Quantum Accelerated Causal Tomography:
Circuit Considerations For Applications In Bioinformatics and AGI[*]

Tamal Acharya[†]

QWorld Association. and
Purdue University, West Lafayette, USA.

Akash Kundu[‡]

QWorld Association.
Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, and
Joint Doctoral School, Silesian University of Technology, Gliwice, Poland

Aritra Sarkar[§]

QWorld Association. and
Quantum Machine Learning group, QuTech,
Delft University of Technology, Delft, The Netherlands
(Dated: January 24, 2023)

In this work, we study quantum computing algorithms for accelerating causal inference. Specifically, we extend the formulation of causal hypothesis testing presented in [*Nat Commun* 10, 1472 (2019)] for practical scenarios. Through an implementable algorithm, we show that the error probability introduced in the previous work requires modification. The practical scenario which is followed by a theoretical description is constructed as a scalable quantum gate-based algorithm on IBM Qiskit. We present the circuit construction of the oracle embedding the causal hypothesis and assess the associated gate complexities. Additionally, our demonstration on a simulator platform validate the predicted speedup. We discuss applications of this framework for causal inference use cases in bioinformatics and artificial general intelligence.

I. INTRODUCTION

Despite the huge success of machine learning (ML) algorithms based on deep neural networks, these systems are inscrutable black-box models. This hampers users' trust in the system and obfuscated the discovery of algorithmic biases arising from flawed generated processes that are prejudicial to certain inputs (e.g. racial discrimination). Explainable artificial intelligence (XAI) focuses on human understanding of the decision from the learned solution, as white-box models. These models provide results that are understandable for domain experts, thus providing transparency, interpretability, and explainability. XAI algorithms provide a basis for justifying decisions, tracking and thereby verifying them, improving the algorithms, and exploring new facts. There has been relatively slow progress in XAI, despite realizing its importance as we increasingly automate critical systems. Early advances in XAI were based on symbolic reasoning systems and truth maintenance systems. To achieve causal reasoning, rule-based learning and logic-based inference systems were proposed. Methods to address inherent opaque modern methods like deep learning-based neural networks and genetic algorithms include layer-wise relevance propagation and local interpretability. There

exist other ML algorithms (e.g. decision trees, Bayesian classifiers, additive models) that generate interpretable models i.e. the model components (e.g., the weight of a feature, a path in a decision tree, or a specific rule) can be directly inspected to understand the predictions. However, these models are not general and scalable to compete with the adoption and impact of neural networks. On the other hand, symbolic reasoning systems were abandoned owing to the difficulty in scaling these systems for a large number of parameters.

The capability of *quantum computation allows us to scale symbolic reasoning models by encoding the classical rules as a superposition of quantum states or processes* [1]. Thus, quantum mechanics provide enhanced ways to identify causal links. Certain quantum correlations can be used to infer classical causal relationships [2, 3]. This could overcome the apprehension of existing classical approaches being pursued in XAI.

In this article, we will explore how we can distinguish quantum processes by their causal structure. Specifically, we study the construction proposed in [4] towards a quantum circuit implementation on the IBM Qiskit quantum programming language. In doing so, we uncover: (i) the implementation aspects of the causal oracle, (ii) the gate and qubit complexity of the full algorithm, and (iii) the dependence of the error probability on the set of processes/hypothesis being tested. While the current technology readiness level of quantum systems prevents us from demonstrating this causal reasoning within a broader application framework, we present the quantum kernel that can be readily embedded within a software

pipeline for applications in bioinformatics and artificial general intelligence (AGI). In particular, it will be useful in XAI pipelines [5, 6].

The remaining article is organized as follows. In Section [II] we briefly review the specifics of causal reasoning and some of the well-studied techniques. A discussion on the basic concepts and quantum advantage of causal hypothesis testing is given in Section [III]. In the following Section [IV] we describe the problem formulation; containing the main findings of the article. Here we define a model implementation on Qiskit followed by a correction factor introduced in the error probability based on our empirical results; which we call *practical error probability*. Finally, in Section [V] we discuss some potential use cases in bioinformatics and artificial general intelligence. The corresponding quantum resources of gates and qubits are assessed for realistic cases. Section [VI] concludes the article.

II. OVERVIEW OF CAUSAL INFERENCE

Causal inference refers to an intellectual discipline that considers the assumptions, study designs, and estimation strategies that allow researchers to draw conclusions on the cause-effect relationships between data. The dominant perspective on causal inference in statistics has philosophical underpinnings that rely on the consideration of counterfactual states. In particular, it considers the outcomes that could manifest given exposure to each of a set of treatment conditions (dynamics of a specific causal variable). Causal effects are defined as comparisons between these potential outcomes. Causal inference consists of a family of statistical methods whose purpose is to answer the question of why something happens. Standard approaches in statistics, such as regression analysis, are concerned with quantifying how changes between two variables are associated, with no directional sense. In contrast to that, causal inference methods are used to determine whether changes in a variable X cause changes in another variable Y or vice-versa. If X is causally related to Y, then Y's change can (at least partially) be explained in terms of X's change.

A. Challenges of performing causal inference

Causal models are based on the idea of potential outcomes. The fundamental problem for causal inference is that, for any individual unit (a physical object at a particular point in time), we can observe only one of the potential outcomes, either Y(1) (under treatment) or Y(0) (under no treatment). As we observe the value of the potential outcome under only one of the possible treatments [7], namely, the treatment actually assigned, hence the potential outcome under the other treatment is missing.

