

機械学習による符号問題の最適化

森 勇登 (京大理)

共同研究者：柏 浩司、大西 明 (京大基研)

Y.M., K. Kashiwa, and A. Ohnishi, arXiv:1705.05605

Y.M., K. Kashiwa, and A. Ohnishi, arXiv:1709.03208 2017/11/11 千葉工大セミナー

Content

1. Introduction
2. 機械学習
3. 符号問題
4. 経路最適化法
5. 複素 ϕ^4 模型
6. まとめ

1. Introduction

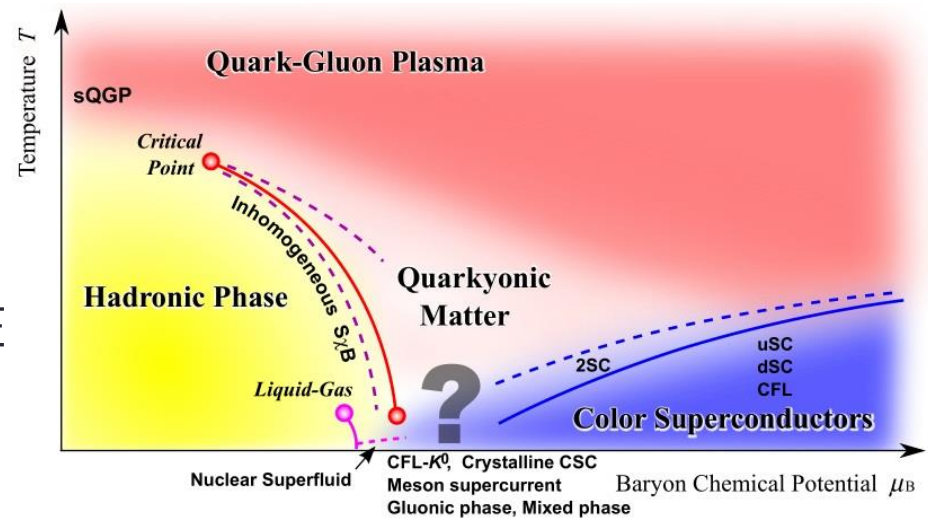
本研究の目的 (目標)

符号問題の解決

有限密度QCDに存在

道具

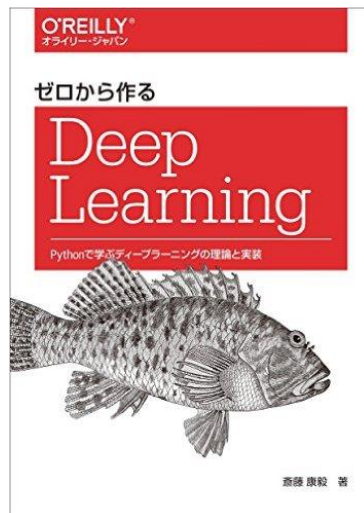
- 機械学習
- コーシー (-ポアンカレ) の積分定理



K. Fukushima and T. Hatsuda,
Rept.Prog.Phys. 74 (2011) 014001

2. 機械学習

- 近年機械学習が理論物理学に適用され始めている
 - ニューラルネットワークによる多体系の計算・学習
 - 相転移の探索, etc...
- 以下で機械学習について少し説明する
- 参考書のススメ



「ゼロから作るDeep Learning —
Pythonで学ぶディープラーニングの
理論と実装」
作者: 斎藤康毅
出版社/メーカー: オライリージャパン

2. 機械学習 | 定義

定義 (Tom M. Mitchell, 1997)

「コンピュータプログラムが、ある種のタスク **T** と評価尺度 **P** において、経験 **E** から学習するとは、タスク **T** におけるその性能を **P** によって評価した際に、経験 **E** によってそれが改善されている場合である」

個人的な解釈

データを用いた関数の最適化 (フィッティング)

