

高速にシミュレート可能な パラメータ付き量子回路による学習

真鍋 秀隆^{1,a)} 原田 健自¹

概要: NISQ 上で実行可能なタスクとして量子回路学習などの学習アルゴリズムがあり、古典コンピュータによる学習の性能を上回る可能性があるとして注目を集めている。ところが、計算量の制約から、多数の量子ビットを用いた場合において量子回路学習がどのような性能、性質を持つかはあまり分かっていない。本研究では、古典コンピュータで高速にシミュレート可能なテンソルネットワーク構造を持つパラメータ付き量子回路による学習を考え、大規模量子系における量子回路学習タスクの性能や性質について古典シミュレーションの領域から考察する。

Quantum circuit learning with efficiently simulatable parameterized circuit

1. はじめに

近年、NISQ(Noisy Intermediate-Scale Quantum device)[1] と呼ばれる、誤り訂正機能のない数百~数千量子ビットからなる量子コンピュータの開発が進んでいる。量子コンピュータは量子ビットと呼ばれる基本単位を用いて量子力学的な処理を実行できるため、古典コンピュータと比べて指数関数的な速度で情報を処理できる。NISQ は誤り訂正機能がないため Shor の素因数分解アルゴリズムのような計算量理論的な優位性の保証のあるタスクを実行することはできないが、特定のタスクにおいてはスーパーコンピュータを用いた古典計算よりも高速に計算できる、いわゆる量子超越性 [2] が実証されている。

NISQ 上で動作するタスクのひとつに量子機械学習がある。その中でも、量子回路学習 [3] はニューラルネットワークに着想を得た教師あり学習アルゴリズムであり、古典機械学習では学習不可能なデータセットを学習できる可能性があるとして注目されている。ところが、どのような量子カーネルを用いればよいか、どのように量子ゲートを配置すれば高い性能を発揮できるかについてはよく分かっていない。さらに、古典シミュレーションの制限により、大規模な量子系に対して量子回路学習がどの程度の性能を発揮

できるか厳密に調べることができない。

しかしながら、量子多体系の研究分野においてよく用いられる計算手法であるテンソルネットワーク法の考え方を利用すれば大規模量子系であっても古典シミュレート可能な量子回路を構成することができることが知られている [4], [5]。本研究報告では、そのような高速にシミュレート可能なパラメータ付き量子回路を用いて量子回路学習を実行することで、大規模量子系において達成しうる学習性能や学習そのものの性質について古典シミュレーションの立場から研究することを考える。

量子回路学習の性質について、近年エクスプレシビリティ (Expressibility) とエンタングル容量 (Entangling capability) という指標が提案されており [6], [7], [8], 量子回路学習の性能と相関があることが分かっている [9]。これらの量は量子回路にテンソルネットワーク構造を仮定すれば高速に計算ができる。そのため本稿では、テンソルネットワーク構造を利用した量子回路学習についてその学習性能を調べ、さらに Expressibility や Entangling capability の観点からその性質について考察する。

以下では、2章で量子回路学習の説明を行い、3章でテンソルネットワーク構造を用いた大規模量子回路学習について具体的な性能を示す。4章でパラメータ付き量子回路の性能指標であるエクスプレシビリティ、エンタングル容量について紹介し、5章ではテンソルネットワーク構造を

¹ 京都大学大学院情報学研究科

^{a)} manabe@acs.i.kyoto-u.ac.jp

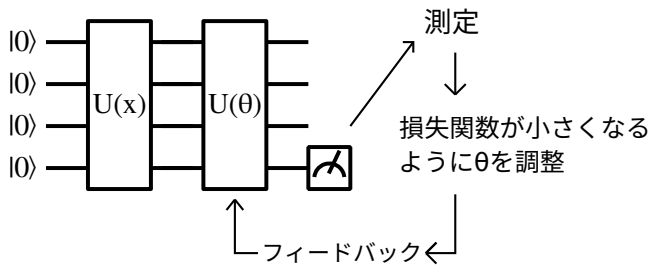


図 1 量子回路学習の概念図.

持つ量子回路についてそれらの値を数値実験により求め、性能との関連性について議論する.

2. 量子回路学習

2.1 量子回路学習アルゴリズム

量子回路学習は教師あり学習の一種である [3]. 入力データの集合 $\{\mathbf{x}_i\}$ と対応する教師データ $\{\mathbf{f}(\mathbf{x}_i)\}$ が与えられたとき, 出力 $\mathbf{y}_i = \mathbf{y}(\mathbf{x}_i, \theta)$ が $\mathbf{f}(\mathbf{x}_i)$ に近づくようにパラメータ θ を調整する. 量子回路学習においては, パラメータ θ は量子ゲートを制御する古典パラメータである. 以下に量子回路学習のアルゴリズムを要約する.

