

# 生成モデルを用いた量子状態トモグラフィーに基づく ノイズレスな期待値の推定

前蔵 遼<sup>3,1,a)</sup> 鈴木 泰成<sup>1,2,b)</sup> 吉岡 信行<sup>4,2,c)</sup> 徳永 裕己<sup>1,d)</sup>

**概要:** 近年、NISQ デバイスを用いた量子化学計算や量子機械学習などの様々な応用が研究されている。しかし、NISQ デバイスでは大量の量子ビットを必要とする量子誤り訂正を行うことができず、計算中のノイズの影響が大きいという問題がある。そこで、統計的な処理を行うことでエラーを低減でき、NISQ デバイスでも実行可能な量子エラー抑制が注目されている。本研究では、生成モデルを用いてノイズのある測定データから量子状態を再構成し、ノイズレスなオブザーバブルの期待値推定を行った。

## Estimation of noiseless expectation value by quantum state tomography with generative model

**Abstract:** To obtain reliable computational results with noisy quantum computers, we need to suppress computational errors stemming from high error rates of qubits. Thus, quantum error mitigation methods are eagerly explored since they can suppress physical errors without requiring additional qubits. The virtual distillation method is expected as an efficient error mitigation method. However, the applicable scope of the virtual distillation method is restricted since it requires several copies of quantum states at the same time, which is difficult to achieve with the current devices. In this paper, we propose a novel method for estimating the noiseless expectation values from several measurement results of noisy quantum states. The key of our idea is to train neural-network quantum states to reproduce the measurement distribution of the unknown quantum states, and apply virtual distillation method to the trained neural-network quantum states.

### 1. Introduction

量子計算機は物質の量子性を活用した計算機であり、いくつかの問題について通常の計算機よりも指数関数的な高速化を実現すると期待されているため、実用的な規模の量子計算機の開発が世界で盛んに行われている [1]。量子計算機の開発における課題の一つに、量子ビットの持つ大きな誤り率がある。この誤り率のため、素朴に量子計算機を動かすだけでは信頼性のある計算結果を得ることは難しい。従って、量子計算機を効率的に活用するには、誤り率を実効的に削減する工夫が重要となる。量子誤り抑制は、誤り率の大きな量子計算機の結果を組み合わせることで、信頼

性のある計算結果を得るための手法の総称である。量子誤り抑制は従来の量子誤り訂正に比べ、計算の繰り返しにより小さなオーバーヘッドで計算結果に残る誤差を除去することができる [2], [3]。特に、量子誤り抑制手法の一つである仮想蒸留法は、計算の対象となる量子状態のコピーを複数用意し derangement と呼ばれる操作を行って測定することで、計算結果に生じるバイアスをコピーの数に対し指数関数的に削減できる手法として注目を集めている [4], [5]。しかし、提案された仮想蒸留法では、生成された量子状態のコピーを、同時に量子計算機上に実現する必要があることから、サイズの大きな量子計算機が必要となる。また、仮想蒸留法で必要となる derangement の操作自体にもエラーが生じるため、利用できるコピーの数には限界があるという点が仮想蒸留法の欠点となっていた。

本研究では、ニューラルネットワークを用いた量子状態の表現、つまり、ニューラルネットワーク量子状態 (Neural network quantum state, NQS) [6], [7], [8] を用いて量子状態を一旦古典情報の形に学習し、ニューラルネットワーク

<sup>1</sup> NTT コンピュータ&データサイエンス研究所

<sup>2</sup> JST さきがけ

<sup>3</sup> 東京大学大学院 総合文化研究所

<sup>4</sup> 東京大学大学院 工学系研究科

a) ryomaekura@g.ecc.u-tokyo.ac.jp

b) yasunari.suzuki.gz@hco.ntt.co.jp

c) nyoshioka@ap.t.u-tokyo.ac.jp

d) yuuki.tokunaga.bf@hco.ntt.co.jp

で表現された量子状態に仮想蒸留法と同じアイデアの手法を適用することで、低コストにノイズの影響を削減できる手法を提案する。ニューラルネットワークを用いて量子状態を表現するには一般に量子ビット数に対して指数関数的なサイズのメモリを必要とするものの、多体量子状態の中でも物理的に興味深い性質を示すものの多くは、多項式的な数の変分パラメータで構成されるニューラルネットワーク量子状態によって効率的に表現できることが知られている [7], [8], [9], [10]。NQS で表現された状態は古典的な情報であるから容易にコピーが可能であり、モンテカルロサンプリングを通じて物理量を効率的に評価することができる。本研究では、NQS によって表現された量子状態に対しても仮想状態蒸留法を効率的に実施することができ、ノイズの影響を排除した期待値が効率的に得られることを示す。我々は実際に本手法を数値的に検証し、2 量子ビットのノイズが生じたベル状態について実際に期待値のノイズが除去できることを示した。

## 2. Preliminary

### 2.1 制限ボルツマンマシン

制限ボルツマンマシン (Restricted Boltzmann machine, RBM) は図 1 のように  $n$  個の可視ユニット  $v_i \in \{0, 1\}$  と  $m$  個の不可視ユニット  $h_j \in \{0, 1\}$  を持つユニットと 2 つ変数間の結合を持つ、2 部グラフで構成されるニューラルネットワークの一種である [11], [12], [13]。RBM のエネルギー関数は以下のように定義される。