The inadequacy of the causal models is due to their failure to include relevant spatial and structural information in a way that does not render the model non-explanatory, unmanageable, or inconsistent with the basic assumptions of causal graph theory. Making valid causal inferences is challenging because it requires high-quality data and adequate statistical methods. One prominent example of common non-causal methodology is the erroneous assumption of correlative properties as causal properties. Correlated phenomena may not carry inherent causation. Regression models are designed to measure variance within data relative to a theoretical model. This suggests that there is nothing in the data that presents high levels of covariance that have any meaningful relationship among them. This suggests the absence of a proposed causal mechanism with predictive properties or a random assignment of treatment. The presupposition that two correlated phenomena are inherently related is a logical fallacy, known as spurious correlation.

- **Causation does not imply association** For example, we want to compare the impact of an academic degree on the income of a middle-aged individual. The person might have attended the academic degree or might not have. To calculate the causal effect of having an academic degree, we need to compare the output in both situations, which is not possible. This dilemma is a fundamental problem of causal inference. The challenge in causal inference is that *all potential outcomes are not observed*, we only observe one. Another example is described by the situation of – a black cat ran under the fence and I tripped and fell over. We could have tripped anyway. It had nothing to do with the cause of the cat running under the fence, i.e., *associated events do not imply causal connection*.

- **Correlation does not prove causality**
  Causal inference is usually a missing data problem [8] and we tend to make *assumptions to make up for the missing causes/variables*. An example is a correlation between people eating ice cream and people drownings. It could indicate that eating ice cream affects drowning. The actual correlation is between the season (summer) and these otherwise unrelated things. In this case, the missing cause is the season. Another example is the correlation between higher SAT scores and a greater number of books in the house of the student taking the tests. Causal inference would imply that the number of books directly affects the SAT scores, when in reality, they are both affected by something else (in this case most likely a higher average intelligence in the household). Causal inference is the process of drawing a conclusion about a causal connection based on the occurrence of an effect.

The main difference between causal inference and inference of association is that the former analyzes the response of the effect variable when the cause is changed,

and in the latter, inferences about the strength of association between variables are made using a random multivariate sample of data drawn from the population of interest. Thus, in causal inference, we have access to the additional information about the variable/s on which the intervention has occurred between two sets of associated data samples. Another problem of causal inference is that the cause and the effect may have occurred by chance rather than by intention.

The causal effect of a drug on systolic blood pressure, one month after the drug dosage vs no exposure to the drug dosage can be represented as a comparison of systolic blood pressure that would be measured at the time of given exposure to the drug with the systolic blood pressure that would be measured at the same point in time in the absence of exposure to the drug. The challenge for causal inference is that we are not generally able to observe both of these states: at some point in time when we are measuring the outcomes, each individual either has had drug exposure or has not. From the data perspective, there are some open problems to review regarding the great potential of learning causality with data. Such as the study of heterogeneous groups, learning causality with imbalanced data, and learning causality with complex variables.

B. Classical techniques in causal inference

Causal inference is conducted via the study of systems where the measure of one variable is suspected to affect the measure of another. Causal inference is conducted with regard to the scientific method. The first step of causal inference is to formulate a falsifiable null hypothesis, which is subsequently tested with statistical methods. Frequentist statistical inference uses statistical methods to determine the probability that the data occur under the null hypothesis by chance. Bayesian inference is used to determine the effect of an independent variable.

Common frameworks for causal inference include the causal pie model (component-cause) [9], Pearl's structural causal model (causal diagram and do-calculus) [10], structural equation modeling, and Rubin's causal model (potential-outcome) [11].

The most frequently used causal models can be grouped into two kinds: causal Bayesian networks and structural equation models (which are distinct but closely related). Causal graph models combine mathematics and philosophy: the mathematical elements are Directed Acyclic Graphs (DAGs) and probability theory (with focus on conditional independence); the philosophical elements are assumptions about the relationship between causation and probability [12].

C. Algorithmic machine learning

An alternative approach to causal inference based on algorithmic generative models is currently gaining popularity. [13] describes the process of performing causal deconvolution using this technique. This paper talks about the different generating mechanisms by which complex data is produced. The authors introduced a universal, unsupervised, and parameter-free model-oriented approach based upon algorithmic probability that decomposes an observation into its most likely algorithmic generative sources. It uses causal calculus to infer model representations. This research suggests that cellular automata offer an optimal testbed as they are discrete dynamical systems that are able to illustrate an algorithm's inner workings because of their visual nature.

The approach is similar to the Data Generating Process (DGP) in statistics which posits that each set of observations is just a snapshot that we see and is just one instance of the many that could have been possible. This approach also tells us that each and every data point is being generated from a data-generating process which can be taken as coming from some distribution of data and is just a sample of the total population. We try to understand the DGP while analyzing that data and try to find out what could have been the sources that generated those data. The research introduces algorithmic probability and the coding theorem. The deconvolution algorithm breaks a dataset into groups that do not share certain features, i.e. essentially causal clustering and algorithmic partition by probable generative mechanism. These methods are conceptually different from traditional clustering and partition in machine learning approaches.

Ultimately the purpose of this method is to distinguish components that are generated similarly from those that are generated differently. In information theoretic terms, the question the method answers is, what are the elements (e.g. nodes or edges) that can break a network into the components that maximize their algorithmic information content, i.e., those elements that preserve the information about the underlying programs generating the data. It basically preserves maximum information while removing redundant components. This new conceptual framework involves parameter-free methods to tackle the problem of information decomposition by means of causal deconvolution based on the fundamental theory of algorithmic probability which defines optimal inductive inference.