- 初期状態 $|0\rangle^{\otimes n}$ に対して, 入力データに依存した量子カーネル $U(\mathbf{x}_i)$ をかけることで, 古典入力データ \mathbf{x}_i を量子状態 $|\psi_{\text{in}}(\mathbf{x}_i)\rangle$ にエンコードする.

$$|\psi_{\text{in}}(\mathbf{x}_i)\rangle = U(\mathbf{x}_i) |0\rangle^{\otimes n} \quad (1)$$

- 古典的に制御可能なパラメータ θ に依存する量子ゲート $U(\theta)$ をかける. このとき状態は $|\psi_{\text{out}}(\mathbf{x}_i, \theta)\rangle$ になるとする.

$$|\psi_{\text{out}}(\mathbf{x}_i, \theta)\rangle = U(\theta) |\psi_{\text{in}}(\mathbf{x}_i)\rangle \quad (2)$$

- POVM の組を $\{M_j\}_{j \in \Omega}$, (実数値の) 測定値を $\{m_j\}_{j \in \Omega}$ としたとき, 測定結果が j となる確率に m_j をかけたものを j 番目の出力とする.

$$\mathbf{y}(\mathbf{x}_i, \theta)_j = \langle \psi_{\text{out}}(\mathbf{x}_i, \theta) | m_j M_j | \psi_{\text{out}}(\mathbf{x}_i, \theta) \rangle \quad (3)$$

- 教師データと出力に依存して定義される損失関数 $L(\mathbf{f}(\mathbf{x}), \mathbf{y}(\mathbf{x}, \theta))$ に対して, これを小さくする方向にパラメータ θ を調整する. 以上のステップを, L が収束するまで繰り返す.

なお, 本研究においては損失関数 (loss) は二乗誤差

$$L(\mathbf{f}(\mathbf{x}), \mathbf{y}(\mathbf{x}, \theta)) = \sum_i \sum_j \{ \mathbf{f}(\mathbf{x}_i)_j - \mathbf{y}(\mathbf{x}_i, \theta)_j \}^2 \quad (4)$$

を用いる. 以上の量子回路学習の概念図を図 1 に示す.

2.2 量子回路学習の優位性と考慮すべき点

量子回路学習は, 指数関数的な数の基底関数を用いて古典データを量子状態にエンコードして計算するため, 古典

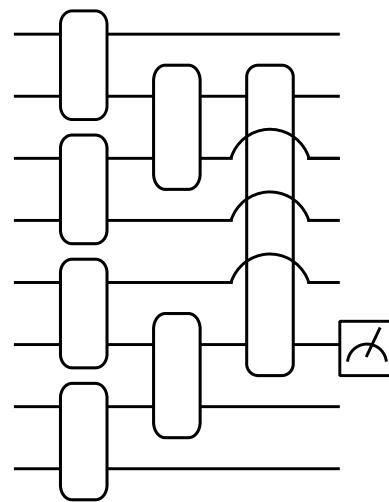


図 2 ツリーテンソルネットワーク型量子回路の例.

機械学習に対し優れた学習性能を発揮することが期待されている. ところが, 単に上記のアルゴリズムを実行したからといって必ずしも量子的な優位性があるとは限らない. 例えば量子カーネルをかけた後の量子状態同士の積が簡単に計算できるならば量子回路学習のタスクは古典カーネル法を用いて解くことができってしまう [10], [11]. 他にも, 浅い量子回路で生み出される量子状態のエンタングルメントが小さい場合は量子回路の期待値を古典シミュレーションにより高速に求めることができってしまう.

また, NISQ を用いた勾配計算アルゴリズムはパラメータ数の二乗の計算量オーダーになってしまう上, ノイズが入るため正確な勾配を求めようとすると多くの繰り返し回数を必要とする. さらに, 量子回路を用いた学習において勾配が消失してしまう Barren plateaus 問題 [12] も避けては通れず, 誤差逆伝播法を用いた高精度かつ効率の良い勾配計算が可能な古典コンピュータによる学習に優位性を取るのが難しい.

3. テンソルネットワーク構造を用いた量子回路学習

3.1 テンソルネットワーク構造

量子回路学習がどの程度の性能あるいはどのような性質を持ち, 本当に古典コンピュータを用いた学習に対して優位性があるのかを調べるのは非常に重要な問題である. ところが, そもそも一般に n 量子ビット系の量子コンピュータを古典コンピュータでシミュレートするには時間計算量, 空間計算量ともに $O(2^n)$ だけかかってしまうので, 実際に大規模量子回路学習に耐えうる NISQ デバイスが開発されていない現状ではその性能を推察することは難しい.

そこで, 量子回路をあえて古典シミュレーションしやすい形に限定することで, 仮想的に大規模量子回路学習を実行することが考えられている. 具体的には, 量子回路の構造として, 量子多体系の分野の計算手法としてよく用いら

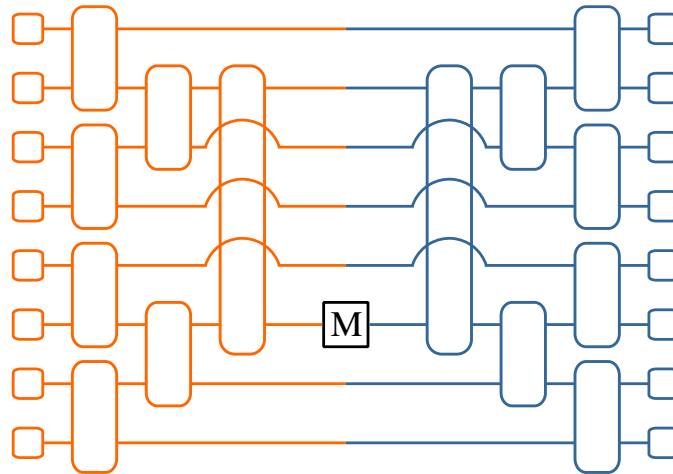


図 3 ツリーテンソルネットワーク型量子回路の期待値計算.

