

機械学習を用いた量子状態異常検知

Quantum State Anomaly Detection Using Machine Learning

原 聡^{*1} 小野 貴史^{*2} 岡本 亮^{*3} 鷲尾 隆^{*4} 竹内 繁樹^{*3}
 Satoshi Hara Takafumi Ono Ryo Okamoto Takashi Washio Shigeki Takeuchi

^{*1}IBM 東京基礎研究所 ^{*2}ブリストル大学 ^{*3}京都大学 ^{*4}大阪大学
 IBM Research – Tokyo Bristol University Kyoto University Osaka University

Density matrix describes the state of quantum systems. In quantum informatics, it is crucial to verify whether the density matrix of the target system is in an aimed state. We present a machine learning approach to find matrices that are away from the aimed state. In the method, we use a sparse regularization method to extract anomalies from intrinsic fluctuations of matrices.

1. はじめに

精緻な処理が必要となる量子情報処理では所望の量子状態からのずれ、すなわち異常の検出が重要な課題である。しかし、異常の検知は量子系が持つ本質的なゆらぎや統計的なゆらぎのために困難な課題の一つでもある。通常、量子系の状態を表現する密度行列は量子トモグラフィ [1] を用いて推定される。この際、実験で得られる観測の総数は限られているために、推定された密度行列には統計的なゆらぎが生じる。この統計的なゆらぎと量子系の変化とを区別することが異常検知における技術的な課題である。

本稿では機械学習を用いて密度行列の異常検知を行う方法 ED³ (Erroneous Deviation Detector for Density matrices) [2] を紹介する。ED³ は密度行列の各要素の絶対値が変化するような量子系の異常を検知する手法である。このような異常にはデコヒーレンスなど実用上重要な例が数多く含まれる。量子もつれ光子対の量子トモグラフィにおいて ED³ と単純な手法とを比較したところ、シミュレーション・実験データの両方で異常検知性能の大きな向上が見られた。ED³ は密度行列を扱う問題には全て適用できるため、量子光学系に限らず超伝導回路やイオントラップを用いた系にも適用できる。このため、ED³ は密度行列の異常が問題となる多くの局面において有用な解決手段を提供すると考えている。

2. 密度行列の異常検知手法

量子トモグラフィにより K 個の観測密度行列 $\hat{\rho}_1, \hat{\rho}_2, \dots, \hat{\rho}_K$ が得られているとする。以下では k 番目の観測密度行列の (i, j) 要素の絶対値を $|\hat{\rho}_{k,ij}|$ とし、これら絶対値を要素として持つ観測密度行列を $|\hat{\rho}_k| \in \mathbb{R}^{d \times d}$ と表記する。本稿では、これら K 個の絶対値の観測密度行列 $|\hat{\rho}_1|, |\hat{\rho}_2|, \dots, |\hat{\rho}_K|$ から異常な行列のインデックス集合 $S \subseteq \{1, 2, \dots, K\}$ を探す問題を対象とする。ただし、異常行列のインデックス集合 S の補集合 S^c においては行列 $|\hat{\rho}_{k'}|$ ($k' \in S^c$) は全て正常であると、また異常行列は K 個の行列のごく一部、つまり $|S| \ll K$ であると仮定する。このような問題は機械学習において異常検知や外れ値検知として知られている [3]。

異常検知を最も直接的に実現する方法は、計測された密度行列を”正しい”密度行列と比較することである。量子トモグラフィ

装置を物理モデルに基づいて表現することができれば、その物理モデルから”正しい”密度行列を計算することができる。そしてその”正しい”密度行列と各観測密度行列がどの程度異なるかを評価することで異常な行列を見つけることができる。このためには、量子トモグラフィ装置内の光路に関する正確なパラメータを用いて物理モデルを計算する必要がある。しかし、現実的には実験装置の正確なパラメータを知ることは困難であり、その近似値を用いざるをえない。つまり、近似的なパラメータから計算された”正しい”密度行列は、量子トモグラフィにより実際に得られる観測密度行列とは必ずしも一致せず、妥当な異常検知ができるとは限らない。そのため、以下では物理モデルを仮定せず、統計的な手法を用いて異常検知を行うことを考える。