D. Quantum computation and algorithmic information

The synergy between quantum computation and algorithmic information has been studied extensively in [14]. Two main directions were explored, that can be applied for causal inference.
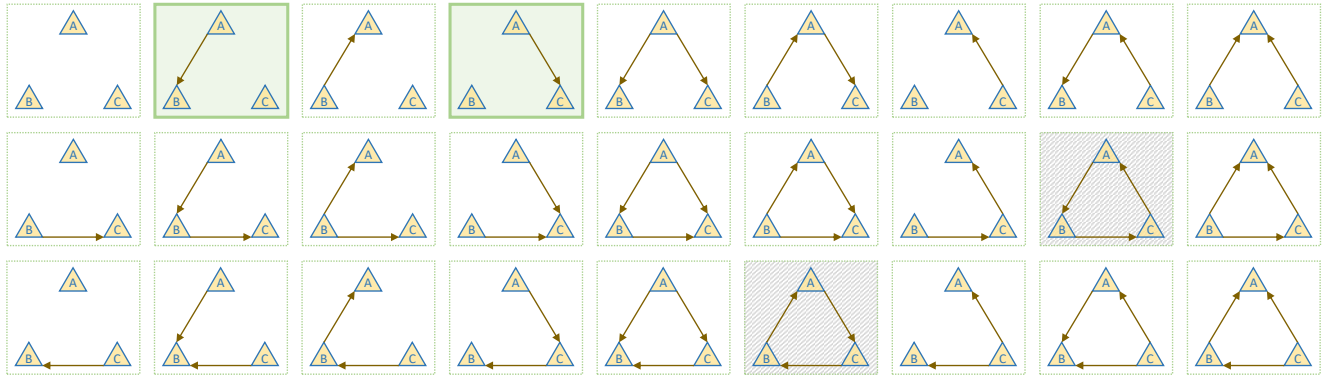
FIG. 1. All possible causal relations for 3 variables. Blocks shaded grey have causal loops (and are typically not considered). The set of blocks shaded green indicates a case of causal hypothesis testing.

In [1] a global/objective view is presented, which involves quantum automata for algorithmic information. A framework for causal inference based on algorithmic generative models is developed. This technique of quantum-accelerated experimental algorithmic information theory (QEAIT) can be ubiquitously applied to diverse domains. Specifically for genome analysis, the problem of identifying bit strings capable of self-replication is presented. A new quantum circuit design of a quantum parallel universal linear bounded automata (QPULBA) model is developed that enables a superposition of classical models/programs to be executed, and their properties can be explored. The automaton prepares the universal distribution as a quantum superposition state which can be queried to estimate algorithmic properties of the causal model.

In [15] a local/subjective view is presented, which involves universal reinforcement learning in quantum environments. This theoretical framework can be applied to automated scientific modelling. A universal artificial general intelligence formalism is presented that can model quantum processes. The developed quantum knowledge-seeking agent (QKSA) is an evolutionary general reinforcement learning model for recursive self-improvement. It uses resource-bounded algorithmic complexity of quantum process tomography (QPT) algorithms. The cost function determining the optimal strategy is implemented as a mutating gene within a quine. The utility function for an individual agent is based on a selected quantum distance measure between the predicted and perceived environment.

These techniques motivate our research in this article. Unlike quantum/classical data-driven machine learning, these primitives of QEAIT and QKSA preserve the explanatory power of the model the learning converges to by exploring the space of programs on an automata model. The QPULBA model (in QEAIT) and QPT algorithms (in QKSA) can be generalized to a causal oracle and causal tomography respectively.

III. CAUSAL HYPOTHESIS TESTING

A canonical approach in causal inference is to formulate different hypotheses on the cause–effect relations and test them against each other. This technique is typically used when there is some knowledge of the characteristic of the phenomenon that is being tested. The complete search space of directed graphs between events grows exponentially. An exhaustive example of 3 variable causal relations is shown in Fig. [1]. In causal hypothesis testing (CHT) a subset of these graphs is considered as the set of hypotheses being tested against each other. For example, in Fig. [1], we can consider, two hypotheses (shaded in green):

- A causes B, C is independent
- A causes C, B is independent.

A. Quantum advantage in classical causal hypothesis testing

Quantum information enables a richer spectrum of causal relations that is not possible to access via classical statistics. Most research in this direction is towards exploring causality in the quantum context [16–21]. Our focus in this work is specifically using the quantum formulation to provide a *computational advantage with respect to a classical technique on classical data.*

This problem is studied extensively in [4]. The research analyzes the task of identifying the effect of a given input variable. The constructed quantum strategy reduces the error probability by an exponential amount, doubling the decay rate of the error probability with the number of accesses to the relevant variables. This decay rate is the highest achievable rate allowed by quantum mechanics, even if one allows for exotic setups where the order of operations is indefinite. The key ingredient of the quantum speedup is the ability to run multiple equivalent experiments in a quantum superposition. This can be used for the detection of a causal link between two variables, and
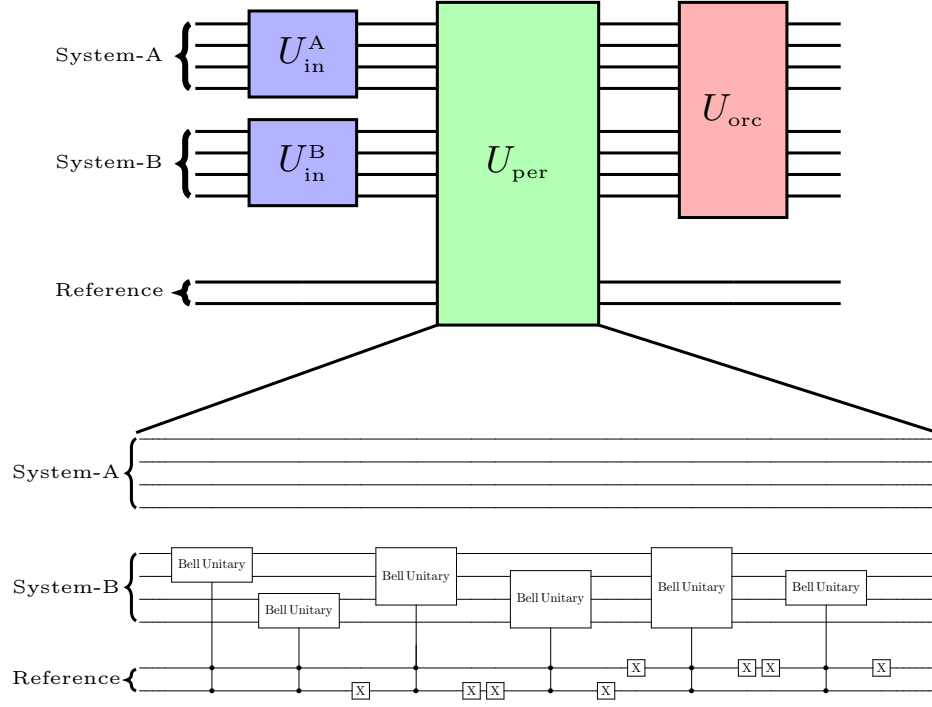
FIG. 2. The illustration of the model implementation. Where we also show the decomposition of U_{per} for $N_{\mathrm{A}} = N_{\mathrm{B}} = 4$.