れるテンソルネットワーク法の Ansatz(仮説) を採用することで実現できる.

例えば図 2 のような量子回路はツリーテンソルネットワーク型と呼ばれる [4], [5]. このとき先端の 1 量子ビット測定の結果の期待値計算は図 3 のようになり, 多項式時間で計算できる.

もちろん, 古典シミュレーション可能であることから必然的に量子性あるいは量子超越性は存在しないが, このような古典シミュレーション可能なクラスに制限された量子回路学習を考えるのはいくつかの利点がある.

- 大規模量子回路学習が達成しうる最低限の学習性能が分かる. テンソルネットワーク構造を持つ量子回路も実機に実装可能であるのだから, 更に学習モデル(回路)を複雑にしたときに得られる理想的な損失関数の最小値はそこから必ず減少する. これにより, 大規模量子回路学習が具体的にどの程度のパフォーマンスを得られるかの目安, ベースラインを設定することができる.
- 古典コンピュータによる事前学習. テンソルネットワーク構造を仮定して古典コンピュータによって量子回路学習をシミュレートする際は誤差逆伝搬法が使えるため, 精度の良い勾配計算と性能の高い学習が可能である. そのようにして得られたパラメータは実機上で量子回路学習を実行する際の初期値として活用できる可能性がある.
- 大規模量子回路学習の性質の調査. 後述するエクスプレシビリティやエンタングル容量などの性質が大規模量子回路学習においてどのような役割を果たすのかについて, 古典計算の側から推察することができる. また, テンソルネットワークのエンタングルメント表現能力などの性質が量子回路学習の性質に深い関係がある可能性があり, その点についても研究を進めることができる.

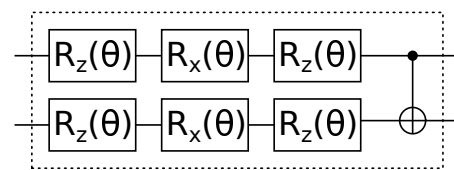


図 4 使用する 2 量子ビットゲート. 点線の部分を 1~3 回繰り返す.

3.2 ツリーテンソルネットワーク型量子回路を用いた学習

本節ではテンソルネットワーク構造を用いた 64 量子ビット, 256 量子ビットの量子回路学習の学習性能についての結果を記す.

3.2.1 データセットと量子カーネルの準備

学習データはそれぞれ 8×8 , 16×16 に圧縮された MNIST データセットのラベル 0, 1 を各 1000 枚ずつ用意し, 二値分類のタスクとする. 以下の量子カーネル

$$|\psi_{\text{in}}(x)\rangle = R_y(x_1)|0\rangle \otimes R_y(x_2)|0\rangle \otimes \dots \otimes R_y(x_n)|0\rangle \quad (5)$$

を用いることで古典データ $\{x_i\}$ を量子状態に埋め込む. ただし, R_y は y 軸まわりの回転ゲートで,

$$R_y(\theta) = \begin{pmatrix} \cos \theta/2 & -\sin \theta/2 \\ \sin \theta/2 & \cos \theta/2 \end{pmatrix} \quad (6)$$

である.

3.2.2 量子回路の準備

パラメータ付き量子回路として図 2 のツリーテンソルネットワーク構造を用いる. ツリーテンソルネットワークを構成する各 2 量子ビットゲートとしては図 4 のように点線で囲まれた 2 量子ビットゲートを 1~3 回繰り返すことで構成する. なお, 図中の R_x, R_z はそれぞれ x 軸, z 軸まわりの回転ゲートで, パラメータ θ はおのおの独立である.

3.2.3 測定の準備

ツリーテンソルネットワークの構造上, 最後に測定する量子ビットの位置は必然的にその頂点(図 2 では 2 番目, 6 番目の量子ビット)に限られる. 今回の学習ではそれら頂点のうちどちらかを測定する. 測定する物理量を表す

繰り返し回数	1	2	3
64 量子ビット	262.3206	30.2593	37.2234
256 量子ビット	253.4984	15.4703	20.1904

繰り返し回数	1	2	3
64 量子ビット	82.56%	99.56%	99.31%
256 量子ビット	87.5%	100%	99.75%

POVM について、 n 個の量子ビットのうちの i 番目の量子ビットを測定するとき、 M_1, m_1 を

$$M_1 = \mathbb{I}^{\otimes(i-1)} \otimes \frac{(\mathbb{I} + Z)}{2} \otimes \mathbb{I}^{\otimes(n-i-1)} \quad (7)$$

$$m_1 = 1 \quad (8)$$

とし、 M_0, m_0 を

$$M_0 = \mathbb{I} - M_1 \quad (9)$$

$$m_0 = 0 \quad (10)$$

とすると、物理量 $M = M_1$ の期待値を計算すれば必ず区間 $[0, 1]$ に入り、 M の期待値をそのままラベル 0, 1 の分類に対応させることができる。つまり、ラベル 1 の学習データに対しては物理量 M の期待値が 1 に近づくように、ラベル 0 の学習データに対しては物理量 M の期待値が 0 に近づくように学習すれば十分となる。