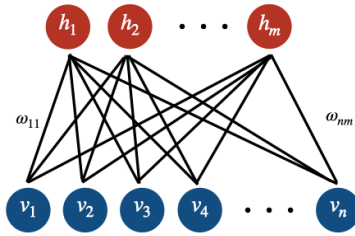


図 1 RBM の構造

エネルギー関数は以下のように定義される。

$$E(\mathbf{v}, \mathbf{h}, \boldsymbol{\theta}) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_i \sum_j \omega_{ij} v_i h_j \quad (1)$$

ここで、 $\boldsymbol{\theta} = (\{a_i\}_i, \{b_j\}_j, \{\omega_{ij}\}_{ij})$  は RBM を特徴づける学習可能なパラメータである。RBM を上記で定義されるエネルギーを持ったスピン系だとみなすと、状態  $(\mathbf{v}, \mathbf{h})$  が実現される確率  $p(\mathbf{v}, \mathbf{h})$  は以下のように書くことができる。

$$p_{\boldsymbol{\theta}}(\mathbf{v}, \mathbf{h}) = \exp(-E(\mathbf{v}, \mathbf{h}, \boldsymbol{\theta})) / Z_{\boldsymbol{\theta}} \quad (2)$$

ただし、 $Z$  はこのモデルの分配関数であり、 $Z_{\boldsymbol{\theta}} = \sum_{\mathbf{v}, \mathbf{h}} e^{-E(\mathbf{v}, \mathbf{h}, \boldsymbol{\theta})}$  である。ここで、可視ユニットからスピン列  $\mathbf{v}$  を得る確率を、 $\mathbf{h}$  に関して周辺化した確率とみなすと

以下のように計算できる。

$$p_{\boldsymbol{\theta}}(\mathbf{v}) = \sum_{\mathbf{h}} p_{\boldsymbol{\theta}}(\mathbf{v}, \mathbf{h}) \quad (3)$$

上記における分配関数の計算は 2 部グラフで表現表現できないような一般のボルツマンマシンでは、不可視ユニットの数に対して指数的に増加する配列すべてについて総和を取る必要があるため困難である。一方で、RBM では、上記の部分和の表式を解析的に求めることが可能であることから、規格化されていない確率振幅  $\tilde{p}_{\boldsymbol{\theta}}(\mathbf{v}) = Z_{\boldsymbol{\theta}} p_{\boldsymbol{\theta}}(\mathbf{v})$  は、可視ユニットや不可視ユニットの数に対して多項式的な時間で計算することができる。従って、RBM に対しては以下の二つの操作を可視ユニットと不可視ユニットの数に対して多項式時間  $\text{poly}(n, m)$  で行うことができる。

- サンプルング:  $p_{\boldsymbol{\theta}}(\mathbf{v})$  の確率で添え字  $\mathbf{v} \in \{0, 1\}^n$  を得る。
- 振幅の計算: 与えられた  $\mathbf{v} \in \{0, 1\}^n$  について、規格化されていない確率  $\tilde{p}_{\boldsymbol{\theta}}(\mathbf{v})$  の値を計算する。

不可視ユニットの数を増やし自由度を大きくし、適切な  $\boldsymbol{\theta}$  を選ぶことで、RBM は多様な確率分布をサンプリング可能な形で効率的に表現することができる。

$n$  ビット列に関するある確率分布  $\{q(\mathbf{v})\}_{\mathbf{v}} (\mathbf{v} \in \{0, 1\}^n)$  があり、我々は  $\{q(\mathbf{v})\}_{\mathbf{v}}$  の確率分布から得られた  $S$  個のデータ  $D = (\mathbf{v}^{(1)} \dots \mathbf{v}^{(S)})$  を保持しているとする。この時、 $D$  を用いて  $n$  個の可視ユニットを持つ RBM の学習パラメータ  $\boldsymbol{\theta}$  を最適化し、RBM から得られる確率分布  $\{p_{\boldsymbol{\theta}}(\mathbf{v})\}_{\mathbf{v}}$  と対象となる確率分布  $\{q(\mathbf{v})\}$  を出来るだけ一致させるというタスクを学習と呼び、 $D$  を訓練データと呼ぶ。ある学習パラメータ  $\boldsymbol{\theta}$  のもとで、RBM から  $S$  個の独立なサンプリングを可視層から行った結果としてデータセット  $D$  を観測する尤度関数およびその対数は下記のように書ける。

$$p_D(\boldsymbol{\theta}) = \prod_i p(\mathbf{v}^{(i)}) \quad (4)$$

$$\mathcal{L}_D(\boldsymbol{\theta}) = \ln p_D(\boldsymbol{\theta}) = \sum_i \ln p(\mathbf{v}^{(i)}) \quad (5)$$

観測した訓練データの下で尤もらしい学習パラメータを探るには、上記の対数尤度関数を最大化する  $\boldsymbol{\theta}$  を探せばよい。従って、尤度を最大化するために、対数尤度に対する各学習パラメータでの微分が 0 となるように  $\boldsymbol{\theta}$  を最適化する。これを実現する簡便な手法は、現地点でのパラメータを  $\boldsymbol{\theta}^{(i)}$  として、ステップごとに学習パラメータを勾配に従って逐次的に最適化していく勾配最適化法である。

$$\boldsymbol{\theta}^{(i+1)} \leftarrow \boldsymbol{\theta}^{(i)} + \eta \frac{\partial}{\partial \boldsymbol{\theta}} \mathcal{L}_D(\boldsymbol{\theta}) \quad (6)$$

