修士論文 要旨

クリギング法を用いたゼロ次最適化アルゴリズムによる量子状態作成

数理情報学専攻 48-206201　**穴澤 徳明**

指導教員　**山西 健司 教授**

1　背景

量子コンピュータを構成するデバイスは外部からのノイズによる量子誤りの頻度が非常に大きく，量子誤りを訂正し十分な計算が行える状態（エラーフリー）が求められる．Gottesman-Kitaev-Preskill 状態（GKP 状態）[1] という量子状態は，ウィグナー関数と呼ばれる位相空間上の擬確率分布において規則的に多くのシャープな構造を持つ特殊な性質を持つ．これにより GKP 状態は，離散的な確率密度分布を持ちながら光波の連続的な性質を保つため，連続変数型（Continuous Variable：CV）量子計算において量子もつれを生むことが可能であり，また量子誤り訂正能力も備えている．そのため大規模な量子もつれを生成でき，将来的に複雑な量子計算を行える有望な量子状態として期待される．

GKP 状態を準備することでエラーフリーを備えたユニバーサルな量子計算を行う方法が提唱され，この GKP 状態の実現に向けて盛んに研究が行われている．数値実験では任意の状態生成問題に対し多層パーセプトロンを用いたパラメータ最適化への帰着が行われ，誤差逆伝播による求解で GKP 状態との内積が 0.99 以上の量子状態が得られている [2].

本研究ではこのパラメータ最適化問題において，勾配を必要としないゼロ次最適化手法を導入する．これにより状態ベクトルの観測困難性を回避でき，量子デバイスでの実現可能性が高まる．また 1 次勾配を必要しないため，量子状態の差を表す距離の中で微分困難なものも損失関数として扱うことが可能である．

2　問題設定

量子状態はヒルベルト空間内の L2 ノルム 1 を保つベクトルで表現可能であり，$|\psi\rangle$ と表される．これを状態ベクトルと呼ぶ．

$$|\psi\rangle\text{: 状態ベクトル,} \quad c_n \in \mathbb{C}, n = 0, 1, 2, \ldots \tag{1}$$

$$|\psi\rangle = (c_0, c_1, \ldots, c_n, \ldots)^{\mathrm{T}}. \tag{2}$$

量子回路は，入力 $|\psi_0\rangle$ から出力 $|\psi\rangle$ への行列変換で記述可能である．$\theta \in \mathbb{R}^f$ を量子回路のパラメータ列，$U(\theta) \in \mathbb{C}^{N\times N}$ を量子回路に対応する行列変換であるとする．このとき出力は以下を満たす．

$$|\psi\rangle = U(\theta)|\psi_0\rangle. \tag{3}$$

ここで所望のターゲット $|\psi_t\rangle$ を出力するための最適な θ を決定するパラメータ最適化問題を考えることができる．出力とターゲットの量子状態の近さをベクトルの内積で表現する Fidelity を用いる．

$$F(\psi_t, \psi) = |\langle\psi_t|\psi\rangle|^2. \tag{4}$$

最適化問題は以下．

$$\max_{\theta} \quad F(\psi_t, \psi). \tag{5}$$

以下のような損失関数（Fidelity 損失）を構成することが多い．

$$Z(\theta) = 1 - F(\psi_t, \psi). \tag{6}$$

式変形により損失関数を定数差で以下のように取り替えることができ，最小化問題は等価であることが言える．

$$Z(\theta) = |\langle\psi_t|\psi\rangle - 1|. \tag{7}$$

3　既存手法

シミュレーションでは，損失関数 $Z(\theta)$ の θ による勾配を自動微分により計算し，最急降下法であれば学習率 η を決めて以下のような更新式に従い最適化を行うことが可能である．先行研究では実際に深層学習で主に用いられている勾配法の Adam を用い [3]，学習率を調整することで高速化を狙った．いずれのシミュレーションも数百程度のゲートを持つネットワークを用いて 0.99 を超える Fidelity を達成した．

$$\theta \mapsto \theta - \eta\nabla_\theta Z. \tag{8}$$

仮想回路での最適化によって得られたパラメータ設定を実際の回路に利用すると誤差が発生するため，実際の回路で最適化を行う必要がある．しかし既存手法は勾配計算のために観測が困難な $|\psi\rangle$ を用いる必要がある．これは $\nabla_\theta Z$ を必要とするアルゴリズムを用いることで必ず発生する課題である．

4 提案手法

本研究ではパラメータ最適化問題にゼロ次最適化を導入する．ゼロ次最適化とはパラメータ $\tilde{\theta}_i$ と関数値 $y_i = Z(\tilde{\theta}_i)$ のサンプルのみを用い，勾配情報を使わない最適化アルゴリズムである．

$$D = \{(\tilde{\theta}_1, y_1), (\tilde{\theta}_2, y_2), \ldots, (\tilde{\theta}_M, y_M)\}. \tag{9}$$