3.2.4 学習結果

以上の設定で量子回路学習を実行した結果を表 1 と表 2 に示す。2 量子ビットゲートの繰り返し回数が 1 回のときはあまり性能を発揮できないが、繰り返し回数 2 回以降では十分な性能を発揮し、MNIST データセットの画像 0 と 1 の分類タスクについて正答率がほぼ 100%を達成できる。

4. パラメータ付き量子回路の性質と学習性能

上の学習結果で見たように、用いるパラメータ付き量子回路によって最終的な学習性能は大きく異なる場合がある。学習性能からみたときのパラメータ付き量子回路の性質を定量化するものとしてエクスペシビリティとエンタングル容量が提案されている [6], [7], [8].

4.1 量子回路の出力の一般的な記述

パラメータ付き量子回路 $U(\theta)$ が与えられたとき、その量子回路の出力（期待値）は入力状態 ρ と物理量 M に依存して

$$C_{\rho, M}(\theta) = \text{Tr}[U(\theta)\rho U(\theta)^\dagger M] \quad (11)$$

と表される。量子回路学習の枠組みではこの量を目的の値に近づけるように最適化を行う。よって学習に用いるパラメータ付き量子回路 $U(\theta)$ としては、上の写像が与えられた状態と物理量に対して十分に高い表現能力を示すことが必要となる。

しかしながら、状態と物理量の両方に依存する写像を考えるのは複雑になるため、

- 入力状態 ρ を固定して、物理量 M に対して

$$\text{Tr}[U(\theta)\rho U(\theta)^\dagger (\cdot)] \quad (12)$$

- 物理量 M を固定して、入力状態 ρ に対して

$$\text{Tr}[U(\theta)(\cdot)U(\theta)^\dagger M] \quad (13)$$

の 2 種類の写像を考えることにする。

4.2 エクスペシビリティ

「十分に高い表現能力を持つ」ことを定量化するために、パラメータ付き量子回路の分布 $\{U(\theta)\}_{\theta \in \Theta}$ がユニタリ様分布 (Haar measure) の分布とどれだけ近いかを考え、以下の超作用素を導入する。

$$\mathcal{A}_{\Theta}^{(k)} := \int_{\text{Haar}} dU \cdot U^{\otimes k} (\cdot) (U^\dagger)^{\otimes k} - \int_{\Theta} d\theta \cdot U(\theta)^{\otimes k} (\cdot) (U(\theta)^\dagger)^{\otimes k} \quad (14)$$

式 (14) は、 k 次モーメントまでみたときに理想的なユニタリ分布と比べてどの程度逸脱しているかを示す。このとき、状態 ρ 、物理量 M に対するエクスペシビリティをそれぞれ以下で定義する。

$$\mathcal{E}_{\rho}^{(k)} := \left\| \mathcal{A}_{\Theta}^{(k)}(\rho^{\otimes k}) \right\|_{HS}^2 \quad (15)$$

$$\mathcal{E}_M^{(k)} := \left\| \mathcal{A}_{\Theta}^{(k)}(M^{\otimes k}) \right\|_{HS}^2 \quad (16)$$

エクスペシビリティを式変形すると、

$$\left\| \mathcal{A}_{\Theta}^{(k)}(\rho^{\otimes k}) \right\|_{HS}^2 \quad (17)$$

$$= \text{Tr}\{\mathcal{A}_{\Theta}^{(k)}(\rho^{\otimes k})^\dagger \mathcal{A}_{\Theta}^{(k)}(\rho^{\otimes k})\} \quad (18)$$

$$= \int_{\Theta} \int_{\Theta'} d\theta d\theta' \text{Tr}\{U(\theta)\rho U(\theta)^\dagger U(\theta')\rho U(\theta')^\dagger\}^k - \int_{\text{Haar}} \int_{\text{Haar}} dU dU' \text{Tr}\{U\rho U^\dagger U'\rho U'^\dagger\}^k \quad (19)$$

となる。途中で ρ が自己共役演算子であることを使った。結局 frame potential の差

$$\mathcal{F}^{(k)}(\rho^{\otimes k}) - \mathcal{F}_{\text{Haar}}^{(k)}(\rho^{\otimes k}) \quad (20)$$

を計算すればいいことがわかる。ここで $\rho^{\otimes k}$ に対する frame potential は

$$\mathcal{F}_{\Theta}^{(k)}(\rho^{\otimes k}) = \int_{\Theta} \int_{\Theta'} d\theta d\theta' \text{Tr}\{U(\theta)\rho U(\theta)^\dagger U(\theta')\rho U(\theta')^\dagger\}^k \quad (21)$$

で定義される [7], [13]. ところが、全ての k についてこれを計算し比較するのは指標としては不適切であるので、次のように frame potential を Fidelity の確率分布

$$p(F = \text{Tr}\{U(\theta)\rho U(\theta)^\dagger U(\theta')\rho U(\theta')^\dagger\}) \quad (22)$$

