Quantum PC algorithm: data-efficient and nonlinear causal discovery

Yota Maeda
Advanced Research Laboratory, Research Platform
Sony Group Corporation
Tokyo, Japan
Yota.Maeda@sony.com

Ken Arai
Advanced Research Laboratory, Research Platform
Sony Group Corporation
Tokyo, Japan
Ken.Arai@sony.com

Yu Tanaka
Advanced Research Laboratory, Research Platform
Sony Group Corporation
Tokyo, Japan
Yu.Tanaka@sony.com

Yu Terada
Advanced Research Laboratory, Research Platform
Sony Group Corporation
Tokyo, Japan
yu.terada@sony.com

Hiroshi Ueno
Advanced Research Laboratory, Research Platform
Sony Group Corporation
Tokyo, Japan
Hiroshi.C.Ueno@sony.com

Hiroyuki Tezuka
Advanced Research Laboratory, Research Platform
Sony Group Corporation
Tokyo, Japan
Hiroyuki.Tezuka@sony.com

***Abstract*—Causal discovery is the task of finding causal relationships between random variables from observed data. Typically, one assumes that the causal relationships can be represented by a directed acyclic graph (DAG), and makes additional assumptions to ensure that the DAG can be recovered from the observed data. In this study, we propose the *quantum Peter–Clark (qPC) algorithm* for nonlinear causal discovery based on quantum kernel methods. The qPC algorithm takes advantage of the quantum kernel-based conditional independent test. Through the synthetic data experiment, we show that the qPC algorithm outperforms the classical method in the regime of a small number of samples. This suggests that the kernel-based causal discovery can significantly improve performance under such conditions. Our experiments highlight that the proposed algorithm can accurately support classical algorithms in causal discovery, paving the way for future advances with the utilization of quantum computation.**

I. Introduction

The discovery of causal relationships in a set of random variables from observed data is a fundamental task of many empirical sciences. A directed acyclic graph (DAG) typically represents the causal structure among them, in which the directed edges correspond to pairwise causal relations. The Peter–Clark (PC) algorithm [1] is one of the most general algorithms for causal discovery, which recovers causal structures by exploiting the conditional independence (CI) test. Since it is not based on any specific models unlike other methods, it can reveal nonlinear causality between variables from observed data. To treat nonlinear data, the best known and performing CI test is the kernel-based CI test (KCIT) [2]. It adopted the Hilbert–Schmidt norm on a reproducing kernel Hilbert space (RKHS) and generalized the usual CI tests to highly nonlinear situations.

Despite its brevity, however, the PC algorithm is known to have problems with the curse of dimensionality and requires a large amount number of samples. To overcome this problem, there are several studies using quantum kernel methods; the previous work [3] proposed the quantum kernel-based linear non-Gaussian acyclic model (qLiNGAM) and the causal structure could be correctly estimated even in the regime of the small number of samples [4].

However, it is still unclear how quantum kernels may or may not have an advantage over classical kernels in causal discovery. Furthermore, the qLiNGAM is not suitable for nonlinear causal discovery due to the use of linear models. To address the issues, we proposed the quantum kernel-based PC (qPC) algorithm, evaluated the validity of the proposed method and studied the advantage of quantum methods in terms of the analysis of eigenvalues of kernel matrices and KCIT. Our contributions are as follows:

- We proposed the qPC algorithm for nonlinear causal discovery, developing the existing KCIT methods in terms of quantum kernels.
- We applied the qPC algorithm to the fundamental problem on the basic causal graphs (Fig. 1) and showed that it recovered the causal graphs more accurately than the classical PC algorithm in the small sample-size regime, which implies that our method is superior in detecting spurious correlations.
- Through the kernel eigenvalue analysis, we found that the eigenvalue distribution affects the accuracy of causal discovery, which can explain the reason for the high performance of the qPC algorithm.

II. Proposed method: Quantum PC algorithm

We propose to extend the PC algorithm, embedding the data into the RKHS defined by the following quantum kernels. As the kernel function, the quantum kernels are denoted by

979-8-3315-4137-8/24/$31.00 ©2024 IEEE
DOI 10.1109/QCE60285.2024.10322

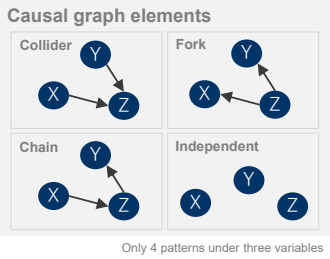

Fig. 1. The basic causal graphs under three variables.