ここで、 $\eta$  は学習率と呼ばれる非負の実数であり、学習の際にヒューリスティックに決定する。勾配最適化を行うにはパラメータの学習に各パラメータで対数尤度を微分した値  $\frac{\partial}{\partial \boldsymbol{\theta}} \mathcal{L}_D(\boldsymbol{\theta})$  の評価が毎ステップ必要となる。一般に、個

の微分値を厳密に求めるためには指数関数的な計算量が必要になってしまう一方で、モンテカルロサンプリングに基づく Contrastive divergence 法 [13] などによって効率的に評価できることが経験的に知られている。この方法で1ステップ勾配最適化を行うことを、1 epoch と呼ぶ。

## 2.2 制限ボルツマンマシンを用いた量子状態の表現

この節では RBM を用いた量子状態の表現方法について解説する [6]。\$n\$-qubit の量子状態は一般に \$2^n \times 2^n\$ の密度行列 \$\rho\$ で表現することができるが、まずは \$\rho = |\psi\rangle\langle\psi|\$ と分解できる純粋状態を考え、さらにパウリ \$Z\$ 基底で表示した \$|\psi\rangle\$ の各要素が非負の実数であるような特殊なケースを考える。この場合、我々が量子状態に対する \$Z\$ 測定の結果、\$n\$-bit 列 \$x\$ を得る確率は \$q(x) = |\langle x|\psi\rangle|^2\$ で与えられる。従って、表現したい量子状態に対して \$Z\$ 測定を行った \$S\$ 個の \$n\$-bit 列を学習データ \$D\$ としたとき、RBM から得られるサンプルの確率分布 \$\{p\_\theta(x)\}\$ の、\$D\$ に対する対数尤度が最大化されるように学習を行うことができる。2.1 で導入した RBM では各要素が非負の実数であるような限られた状態ベクトルしか表現することができないが、いくつかの拡張によって複素係数を持つ任意の純粋状態や、一般の密度行列を表現できるようになる [7], [8]。複素係数をもつ純粋状態への拡張は、下記のように二つの確率分布 \$\{p\_\theta\}\$ と \$\{p\_\mu\}\$ を用いて可能となる。

$$|\psi\rangle = \sum_{\mathbf{v}} \sqrt{p_\theta(\mathbf{v})} \exp(i\phi_\mu(\mathbf{v})/2) |\mathbf{v}\rangle \quad (7)$$

この時、\$\{\phi\_\mu(\mathbf{v})\}\$ は後者の RBM から得られる規格化されていない振幅 \$\tilde{p}\_\mu(\mathbf{v})\$ について、\$\phi\_\mu(\mathbf{v}) = \log \tilde{p}\_\mu(\mathbf{v})\$ として得られる値である。ここで、パラメータ \$\theta, \mu\$ を最適化するには、一般に \$Z\$ 基底で無いような複数のパウリ基底における測定結果を学習データとして、2.1 で導入した最尤推定をそれぞれに対して実行すればよい。混合状態については、興味の対象となる系の可視ユニットに加えて、補助ユニットを加えた拡張ヒルベルト空間上において純粋化された量子状態を学習し、部分トレースを通して対象となる混合状態を表現する。すなわち、任意の \$n\$-qubit の密度行列 \$\rho\$ について、\$\rho = \text{Tr}\_2[|\psi\rangle\_{\text{pf}}\langle\psi|\_{\text{pf}}]\$ となるような高々 \$2n\$ 量子ビットに関する状態 \$|\psi\rangle\_{\text{pf}}\$ が存在することを用いて、\$|\psi\rangle\_{\text{pf}}\$ を RBM で表現することで学習を行う。ただし、\$\text{Tr}\_2[\cdot]\$ は追加した高々 \$n\$-qubit の空間について部分トレースを取る操作である。制限ボルツマンマシンで学習した純粋状態(混合状態)に対しては、以下の二つの操作を可視ユニットと不可視ユニットの数に対して多項式時間で行うことができる。

- サンプリング: \$|\psi\_x|^2\$ (または \$\rho\_{xx}\$) の確率で添え字 \$x \in \{0, 1\}^n\$ を得る。
- 振幅の計算: 与えられた \$x \in \{0, 1\}^n\$ (または \$x, y \in \{0, 1\}^n\$) について、規格化されていない量子状態の振幅 \$\tilde{\psi}\_i\$ (または \$\tilde{\rho}\_{xy}\$) の値を計算する。

一般に任意の量子状態を学習するには指数関数的な量子状態が必要となる。このことは最悪ケースでは不可視ユニットが指数関数個必要となることを示唆しており、任意の \$n\$-qubit の量子状態が空間的にも時間的にも効率的に学習できるわけではない。一方、量子ビットの数が十分に少ないケースでは本手法は有効であるため、少数の量子ビットからなるデバイスのトモグラフィに有効である [7], [8], [9]。また、物性分野において重要なハミルトニアン基底状態・励起状態や非平衡状態の一部は、RBM で効率的に表現できることが知られている [10], [14], [15], [16]。