の k 次モーメントとみなす：

$$\mathcal{F}^{(k)}(\rho^{\otimes k}) = E[F^k] \quad (23)$$

そこで、式 (20) の計算を、パラメータ付き量子回路を用いたときの Fidelity の分布と Haar measure に従うときの Fidelity の分布間の距離で代用する。例えば Wasserstein 計量などが用いられ、本稿では Expressibility の計算にこの量を採用する。

ここまで入力状態 ρ に対して計算してきたが、全く同様のことが物理量 M に対しても成立する。

4.3 エンタングル容量

式 (12), (13) について、パラメータ θ を変えたときに、平均してその写像を構成する演算子

$$U(\theta)\rho U(\theta)^\dagger \quad (24)$$

$$U(\theta)MU(\theta)^\dagger \quad (25)$$

自身がどれだけのエンタングルメントを持つかを定量化する量がエンタングル容量である。例えば式 (24) について、状態が純粋状態 $|\psi\rangle$ のとき

$$U(\theta)|\psi\rangle\langle\psi|U(\theta)^\dagger \quad (26)$$

は純粋状態であり、エンタングルメントエントロピーを計算することによりその写像の持つ量子的な相関の強さがわかる。物理量 M についても事後操作を考えることで $M \geq 0$ と考えて一般性を失わないため、同様の議論で写像のエンタングルメントの強さを計算することができる。

ところが、 ρ あるいは M が混合状態のときは式 (26) は混合状態となり、量子回路学習の文脈において混合状態の持つ量子相関をどのように定義すればよいかは定かではない。今回は混合状態に対する簡易的なエンタングルメントの指標である対数ネガティビティ [14] を用いる。系を部分系 A, B に分けたときの対数ネガティビティは

$$E_N(\rho) = \log \|\rho^{TA}\|_1 \quad (27)$$

で定義される。ここで ρ^{TA} は ρ の部分転置であり、

$$\langle i_A, j_B | \rho^{TA} | k_A, l_B \rangle \equiv \langle k_A, j_B | \rho | i_A, l_B \rangle \quad (28)$$

を任意の各部分系における直交基底において満たす。また、 $\|A\|_1$ は A のトレースノルム

$$\|A\|_1 \equiv \text{Tr} \sqrt{A^\dagger A} \quad (29)$$

を表す。

4.4 実験結果

実際に 4 量子ビット系の量子回路学習を実行し、エクスプレシビリティ、エンタングル容量と学習性能の関係について調べる。

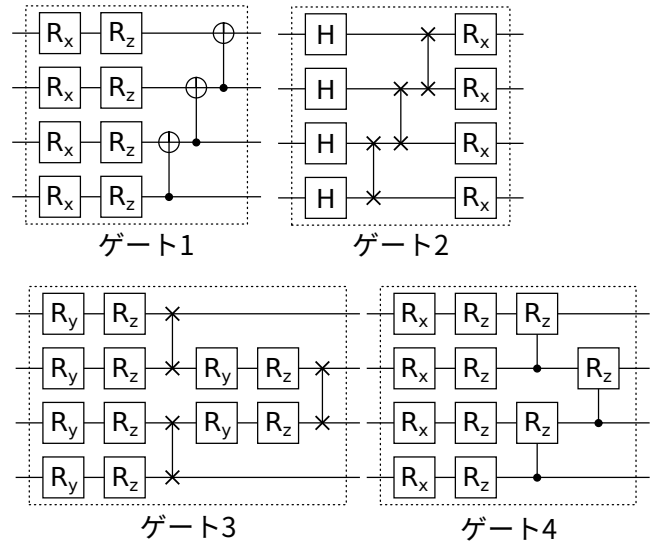


図 5 使用する 4 量子ビットゲート。

4.4.1 データセットの準備

今回は汎化能力ではなく、一般的なデータセットに対してどれだけ loss を下げることができるかを調べたいため、一様にサンプルされた 20 個ランダムなデータセット

$$\{|\psi_i\rangle\}_{i=1}^{20}, \quad |\psi_i\rangle \in (\mathbb{C}^2)^{\otimes n} \quad (30)$$

について、それぞれランダムにラベル 0 と 1 を割り振ったものを学習する。

4.4.2 量子回路、測定の準備

パラメータ付き量子回路として図 5 の 4 種類の量子ゲートを r 回繰り返しかけたものを使用する。測定について、1 番目の量子ビットを、物理量

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad (31)$$

で測定する。

4.4.3 エクスプレシビリティ、エンタングル容量の計算

ここから簡単のため、式 (15) に対応する量を単に Expressibility、式 (16) に対応する量を M-Expressibility、式 (24) に対応する量を単に Entangling capability、式 (25) に対応する量を M-Entangling capability と定義する。

入力状態について今回は $\rho = |0\rangle\langle 0|$ とし、初期状態 $|0\rangle$ からのパラメータ付き量子回路の時間発展を考える。物理量については上に定義した M を用いて $M \otimes \mathbb{I}^{\otimes 3}$ と記述される。Entangling capability の計算は、量子ビットが 1 次元的に並んでいるとした上で、その全ての 2 分割に対するエンタングルメントエントロピー、対数ネガティビティの平均をとっている。

4.4.4 学習結果

図 6 に、それぞれ Expressibility, M-Expressibility, Entangling capability, M-Entangling capability の値と学習後の loss の値の関係をプロットした。図から分かるとお

