m)$$

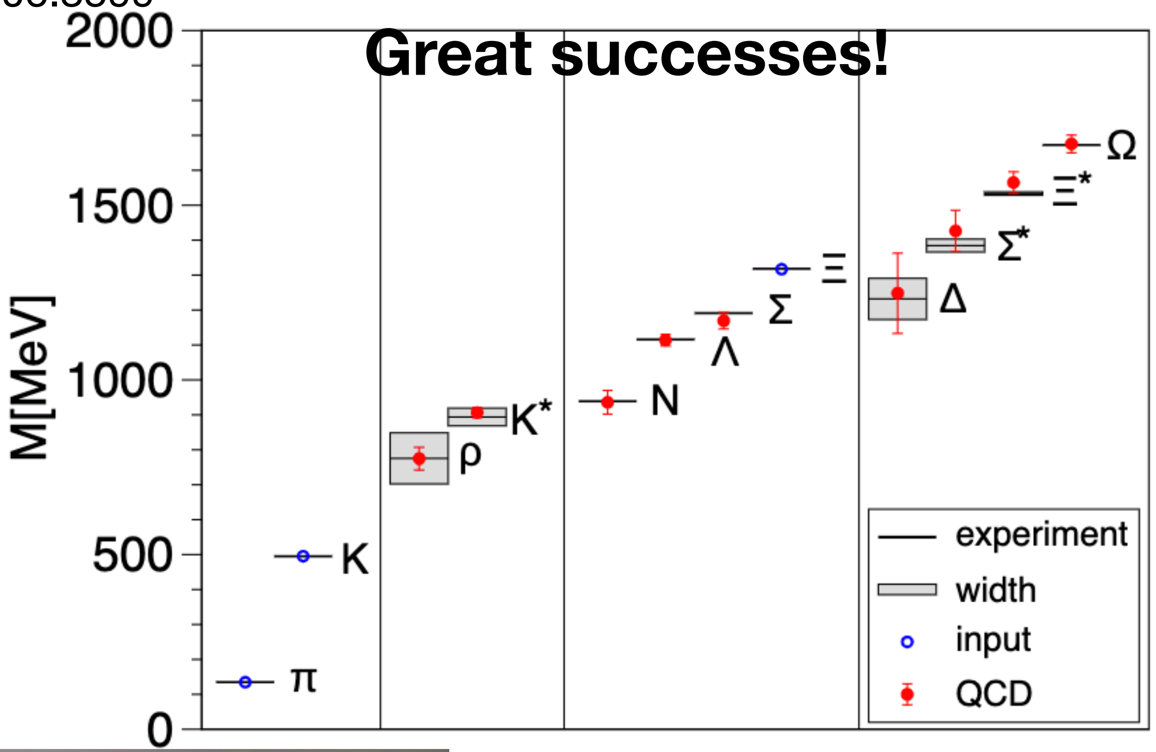
この積分を計算すれば良い。

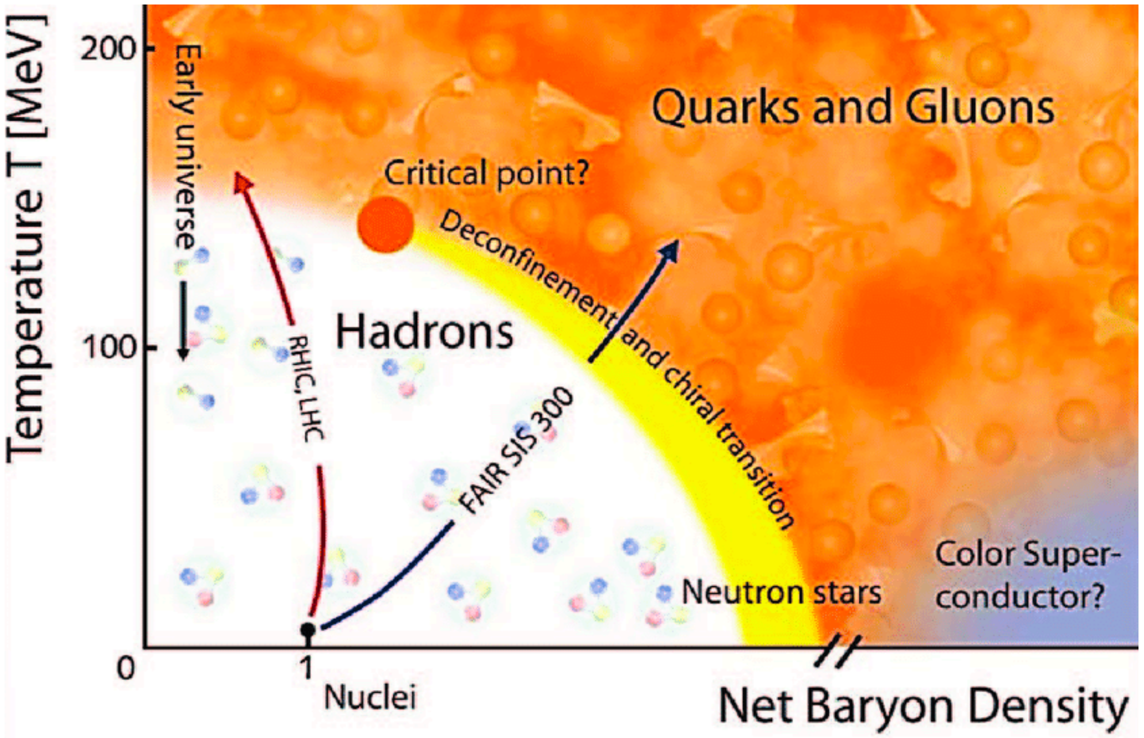
格子QCD

Akio Tomiya

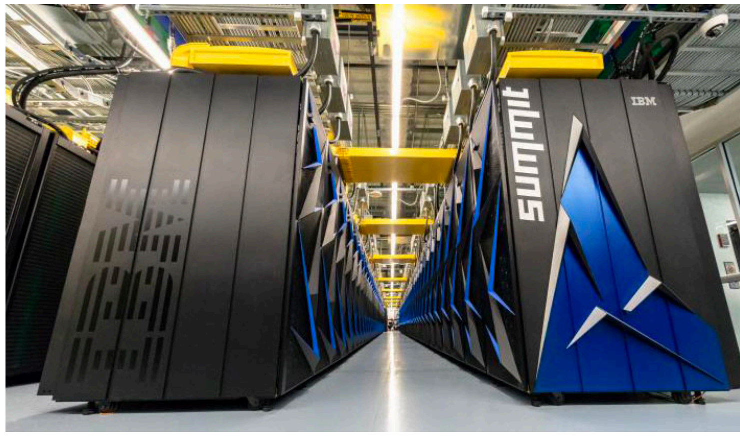
格子QCDは定量的。実験と理論を結ぶ重要、場の理論の計算道具

arXiv:0906.3599





他にわかること？
- 核力ポテンシャル
- エンタングルメントエントロピー
- 形状因子
- パートン分布,　　etc…



格子QCD
小さいサイズならノートパソコンでできます

Juliaの公開コードを作りました！

https://github.com/akio-tomiya/LatticeQCD.jl

お手元のパソコンで簡単にできます

HMC/heatbath/SLMC + Measurements

Fortranのコードと同等のスピード

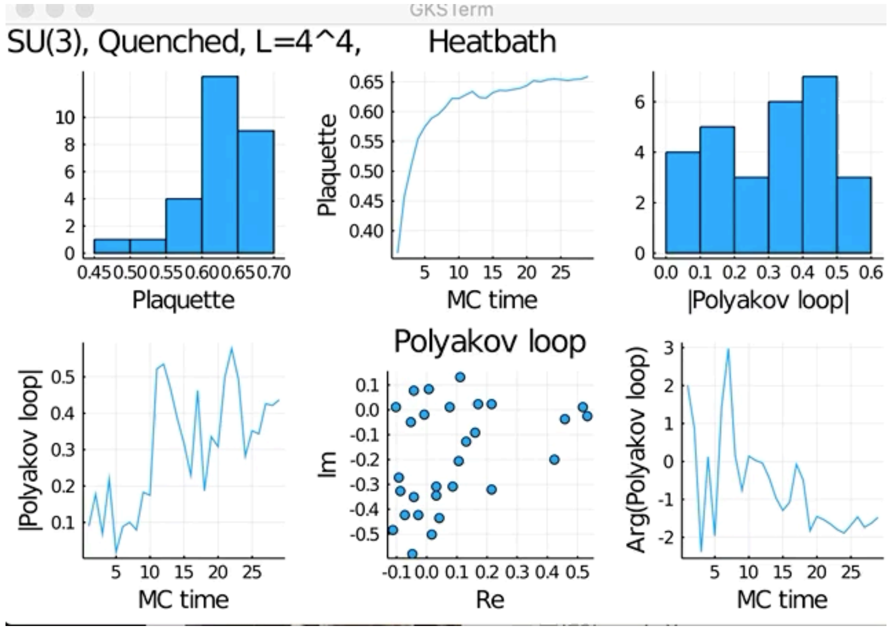
You can start in 3 steps (in 10 min)

1. Download Julia binary
2. Add the package through Julia function
3. Execute!

LatticeQCD.jl


A.T.