## 2.3 仮想蒸留法 (Virtual distillation)

仮想蒸留法は、複数のノイズのある量子状態を用いて、出来るだけノイズが軽減された期待値を計算する手法である [4], [5]。我々は、初期状態 \$|\psi\_0\rangle\$ に対して、あるユニタリ操作 \$U\$ を適用して得られる \$|\psi\rangle\_{\text{ideal}} = U|\psi\_0\rangle\$ に対する、オブザーバブル \$O\$ の期待値 \$\langle O \rangle\_{\text{clean}} = \langle \psi\_0|O|\psi\_0\rangle\$ を正確に知りたいとする。ただし、オブザーバブル \$O\$ は多項式個のパウリ演算子を用いて実数係数 \$\alpha\_k\$ で \$O = \sum\_k \alpha\_k P\_k\$ と展開出来ているとする。期待値を計算するにあたり、我々はノイズの無い量子回路 \$U\$ は実行できず、ノイズの乗った量子回路 \$\mathcal{E}\_U\$ しか実行することが出来ないとする。この時、得られる状態は \$\rho\_{\text{noisy}} = \mathcal{E}\_U(|\psi\_0\rangle\langle\psi\_0|)\$ であり、得られる期待値は \$\langle O \rangle\_{\text{noisy}} = \text{Tr}[\rho\_{\text{noisy}}O]\$ と異なる値になってしまう。

仮想蒸留法は、ノイズに対するいくつかの仮定の下で、ノイズのある量子状態 \$\rho\_{\text{noisy}}\$ のコピー \$m\$ 個を用いて、\$m\$ の値が大きくなるにつれ漸的にバイアスの小さな \$\langle O \rangle\_{\text{clean}}\$ の推定量を得る手法である。\$\rho\_{\text{noisy}}\$ は密度行列の性質より半正定値行列であるからスペクトル分解が可能であり、正規直交基底をなす固有ベクトルと \$\{|\lambda\_i\rangle\}\$ と非負実数の固有値 \$\lambda\_i \geq 0\$ を用いて以下のように展開できる。

$$\rho_{\text{noisy}} = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i| \quad (8)$$

ここで固有値 \$\lambda\_i\$ は \$\lambda\_1 \geq \lambda\_2 \geq \dots \lambda\_{2^n} \geq 0\$ と降順に並んでいるものとする。この時、密度行列を \$m\$ 乗して規格化した状態 \$\frac{\rho\_{\text{noisy}}^m}{\text{Tr}[\rho\_{\text{noisy}}^m]}\$ を蒸留された状態と呼ぶ。蒸留された状態は以下のように書ける。

$$\frac{\rho_{\text{noisy}}^m}{\text{Tr}[\rho_{\text{noisy}}^m]} = \sum_i \frac{\lambda_i^m}{\sum_j \lambda_j^m} |\lambda_i\rangle\langle\lambda_i| \quad (9)$$

蒸留された状態と最大固有値に対応する状態 \$|\lambda\_1\rangle\$ との忠実度は

$$\langle \lambda_1 | \frac{\rho_{\text{noisy}}^m}{\text{Tr}[\rho_{\text{noisy}}^m]} | \lambda_1 \rangle = \frac{\lambda_1^m}{\sum_j \lambda_j^m} \quad (10)$$

となることから、\$m\$ の増加に対して指数関数的に 1 に

近づいていく。従って、ノイズの乗った状態  $\rho_{\text{noisy}}$  の最大固有値ベクトルがノイズの無い量子状態に十分近い ( $|\lambda_1\rangle\langle\lambda_1| \simeq |\psi\rangle\langle\psi|$ ) 場合に、蒸留された量子状態の期待値  $\frac{\rho_{\text{noisy}}^m}{\text{Tr}[\rho_{\text{noisy}}^m]}$  は漸的にノイズの無い状態の期待値に近づいていく。蒸留された量子状態の期待値を以下のように定義する。

$$\langle O \rangle_{\text{VD},m} := \frac{\text{Tr}[\rho_{\text{noisy}}^m O]}{\text{Tr}[\rho_{\text{noisy}}^m]} \quad (11)$$

一般に密度行列を  $m$  乗して規格化する量子回路は物理的には構築出来ない。しかし、 $m$  個のコピーされた量子状態  $\rho^{\otimes m}$  がある場合、以下のようにして式 (11) の分子  $\text{Tr}[\rho_{\text{noisy}}^m O]$  と分母  $\text{Tr}[\rho_{\text{noisy}}^m]$  をそれぞれ評価することが出来る。 $m$  個の  $n$ -qubit 量子状態の計算基底を  $|x_1 \dots x_m\rangle (x_i \in \{0, 1\}^n)$  とする。この時、 $n$ -qubit の計算の添え字を一つずつ左に移動し、一番左の  $x_1$  を一番右に移動するユニタリ操作  $\Lambda$  を derangement と呼ぶ。すなわち、 $\Lambda |x_1 \dots x_m\rangle = |x_2 x_3 \dots x_m x_1\rangle$  である。Derangement operator を用いると、密度行列の  $m$  乗に関する期待値を、 $m$  個のコピーに関するトレースの値として表現しなおすことができる。

$$\text{Tr}[\rho_{\text{noisy}}^{\otimes m} (O \otimes I^{\otimes(m-1)}) \Lambda] = \text{Tr}[\rho_{\text{noisy}}^m O] \quad (12)$$