2.1 平均行列を使う手法 (平均値法)

物理モデルを用いずに正しい密度行列を観測密度行列から近似的に表現することを考える。最も単純な方法の一つとして K 個の観測密度行列の平均値 $|\bar{\rho}| = \frac{1}{K} \sum_{k=1}^K |\hat{\rho}_k|$ を使う方法が考えられる。 K 個の観測密度行列のうち大半は正常な行列だと仮定しているため、この平均行列を正しい密度行列の近似値として使うことができる。そして、各行列のこの平均行列からの偏差 $\tilde{\omega}_k = |\hat{\rho}_k| - |\bar{\rho}|$ に基づいて、異常の度合いを評価する。ここでは量子力学的な差異を測る一般的な方法としてトレース距離 [4] $\tilde{e}_k = \|\tilde{\omega}_k\|_{\text{tr}}$ を用いる。ここで、 $\|\cdot\|_{\text{tr}}$ は行列の特異値の和 (トレースノルム) である。計算された \tilde{e}_k が一定値以上大きければ、密度行列は正しい密度行列からの乖離が大きいため、つまり異常だと判定できる。

平均行列を使う方法は単純で直感的にもわかりやすいが、問題点として平均行列 $|\bar{\rho}|$ の計算に数が少ないとは言え異常な行列が混入して偏りを生じさせていることがあげられる。この偏りが異常検知の性能を低下させてしまう要因となる。

2.2 機械学習を用いた異常検知手法 ED³ [2]

k 番目の絶対値の密度行列の真の値を $|\rho_k|$ とすると、これは”正しい”密度行列 θ とそこからの偏差 ω_k を使って $|\rho_k| = \theta + \omega_k$ と表現できる。また、観測密度行列 $|\hat{\rho}_k|$ はこの $|\rho_k|$ に観測に伴う統計的ゆらぎが加わったものとして解釈できる。このとき、正常な行列については”正しい”密度行列 θ からの偏差については $\omega_k \approx 0$ と想定できる。他方、異常な行列については $\omega_k \neq 0$ となる。そのため、偏差の行列の集合 $\{\omega_1, \omega_2, \dots, \omega_K\}$ は多くが零行列となるスパースな構造を持つと仮定できる。このようにデータの背後にスパースな構造が隠れている場合には機械学

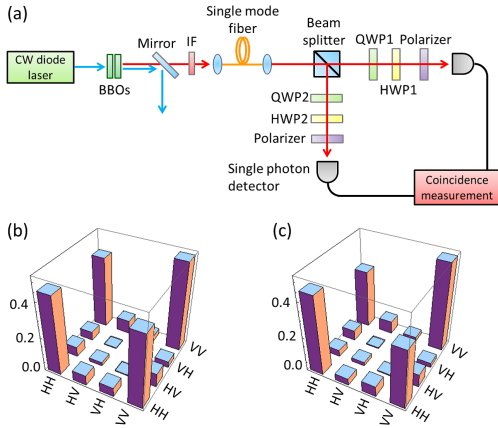


図1 (a) 量子もつれ光子対の量子トモグラフィー装置, (b) 正常な密度行列の各要素の平均のヒストグラム, (c) 異常な密度行列の各要素の平均のヒストグラム

習のスパース推定の技術が有効である [5, 6]. そこで, ED^3 では以下の正則化付きの最小二乗問題を解くことで, このスパースな偏差行列の集合を推定する.

$$\min_{\theta, \{\omega_k\}_{k=1}^K} \frac{1}{2} \sum_{k=1}^K \|\hat{\rho}_k - (\theta + \omega_k)\|_F^2 + \gamma \sum_{k=1}^K \|s \odot \omega_k\|_F. \quad (1)$$