in the identification of the cause of a given variable, as well.

The set of allowed causal relationships depends on the physical theory, which determines which maps can be implemented by physical processes. In classical physics, cause-effect relations are typically represented by conditional probability distributions, while, in quantum theory, they are described by quantum channels, i.e., completely positive trace-preserving maps transforming density matrices of system A into density matrices of system B.

IV. PROBLEM FORMULATION

Contrary to the general case considered in [4], our formulation is towards an implementable quantum circuit that is demonstrated on the Qiskit `statevector` simulator. This formulation is detailed here.

The input variables (causes) are denoted by the set $C = \{c_0, c_1, \ldots, c_{|C|-1}\}$, while the output variables (effects) are denoted by the set $E = \{e_0, e_1, \ldots, e_{|E|-1}\}$. If we consider an equal number of input and output variables to implement a quantum circuit to preserve unitarity, i.e., $|C| = |E| = k$. We maintain the simplification in [4] that all variables are of equal length,

$$d = |c_i| = |e_j|, \tag{1}$$

where $c_i \in C$ and $e_j \in E$. Thus, each variable has 2^d states. Furthermore, we consider that each effect is a

permutation of only one cause, i.e. the map from C to E is a bijective function, one-to-one correspondence, or invertible function.

A. Model implementation

In our implementation, we consider a general case where the access to interventions on the causes and the effects are not equal. As a proof-of-concept, we implement a causal hypothesis testing with control over 1 cause $C = \{c_0\}$ and measurement capability over 2 potential effects $E = \{e_0, e_1\}$. These hypotheses are mutually exclusive, i.e.,

- Hypothesis 0: $e_0 \leftarrow U c_0$, e_1 is independent of c_0

- Hypothesis 1: $e_1 \leftarrow U c_0$, e_0 is independent of c_0

where U is a unitary map between the cause and effect.

To design a quantum circuit, all variables that are potentially inspected as causes need to be assigned a quantum information placeholder. The total such placeholders required are,

$$k = \max\{|C|, |E|\}. \tag{2}$$

For this example, $k = 2$ sets of qubit registers are used, namely system A and system B, both initialized to the all-zero state. The initialization and interaction between these two systems and how it affects the performance of causal tomography are studied.

The practical implementation is decomposed into the following steps:

- The number of times the causal process (oracle) is queried is denoted by N. The larger the N, the lesser the error probability. For the simulations, the value is chosen as $N = N_A = N_B = 4$.

- U_{in}^A, U_{in}^B, represents the initialization of the system A and B respectively. The causes are thus, $c_0 = U_{in}^A |0\rangle^{\otimes d}$ and $c_1 = U_{in}^B |0\rangle^{\otimes d}$. These are used to study the circuit for various cases of initialization. For the simulations, the variable length $d = 1$. For each set of d qubits among the N copies, the same unitary needs to be applied.

- U_{per}, decomposes into parallel strategies which represent the possible ways to group inputs into entangled pairs. For n-qubits, the total possible such linearly independent pair combinations are $r = \dfrac{n!}{(n/2)! \times 2^{n/2}}$, where,

$$n = N \times |C| \times d \tag{3}$$

Since we have control over only 1 cause, $n = 4 \times 1 \times 1 = 4$, and $r = 4!/(2! \times 2^2) = 3$. The number of qubits in the reference to encode these states are $N_{ref} = \lceil \log_2 r \rceil$. In our case, $N_{ref} = \lceil \log_2 3 \rceil = 2$

- U_{orc}, represents an unknown process that induces a causal relation between inputs and outputs. For example, for the two hypotheses considered, it can be an `Identity` gate for hypothesis 0 and a `Swap` gate for hypothesis 1.

The qubit requirement of this model grows as:

$$N \times k \times d + \left\lceil \log_2 \left(\frac{n!}{\frac{n}{2}! \times 2^{\frac{n}{2}}} \right) \right\rceil \tag{4}$$

The model is illustrated through an implementable quantum circuit in Fig. [2] in Appendix. In the previous work, one of the innovations introduced was to entangle the inputs for the parallel strategy instead of initializing with a tensor product state which reduces the exponential measurement resource requirement by correlating the basis. But the cost of implementing the entangled initialization was not analyzed.

Through our model implementation, we are able to show that the operations required to implement the U_{per} grows faster than exponentially with the dimension of the subsystem as illustrated in Fig. [3] but linear in respect to the linearly independent states i.e. r.