2. 機械学習 | 具体例

相転移の探索

J Carrasquilla and RG Melko, Nature Physics, 2017

- モンテカルロ計算は可能
- しかし計算すべき物理量が分からない . . .
- けれども相転移などの構造を知りたい



機械に相を分類してもらう

2. 機械学習 | 具体例

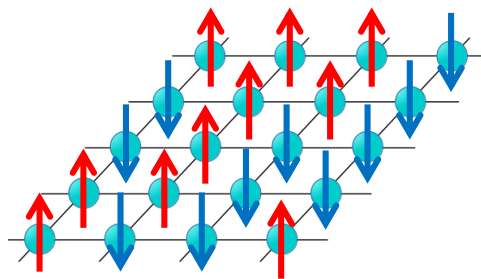
相転移の探索

J Carrasquilla and RG Melko, Nature Physics, 2017

- モンテカルロ計算は可能
- しかし計算すべき物理量が分からない・・・
- けれども相転移などの構造を知りたい

➡ **機械に相を分類してもらう**

欲しいもの (例. イジング模型) :



0 (低温相)
or
1 (高温相)

出典: 「Ising Modelを平易に解説してみる」
<http://enakai00.hatenablog.com/entry/20150106/1420538321>

出典: いらすとや

2. 機械学習 | 具体例

- タスク **T** : configuration X から 0~1 の値を返す
パラメータ W を持った関数 $f(X, W)$
- 経験 **E** : 各相の configuration X_i と
相の情報 $y_i = 0$ or 1
- 評価尺度 **P** : 正しく分類できてるかを定量的に評価
クロスエントロピー

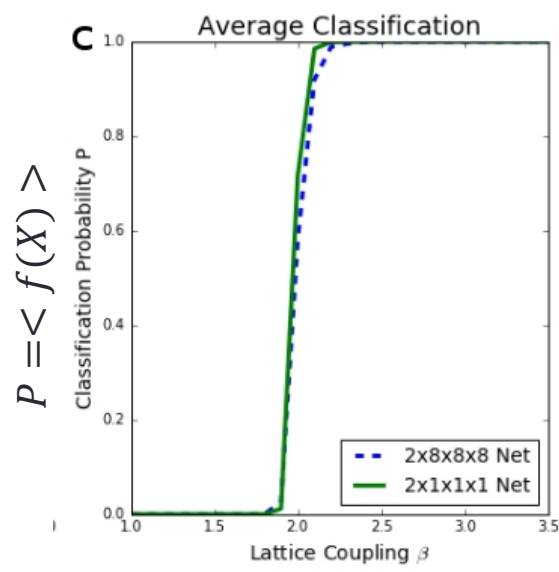
$$S = - \sum_i [y_i \log f(X_i, W) + (1 - y_i) \log(1 - f(X_i, W))]$$

- 学習 . . . S が小さくなるように W を調整  分類が可能に !