ここで $\|\cdot\|_F$ は行列のフロベニウスノルム, \odot はアダマール積, γ は非負のパラメータである. また, 行列 $s \in \mathbb{R}^{d \times d}$ は $s_{ij} = \left(\frac{1}{K} \sum_{k=1}^K \tilde{\omega}_{k,ij}^2\right)^{-1/2}$ により定義される. 式 (1) の第一項は各観測密度行列 $|\hat{\rho}_k|$ に近い $\theta + \omega_k$ を推定すること, 第二項は行列の各要素の大きさを正規化したうえでできる限りスパースでありゼロに近い偏差行列の集合を推定することに対応する. 最後に問題 (1) を解くことによって得られる ω_k からトレース距離 $e_k = \|\omega_k\|_{tr}$ を計算し, e_k が一定値以上大きければ異常と判断する. 式 (1) による推定では, θ と ω_k の同時推定により平均行列を使う場合に比べて”正しい”密度行列 θ の偏りが抑えられること, スパースな偏差行列を推定できることの2点により, 精度の高い異常検知が期待できる.

問題 (1) の各項は凸関数なので全体としても凸関数であり, 各種凸最適化アルゴリズムによって $\theta, \{\omega_k\}_{k=1}^K$ の大域最適解を得ることができる. 本研究では最適化手法として交互方向乗数法 [7] を用いた. このアルゴリズムは問題 (1) の構造を活用することで効率的に大域最適解を求めることができる.

3. 評価実験

ED^3 と平均値法との性能比較実験を行った. 実験では図 1(a) に概要を示した量子トモグラフィー装置を用い, 1 個の観測密度行列を 1000 回の量子もつれ光子対入射の計数から得た. 観測を繰り返すことで, 図 1(b) の正常な密度行列及び図 1(c) の異常な密度行列をそれぞれ複数得た. 異常な密度行列では (1, 4) 及び (4, 1) の非対角要素 (図中 (HH, VV) に相当) の値が 0.421 から 0.339 まで低下している. この異常状態は正常な場合よりも 2 光子間の量子もつれの程度が少ない状態である.

本実験では物理実験から得られたデータ, およびシミュレーションから得られたデータの両方で 2 手法の比較を行った. 両データでの結果は同様であったため, ここでは物理実験データでの結果を報告する. 物理実験データとして, 正常な密度行列

平均値法	95%	90%	85%		
ED^3	95%	90%	85%		

図2 1000 回の実験における AUC の分布

を 250 個, 異常な密度行列を 50 個生成した. ここからランダムに正常な行列を 25 個, 異常な行列を 5 個取り出し計 30 個を 1 つのデータセットとした. これを 1000 回繰り返して 1000 個のデータセットを作り, それぞれで平均値法と ED^3 の性能を比較した. なお, ED^3 のパラメータ γ は事前に 10 個のデータセットで複数の γ の値を試し, その結果が一番良かった値を 1000 個のデータセット実験で使った.

1000 個のデータセットについて異常検知性能を ROC 曲線で評価し, 曲線下の面積 Area Under the Curve (AUC) を 0% から 100% で計算した結果が図 2 である. 平均値法では約 3 割でしか AUC=95% が達成できなかったのに対し, ED^3 では半数以上で AUC=95% が達成できた. この結果から, ED^3 の優位性が確認できる.

4. まとめ

本稿では機械学習を用いて密度行列の異常を検知する方法 ED^3 について紹介した. ED^3 では観測行列の偏差の集合がスパースになることに着目し, 機械学習のスパース正則化の技術を導入することで異常検知性能の向上を達成した. 実際に, ED^3 を使うことで平均値法よりも高い精度で異常が検知できることを確認した.

参考文献

- [1] Daniel FV James, Paul G Kwiat, William J Munro, and Andrew G White. Measurement of qubits. *Physical Review A*, 64(5):052312, 2001.
- [2] Satoshi Hara, Takafumi Ono, Ryo Okamoto, Takashi Washio, and Shigeki Takeuchi. Anomaly detection in reconstructed quantum states using a machine-learning technique. *Physical Review A*, 89(2):022104, 2014.
- [3] Victoria J Hodge and Jim Austin. A survey of outlier detection methodologies. *Artificial Intelligence Review*, 22(2):85–126, 2004.
- [4] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.
- [5] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.
- [6] Ali Jalali, Sujay Sanghavi, Chao Ruan, and Pradeep K Ravikumar. A dirty model for multi-task learning. In *Advances in Neural Information Processing Systems*, pages 964–972, 2010.
- [7] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2011.