上記の式で  $O = I$  とすれば  $\text{Tr}[\rho_{\text{noisy}}^m]$  も同様の変形が可能となる。

上記の形式のトレースは、 $|+\rangle$  状態に初期化された補助量子ビットを加え、この補助量子ビットを制御量子ビットとして動作する  $\Lambda_{k,c} = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes (P_k \otimes I^{\otimes(m-1)}) \Lambda$  を用いると実験的に評価できる。この時、状態  $\rho^{\otimes m}$  に対して  $\Lambda_{k,c}$  を作用して補助量子ビットに対して  $X$  測定を行うと、 $|+\rangle$  で射影される確率は以下ようになる。

$$p_{k,\pm} = \langle \pm | \Lambda_{k,c} (|+\rangle\langle +| \otimes \rho^{\otimes m}) \Lambda_{k,c} | \pm \rangle \quad (13)$$

$$= \frac{1}{2} (1 \pm \text{Tr}[\rho^{\otimes m} (P_k \otimes I^{\otimes(m-1)}) \Lambda]) \quad (14)$$

従って、 $\text{Tr}[\rho^{\otimes m} (P_k \otimes I^{\otimes(m-1)}) \Lambda] = 2p_{k,+} - 1$  として求めることができる。最終的に求めたい  $O$  に関する期待値は下記のように与えられる。

$$\sum_k \alpha_k (2p_{k,+} - 1) = \text{Tr}[\rho^m O] \quad (15)$$

上記の仮想蒸留法は理想的には推定量のバイアスを  $m$  に対して指数関数的に小さくすることができる。一方で、仮想蒸留法を用いることによるデメリットは下記の通りである。

- $m$  の拡大に対して式 (11) の分子分母の期待値は分散は定数のままに指数関数的に小さくなる。このため得られる期待値の分散は指数関数的に大きくなる。
- ノイズが乗った量子状態の最大固有値ベクトル  $|\lambda_1\rangle$  が

ノイズの無い状態  $|\psi\rangle_{\text{ideal}}$  から遠い場合、一定のバイアスが残ってしまう。従って、コヒーレントなエラーに対しては有効性が薄い。

- $m$  個のコピーを同時に用意し制御 derangement operator を作用しなければならないため、素朴なケースに比べ  $m$  倍のサイズの量子計算機が必要となるだけでなく、この操作に起因した無視できない量のノイズが生じる。

### 3. Method

本研究では、NQS を通して蒸留された量子状態の期待値を推定することでバイアスの小さい期待値を与える4つの手法を提案する。これにより、前節の末尾に挙げた欠点のうち、最後の欠点を取り除くことができる。また、NQS からのサンプリングは古典計算機で容易に並列化できるため、指数関数的な分散の拡大の影響も緩和できると期待される。

#### 3.1 手法1：密度行列のトモグラフィによる期待値蒸留

RBM を混合状態を生成するように学習させ、密度行列  $\rho_{\text{train}}$  を表現する RBM を得る。さらに、RBM が表現する密度行列の全要素を取得して規格化することで、RBM が表現している密度行列  $\rho_{\text{train}}$  を直接得ることができる。さらに、密度行列  $\rho_{\text{train}}$  を対角化して

$$\rho_{\text{train}} = \lambda_1 |\lambda_1\rangle\langle\lambda_1| + \lambda_2 |\lambda_2\rangle\langle\lambda_2| + \dots \quad (16)$$

のように分解し、最大固有値ベクトル  $|\lambda_1\rangle$  を得る。RBM が表現する状態が正確であり、かつ、仮想蒸留法の仮定である、密度行列の最大固有値に対応する固有ベクトルがノイズレスな状態ベクトルに近いという条件 ( $|\psi\rangle_{\text{ideal}} \simeq |\lambda_1\rangle$ ) が成り立つ場合、 $|\psi\rangle_{\text{ideal}}$  に十分近い状態  $|\lambda_1\rangle$  を得ることができる。この  $|\lambda_1\rangle$  について

$$\langle O \rangle = \langle \lambda_1 | O | \lambda_1 \rangle \quad (17)$$

を評価し、蒸留された期待値を得る。

この手法は RBM が表現する状態が正確であれば厳密な最大固有値ベクトルを得ることができるが、RBM から規格化された密度行列を計算したのち、対角化を実行する必要があるため、ヒルベルト空間の次元  $d = 2^n$  に対して  $O(d^2)$  のメモリと  $O(d^3)$  の計算量が必要となる。

#### 3.2 手法2：状態ベクトルを用いた期待値推定

混合状態であることが分かっている量子状態に対して、純粋状態のみが表現できる RBM を用いて学習を行う。ここで、もし  $\rho_{\text{noisy}}$  に対して、忠実度を最大化するような  $|\psi\rangle_{\text{train}}$  を学習するようなプロトコルを採用した場合には、RBM が表現する純粋状態は  $\rho_{\text{noisy}}$  の最大固有ベクトルと等しくなる。この状態ベクトルの全要素を取り出して状態

ベクトル  $|\psi_{\text{train}}\rangle$  を再構築し、下記の通りオブザーバブルの期待値を数値的に計算する。

$$\langle O \rangle = \langle \psi_{\text{train}} | O | \psi_{\text{train}} \rangle \quad (18)$$

この手法も RBM から  $2^n$  のオーダーの次元のベクトルを再構成して期待値を評価するため、 $n$  が大きいときには適用できない。また、有限サイズのデータと学習時間で、RBM の状態が  $\rho_{\text{noisy}}$  の最大固有値ベクトルに十分近い状態を学習する保証もない。一方で、ノイズによって生成された混合状態ではなく、最大固有ベクトルに対応する純粋状態を学習するため、必要なメモリと計算量は共に  $O(d)(=O(2^n))$  に抑えられるという利点がある。