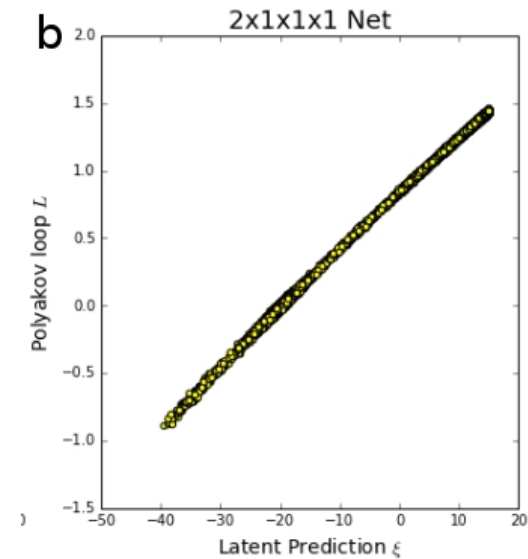
2. 機械学習 | 具体例

- 4D SU(2) gauge theory

S. J. Wetzel and M. Scherzer,
Phys. Rev. B 96 (2017), 184410



相転移点



$f(X)$ から引き出した量



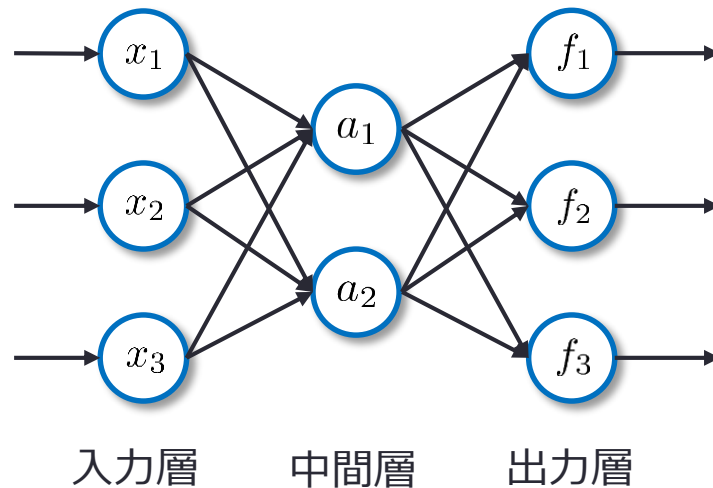
Order parameter

2. 機械学習 | ニューラルネットワーク

ニューラルネットワーク・・・神経回路を模した数学モデル

線型変換と非線型な変換の組み合わせ

順伝播型の模式図



$$a_i = g(W_{ij}^{(1)} x_j + b_i^{(1)})$$

$$f_i = g(W_{ij}^{(2)} a_j + b_i^{(2)})$$

$g(x)$: 活性化関数(tanh等)

※ W , b がパラメータ

中間層のユニット数 $\rightarrow \infty$ で
任意の連続関数を近似できる
(Universal approximation theorem)

G. Cybenko, MCSS 2, 303 (1989)

3. 符号問題

- 分配関数

$$\mathcal{Z} = \int \mathcal{D}x e^{-S[x]}$$

- 物理量

$$\langle \mathcal{O} \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}x \mathcal{O} e^{-S}$$

符号問題 : $S[x] \in \mathbb{C}$

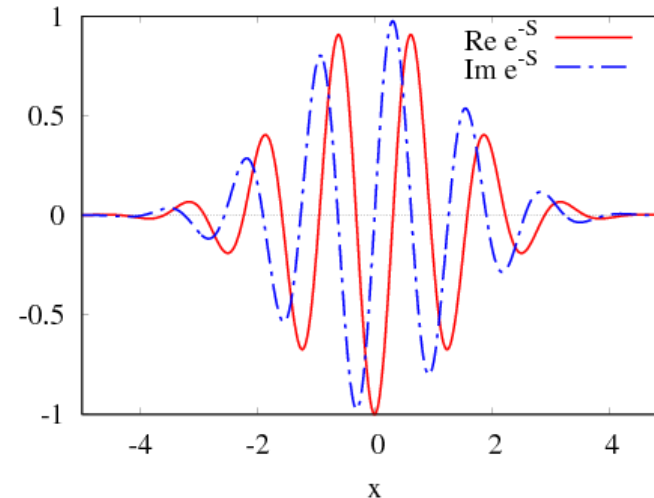
- 打ち消し合いによる
計算精度の悪化

- 確立解釈が困難

phase-reweightingで
解釈可能に

例): $e^{-S(x)} = (x + 10i)^{50} e^{-x^2/2}$

J. Nishimura and S. Shimasaki,
Phys. Rev. D92 (2015), 011501



3. 符号問題 | phase-reweighting

phase-reweighting method

$$\langle \mathcal{O} \rangle = \frac{\int \mathcal{D}x \mathcal{O} e^{-S}}{\int \mathcal{D}x e^{-S}} = \frac{\int \mathcal{D}x \mathcal{O} \frac{e^{-S}}{|e^{-S}|} |e^{-S}|}{\int \mathcal{D}x \frac{e^{-S}}{|e^{-S}|} |e^{-S}|} = \frac{\langle e^{-i\text{Im}S} \mathcal{O} \rangle_{\text{pq}}}{\langle e^{-i\text{Im}S} \rangle_{\text{pq}}}$$

$$\langle \mathcal{O} \rangle_{\text{pq}} = \frac{\int \mathcal{D}x \mathcal{O} |e^{-S}|}{\int \mathcal{D}x |e^{-S}|} \quad \Rightarrow \quad |e^{-S}| \text{ を確率と見なせる}$$