B. Practical error probability

While the number of controllable causes $|C|$ and the encoding length of the causes d depends on the problem formulation, the number of queries, i.e. N and thereby r is a free parameter. It can be chosen based on the available quantum circuit resources of the number of qubits in the quantum processor, the decoherence time, and gate error probability, such that the pragmatic error remains low.

In the quantum case, the coherent superposition of equivalent configuration gives the error probability [4]

$$p_{err} = \frac{r}{2d^N} \left(1 - \sqrt{1 - r^{-2}} \right) \xrightarrow{r \gg 1} \frac{1}{4rd^N}, \tag{5}$$

where r is the number of linearly independent states. It is suggested that the Eq.(5) gives the limiting case error probability.

However, in a more general case, the error probability of causal hypothesis testing should be dependent on the two specific hypotheses.

As an example, if the unitary oracles implementing the two causal hypotheses, $U_{orc}^{H0} = U_{orc}^{H1}$, then $p_{err} = 1$.

To incorporate this dependence on the relative distinguishability of the hypotheses, we introduce a correction factor proportional to the process distance between the two oracles. Hence the modified Eq.(5) takes the form

$$p_{err}^{prac} = \frac{r}{2d^N} \left(1 - \sqrt{1 - r^{-2}} \right) \Delta \left[U_{orc}^{H0}, U_{orc}^{H1} \right] \xrightarrow{r \gg 1} \frac{\Delta \left[U_{orc}^{H0}, U_{orc}^{H1} \right]}{4rd^N}, \tag{6}$$

There are many choices for the distance Δ function between the hypotheses and need to be chosen based on the experimental and theoretical specifications of the application. For this investigation, we have considered three distance measures

1. *Trace distance* which is defined by

$$\Delta = \frac{1}{2} \text{Tr} \left(\rho_{orc} - \rho_{orc}^{alter} \right), \tag{7}$$

where ρ represents the Choi representation of the unitary.

2. *Bures distance* which defined as

$$\Delta = 2 \left(1 - \sqrt{F(\rho_{orc}, \rho_{orc}^{alter})} \right), \tag{8}$$

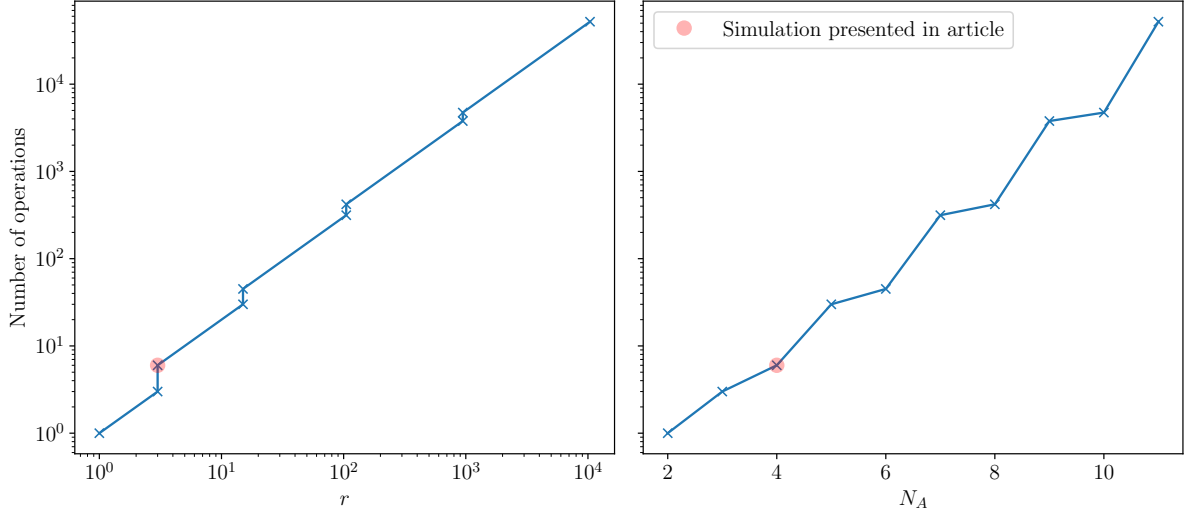where F quantifies the process fidelity between the Choi representation of oracles.

FIG. 3. Illustration of the number of controlled Bell unitary required to implement the U_{per} in respect to a number of linearly independent state r (left-hand side figure) and with the dimension of the subsystem (right-hand side figure).

3. *Hilbert-Schmidt distance* which is defined by

$$\Delta = \text{Tr}\left[\left(\rho_{\text{orc}} - \rho_{\text{orc}}^{\text{alter}}\right)^2\right]. \tag{9}$$

C. Numerical results

To obtain numerical results we consider the setup represented in Fig. [2] where the dimension of the subsystems are the same. IBM's open-source quantum computer simulator `qiskit` was used to simulate the above-mentioned implementation. Throughout the numerical investigation, we choose the oracle unitary as follows:

$$U_{\text{orc}}^{H_{0/1}}(\theta_x, \theta_y) = \left[\text{CSWAP}(c, A_i, B_i)^{\otimes N}\right]\left[\text{RX}(c, \theta_x) \otimes \mathbb{I}(A_i)^{\otimes N} \otimes \text{RY}(B_i, \theta_y)^{\otimes N}\right] \tag{10}$$

The parameter θ_x allows interpolating between the two hypotheses discussed before. To obtain the specific hypothesis cases of `Identity` and `SWAP`, the above parameters are set to $U_{\text{orc}}^{H_0}(0,0)$ and $U_{\text{orc}}^{H_1}(\pi, 0)$ respectively. The dependence of the error probability of distinguishing two hypotheses with respect to the relative difference is enquired by varying the alternate hypothesis to

$$U_{\text{orc}}^{H_{0/1}} = \begin{cases} U_{\text{orc}}^{H_0}(0,0), \\ \\ U_{\text{orc}}^{H_1}(\theta_x,0) \end{cases} \tag{11}$$

The result is illustrated in Fig.[5] and left-hand side of Fig.[4]. It can be identified that the lowest error (or most distinguishability) is achieved for the Identity vs. SWAP case, when the angle is an odd multiple of π, (i.e., the RX gate is a Pauli-X) while for even multiples of π, the hypotheses are not distinguishable, with an error probability of 1.