### 3.3 手法3：状態ベクトルからのサンプリングによる期待値推定

手法2のように純粋状態の RBM の全要素を再構築する方法は正確ではあるが、必要なメモリと計算量は、量子ビット数  $n$  に対して、指数関数的に増加してしまう。一方で、モンテカルロ・サンプリングを活用すると、全計算基底に関する和や、規格化定数  $Z_\theta$  の計算を陽に実行することなく、オブザーバブルに対する期待値が評価できる。

規格化されていない純粋状態  $|\psi\rangle = \sum_i \psi_i |i\rangle$  について、オブザーバブルの期待値は下記のように展開できる。

$$\langle O \rangle = \frac{\langle \psi | O | \psi \rangle}{\langle \psi | \psi \rangle} = \mathbb{E}_{i \sim \{p_i\}} \left[ \sum_j \frac{\psi_j}{\psi_i} O_{i,j} \right] \quad (19)$$

ここで、 $i \sim \{p_i\}$  は  $p_i$  の確率で添え字  $i$  を取り出す操作である。ここで、 $j$  に関する総和は指数関数的なパターンを一般に含むが、オブザーバブルがパウリ演算子の多項式個の和であるとき、 $O$  の各行各列は高々多項式個の非ゼロ要素しか含まないため、効率的に評価を行うことができる。従って、純粋状態の NQS から添え字  $i$  を分布  $\{p_i\}$  からサンプリングして集合  $S$  を作り、これらを、 $\sum_j \frac{\psi_j}{\psi_i} O_{i,j}$  の重みで足し合わせることで、十分大きいサンプル数  $|S|$  で期待値と一致させることができる。これらの重みは NQS への問い合わせで効率的に得ることができる。

$$\langle \psi | O | \psi \rangle / \langle \psi | \psi \rangle \simeq \sum_{i \in S} \sum_j \frac{\psi_j}{\psi_i} O_{i,j} \quad (20)$$

### 3.4 手法4：密度行列からのサンプリングによる期待値推定

手法3では、NQS のサンプリングを通すことで状態全体を再構成することなく、状態に対する期待値を評価した。以下では、密度行列を NQS で学習した際にも、密度行列を再構成することなく、サンプリングを通して蒸留された量子状態に対する期待値が評価できることを示す。

まず、蒸留されていない密度行列の期待値評価について考える。規格化されていない混合状態  $\rho = \sum_{i,j} \rho_{i,j} |i\rangle \langle j|$

を学習した NQS を考える。すると、手法3と同様に以下のように式を展開できる。

$$\langle O \rangle = \frac{\text{Tr}[O\rho]}{\text{Tr}[\rho]} = \mathbb{E}_{i \sim \{p_i\}} \left[ \sum_j \frac{\rho_{j,i}}{\rho_{i,i}} O_{i,j} \right] \quad (21)$$

従って、オブザーバブルが各行各列で高々多項式個しか非ゼロ要素を持たない場合、密度行列に対するサンプリングと、 $\sum_j \frac{\rho_{j,i}}{\rho_{i,i}} O_{i,j}$  の重みの計算を通して期待値の評価を行うことができる。

一方、蒸留された密度行列に対する期待値の評価は、式(11)および式(12)の変形を用いると、以下のように展開できる。

$$\frac{\text{Tr}[\rho^m O]}{\text{Tr}[\rho^m]} = \frac{\text{Tr}[\rho^{\otimes m} (O \otimes I^{\otimes(m-1)}) \Lambda]}{\text{Tr}[\rho^{\otimes m}]} \left( \frac{\text{Tr}[\rho^{\otimes m} \Lambda]}{\text{Tr}[\rho^{\otimes m}]} \right)^{-1} \quad (22)$$

この時、右辺の左側の分数は量子状態  $\rho^{\otimes m}$  に対するオブザーバブル  $(O \otimes I^{\otimes(m-1)}) \Lambda$  の評価であり、右側の分数は同じ状態のオブザーバブル  $\Lambda$  に対する評価である。Derangement 演算子は各行各列に一つしか非ゼロ要素を持たない演算子であるため、これら二つのオブザーバブルは、 $O$  が各行各列に多項式個しか非ゼロ要素を持たない場合、同様に多項式個しか非ゼロ要素を持たない。従って、式(21)と同様の式展開を行うことができる。この時、 $\rho^{\otimes m}$  の量子状態から  $nm$ -bit の添え字をサンプリングする操作は  $\rho$  から  $m$  回  $n$ -bit の添え字を独立にサンプリングして結合すればよい。また、 $\rho^{\otimes m}$  の各行各列の成分は、 $\rho$  の成分の  $m$  個の積として計算することができる。従って、 $\rho^{\otimes m}$  に対する評価は、 $\rho$  に対する RBM の操作を  $m$  倍行うことで実施することができる。この方法は NQS に対する操作が  $m$  倍になるという欠点はあるものの、純粋状態が混合状態をうまく学習するという仮定も不要であり、量子ビット数  $n$  に対して多項式的な空間や時間でも不要となる。