モンテカルロ計算が可能に

平均位相因子 $\langle e^{-i\text{Im}S} \rangle_{\text{pq}}$

- ・・・小さい時符号問題は深刻
(一般に体積大で小さい)
0/0 に近い計算が必要

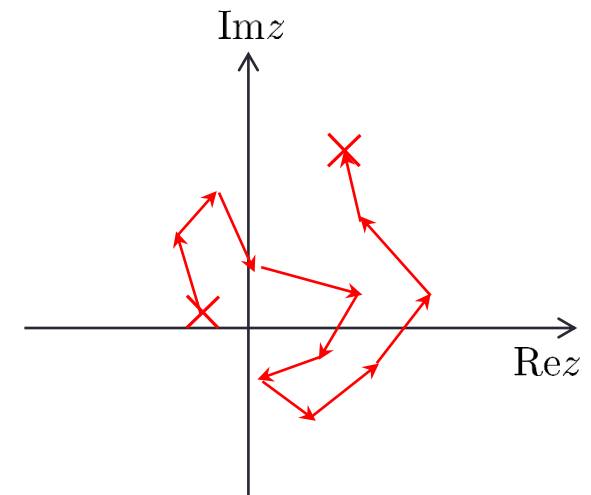
3. 符号問題 | 注目されている手法

● Complex Langevin

- 変数を複素化 $x \in \mathbb{R}^n \rightarrow z \in \mathbb{C}^n$
- 複素平面上で Random walk

$$\frac{d}{dt}z(t) = -\frac{\partial S[z]}{\partial z} + \eta(t), \quad \langle \eta(t)\eta(t') \rangle = 2\delta(t-t')$$

- 符号問題は存在しないが
間違っただけの解を出すことも



G. Parisi, Phys. Lett. B131 (1983) 393
G. Aarts, *et al.*, Phys. Rev. D81 (2010) 054508

● Lefschetz thimble

- 変数を複素化 $x \in \mathbb{R}^n \rightarrow z \in \mathbb{C}^n$
- 元と等価な経路の上で積分

次ページで説明

E. Witten, AMS/IP Stud. Adv. Math. 50 (2011) 347-446
M. Cristoforetti, Phys. Rev. D86 (2012) 074506
H. Fujii, *et al.*, JHEP 1310 (2013) 147

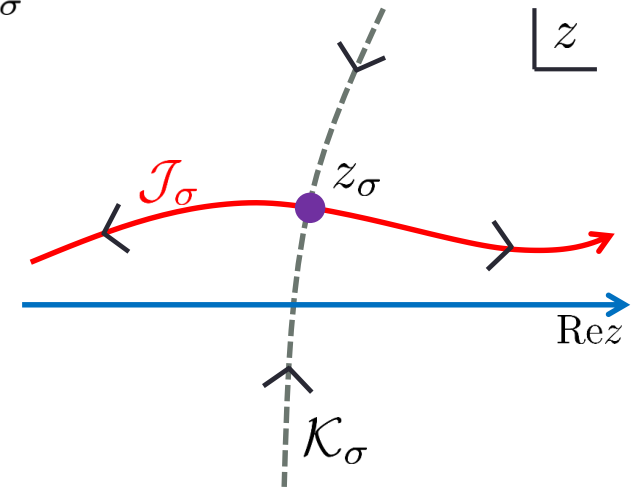
3. 符号問題 | Lefschetz thimble

Flow方程式

$$\frac{d}{dt}z(t) = \overline{\left(\frac{\partial S[z]}{\partial z}\right)}, \text{ 停留点: } \left.\frac{\partial S}{\partial z}\right|_{z_\sigma} = 0$$

に対してthimbleを以下で定める

$$\begin{cases} \mathcal{J}_\sigma = \{z(t) \mid t \in \mathbb{R}, z(-\infty) = z_\sigma\} \\ \mathcal{K}_\sigma = \{z(t) \mid t \in \mathbb{R}, z(+\infty) = z_\sigma\} \end{cases}$$



元の積分領域 ($\mathbb{R}^n$) と等価な領域は

$$\mathcal{C} = \sum_{\sigma} n_{\sigma} \mathcal{J}_{\sigma}, \quad n_{\sigma} = \langle \mathbb{R}^n, \mathcal{K}_{\sigma} \rangle$$