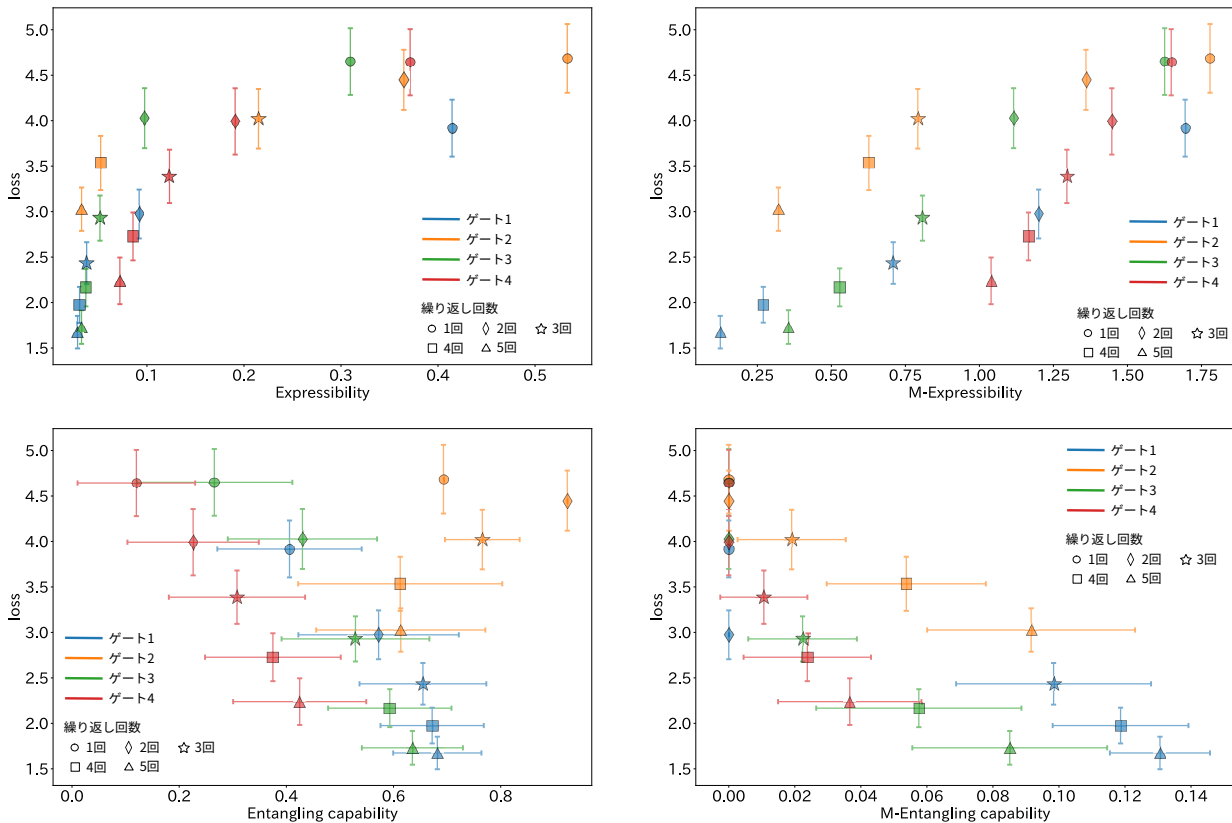


図 6 4 量子ビット量子回路学習における諸量と学習後の loss の関係.

り、どの指標についても loss との間に相関がみられた。特に M-Entangling capability は Entangling capability とは異なり測定する物理量の影響を考慮しており、全体として分離可能状態となってしまう回路を区別できている。

ところが、各ゲートごとに見れば強い相関を示すが、全てのゲートを合わせると弱い相関となり、これら指標と学習性能が必ずしも 1 対 1 対応しないことは注意しなければならない。

5. テンソルネットワーク構造を持つ量子回路の性質と学習性能

5.1 ツリーテンソルネットワーク構造に対する諸量の計算

テンソルネットワーク構造を持つ量子回路についても、性能の指標として同様にエクスペシビリティとエンタングル容量を考えることができる。

特に、ツリーテンソルネットワーク型量子回路においては、入力状態 $\rho = |0\rangle\langle 0|$ に対する Expressibility, Entangling capability の計算と、物理量 M として式 (31) で表される 1 量子ビットの測定を行ったときの M-Expressibility, M-Entangling capability について、量子ビット数に対して多項式時間で計算できる。

以下では、大規模量子回路学習において回路にツリーテンソルネットワーク構造を仮定し、その学習性能とエク

プレシビリティ、エンタングル容量の関係調べる。

5.2 実験結果

学習データセットとして主成分分析を用いて 8 次元に圧縮した MNIST データセットのラベル 0 と 1 の二値分類、ラベル 3 と 5 の二値分類を行う。データセットは各ラベルにつき 1000 枚とした。また、量子カーネルとして式 (5) を用いる、パラメータ付き量子回路として、

- 2 量子ビットゲートで構成された合計 3 層のツリーテンソルネットワーク構造
- 4 量子ビットゲートで構成された合計 2 層のツリーテンソルネットワーク構造