In the following demonstration, we test the dependence of the error probability on arbitrary unitaries applied to a specific hypothesis when they were originally maximally distinguishable. For this, we set the parameters to

$$U_{\text{orc}}^{H_{0/1}} = \begin{cases} U_{\text{orc}}^{H_0}(0,0), \\ \\ U_{\text{orc}}^{H_1}(\pi, \theta_y) \end{cases} \tag{12}$$

The result is shown in Fig.[5] and right-hand side of Fig.[4]. It is seen that the error probability is independent of local operations on a specific hypothesis.

In the Fig. [5] for setting of oracle (11) we observe that the error probability varies as

$$p_{\text{err}}^{\text{prac}} = \begin{cases} p_{\text{err}}^{\text{max}} & \text{for } \theta = (2m+1)\pi/2, \\ p_{\text{err}}^{\text{min}} & \text{for } \theta = m\pi, \\ p_{\text{err}}^{\text{min}} < p_{\text{err}}^{\text{prac}} < p_{\text{err}}^{\text{max}} & \text{otherwise,} \end{cases}$$

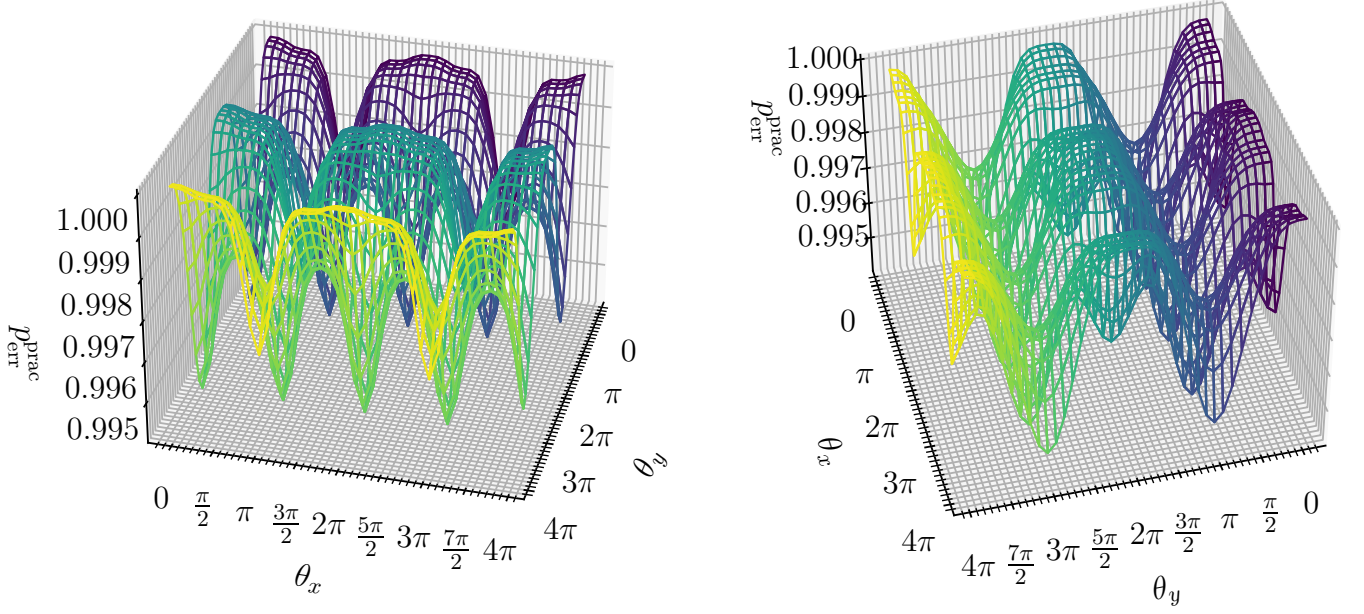and in the Fig. [5] for setting of oracle (12) we observe

FIG. 4. On the left-hand side we illustrate the variety of practical case error probability with respect to θ_x. On the right-hand side we show the variation of practical case error probability with respect to oracle angle θ_y.

that the error probability varies as

$$p_{\text{err}}^{\text{prac}} = \begin{cases} p_{\text{err}}^{\max} & \text{for } \theta = 2m\pi, \\ p_{\text{err}}^{\min} & \text{for } \theta = (2m+1)\pi, \\ p_{\text{err}}^{\min} < p_{\text{err}}^{\text{prac}} < p_{\text{err}}^{\max} & \text{otherwise}, \end{cases}$$

where $m = 0, 1, 2, \ldots$. Through the numerical simulation we show that the $p_{\text{err}}^{\max} = 1.0$, when two hypothesis coincide i.e. $\theta_x = 0$, $\theta_y = 0$; and $p_{\text{err}}^{\min} = 0.994$, when we consider the alternate hypothesis $\theta_x = \pi$, $\theta_y = 0$.