※コーシーの積分定理より
物理量の期待値等に影響はない

3. 符号問題 | Lefschetz thimble

- 重要な性質：
thimble毎に作用の虚部は一定

$$\frac{d}{dt}S(z(t)) = \left| \frac{\partial S}{\partial z} \right|^2 \in \mathbb{R}$$

 thimble上で積分すれば
打ち消し合いを防ぐことができる

- 欠点：
 - thimbleの構造(n_σ 等)を調べるのが面倒
(cf. generalized Lefschetz thimble (A. Alexandru *et al.* JHEP 1605 (2016) 053))
 - 積分測度もFlow方程式から求める必要あり

4. 経路最適化法

Y.M., K. Kashiwa, and A. Ohnishi,
arXiv:1705.05605

打ち消し合いが小さくなるような経路を**変分的に求める**

つまり符号問題を最適化問題とみなす

(例) 1変数の場合

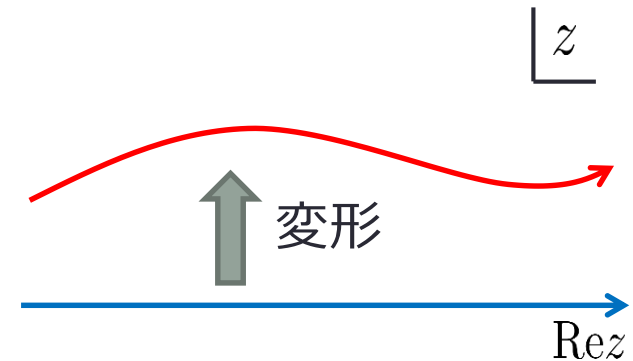
- 試行関数(積分経路)

$$z(\cdot) : \mathbb{R} \rightarrow \mathbb{C}$$

- 目的関数(最小化すべき関数)

$$\mathcal{F}[z(t)] = \frac{1}{2} \int dt \underbrace{|e^{i\theta(t)} - e^{i\theta_0}|^2}_{\text{打ち消し合い}} \times \underbrace{|J(t)e^{-S(z(t))}|}_{\text{重み}}$$

$$J(t) = \frac{\partial z(t)}{\partial t}, \quad \theta(t) = \arg \left(\frac{\partial z}{\partial t} e^{-S(z(t))} \right), \quad \theta_0 = \arg(\mathcal{Z})$$



4. 経路最適化法

Y.M., K. Kashiwa, and A. Ohnishi,
arXiv:1705.05605

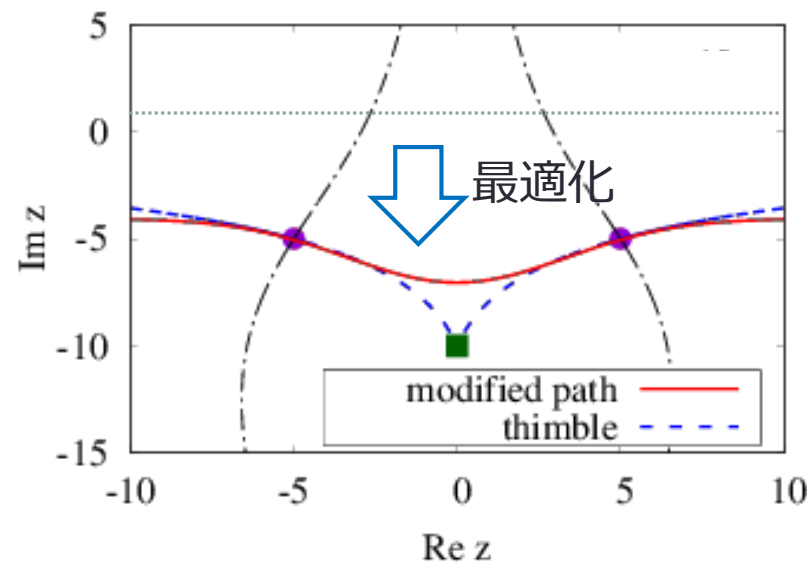
分配関数 J. Nishimura and S. Shimasaki,
Phys.Rev. D92 (2015), 011501