## 4. Result

### 4.1 数値計算の設定

学習データは「各量子ビットのパウリ測定結果」、「測定に用いたパウリ演算子」、「パウリ演算子の組み合わせ」の三つからなる。このため、訓練データを作成するには測定に用いるパウリ演算子の集合を選ぶ必要がある。今回の評価では簡単のため、測定に用いるパウリ演算子の組み合わせは  $n$ -qubit 系であれば  $\{X, Y, Z\}^{\otimes n}$  の要素である  $3^n$  パターンとした。今回は訓練データのサイズは  $S = 10^4$  として、データセットを  $3^n$  パターンの測定で当分割することで設定した。訓練データの作成は Qiskit [17] のノイズシミュレータ上で、各量子ゲートには一様な depolarizing error が 1% で生じる量子回路を構築し、 $|\psi_0\rangle = |00\rangle, U = \Lambda_{1,2}(H \otimes I)$  として測定データの生成を行った。ただし、 $\Lambda_{1,2}$  は 1-qubit

目を制御量子ビット、2-qubit 目を対象量子ビットとする CNOT ゲートである。この時、理想的な量子状態はベル状態  $|\psi\rangle_{\text{ideal}} = (|00\rangle + |11\rangle)/\sqrt{2}$  となる。評価においてはオブザーバブルは  $O = X_1 X_2$  を用いた。このオブザーバブルは理想的には  $\langle O \rangle_{\text{clean}} = 1$  を与える。RBM を用いた量子状態の学習には QuCumber [9] を用いた。RBM の学習におけるパラメータは以下のようにした。学習レートは  $\eta = 10$  として、10 Epoch おきに 0.9 倍となる設定にした。CD 法におけるステップ数は  $k = 100$ 、勾配を計算する際の positive batch size, negative batch size は両方ともに 1000 とした。不可視ユニット数と、密度行列を RBM で学習する際の補助ユニットの数は両方ともに 6 とした。NQS からのサンプリングを必要とする手法 3, 4 では期待値の計算のために量子状態からサンプリングを行う必要がある。このサンプル数は今回の計算では  $10^4$  で固定とした。また、手法 4 におけるコピー数は  $m = 2$  とした。

## 4.2 数値結果

まず、学習データを用いて RBM の学習を行った。図 2 は各 Epoch 数における、ノイズの無い理想的な量子状態  $|\psi\rangle_{\text{ideal}}$  に対する量子状態の忠実度を表している。純粋状態を扱う RBM の忠実度は緑色、密度行列を扱う RBM の忠実度は青色でプロットされている。結果より、密度行列

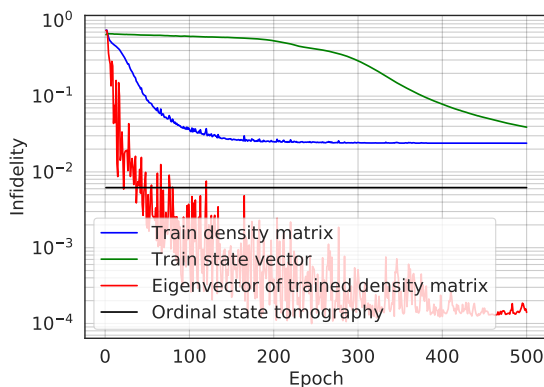


図 2 学習に伴う忠実度の改善

を用いた学習は早い段階で忠実度が高い状態に到達し、その後の改善は緩やかである。一方で、純粋状態を用いた学習は忠実度の高い状態に到達するまで長い時間の学習を必要としている。これは、密度行列を用いた RBM のモデルは学習可能な探索空間が広いため、純粋状態に比べて比較的早期に最適なパラメータに到達できるからだと期待される。

次に学習が進む間の各 Epoch で、提案した手法から得られるオブザーバブル  $X_1 X_2$  の期待値が、理想的なオブザーバブルの値がからどの程度ずれているかの絶対値を評価した。この期待値は理想的には +1 が得られるはずであ

る。図では、ノイズのある量子状態  $\rho_{\text{noisy}}$  の期待値を直接測定したものをオレンジ色の破線で、密度行列を NQS として学習し、そのまま期待値を評価したものを青線で表示している。赤色と紫色のプロットは密度行列を用いた方法である手法 1, 4 に対応する。黄色と緑色のプロットは、それぞれ状態ベクトルを用いた手法 2, 3 の結果に対応している。今回実施した中では、密度行列を再構成し対角化

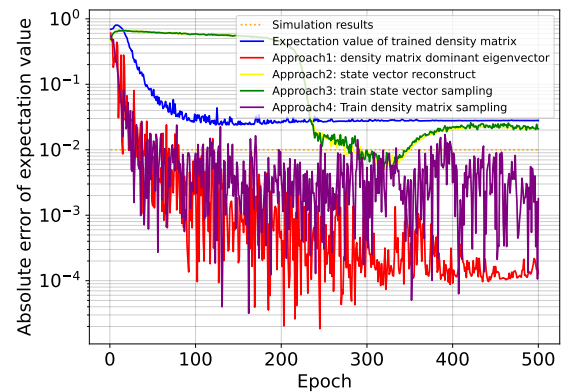


図 3 学習に伴う  $O = X_1 X_2$  の場合の期待値の絶対誤差のプロット

により最大固有値ベクトルを求める手法 1 と、密度行列に対して仮想状態蒸留と同等の評価をサンプリングを通して行う手法 4 の二つが収束の速さと収束する値の正確さの二つで、素朴な評価に比べて高い性能を示した。密度行列を直接評価した青色のプロットの収束値は、ノイズのある量子状態を直接評価したオレンジの破線よりも性能が悪いことを踏まえると、最大固有値ベクトルを取り出す手法は有効であることが分かる。特に、密度行列からサンプリングを通して蒸留された期待値を評価する方法は、量子状態が効率的に学習できる場合には多項式的な時間と空間で評価が可能のため、実用性が最も高いと期待される。一方、純粋状態を用いて直接学習する手法は学習の途中で素朴な評価であるオレンジの破線を超えることはあるものの、最終的な収束値は素朴な評価よりも悪い値に収束しており、純粋状態を直接学習する方法は想定していたような効果をもたらさないことが分かる。