It should be noted that the construction we used for numerical simulation has the following parameter values: $r = 3$, $d = 2$ and $N = 4$. In this setting, the limiting case probability that is given in ref.[4] and expressed in Eq.(5) is constant at

$$p_{\text{err}} = 0.00536. \tag{13}$$

We observe that the results for the limiting case and the numerically simulated error probability are distinct. This is because the p_{err} was introduced purely theoretically where the authors evaluated it for a maximally mixed alternate hypothesis. In a practical case (numerical simulation of the quantum circuit) it is not feasible to construct a maximally mixed oracle. This leads us to introduce the realistic case error probability. In the practical scenario (see oracle presented in (10)) the alternate hypothesis is implementable in currently widely available hardware and through the simulation, we show that the limiting case error probability in the practical scenario requires a slight modification. This modification is reflected through the introduction of the distance between the oracles under consideration i.e. $\Delta \left[U_{\text{orc}}^{\text{alter}}, U_{\text{orc}} \right]$. Based on this in Eq.(6) we introduce a modified version of limiting
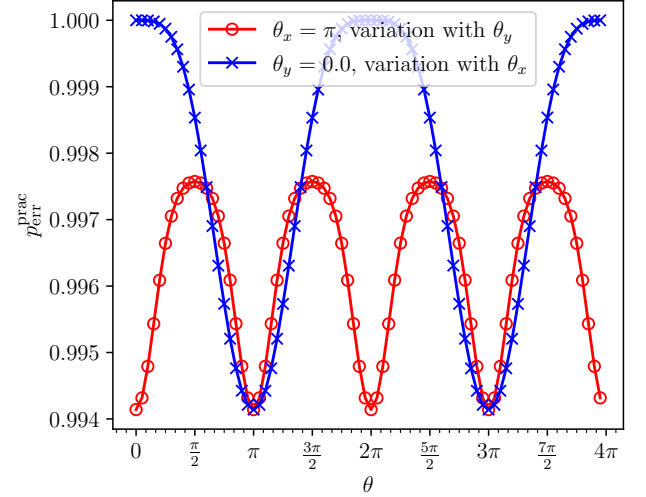

FIG. 5. Illustration of variation of practical case error probability with respect to θ_x, when $\theta_y = 0.0$ (blue line) and θ_y, when $\theta_x = \pi$ (red line).

case error probability that can be evaluated realistically on quantum hardware.

Next, it is logical to examine the variation of process distance with respect to the oracle parameter θ_y. In Fig. [6] we illustrate the distance between the null hypothesis $U_{\text{orc}}^{H_0}(0, 0)$ and the alternate hypothesis $U_{\text{orc}}^{H_1}(\pi, \theta_y)$ for a class of distance measure.

For the sake of the demonstration, we present the observation for *Trace distance*, *Bures distance* and the
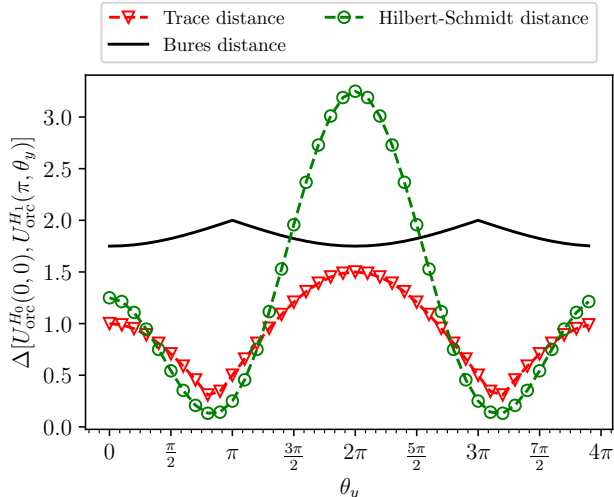
FIG. 6. Illustration variation of different distance measure with respect to the oracle angle θ_y.

Hilbert-Schmidt distance. It can be seen that the characteristics of *Trace* and *Hilbert-Schmidt* are similar, due to the fact that both of them fundamentally depends on the difference in the Choi representation of the oracles, whereas *Bures distance* depends on the process distance.

V. APPLICATION FRAMEWORK

With the quantum kernel presented above, in this section, we will develop an application framework in the context and discuss two consequential applications.

A. Bioinformatics

Applications of causal inference are widespread in bioinformatics. Specifically, inferring a causal network of the variables is practised in medical diagnostics and genomics.

For example, causal discovery in Alzheimer's pathophysiology is studied in ref. [22] with 9 variables (13 with longitudinal data). Similarly, for detecting causal regulatory interactions between genes, tools like `Scribe-py` [23] are currently used. Exemplary use cases of (i) transcription expression dynamics hierarchy of C. elegans' early embryogenesis and (ii) core regulatory network responsible for myelopoiesis are used for this research, with the latter graph consisting of 10 nodes.

The current generation of quantum processors supports 100s of qubits and is expected to scale to 1000s in a few years. However, the challenge is the limitation of the decoherence time and gate errors, which bounds the runtime of the algorithm that can be effectively executed. For a causal graph in the order of 10 causes/effects, a causal specificity bits of $d = 1$ and with $N = 100$, the estimation presented in Eq.(4) is 6262 qubits. Keeping in mind the potential of near-term devices, pragmatic industrial cases of causal tomography will remain outside the reach of quantum computing in the near term.

B. Artificial General Intelligence

In the long term, quantum accelerated causal inference will be beneficial for artificial general intelligence. Quantum accelerated AGI is still in its infancy. In ref. [24], the authors proposed the AIXI-q reinforcement learning agent empowered with quantum counting. Meanwhile, ref. [1] proposed an exhaustive enumeration of all causal oracles (or, alternatively, all bounded-size programs of a Turing machine). These techniques can develop synergies with the quantum accelerated causal tomography circuit as developed in this article.

In theoretical physics, automated science tools, specifically in the context of causal set theory, will also find the causal tomography framework of crucial use.