$$\mathcal{Z} = \int dx (x + 10i)^{50} \exp(-x^2/2)$$

積分経路

$$z(t) = t + i \left(c_1 \exp\left(-\frac{c_2^2 t^2}{2}\right) + c_3 \right)$$

- c_i を最急降下法で最適化

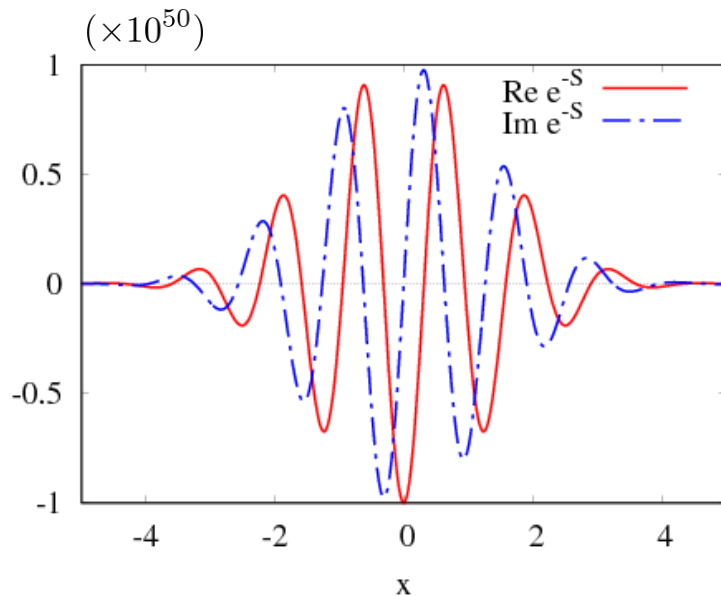
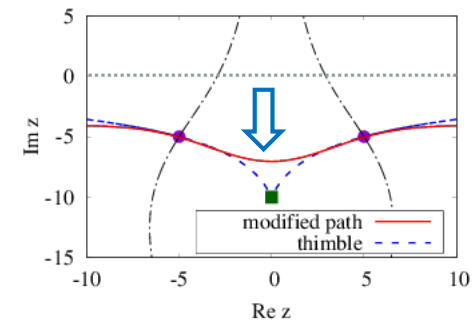


4. 経路最適化法

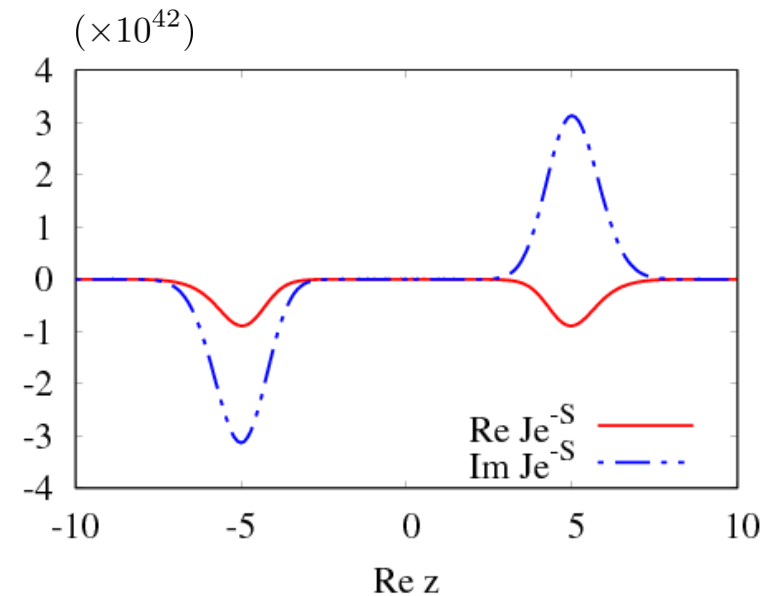
Y.M., K. Kashiwa, and A. Ohnishi,
arXiv:1705.05605

分配関数 J. Nishimura and S. Shimasaki,
Phys.Rev. D92 (2015), 011501

$$\mathcal{Z} = \int dx (x + 10i)^{50} \exp(-x^2/2)$$



最適化



4. 経路最適化法 | 機械学習との対応

多変数の場合は目的関数 $\mathcal{F}[z(t)]$ も
モンテカルロ法で評価

この場合に、経路最適化法で実際に行うこと：

- | | |
|---|-----------------|
| 1. 積分経路を試行関数で与える | ↔ タスク T |
| 2. 符号問題の程度を表す目的関数を用意 | ↔ 評価尺度 P |
| 3. モンテカルロ法で媒介変数 t に対する
配位を(大量に)生成 | ↔ 経験 E |
| 4. 与えられた配位に対して、目的関数を
最小化するように試行関数を変分 | ↔ 学習 |