## 5. Discussion

本研究ではニューラルネットワークを用いた量子状態を用いて量子状態を学習し、学習したニューラルネットワークに対して仮想蒸留と同等の操作を行うことで、ノイズのある量子状態からノイズが除去された期待値を計算する手法を提案した。本手法はその実行方法から 4 つに分類されるが、量子状態を密度行列を表現する RBM で学習し、さらに RBM からのサンプリングを通して期待値を計算する手法は量子状態が RBM で効率的に学習できる限り、指数関数的な空間や時間を必要とせず期待値の計算が可能であることが分かった。この手法はひとたび量子状態が学習

できれば、仮想蒸留法と異なり実態として複数のコピーや derangement のような操作が無くとも仮想蒸留法と同等の結果が得られるという利点がある。

## 参考文献

- [1] Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J. C., Barends, R., Biswas, R., Boixo, S., Brandao, F. G., Buell, D. A. et al.: Quantum supremacy using a programmable superconducting processor, *Nature*, Vol. 574, No. 7779, pp. 505–510 (2019).
- [2] Temme, K., Bravyi, S. and Gambetta, J. M.: Error mitigation for short-depth quantum circuits, *Physical review letters*, Vol. 119, No. 18, p. 180509 (2017).
- [3] Endo, S., Benjamin, S. C. and Li, Y.: Practical quantum error mitigation for near-future applications, *Physical Review X*, Vol. 8, No. 3, p. 031027 (2018).
- [4] Koczor, B.: Exponential Error Suppression for Near-Term Quantum Devices, *Phys. Rev. X*, Vol. 11, p. 031057 (online), DOI: 10.1103/PhysRevX.11.031057 (2021).
- [5] Huggins, W. J., McArdle, S., O’Brien, T. E., Lee, J., Rubin, N. C., Boixo, S., Whaley, K. B., Babbush, R. and McClean, J. R.: Virtual Distillation for Quantum Error Mitigation, *Phys. Rev. X*, Vol. 11, p. 041036 (online), DOI: 10.1103/PhysRevX.11.041036 (2021).
- [6] Carleo, G. and Troyer, M.: Solving the quantum many-body problem with artificial neural networks, *Science*, Vol. 355, No. 6325, pp. 602–606 (online), DOI: 10.1126/science.aag2302 (2017).
- [7] Torlai, G., Mazzola, G., Carrasquilla, J., Troyer, M., Melko, R. and Carleo, G.: Neural-network quantum state tomography, *Nature Physics*, Vol. 14, No. 5, pp. 447–450 (online), DOI: 10.1038/s41567-018-0048-5 (2018).
- [8] Torlai, G. and Melko, R. G.: Latent Space Purification via Neural Density Operators, *Phys. Rev. Lett.*, Vol. 120, p. 240503 (online), DOI: 10.1103/PhysRevLett.120.240503 (2018).
- [9] Beach, M. J. S., Vlugt, I. D., Golubeva, A., Huembeli, P., Kulchytskyy, B., Luo, X., Melko, R. G., Merali, E. and Torlai, G.: QuCumber: wavefunction reconstruction with neural networks, *SciPost Phys.*, Vol. 7, p. 9 (online), DOI: 10.21468/SciPostPhys.7.1.009 (2019).
- [10] Carleo, G., Nomura, Y. and Imada, M.: Constructing exact representations of quantum many-body systems with deep neural networks, *Nature Communications*, Vol. 9, No. 1, pp. 1–11 (online), DOI: 10.1038/s41467-018-07520-3 (2018).
- [11] Smolensky, P.: *Information Processing in Dynamical Systems: Foundations of Harmony Theory*, p. 194–281, MIT Press (1986).
- [12] Hinton, G. E., Osindero, S. and Teh, Y.-W.: A Fast Learning Algorithm for Deep Belief Nets, *Neural Computation*, Vol. 18, No. 7, pp. 1527–1554 (online), DOI: 10.1162/neco.2006.18.7.1527 (2006).
- [13] Hinton, G. E.: Training Products of Experts by Minimizing Contrastive Divergence, *Neural Computation*, Vol. 14, No. 8, pp. 1771–1800 (online), DOI: 10.1162/089976602760128018 (2002).
- [14] Yoshioka, N. and Hamazaki, R.: Constructing neural stationary states for open quantum many-body systems, *Phys. Rev. B*, Vol. 99, p. 214306 (2019).
- [15] Yoshioka, N., Mizukami, W. and Nori, F.: Solving quasiparticle band spectra of real solids using neural-network quantum states, *Communications Physics*, Vol. 4, No. 1, p. 106 (online), DOI: 10.1038/s42005-021-00609-0 (2021).
- [16] Nomura, Y., Yoshioka, N. and Nori, F.: Purifying Deep Boltzmann Machines for Thermal Quantum States, *Phys. Rev. Lett.*, Vol. 127, p. 060601 (2021).
- [17] ANIS, M. S. et al.: Qiskit: An Open-source Framework for Quantum Computing (2021).