Y. Nagai



格子QCD

Remarks

Akio Tomiya

量子ゲージ論の格子化は数値計算のためだけではない

場の理論の1つの定義

場の理論の経路積分を離散化してるので、定義になってる。

リーマン和で積分を定義してるようなもの。ハミルトン形式もうまくいく。

$$\sum_i f(x_i)\delta x_i \to \int dx f(x)$$

J. Kogut+ 1975
M. Creutz 1976
M. Luscher 1977
K. Osterwalder+ 1977

閉じ込めやメソンの性質の探索

Wilson は格子正則化した非可換ゲージ理論の強結合展開から閉じ込めが理解できることを見た

さらにメソンも導入して理論の構造も理解できる。摂動を超えて理解できる例。

K. Wilson 1974
N. Kawamoto+ 1981

物性にもつかえる

Staggered fermion: グラフェン

Wilson fermion: トポロジカル絶縁体

A. Sekine+ arXiv: 1301.4424
M. Hirotsu arXiv: 1303.2886
A. G. Grushin arXiv: 1909.02983

双対性の計算

くりこみ群の上流で格子化した作用を持ってきて、違った場をintegrating out

すると双対な理論が得られる

T. Furusawa+ arXiv: 1810.11808

数学における指数定理(Atiyah-Patodi-Singer 等)の証明

正則化したディラック演算子(共変微分)の固有モードとゲージ場のトポロジーが理解できる

境界付き多様体の上のゲージ場のトポロジーとディラック演算子の0モードの関係

H. Fukaya+ 2017 etc

スメアリング

ゲージ対称性をたもった平滑化

M. Albanese+ 1987
R. Hoffmann+ 2007
C. Morningster+ 2003

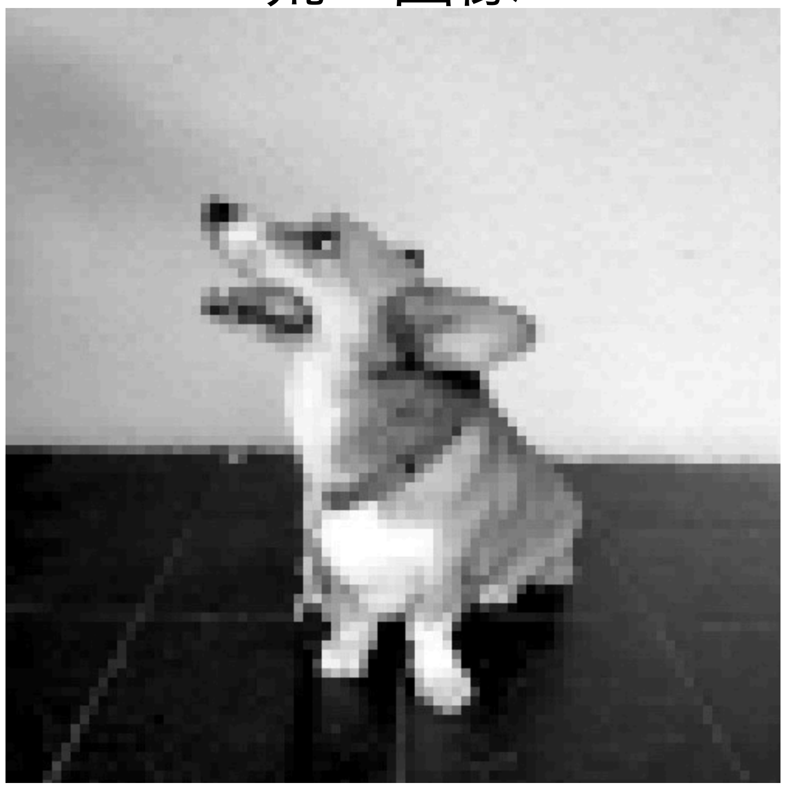
スメアリング

平滑化すると性質が良くなる

Akio Tomiya

例

荒い画像



ピクセルの数値微分がガタガタ

ガウシアンフィルタ

$$\frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \longrightarrow$$

平滑化した画像

ピクセルの数値微分がなめらか

細かい構造は失われるが
大域的構造(トポロジーとか)はOK

トポロジーを調べるには、微分の積分をする(ガウスボネ)

格子QCDでも同じことできる？
ゲージ対称性を保ちながら。

2タイプの平滑化:

APE-type smearing (こっちだけ説明)

Stout-type smearing

M. Albanese+ 1987
R. Hoffmann+ 2007
C. Morningster+ 2003

スメアリング
ゲージ対称性を保つ平滑化操作

M. Albanese+ 1987
R. Hoffmann+ 2007

APE-type smearing, $\alpha \in \mathbb{R}_{>0}$ **は定数**

$$U_\mu(n) \to U_\mu^{\mathrm{fat}}(n) = \mathcal{N}\left[(1-\alpha)U_\mu(n) + \frac{\alpha}{6}V_\mu^\dagger[U](n)\right]$$

規格化

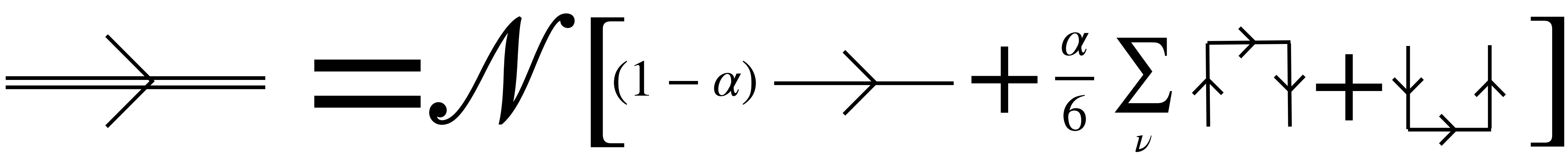
$$\mathcal{N}[M] = \frac{M}{\sqrt{M^\dagger M}}$$

もしくは射影

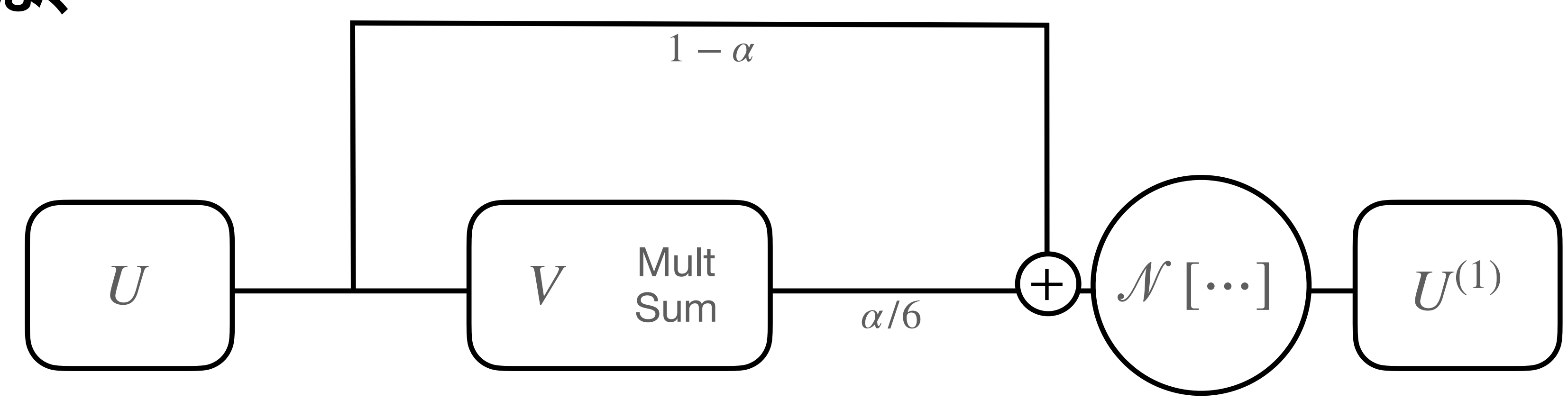
$$V_\mu^\dagger[U](n) = \sum_{\mu\neq\nu} U_\nu(n)U_\mu(n+\hat{\nu})U_\nu^\dagger(n+\hat{\mu}) + \cdots$$

$V_\mu^\dagger[U](n)$と$U_\mu(n)$は同じゲージ変換性を持つ
→$U_\mu^{\mathrm{fat}}[U](n)$も同じゲージ変換性を持つ

図的に書くと



計算グラフ的には、



スメアリング

スメアリングは2つの操作に分解できる

APEtype もstout typeもをまとめて、以下のように書ける

$$U_\mu^{\mathrm{fat}}(n) = \begin{cases} z_\mu(n) = w_1 U_\mu(n) + w_2 \mathcal{G}[U] & \text{ゲージ対称性を保った和} \\ U_\mu^{\mathrm{fat}}(n) = \mathcal{N}(z_\mu(n)) & \text{局所的な関数} \end{cases}$$