機械学習の知識を使える？

4. 経路最適化法 | ニューラルネットワーク

- 格子上の場の理論・・・多変数
- 積分経路に適切な関数形は非自明



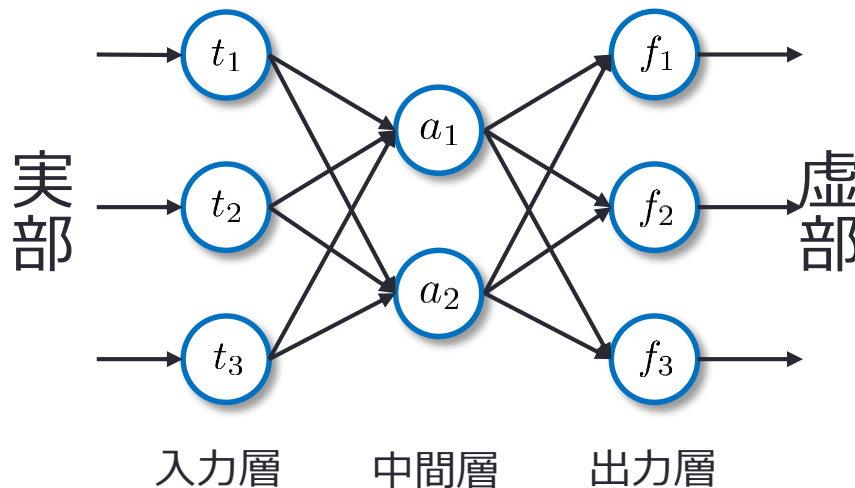
ニューラル
ネットワーク

{ input : 実部
output : 虚部

$$z_i(t) = t_i + i(\alpha_i f_i(t) + \beta_i)$$

今回は活性化関数に $\tanh(x)$ を採用

すること：ニューラルネットワーク中の
パラメータと α , β の学習



5. 複素 ϕ^4 模型

● 2D Lattice

$$S = \sum_x \left[(4 + m^2) \phi_x^* \phi_x + \lambda (\phi_x^* \phi_x)^2 - \sum_{\nu=0}^1 \underbrace{(\phi_x^* e^{-\mu \delta_{\nu,0}} \phi_{x+\hat{\nu}} + \phi_{x+\hat{\nu}}^* e^{+\mu \delta_{\nu,0}} \phi_x)}_{\in \mathbb{C}} \right]$$

● 積分変数

$$\phi = \frac{1}{\sqrt{2}}(x_1 + ix_2) \cdots \text{積分変数 } x_1, x_2 \text{ について作用 } S \text{ は解析的}$$

→ x_1, x_2 それぞれを複素化

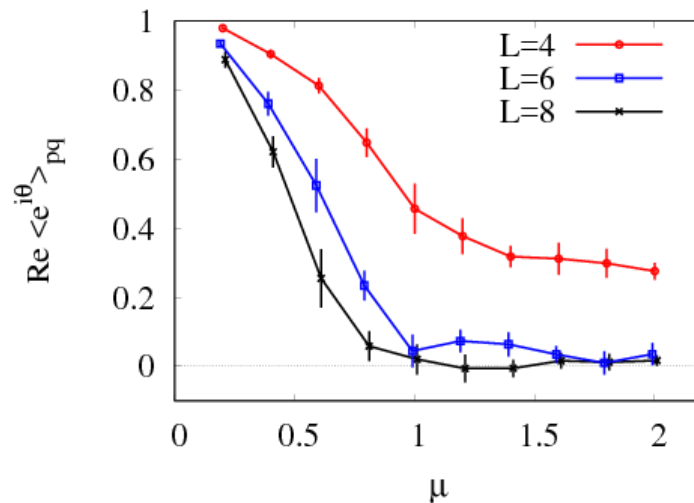
$$z_i(t) = t_i + i(\alpha_i f_i(t) + \beta_i)$$