$k_Q(\mathbf{x}, \mathbf{x}') = \mathrm{Tr}[\rho(\mathbf{x})\rho(\mathbf{x}')]$, where the input $\mathbf{x}$ is encoded into the quantum circuit generating the state $\rho(\mathbf{x})$ as parameters. This suggests a way to encode the data into quantum circuits as a natural extension of the Gaussian kernel. Our proposed quantum circuit can encode multiple random variables as parameters of a single-qubit rotation in each qubit different from the implementation [3]. Our proposed quantum circuit has hyperparameters that are analogous to the widths of the Gaussian kernels. We remark that these parameters can be adjusted so that the eigenvalues are suitable for KCIT.

III. Experiment

We conducted a series of causal discovery tests among various types of probability distributions, consisting of two stages, the data preparation and the test. In the data preparation, the samples were generated based on a certain causal relationship. Here, we employed the four fundamental patterns of the causal graphs shown in Fig. 1, which cover all relations between three nodes and comprise general causal graphs [1]. The random variables assigned to the source nodes (*e.g.*, X and Y in Collider) followed the Gaussian distributions, and they are the inputs of the random variable assigned to the node into which the directed edges from the source nodes go. Then, the random variables were calculated based on certain functions with the Gaussian noise (*e.g.*, $Z = f(X, Y) + \epsilon$). In the experiments, we recovered the DAG (*i.e.*, the causal relationship) using the conventional PC algorithm and the proposed qPC algorithm. The quantum state $\rho(\mathbf{x})$ used in the quantum kernel was generated by the quantum circuit consisting of single-qubit rotation gates for data embedding and controlled gates, *e.g.*, CX-gates, for generating entanglement.

We show the comparison results between the PC and qPC algorithms in Fig. 2. The accuracy represents the ratio of correctly recovering the causal graph relative to the number of experiments. For (b) Fork and (c) Chain, the qPC algorithm outperformed the PC algorithm in the regime of the small number of samples, whereas there are no significant differences in (a) Collider and (d) Independent.

IV. Discussion

To recover the Folk and Chain structures, the PC algorithm is required to apply the adequate kernel to reject the null hypothesis in the independence test. The qPC algorithm resolves this problem by applying the quantum kernel suitable

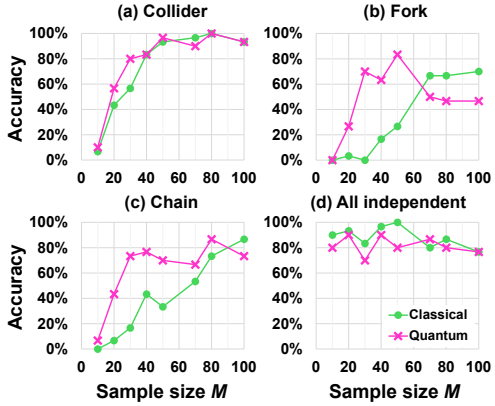

Fig. 2. The accuracy of the PC and qPC algorithms for the causal discovery test with basic causal graphs shown in Fig. 1, where we run 30 independent trials.

for the synthetic data generated from quantum circuits, even in a low-dimensional RKHS associated with quantum kernels. These observations inspire kernel optimization in KCIT, which involves managing the eigenvalues of kernel matrices. While Wang [5] recently proposed a method to optimize the parameters of Gaussian kernels in score-based causal discovery, our analysis introduces a different indicator in the qPC algorithm, applicable for model-free settings.

V. Conclusion

We proposed the qPC algorithm for nonlinear causal discovery applicable to various complex systems, without assuming any underlying model structures. Experiments on synthetic data generated from basic causal graphs showed that the qPC algorithm recovers the causal graphs more accurately than the PC algorithm, in the regime of the small number of samples.

The PC algorithm is often employed with classical statistical tests, relying on the central limit theorem and the law of large numbers. These asymptotic theories require large numbers of samples for robust analysis. It is therefore challenging to apply such statistical methods to cases of high sampling costs or rare events. Our results support that the quantum kernel method can address these issues, highlighting a new aspect of the quantum advantage of reducing sample size.

References

[1] P. Spirtes, C. Glymour, and R. Scheines, *Causation, prediction, and search*. MIT press, 2001.

[2] K. Zhang, J. Peters, D. Janzing, and B. Schölkopf, "Kernel-based conditional independence test and application in causal discovery," in *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence*, ser. UAI'11. Arlington, Virginia, USA: AUAI Press, 2011, p. 804–813.

[3] H. Kawaguchi, "Application of quantum computing to a linear non-gaussian acyclic model for novel medical knowledge discovery," *Plos one*, vol. 18, no. 4, p. e0283933, 2023.

[4] Y. Maeda, H. Kawaguchi, and H. Tezuka, "Estimation of mutual information via quantum kernel method," *arXiv:2310.12396*, 2023.

[5] W. Wang, B. Huang, F. Liu, X. You, T. Liu, K. Zhang, and M. Gong, "Optimal kernel choice for score function-based causal discovery," in *Forty-first International Conference on Machine Learning*, 2024. [Online]. Available: https://openreview.net/forum?id=DYd4vyyhUu