スメアリング

スメアリング ～ ニューラルネット！

AT Y. Nagai arXiv: 2103.11965

APEtype もstout typeもをまとめて、以下のように書ける

$$U_\mu^{\mathrm{fat}}(n) = \begin{cases} z_\mu(n) = w_1 U_\mu(n) + w_2 \mathcal{G}[U] \\ U_\mu^{\mathrm{fat}}(n) = \mathcal{N}(z_\mu(n)) \end{cases}$$

ゲージ対称性を保った和

局所的な関数

ニューラルネットを思い出すと、同じ構造をもつ

$$u_i(x_j) = \begin{cases} z_i^{(l)} = \sum_j w_{ij}^{(l)} x_j + b_i^{(l)} \\ u_i = \sigma^{(l)}(z_i^{(l)}) \end{cases}$$

アフィン変換

要素ごと (局所的)
の関数

(ニューラルネットのi の添字は、n やμ というインデックスに対応。ニューラルネットは物理屋の言葉で"スカラー場"の発展)

多段のスメアリング = ディープラーニング (ただし従来はパラメータ固定)

実は、スメアリングも普通のニューラルネットと同様に"学習"できる

Gauge covariant neural network

ゲージ共変ネットワーク

AT Y. Nagai arXiv: 2103.11965

ゲージ共変ネットワーク
トレーニングできるスメアリング

AT Y. Nagai arXiv: 2103.11965

ゲージ共変ネットワーク = トレーニングできるスメアリング

$$U_\mu^{(l+1)}(n)\left[U^{(l)}\right] = \begin{cases} z_\mu^{(l+1)}(n) = w_1^{(l)} U_\mu^{(l)}(n) + w_2^{(l)} \mathcal{G}_{\bar\theta}^{(l)}[U] \\ \mathcal{N}(z_\mu^{(l+1)}(n)) \end{cases}$$

(重みw は場所と向きに依存しても良い = 全結合的になる。対称性の情報は使わない)

$$U_\mu^{\mathrm{NN}}(n)[U] = U_\mu^{(3)}(n)\left[U_\mu^{(2)}(n)\left[U_\mu^{(1)}(n)\left[U_\mu(n)\right]\right]\right]$$

良い性質: ゲージ対称性、格子上の並進対称性、格子上の回転対称性が明白

(畳み込みニューラルネットと同様に、系にある対称性の情報を活用)

$$U_\mu(n) \mapsto U_\mu^{\mathrm{NN}}(n) = U_\mu^{\mathrm{NN}}(n)[U]$$

ゲージ場を入れたら、ゲージ場が出てくる複合関数、対称性を保つ

ゲージ共変ネットワーク
トレーニングは、(拡張した)backprop でできる

AT Y. Nagai arXiv: 2103.11965

ゲージ不変なloss function を組むには、ゲージ不変な作用を使えば良い

$$S^{\rm NN}[U] = S\left[U_\mu^{\rm NN}(n)[U]\right]$$

S:ゲージ作用 or フェルミオン作用
(Behler-Parrinello type neural net)
(DLAP2020の永井さんのスライド参照)

Loss function $$L_\theta[U] = f\left(S^{\rm NN}[U]\right)$$

f は例えば mean-square,
ミニバッチにもできる

トレーニング: gradient descent を使える。(Adamなどももちろん OK)

$$\theta^{(l)} \leftarrow \theta^{(l)} - \eta \frac{\partial L_\theta[U]}{\partial \theta^{(l)}}$$

$\theta^{(l)}$ は第l層のパラメータ

Lの微分は、行列微分が出てくるのでbackprop (delta rule)を拡張する必要あり

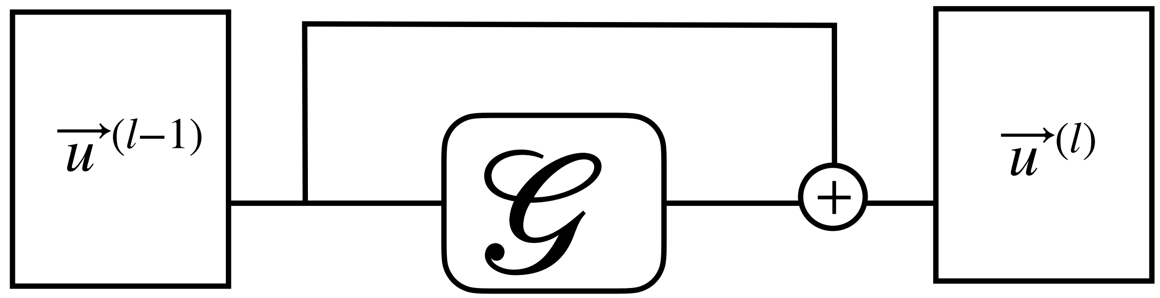
$$\frac{\partial L_\theta[U]}{\partial \theta^{(l)}} = \frac{\partial L}{\partial f}\frac{\partial f}{\partial S^{\rm NN}}\underline{\frac{\partial S^{\rm NN}}{\partial U^{(l+1)}}}\frac{\partial U^{(l+1)}}{\partial z^{(l+1)}}\frac{\partial z^{(l+1)}}{\partial \theta^{(l)}}$$

この部分は実はよく知られたsmeared force と同じ
(-> Extended delta rule, skipped)

ゲージ共変ネットワーク

Neural ODE of Cov-Net = “gradient flow”

Res-Net



arXiv: 1512.03385

↓ Layer の連続極限

Neural ODE

$$\frac{d\vec{u}^{(t)}}{dt} = \mathcal{G}(\vec{u}^{(t)})$$

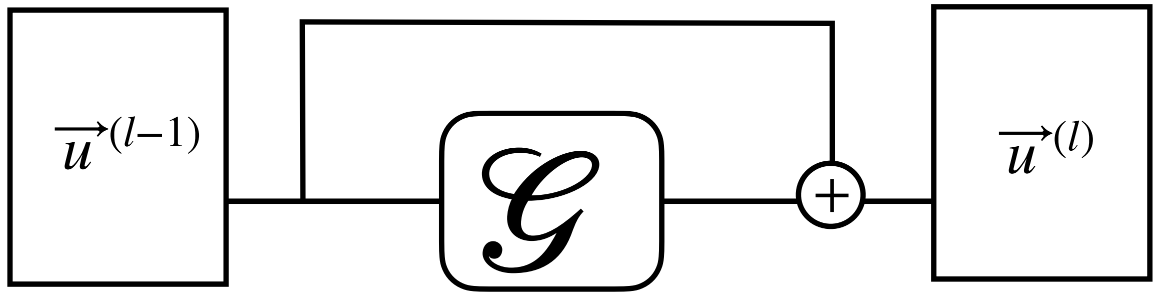
arXiv: 1806.07366
(Neural IPS 2018 のbest paper)

ゲージ共変ネットワーク

Neural ODE of Cov-Net = “gradient flow”

Akio Tomiya

Res-Net



arXiv: 1512.03385

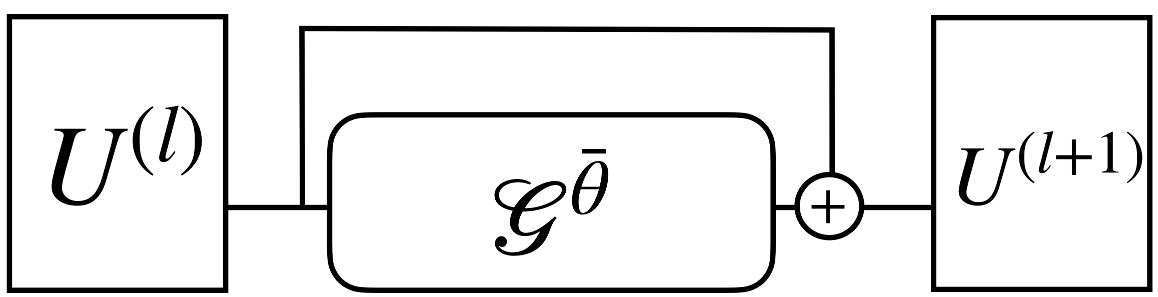
Layer の
連続極限

Neural ODE

$$\frac{d\vec{u}^{(t)}}{dt} = \mathcal{G}(\vec{u}^{(t)})$$

arXiv: 1806.07366
(Neural IPS 2018 のbest paper)

Gauge-cov net



AT Y. Nagai arXiv: 2103.11965

Layer の
連続極限

Neural ODE
for Gauge-cov net

$$\frac{dU_\mu^{(t)}(n)}{dt} = \mathcal{G}^{\bar{\theta}}(U_\mu^{(t)}(n))$$

Gradient flow (と同じ形)

“連続版のStout はGradient flow”

2010 M. Luscher

ゲージ共変ネットワーク
ここまでのまとめ

Akio Tomiya

	対称性	パラメータ固定	レイヤーの連続化	学習
今までのニューラルネット	畳み込み: 並進対称性	畳み込み: 画像フィルタ 境界検出、平滑化	Res-Net -> Neural ODE	デルタルール
ゲージ共変ネット AT Y. Nagai arXiv: 2103.11965	ゲージ対称性、 並進対称性、 回転対称性	スメアリング (平滑化)	“Gradient” flow	連続群に値を 取る関数用の デルタルール