以下の計算では $m = 1, \lambda = 1$ の場合を考える

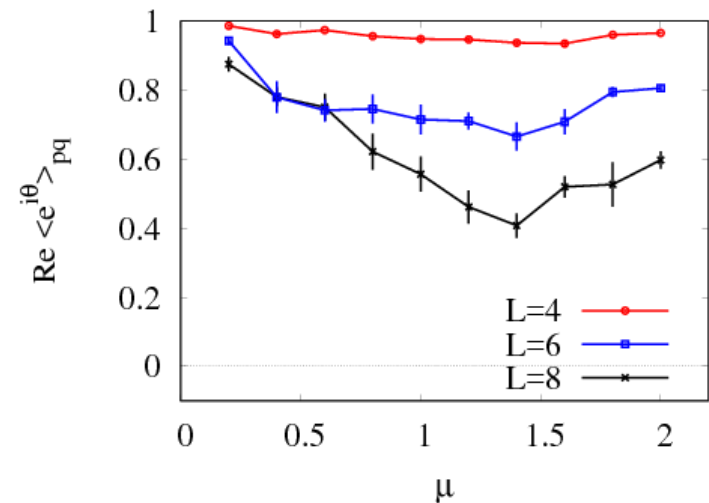
5. 複素 ϕ^4 模型 | 結果

Y.M., K. Kashiwa, and A. Ohnishi,
arXiv:1705.05605

● 平均位相因子



最適化



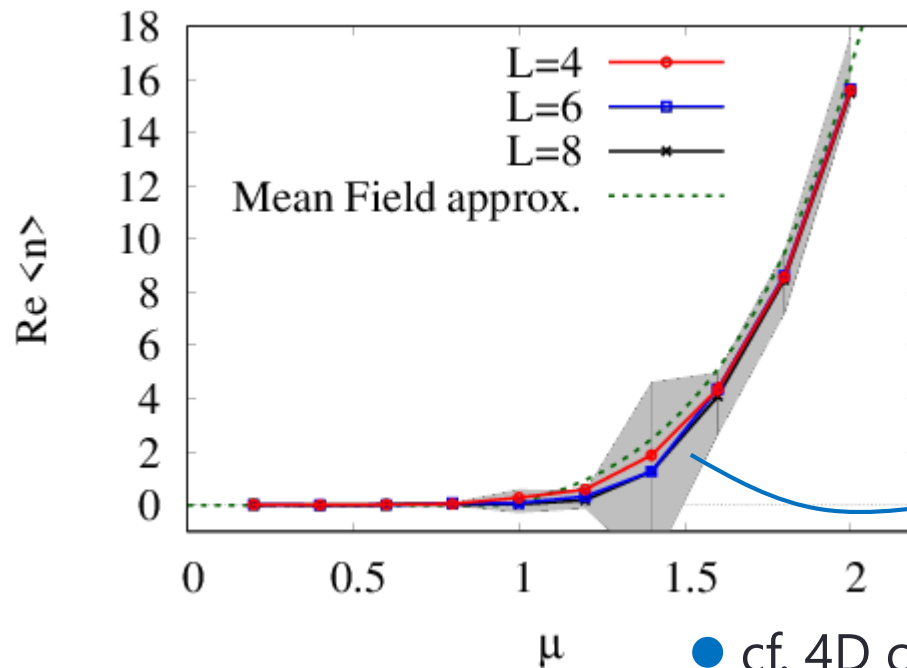
学習によって改善
しかし体積依存性が残る...

学習自体の最適化が必要？

5. 複素 ϕ^4 模型 | 結果

Y.M., K. Kashiwa, and A. Ohnishi,
arXiv:1705.05605

● 密度



● 平均場近似と
consistent

● 最適化無しの場合と
誤差の範囲内で一致

● : 最適化無し (L=8)

● cf. 4D complex ϕ^4

Complex Langevin:

G. Aarts, Phys. Rev. Lett. 102, 131601 (2009)

Lefschetz thimble :

H. Fujii, *et al.*, JHEP 1310, 147 (2013)

6. まとめ.

- 符号問題に取り組むために経路最適化法を提案
 - 試行関数を用いて経路を表現
 - 符号問題を最適化問題とみなす
- 1変数の系、2D lattice 複素 ϕ^4 模型に適用
 - 最適化によって平均位相因子を大きくできる
 - ニューラルネットワークを使うことで積分経路の関数形を陽に仮定する必要はなくなる