を用いる。2 量子ビットゲートについては図 4 を 5 回繰り返したものを、4 量子ビットゲートについては図 5 のゲート 1 を 5 回繰り返したものを使用する。なお、これらの 5 回繰り返した 2 量子ビット、4 量子ビットゲートはほぼ一様ランダムなユニタリ分布を示す。

結果を表 3 にまとめた。Entangling capability と M-Entangling capability については 2 量子ビットゲートを用いたときよりも 4 量子ビットゲートを用いたときの方が高く、それに伴って loss の値も下がっている。ところが、Expressibility, M-Expressibility については関係はよくわからない。実際、今回の物理量の設定においては、量

表 3 テンソルネットワーク構造を用いたときの 8 量子ビット回路学習後の loss と諸量の関係. ただし, Exp(Expressibility), M-Exp(M-Expressibility), Ent(Entangling capability), M-Ent(M-Entangling capability) とする.

	2 量子ビットゲート	4 量子ビットゲート
loss(0 と 1)	46.1413	33.9576
テスト精度 (0 と 1)	99.55%	99.68%
loss(3 と 5)	236.2590	209.3345
テスト精度 (3 と 5)	89.20%	92.56%
Exp	0.2481	0.3604
M-Exp	1.1572	0.4450
Ent	0.7415	1.2266
M-Ent	0.1071	0.1447

表 4 テンソルネットワーク構造を用いたときの 16 量子ビット回路学習後の loss と M-Entangling capability の関係. ただし M-Ent(M-Entangling capability) とする.

	2 量子ビットゲート	4 量子ビットゲート
loss(0 と 1)	34.0585	17.2008
テスト精度 (0 と 1)	98.93%	99.35%
loss(3 と 5)	325.8527	248.1283
テスト精度 (3 と 5)	78.55%	84.30%
M-Ent	0.1058	0.1396

表 5 テンソルネットワーク構造を用いたときの 64 量子ビット回路学習後の loss と M-Entangling capability の関係. ただし M-Ent(M-Entangling capability) とする.

	2 量子ビットゲート	4 量子ビットゲート
loss(0 と 1)	36.4093	11.7440
テスト精度 (0 と 1)	99.5%	99.25%
loss(3 と 5)	254.2854	147.5123
テスト精度 (3 と 5)	80.58%	90.67%
M-Ent	0.1091	0.1434

子ビット数を増やしていくと Fidelity の分布 (22) は鋭いピークを持つようになり, 単純な分布間の距離の比較は難しくなっていく.

同様の実験として, 4×4 , 8×8 に圧縮した MNIST データセットを用いた 16 量子ビット系, 64 量子ビット系量子回路学習を行い, その結果と回路の M-Entangling capability との関係を表 4 と 5 にまとめた. どちらも 8 量子ビット系と同様に, 4 量子ビットゲートを用いたほうが 2 量子ビットゲートを用いたときより学習性能が高く, また M-Entangling capability も高い.

6. 考察

本稿では, 大規模量子回路系における量子回路学習について調査するために, 古典コンピュータでシミュレート可能な量子回路のクラスを考え, それらの性能, 性質を具体的な数値計算により考察した. それにより, 大規模量子回路学習が最低限出するベースラインとなる性能を示した.

また, 状態, 物理量に依存するエクスペシビリティ,

エンタングル容量を考え, それらが少数量子ビット系においては学習性能と相関を持つこと確認した. これらの結果により, 状態, 物理量に依存するエクスペシビリティ, エンタングル容量がある種の量子回路の評価の指標になり得ることが示唆される. また, さらにツリーテンソルネットワーク構造を持つ量子回路について物理量に依存したエンタングル容量を計算し, 学習性能と比較した.

今後の課題として, これら指標を元にしたより効率の良い回路構成の提案や, 大規模量子系におけるより適切な性能の指標の構成などが挙げられる.

謝辞 本研究で使用した量子回路学習ライブラリは 2020 年度未踏ターゲット事業の支援を受けて作成されました [15].