次のページから応用編

(スメアリングを可変にしたところで役に立つのか)

Demonstration
(応用の一例)

AT Y. Nagai arXiv: 2103.11965

格子QCD

Akio Tomiya

格子化した経路積分 = 1000次元積分: 台形則は宇宙年齢でも計算時間足りない

K. Wilson 1974

$$S = \int d^4x \left[+ \frac{1}{2} \mathrm{tr}\, F_{\mu\nu} F_{\mu\nu} + \bar{\psi}\left(\partial\!\!\!/ - \mathrm{i}g A\!\!\!/ + m\right)\psi \right]$$

格子正則化 $\Longrightarrow$

$U_\mu = \mathrm{e}^{aigA_\mu}$

$$S[U, \psi, \bar{\psi}] = a^4 \sum_n \left[-\frac{1}{g^2} \mathrm{Re}\ \mathrm{tr}\ U_{\mu\nu} + \bar{\psi}\left(D\!\!\!\!/ + m\right)\psi \right]$$

a は格子間隔 (カットオフa^{-1})

$$\mathrm{Re}\ U_{\mu\nu} \sim \frac{-1}{2} g^2 a^4 F_{\mu\nu}^2 + O(a^6)$$

どちらも(低エネルギーでは)同じ期待値を返す

(無限個のIrrelevant operatorsしか違わない)

$$\langle \mathcal{O} \rangle = \frac{1}{Z} \int \mathcal{D}U \mathcal{D}\bar{\psi} \mathcal{D}\psi e^{-S} \mathcal{O}(U) = \frac{1}{Z} \int \mathcal{D}U e^{-S_{\mathrm{gauge}}[U]} \det(D + m) \mathcal{O}(U)$$

$$= \frac{1}{Z} \int \mathcal{D}U e^{-S_{\mathrm{eff}}[U]} \mathcal{O}(U)$$

$$\mathcal{D}U = \prod_{n \in \{\mathbb{Z}/L\}^4} \prod_{\mu=1}^{4} dU_\mu(n)$$

>1000 次元。台形則やシンプソン則などの

ニュートンコーツ型の数値積分は計算できない

(誤差をコントロールしようとすると計算時間が宇宙年齢を超える)

格子QCD

格子化した経路積分 〜 統計力学。統計力学の計算法が使える

K. Wilson 1974
M. Creutz 1980

格子QCD (経路積分表示の場の量子論) $S_{\text{eff}} = \frac{1}{g^2}\sum_n \sum_{\mu<\nu} \text{tr}\, U_\mu(n)U_\nu(n+\hat{\mu})U_\mu^\dagger(n+\hat{\nu})U_\nu^\dagger(n) + \cdots$

$U_\mu(n) \in \text{SU}(N)$

$$\langle \mathcal{O} \rangle = \frac{1}{Z}\int \mathcal{D}U e^{-S_{\text{eff}}[U]}\mathcal{O}(U)$$

$\int \mathcal{D}U$ 可能なすべてのUの配位の和

イジング模型 (統計力学) $\frac{1}{T}H[\sigma] = -\frac{1}{T}\sum_{<i,j>} \sigma_i\sigma_j$

$\sigma_i \in \mathbb{Z}_2$

$$\langle \mathcal{O} \rangle = \frac{1}{Z}\sum_{\{\sigma\}} e^{-\frac{1}{T}H[\sigma]}\mathcal{O}(\sigma)$$

$\sum_{\{\sigma\}}$ 可能なすべてのσの配位の和

見た目も構造もだいたい同じ

場の量子論といえど、統計力学の計算手法がつかえる

1. 高温展開(QCDの強結合展開), 低温展開(QCDの弱結合展開)
2. くりこみ群 (歴史的には場の理論→統計系?)
3. 数値計算、特にモンテカルロ法

以下ではモンテカルロ法に注目

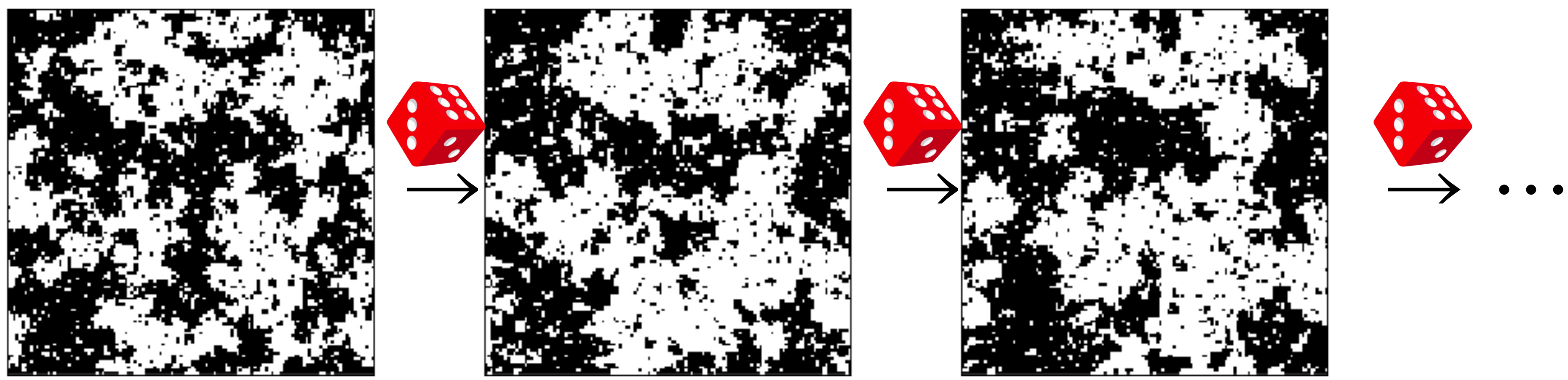
格子QCD
モンテカルロ積分の誤差は次元性によらない

M. Creutz 1980

$$\langle \mathcal{O} \rangle = \frac{1}{Z} \int \mathcal{D}U e^{-S_{\mathrm{eff}}[U]} \mathcal{O}(U) \qquad S_{\mathrm{eff}}[U] = S_{\mathrm{gauge}}[U] - \log\det(\mathbb{D}[U] + m)$$

モンテカルロ法: 確率 "$P[U] = \frac{1}{Z} e^{-S_{\mathrm{eff}}[U]}$"に従って場の配位を作って平均すれば積分できる