VI. CONCLUSION

In this article, we extend the previously introduced [4] causal hypothesis testing formulation for the practical scenario. This lead us to develop a scalable quantum gate-based algorithm which can be implemented in the available near-term quantum devices. Through the algorithm formulation, we empirically show that the limiting case error probability that is represented in ref. [4] requires modification. In our work, this modification is done by introducing process distance between the causal hypotheses, to the formulation of error probability. We term this modified version of error probability as *practical case error probability* which is stricter than the limiting case. Additionally, our implementation enables an estimation of the pragmatic gate complexity of the causal tomography entangled pair indexing.

Furthermore, the proposed algorithm is implemented using the open-source quantum programming and simulation platform `qiskit`. As in ref. [4] it is shown that the quantum advantage holds for generalized probabilistic theory so the practical case that we present here is optimal in any scenario.

Our motivation for this project is driven by the increasing focus on causal inference in machine learning. Besides monitoring the information flow in future quantum communication networks, as discussed in ref.[4], causal tomography is crucial for understanding the bounds of general intelligence and for bioinformatics use cases. In our future work, we aim to apply our developed quantum accelerated causal tomography framework for these applications.

[1] A. Sarkar, Z. Al-Ars, and K. Bertels, Estimating algorithmic information using quantum computing for genomics applications, Applied Sciences **11**, 2696 (2021).

[2] K. Ried, M. Agnew, L. Vermeyden, D. Janzing, R. W. Spekkens, and K. J. Resch, A quantum advantage for inferring causal structure, Nature Physics **11**, 414 (2015).

[3] J. F. Fitzsimons, J. A. Jones, and V. Vedral, Quantum correlations which imply causation, Scientific reports **5**, 1 (2015).

[4] G. Chiribella and D. Ebler, Quantum speedup in the identification of cause–effect relations, Nature communications **10**, 1 (2019).

[5] A. Lavin, H. Zenil, B. Paige, D. Krakauer, J. Gottschlich, T. Mattson, A. Anandkumar, S. Choudry, K. Rocki, A. G. Baydin, *et al.*, Simulation intelligence: Towards a new generation of scientific methods, arXiv preprint arXiv:2112.03235 (2021).

[6] Y. Maruyama, Categorical artificial intelligence: the integration of symbolic and statistical ai for verifiable, ethical, and trustworthy ai, in *International Conference on Artificial General Intelligence* (Springer, 2021) pp. 127–138.

[7] A treatment is an action that can be applied or withheld from that unit.

[8] R. Guo, L. Cheng, J. Li, P. R. Hahn, and H. Liu, A survey of learning causality with data: Problems and methods, ACM Computing Surveys (CSUR) **53**, 1 (2020).

[9] K. J. Rothman and S. Greenland, Causation and causal inference in epidemiology, American journal of public health **95**, S144 (2005).

[10] J. Pearl *et al.*, Models, reasoning and inference, Cambridge, UK: CambridgeUniversityPress **19** (2000).

[11] G. W. Imbens and D. B. Rubin, *Causal inference in statistics, social, and biomedical sciences* (Cambridge University Press, 2015).

[12] P. Spirtes, C. N. Glymour, R. Scheines, and D. Heckerman, *Causation, prediction, and search* (MIT press, 2000).

[13] H. Zenil, N. A. Kiani, A. A. Zea, and J. Tegnér, Causal deconvolution by algorithmic generative models, Nature Machine Intelligence **1**, 58 (2019).

[14] A. Sarkar, Applications of quantum computation and algorithmic information: for causal modeling in genomics and reinforcement learning, (2022).

[15] A. Sarkar, Z. Al-Ars, H. Gandhi, and K. Bertels, Qksa: Quantum knowledge seeking agent–resource-optimized reinforcement learning using quantum process tomography, arXiv preprint arXiv:2112.03643 (2021).

[16] F. Costa and S. Shrapnel, Quantum causal modelling, New Journal of Physics **18**, 063032 (2016).

[17] C. Giarmatzi, A quantum causal discovery algorithm, in *Rethinking Causality in Quantum Mechanics* (Springer, 2019) pp. 125–150.

[18] M. A. Javidian, V. Aggarwal, F. Bao, and Z. Jacob, Quantum entropic causal inference, in *Quantum Information and Measurement* (Optical Society of America, 2021) pp. F2C–3.

[19] G. Bai, Y.-D. Wu, Y. Zhu, M. Hayashi, and G. Chiribella, Quantum causal unravelling, arXiv preprint arXiv:2109.13166 (2021).

[20] G. Bai, Y.-D. Wu, Y. Zhu, M. Hayashi, and G. Chiribella, Efficient algorithms for causal order discovery in quantum networks, arXiv preprint arXiv:2012.01731 (2020).

[21] J. Bavaresco, M. Murao, and M. T. Quintino, Strict hierarchy between parallel, sequential, and indefinite-causal-order strategies for channel discrimination, Physical review letters **127**, 200504 (2021).

[22] X. Shen, S. Ma, P. Vemuri, and G. Simon, Challenges and opportunities with causal discovery algorithms: application to alzheimer's pathophysiology, Scientific reports **10**, 1 (2020).

[23] X. Qiu, A. Rahimzamani, L. Wang, B. Ren, Q. Mao, T. Durham, J. L. McFaline-Figueroa, L. Saunders, C. Trapnell, and S. Kannan, Inferring causal gene regulatory networks from coupled single-cell expression dynamics using scribe, Cell systems **10**, 265 (2020).

[24] E. Catt and M. Hutter, A gentle introduction to quantum computing algorithms with applications to universal prediction, arXiv preprint arXiv:2005.03137 (2020).