参考文献

- [1] Preskill, J.: Quantum Computing in the NISQ era and beyond, *Quantum*, Vol. 2, p. 79 (2018).
- [2] Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J., Barends, R., Biswas, R., Boixo, S., Brandao, F., Buell, D., Burkett, B., Chen, Y., Chen, Z., Chiaro, B., Collins, R., Courtney, W., Dunsworth, A., Farhi, E., Foxen, B. and Martinis, J.: Quantum supremacy using a programmable superconducting processor, *Nature*, Vol. 574, pp. 505–510 (2019).
- [3] Mitarai, K., Negoro, M., Kitagawa, M. and Fujii, K.: Quantum Circuit Learning, *Physical Review A*, Vol. 98 (2018).
- [4] Grant, E., Benedetti, M., Cao, S., Hallam, A., Lockhart, J., Stojevic, V., Green, A. G. and Severini, S.: Hierarchical quantum classifiers, *npj Quantum Information*, Vol. 4, No. 1 (2018).
- [5] Huggins, W., Patil, P., Mitchell, B., Whaley, K. B. and Stoudenmire, E. M.: Towards quantum machine learning with tensor networks, *Quantum Science and Technology*, Vol. 4, No. 2, p. 024001 (2019).
- [6] Sim, S., Johnson, P. and Aspuru-Guzik, A.: Expressibility and Entangling Capability of Parameterized Quantum Circuits for Hybrid Quantum - Classical Algorithms, Vol. 2 (2019).
- [7] Holmes, Z., Sharma, K., Cerezo, M. and Coles, P. J.: Connecting ansatz expressibility to gradient magnitudes and barren plateaus (2021).
- [8] Nakaji, K. and Yamamoto, N.: Expressibility of the alternating layered ansatz for quantum computation, *Quantum*, Vol. 5, p. 434 (2021).
- [9] Hubregtsen, T., Pichlmeier, J., Stecher, P. and Bertels, K.: Evaluation of parameterized quantum circuits: on the relation between classification accuracy, expressibility, and entangling capability, *Quantum Machine Intelligence*, Vol. 3, No. 1, p. 9 (2021).
- [10] Havlíček, V., Córcoles, A., Temme, K., Harrow, A., Kandala, A., Chow, J. and Gambetta, J.: Supervised learning with quantum-enhanced feature spaces, *Nature*, Vol. 567, pp. 209–212 (2019).
- [11] Huang, H.-Y., Broughton, M., Mohseni, M., Babbush, R., Boixo, S., Neven, H. and McClean, J. R.: Power of data in quantum machine learning, *Nature Communications*, Vol. 12, No. 1, p. 2631 (2021).
- [12] Mcclean, J., Boixo, S., Smelyanskiy, V., Babbush, R.

and Neven, H.: Barren plateaus in quantum neural network training landscapes, *Nature Communications*, Vol. 9 (2018).

- [13] Roberts, D. A. and Yoshida, B.: Chaos and complexity by design, *Journal of High Energy Physics*, Vol. 2017, No. 4, p. 121 (2017).
- [14] Vidal, G. and Werner, R. F.: Computable measure of entanglement, *Phys. Rev. A*, Vol. 65, p. 032314 (2002).
- [15] : <https://github.com/wotto27oct/QTensorNet>.

付 録

A.1 ツリーテンソルネットワーク型量子回路における M-Entangling capability の計算

ツリーテンソルネットワーク型量子回路における物理量 M に対する M-Entangling capability の計算の概念図を図 A.1 に示す. 対数ネガティビティの式を書き下すと

$$E_N(\rho) = \log \|\rho^{TA}\|_1 \quad (\text{A.1})$$

$$= \log \text{Tr} \sqrt{(\rho^{TA})^\dagger \rho^{TA}} \quad (\text{A.2})$$

であるが, 式 (A.2) より ρ^{TA} にユニタリ変換を左右どちらからかけても不変となる. よって $E_N(\rho)$ の計算は, 図 A.1 の右図のように, $O(\log n)$ 個のテンソルからなる 2^k 次元の小さな密度演算子 ρ_1^{TA} と恒等演算子 $I^{\otimes(n-k)}$ のテンソル積からなる $\rho_1^{TA} \otimes I^{\otimes(n-k)}$ の対数ネガティビティの計算に等しい. エルミート演算子のトレースノルムは固有値の絶対和に等しく, さらに $n-k$ 個の恒等演算子のテンソル積は固有値の $2^{(n-k)}$ 重縮退を引き起こすだけなので, 結局 ρ_1^{TA} の固有値を求めれば十分となる.

また, 各量子系として d 次元量子系を考えたとき, $k = O(d/2 \cdot \log n)$ であるので, ρ_1^{TA} の次元は $O(2^{d/2 \cdot \log n}) = O(n^{d/2})$ となり, ρ_1^{TA} の固有値は多項式時間で計算できる.

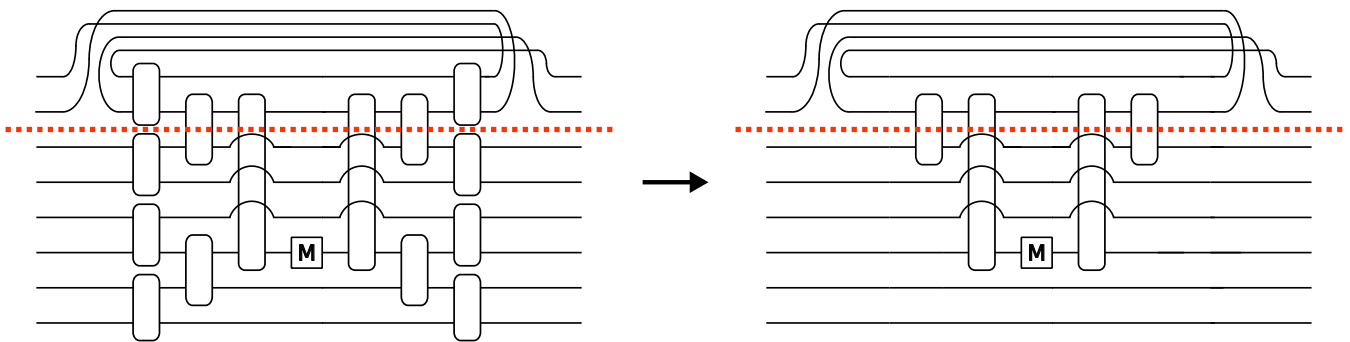


図 A.1 ツリーテンソルネットワーク型量子回路における M-Entangling capability の計算.