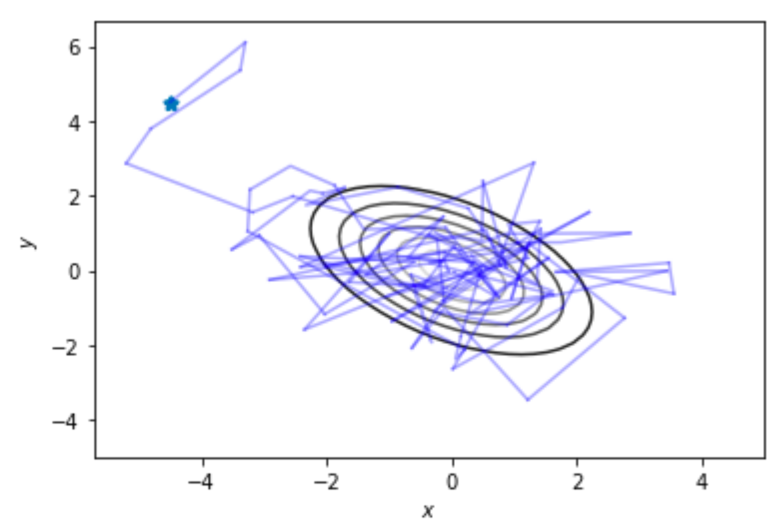
U_1 U_2 U_3



HMC: Hybrid (Hamiltonian) Monte-Carlo
デファクトスタンダードのアルゴリズム

大体メトロポリス法

$$S(x, y) = \frac{1}{2}(x^2 + y^2 + xy)$$



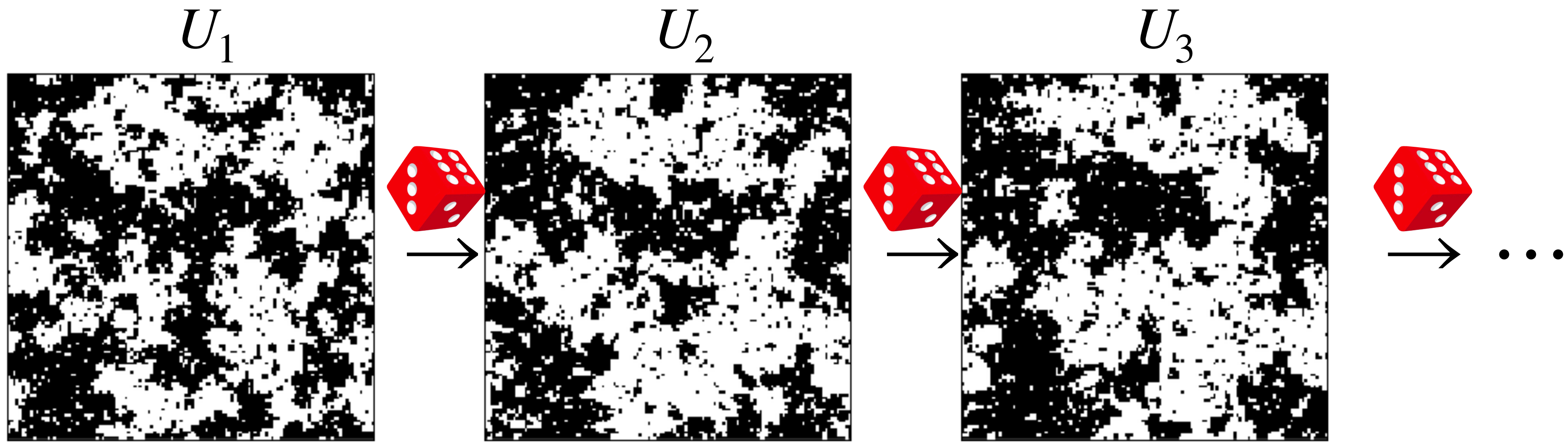
格子QCD

モンテカルロ法の誤差はサンプルの数で決まる

M. Creutz 1980

$$\langle \mathcal{O} \rangle = \frac{1}{Z}\int \mathcal{D}U e^{-S_{\text{eff}}[U]}\mathcal{O}(U) \qquad S_{\text{eff}}[U] = S_{\text{gauge}}[U] - \log\det(\mathbb{D}[U] + m)$$

モンテカルロ法: 確率 “$P[U] = \frac{1}{Z}e^{-S_{\text{eff}}[U]}$”に従って場の配位を作って平均すれば積分できる



積分の誤差は、サンプリングの回数で完全に決まる。定量的。

$$\langle \mathcal{O} \rangle = \frac{1}{N_{\text{sample}}}\sum_{k}^{N_{\text{sample}}} \mathcal{O}[U_k] \pm O\left(\frac{1}{\sqrt{N_{\text{sample}}}}\right)$$

Demonstration

Akio Tomiya

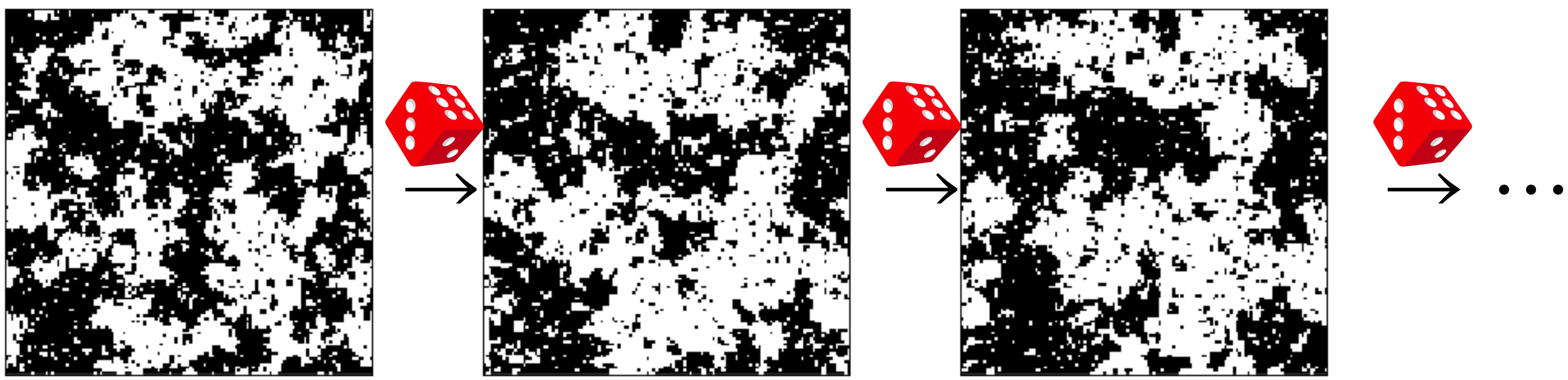
HMC = Metropolis test + molecular dynamics

M. Creutz 1980

$$\langle \mathcal{O} \rangle = \frac{1}{Z} \int \mathcal{D}U e^{-S_{\text{eff}}[U]} \mathcal{O}(U) \qquad S_{\text{eff}}[U] = S_{\text{gauge}}[U] - \log\det(\not{D}[U] + m)$$

モンテカルロ法: 確率 “$P[U] = \frac{1}{Z} e^{-S_{\text{eff}}[U]}$”に従って場の配位を作って平均すれば積分できる

U_1 U_2 U_3



アルゴリズムが厳密(＝期待値が常に正しく計算できる)でないと誤差が未知のため実験と比較できない。アルゴリズムの厳密性は証明可能。

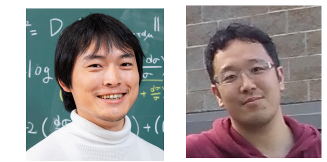
一方、ニューラルネットは近似器。厳密性…?

Demonstration
機械学習を使った場の理論の配位生成

徐々に良くなってきている

Restricted Boltzmann machine + HMC: 2次元スカラー場 A. Tanaka, AT 2017

機械学習+ 格子場の配位生成の最初の挑戦。相転移点で物理量の分布が偏る。厳密でもない。



GAN (Generative adversarial network): 2次元スカラー場

J. Pawlowski+ 2018
G. Endrodi+ 2018

計算結果は良さそう。厳密性の証明はなし(無理？)

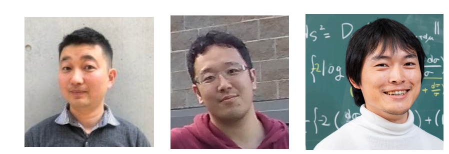
Flow based model: 2次元スカラー場, 純U(1), 純SU(N)ゲージ場 MIT+ Google Brain 2019 …

ニューラルネットットを使って自明化写像を模倣 (可逆なニューラルネット。計算しやすいJacobianをもつ)

厳密アルゴリズム、動的フェルミオンなし。非可換ゲージ対称性も対角化して取り扱う。2次元で実演。

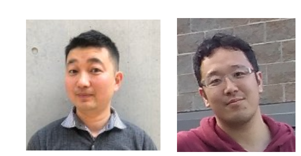
Self-learning Monte Carlo for lattice QCD arxiv 2010.11900 Y. Nagai, AT, A. Tanaka

動的フェルミオンの入った、4次元の非可換ゲージ理論
線形モデルを使ってゲージ不変な有効作用を作る
厳密アルゴリズム、計算が重い(フェルミオンを対角化)



Self-learning Hybrid Monte Carlo for lattice QCD (以下で見る) arxiv 2103.11965 Y. Nagai, AT

動的フェルミオンの入った、4次元の非可換ゲージ理論
共変ニューラルネットットを使ってゲージ不変な有効作用を作る
厳密アルゴリズム、計算が軽い



Demonstration

従来アルゴリズム: HMC = Hybrid (or Hamiltonian) Monte Carlo

S. Duane + 1987
M. Luscher+ 2011

HMC: 仮想的な運動量を入れて”確率的量子化”, 厳密アルゴリズム

$$H = \frac{1}{2}\sum_{\mu,n} \mathrm{tr}[\pi_\mu^2(n)] + S[U]$$

$\pi_\mu(n)$仮想的なランダム運動量、手で入れる

作用S をポテンシャルとみなす

$$P\left[U' \middle| U\right] = P_{\mathrm{acc}}\left(H[U,\pi], H[U',\pi']\right) T_{MD}\left[(U,\pi) \to (U',\pi')\right] P_{\mathrm{Gauss}}(\pi)$$

T_{MD}: **Molecular dynamics = 仮想的なハミルトン力学**

$$U_\mu(n) \leftarrow \mathrm{e}^{\mathrm{i}\pi_\mu(n)} U_\mu(n)$$

$$\pi_\mu(n) \leftarrow \pi_\mu(n) + \epsilon F_\mu(n)$$

$$F_\mu(n) \sim \frac{\partial S}{\partial U_\mu(n)}$$

(ゲージ対称性を拘束条件とする運動方程式)

メトロポリス法はエネルギー保存するアップデートをすると良い

運動方程式の解はエネルギー保存する、動的フェルミオンもOK

手で加えた仮想運動量は結果に影響しない

30年来、デファクトスタンダードのアルゴリズム

Demonstration

Akio Tomiya

SLHMC = Self-learning (自己学習) HMC、厳密

SLHMC: 運動方程式に使うポテンシャルをニューラルネットで変分近似

Yuki Nagai+ 1807.04955
Yuki Nagai+ 1909.02255

$$H = \frac{1}{2}\sum_{\mu,n} \mathrm{tr}[\pi_\mu^2(n)] + S[U]$$

$\pi_\mu(n)$仮想的なランダム運動量、手で入れる

作用S をポテンシャルとみなす

$$P\left[U'\middle|U\right] = P_{\mathrm{acc}}\left(H[U,\pi], H[U',\pi']\right) T^\theta_{MD}\left[(U,\pi)\to(U',\pi')\right] P_{\mathrm{Gauss}}(\pi)$$