以下に各アルゴリズムの概略を示す．

4.1 ベイズ最適化 (BO)[4]

最小値を取る確率最大の点を逐次的に探索して最適化を行う．入力はカーネル K, 近似関数 $\hat{Z}(\theta)$, 探索空間 Θ, 訓練データ D, 損失関数 $Z(\theta)$．出力は $Z(\theta)$ を最小化する $\theta_{\mathrm{opt}} \in \Theta$．まずカーネルを用いて各点の平均，分散を計算する．（よくガウスカーネルを用いる．）

$$\mu(\cdot \mid D), \sigma(\cdot \mid D) \leftarrow K \mid D. \tag{10}$$

次に，獲得関数最大化により次の探索点を決定する．

$$\hat{\theta} = \arg\min_{\Theta} \hat{Z}(\mu(\theta \mid D), \sigma(\theta \mid D)). \tag{11}$$

以上を収束するまで繰り返す．

4.2 Direct Optimization by Kriging method (DO by Kriging)

入力は探索空間 Θ, 訓練データ D, 損失関数 $Z(\theta)$, ハイパーパラメータ a_s, a_r．出力は BO と同様．BO の獲得関数の代わりに，クリギング法 [5] による内挿で近似関数 $\hat{Z}(\theta)$ を計算する．

$$h_{i,j} = \left\| \tilde{\theta}_i - \tilde{\theta}_j \right\| \tag{12}$$

$$g_{i,j} = \frac{1}{2}(y_i - y_j)^2 \tag{13}$$

$$\Sigma_{i,j} = a_s \exp\left\{ -\left(\frac{g_{i,j}}{a_r}\right)^2 \right\} \tag{14}$$

$$b_i(\theta) \leftarrow a_s \exp\left\{ -\left(\frac{\left\|\theta - \tilde{\theta}_i\right\|}{a_r}\right)^2 \right\} \tag{15}$$

$$\hat{Z}(\theta) \leftarrow \bar{y} + b(\theta)^{\mathrm{T}} \Sigma^{-1} (\mathbf{y} - \bar{y}\mathbf{1}) \tag{16}$$

次に，近似関数の最小値を探索する．

$$\hat{\theta} = \arg\min_{\Theta} \hat{Z}(\theta) \tag{17}$$

以上を収束するまで繰り返す．

4.3 Gradient Descent by Kriging method (GD by Kriging)

DO by Kriging で近似関数最小化の代わりに，近似勾配 $\nabla_\theta \hat{Z}$ を用いて勾配降下法を行う．

$$\hat{\theta} = \theta - \eta \nabla_\theta \hat{Z}(\theta) \tag{18}$$

4.2, 4.3 のクリギング内挿はサンプルサイズ $|D|$ の増加に従って計算量が増加してしまう．以下のような規則に従って $D' \subset D$ を選ぶことにする．サンプル $(\tilde{\theta}_i, y_i)$ のスコア J_i は初期値として 1 が与えられる．

$$J_i = 1, \quad (^\forall i). \tag{19}$$

次に，D' は D の中から重み J_i に従って一定数選ばれる．そして，各ステップ T でスコアを以下のように更新する．T における損失関数を Z_T とする．

$$\begin{aligned} J_i \mapsto \max(1, J_i + |D|(Z_T - Z_{T-1})), \\ (\text{if} \quad (\tilde{\theta}_i, y_i) \in D'). \end{aligned} \tag{20}$$

これにより，Z_T が減少した場合はその時に選ばれたサンプルのスコアは上昇し，Z_T が増加した場合はスコアが減少する．

5 数値実験

$|\psi\rangle$ の観測困難性を克服できるゼロ次最適化の中では GD by Kriging が最大 Fidelity を達成した．

表 1. 既存手法と提案手法の比較

手法	最大 Fidelity	$\lvert\psi\rangle$ の必要性
誤差逆伝播	0.6	○
BO[4]	0.45	×
DO by Kriging	0.47	×
GD by Kriging	0.52	×

参考文献

[1] A. Kitaev D. Gottesman and J. Preskill. Encoding a qubit in an oscillator. 2001.

[2] Juan Miguel Arrazola, Thomas R. Bromley, Josh Izaac, Casey R. Myers, Kamil Brádler, and Nathan Killoran. Machine learning method for state preparation and gate synthesis on photonic quantum computers. 2018.

[3] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2014.

[4] Peter I. Frazier. A tutorial on bayesian optimization, 2018.

[5] Georges Matheron. The intrinsic random functions and their applications. *Advances in applied probability*, Vol. 5, No. 3, pp. 439–468, 1973.