T^θ_{MD}**: ニューラルネットを使った作用でのMolecular dynamics**

$$U_\mu(n) \leftarrow \mathrm{e}^{\mathrm{i}\pi_\mu(n)} U_\mu(n)$$

$$\pi_\mu(n) \leftarrow \pi_\mu(n) + \epsilon F^\theta_\mu(n)$$

$$F^\theta_\mu(n) \sim \frac{\partial S^{(NN)}}{\partial U_\mu(n)}$$

(ゲージ対称性を拘束条件とする運動方程式)

メトロポリス法はTの詳細を気にせず、エネルギー保った時間反転対称なアルゴリズムでOK. ニューラルネットを使った作用でも良し。

これは厳密アルゴリズム. もし$S^{(NN)}$とS が違いすぎると受理されず、効率が0
逆に学習がうまく行けば、受理される。ニューラルネットのテストになる

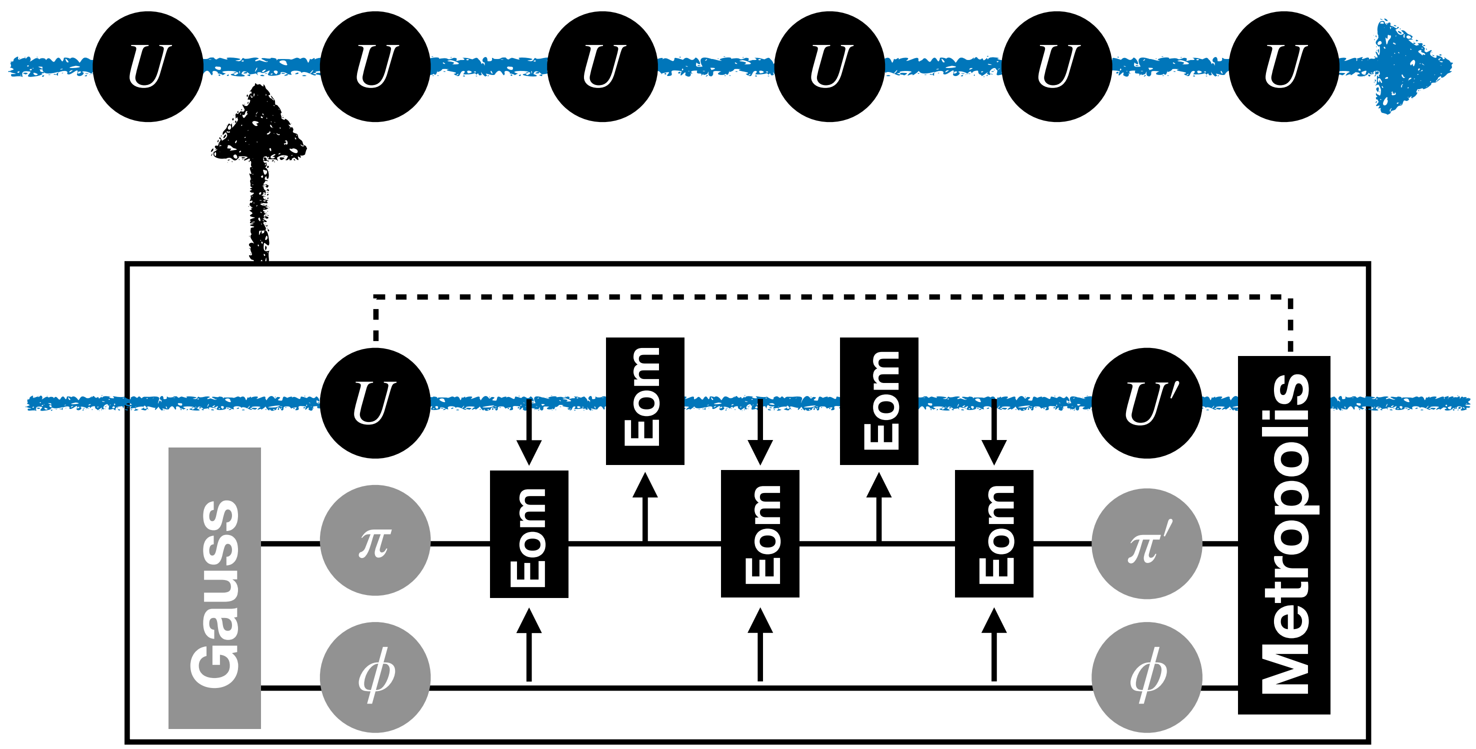
Demonstration

HMC とSLHMC の比較

Akio Tomiya

arXiv: 2103.11965 and reference therein

HMC



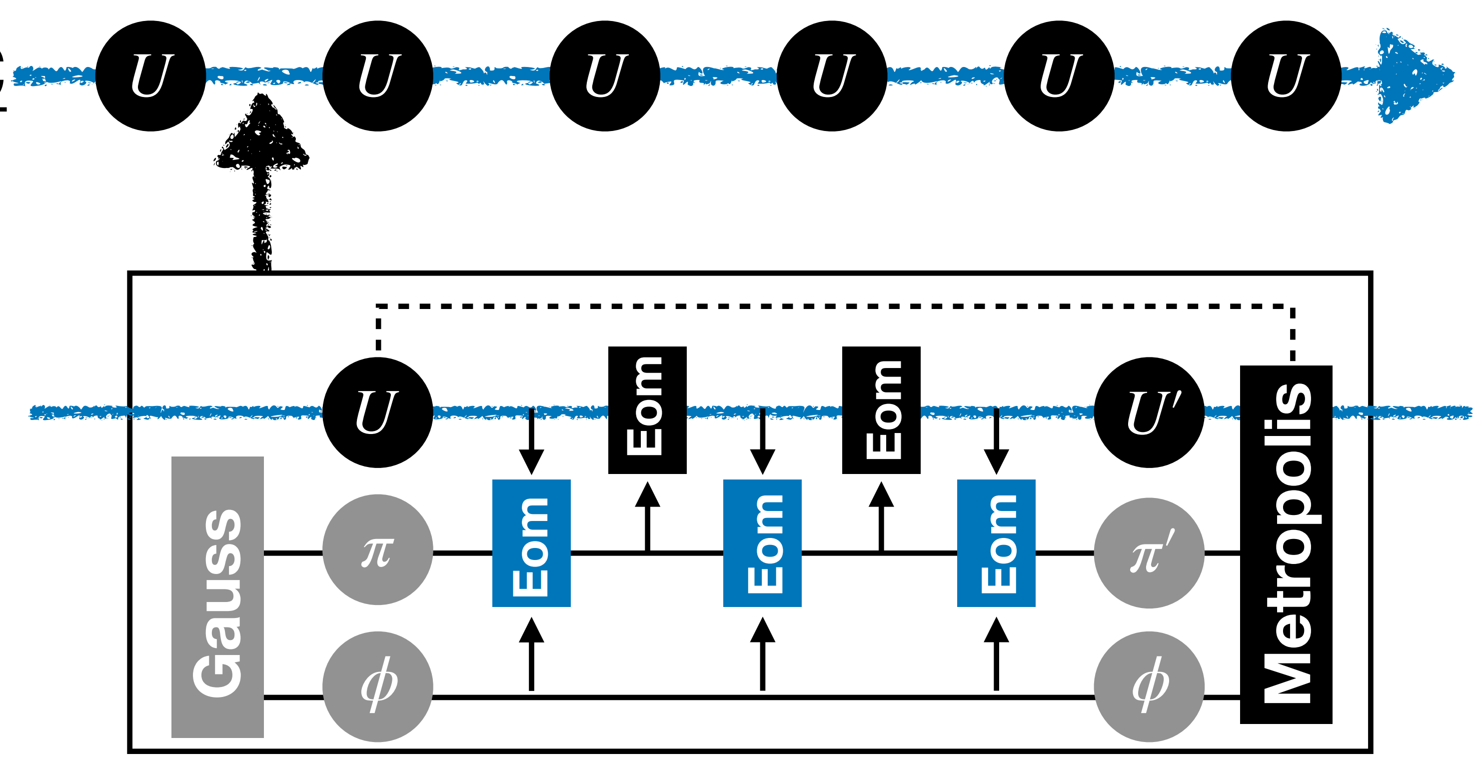
Eom **Metropolis**

Both use

$$H_{\mathrm{HMC}} = \sum \pi^2 + S_{\mathrm{g}} + S_{\mathrm{f}}$$

運動方程式(Eom) を数値的に解くときに誤差が入るが、メトロポリステストで厳密化

SLHMC



Metropolis **Eom**

$$H = \sum \pi^2 + S_{\mathrm{g}} + S_{\mathrm{f}}[U]$$

Eom

$$H = \sum \pi^2 + S_{\mathrm{g}} + S_{\mathrm{f}}[U^{\mathrm{NN}}[U]]$$

作用をニューラルネットで近似してるが厳密

Demonstration

Lattice setup

arXiv: 2103.11965

物理系 **Two color QCD (plaquette + staggered (not rooted))**

アルゴリズム **SLHMC, HMC (比較用)**

パラメータ **L=4, m = 0.3, beta = 2.7**

対象系の作用 $$S[U] = S_{\mathrm{g}}\left[U\right] + S_{\mathrm{f}}\left[\phi, U; m = 0.3\right],$$ **Metropolis Test用**

ニューラルネット有効作用 $$S_{\theta}[U] = S_{\mathrm{g}}\left[U\right] + S_{\mathrm{f}}\left[\phi, U_{\theta}^{\mathrm{NN}}[U]; m_{\mathrm{h}} = 0.4\right],$$ **MD用**

(問 MDの作用が違ってもMetropolis Testとおるか?)

(Sは示量変数なのでmがちょっとでも違うとSの値がぜんぜん違う, 学習で差が減らせるか？)

計算する量 **Plaquette (エネルギー密度), Polyakov loop (閉じ込め相転移の秩序変数)**
Chiral condensate (カイラル相転移の秩序変数) $\langle \bar{\psi}\psi \rangle$

コード **Juliaで書いたコード** LatticeQCD.jl (をもとに改造したもの) AT+ (in prep)

Demonstration

Network: trainable stout (plaq+poly)

arXiv: 2103.11965

ニューラルネットの構成
(Stout smearingに
Polyakov loop+plaq を入れた
回転対称性を減らしてある)

ベンチマークなので
適当に選んだ。
もっと工夫できる。

$$\Omega_\mu^{(l)}(n) = \rho_{\mathrm{plaq}}^{(l)} O_\mu^{\mathrm{plaq}}(n) + \begin{cases} \rho_{\mathrm{poly},4}^{(l)} O_4^{\mathrm{poly}}(n) & (\mu = 4), \\ \rho_{\mathrm{poly,s}}^{(l)} O_i^{\mathrm{poly}}(n), & (\mu = i = 1,2,3) \end{cases}$$

全部の ρ が重み
Oはループ演算子

$$Q_\mu^{(l)}(n) = 2[\Omega_\mu^{(l)}(n)]_{\mathrm{TA}}$$

TA: トレースを引く、アンチエルミート化

$$U_\mu^{(l+1)}(n) = \exp(Q_\mu^{(l)}(n)) U_\mu^{(l)}(n)$$

$$\underline{U_\mu^{\mathrm{NN}}(n)[U]} = U_\mu^{(2)}(n)\left[U_\mu^{(1)}(n)\left[U_\mu(n)\right]\right]$$

2層のstout smear型
全部で6パラメータ

ニューラルネットで
パラメトライズした作用:

$$S_\theta[U] = S_{\mathrm{g}}[U] + S_{\mathrm{f}}[\phi, \underline{U_\theta^{\mathrm{NN}}[U]}; m_{\mathrm{h}} = 0.4],$$

作用はゲージ場の関数
ニューラルネットとして
パラメトライズ

Loss function:

$$L_\theta[U] = \frac{1}{2}\left| S_{\mathrm{f}}[U^{\mathrm{NN}}[U], \phi, m = 0.4] - S_{\mathrm{f}}[U, \phi, m = 0.3] \right|^2,$$

Training strategy:

1.HMC でパラメータを調整 (online training+SDG)

2.固定したパラメータで SLHMC

(SLHMCやりながら学習もできるが今回はわけた)

Demonstration

結果: Loss は学習でちゃんと減ってる

arXiv: 2103.11965

Loss function:

$$L_\theta[U] = \frac{1}{2}\left| S_{\mathrm{f}}[U^{\mathrm{NN}}[U], \phi, m = 0.4] - S_{\mathrm{f}}[U, \phi, m = 0.3] \right|^2,$$

Prior HMC run (training)

$$\frac{\partial S}{\partial \rho_i^{(l)}} = 2\,\mathrm{Re} \sum_{\mu', m} \mathrm{tr}\left[U_{\mu'}^{(l)\dagger}(m) \Lambda_{\mu', m} \frac{\partial C}{\partial \rho_i^{(l)}} \right]$$

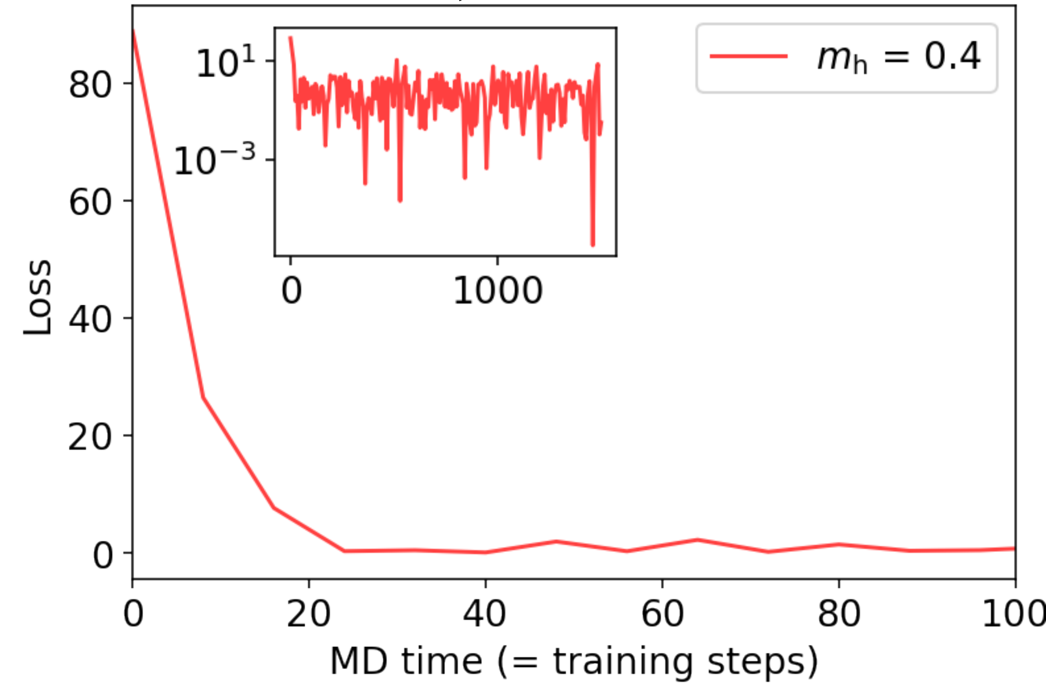
C: ΩからUを1つ除いた物、共変

Λ: Uの多項式、共変

$$\theta \leftarrow \theta - \eta \frac{\partial L_\theta(\mathcal{D})}{\partial \theta},$$

$$\frac{\partial L_\theta(\mathcal{D})}{\partial w_i^{(L-1)}} = \frac{\partial L_\theta(\mathcal{D})}{\partial S_\theta} \frac{\partial S_\theta}{\partial w_i^{(L-1)}}$$

Loss の履歴, ちゃんと減ってる



Layer	Loop	Value of ρ
1	Plaquette	-0.011146476388409423
2	Plaquette	-0.011164492428633698
1	Spatial Polyakov loop	-0.0030283193221172216
2	Spatial Polyakov loop	-0.0029984533773388094
1	Temporal Polyakov loop	0.0042480217272233112
2	Temporal Polyakov loop	0.0041952533380373369

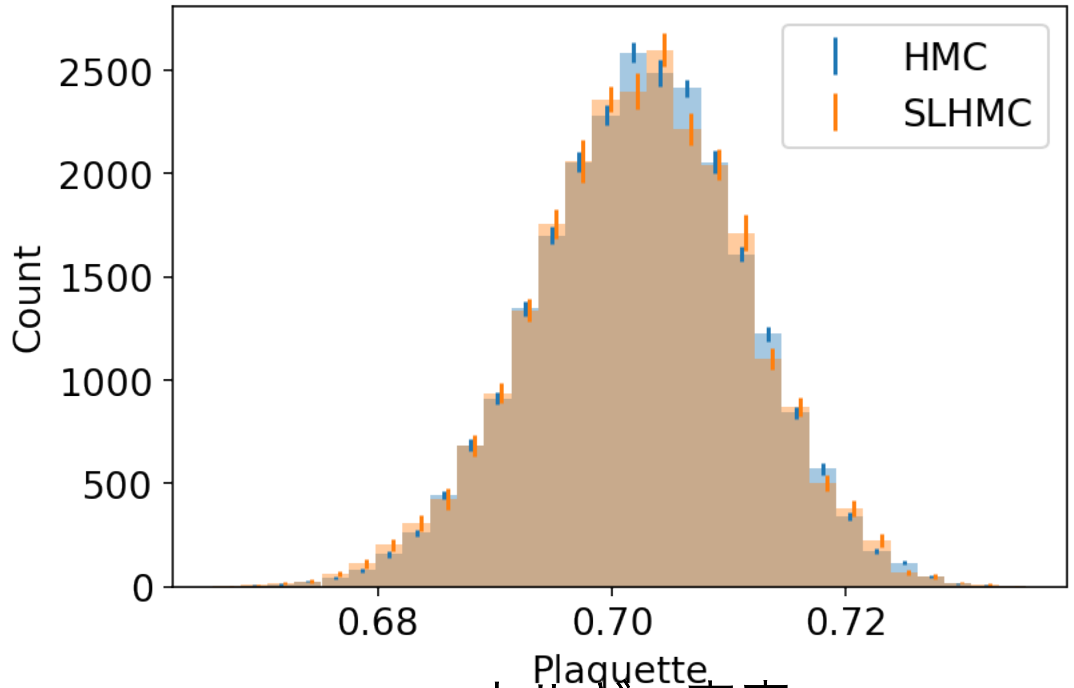
この値を入れて、SLHMCを実行する

Demonstration

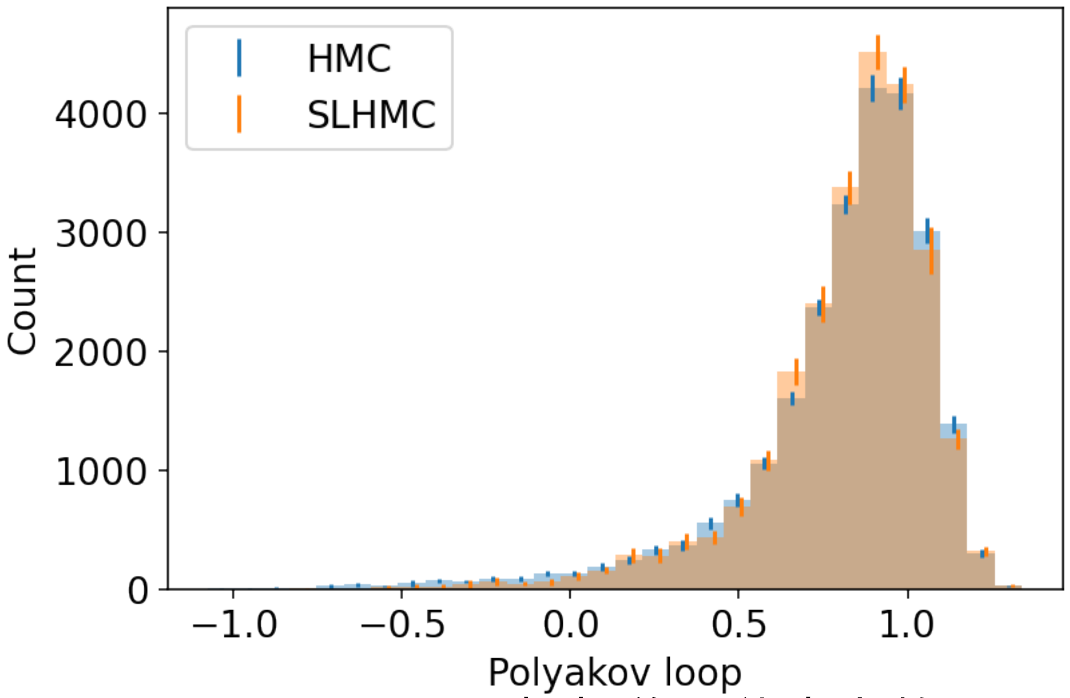
Akio Tomiya

自己学習HMC: ニューラルネットを使った作用でも厳密

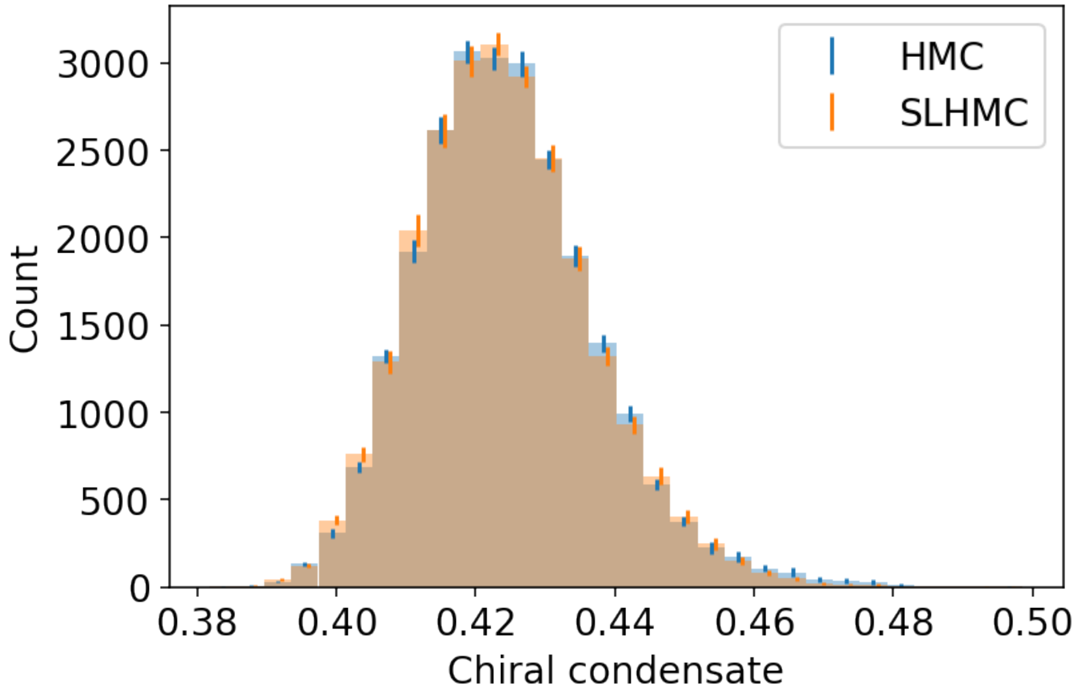
arXiv: 2103.11965 + 統計



エネルギー密度



閉じ込め相転移の秩序変数



カイラル相転移の秩序変数

期待値: 完璧に合ってる

Algorithm	Observable	Value
HMC	Plaquette	0.7025(1)
SLHMC	Plaquette	0.7023(2)
HMC	|Polyakov loop|	0.82(1)
SLHMC	|Polyakov loop|	0.83(1)
HMC	Chiral condensate	0.4245(5)
SLHMC	Chiral condensate	0.4241(5)

Acceptance = 40%

Summary and future work 1/2

まとめ: ゲージ対称性を保つニューラルネットの提案とその応用をした

arXiv: 2103.11965

- 畳み込みレイヤー = 学習で調整する画像フィルター、対称性を尊重
- 共変ニューラルネット = 調整できる”スメアリング”(平滑化)、パラメータによっては平滑化するとは限らない
  - 学習のために拡張したデルタルールを作成(skipped)。smear機能を持つ既存の格子QCDのコードならすぐに実装可
  - ゲージ不変なloss function
  - U(1)ゲージ理論、APE型を考えて格子間隔aで展開、weight sharing をやめる
    → 普通の全結合ニューラルネット(skipped)。普通のニューラルネットの拡張。
  - 共変ニューラルネットのニューラルODE = “gradient flow”
    (ただし、なにかのgradient として書ける必要はない。書けても良い)
- 自己学習HMC = HMC+ニューラルネットでパラメータ付した作用を使ったmolecular dynamics、厳密アルゴリズム。
- トレーニング: 6パラメータしかないが、loss がO(1)に減少。学習できた。
- 自己学習HMC の結果は、HMCの結果と一致。4次元の(フェルミオンの量子効果の入った)非可換ゲージ理論の正しい配位が作れた

Summary and future work 2/2

Akio Tomiya

Possible future works

arXiv: 2103.11965

- 共変ネット: どんな関数が近似できるのか。深層化した時に万能近似性を持つか？
- 共変ネット: 機械学習の文脈の応用はないか？ (T. Cohen et al は離散ゲージ対称性を持つデータを使った。連続ゲージ対称性を持ったタスク？)
- 共変ネット: ゲージ配位を細かい配位にコンバートできないか？ ニューラルネットでの画像の高精細化はできる
- 共変ネット: (A. Tanaka AT 2016) のように、新たなオーダーパラメータの抽出につかえないか。topological chargeの推定もできる(Kitazawa+ 2020)？
- 共変ネット: GAN などを作れる？ RBM化したり、flow based と組み合わせられるか？
- 共変ネット: AdS/DL 的な解釈を持つか？ ゲージ場を取り扱えるので相性は良さそう。
- 共変ネット: HISQ(Highly improved staggered quark, 2段スメア) よりよい1段階スメアフェルミオンの構成？
- 共変ネット: ニューラルネット~Gradient flowとわかった。逆にニューラルネットの解析に場の理論の手法？(J. Halverson+ 2020 との関係？)。生成分布がわかってるので、loss が厳密にわかるはず。
- SLHMC: SU(N>2) に拡張する。一部実装にSU(2)の特殊性を使ってるがstout と同じようにSU(3)化できる。
- SLHMC: 違うフェルミオン作用の近似(たとえばDW/OV)。ゲージ場作用の方も同じトリックを使えるので改良？
- SLHMC: メトロポリステストの受理率を上げるニューラルネットの構成？
- SLHMC: もっと大きな系でtopological charge を測る。トポロジーを変える作用？