Politecnico di Torino

Politecnico di Torino

Master's Degree in Quantum Engineering

A.a. 2025/2026

Graduation Session December 2025

Quantum Causal Discovery

Supervisor: Davide Girolami

Candidate: Xuwei Ou

Abstract

Transfer entropy plays an important role in causal inference and causal discovery in time series analysis. As an information-theoretic measure, transfer entropy can detect nonlinear causal structures without any prior model assumptions. The transfer-entropy-based causal analysis scenario thus has a broader range of applications than the model-based methods like Granger causality tests. This study generalizes the definition of transfer entropy to time-parameterized quantum processes. We define a basis-dependent operational measure of one-way information flow through quantum channels, called quantum transfer entropy. Quantum transfer entropy provides a new characteristic of quantum channels. A nonzero quantum transfer entropy indicates the directed causal influence from the source to the target. In addition, we study controlled-Z channels and find that the quantum transfer entropy is a linear function of the total information of the control qubit—including both classical Shannon information and coherence in a given reference basis —thus providing an operational interpretation of information flow through quantum control channels. We further quantify the effect of a third party by studying the extreme case of the Toffoli gate, where the two control qubits are equivalent and symmetric with respect to the target qubit. This effect is characterized by the quantum transfer entropy conditioned on the third party (or the environment). We illustrate several bounds on the conditional quantum transfer entropy for the Toffoli gate in terms of local systems. These bounds allow us to estimate the conditional transfer entropy without any knowledge of the environment qubit. Finally, we employ the proposed quantum transfer entropy for causal discovery in a synthetic time-series dataset.

Table of Contents

List of Tables

List of Figures

Chapter 1

Causation in Classical Information Systems

1.1 Measures of Information Flow

Schreiber [1] proposed an information-theoretic measure to quantify the causal relationship between two time series. This measure is conventionally regarded as an indicator of information transmission from one time-dependent stochastic process to another. The so-called transfer entropy intuitively characterizes the direction and magnitude of information flow within a communication network [1, 2, 3].

In the following, we introduce the notations used throughout this work. A time series is denoted by a capital letter, e.g., X. The random vector of discrete variables of X (or the coarse-grained random vector of continuous X) from time i to j is represented as $X_{i:j}$. Accordingly, X_t denotes the random variable of X at time t.

Definition 1.1.1 (Shannon Entropy[4]). *The Shannon entropy of a discrete random variable $X \sim p_x$ is defined as the following function of distribution $\{p_x\}$:*

$$H(X) := -\sum_x p_x \log p_x. \tag{1.1}$$

Definition 1.1.2 (Transfer entropy [1]). *Given two time series X_t and Y_t, the transfer entropy from X to Y, denoted by $\mathcal{TE}(X \to Y)$ is defined as the mutual infromation between the past of X and the present of Y, conditioned by the past of Y:*

$$\mathcal{TE}(X \to Y) := I(X_{0:t-1} : Y_t | Y_{0:t-1}). \tag{1.2}$$

The (classical) transfer entropy $\mathcal{TE}(X \to Y)$ quantifies the amount of information that the past of X provides about the future of Y beyond what is already

contained in the past of Y.

Proposition 1.1.3 (Properties of transfer entropy)**.**

1. *Asymmetry. Generally,* $\mathcal{TE}(X \to Y) \neq \mathcal{TE}(Y \to X)$.
2. *Upper and lower bounds.* $0 \leq \mathcal{TE}(X \to Y) \leq H(X)$*, where H(X) is the Shannon entropy of* X.
3. *Model-free. There's no assumption, e.g. linearity or Gaussianity, for transfer entropy.*
4. *Equivalence to Granger Causality. For Gaussian variables, transfer entropy and Granger causality are equivalent.*

1.2 Causality Modeling

Correlations do not imply causation; causal analysis requires a more subtle and rigorous investigation. As proposed by Pearl in 2000 [5], the so-called Ladder of Causation framework consists of three levels: association, intervention, and counterfactual analysis [5, 6]. The first level, association, examines the raw correlations between random variables, which may still include spurious relationships. The second level, intervention, involves designed experiments in which the empirical variables of interest can be fully controlled, thereby revealing the causal effects of those variables. Finally, the counterfactual level seeks to uncover the underlying causal structure of reality based on observed data.

The term causality modeling refers to the geometric structure of interactions among a finite set of random variables. It is typically represented by a directed acyclic graph (DAG) [6, 7, 8], where the variables are depicted as nodes, and arrows indicate the direction of causal influence from one node to another.

Definition 1.2.1 (Classical Causal Model [6])**.** *A classical causal model is given by*

(1) a causal strcture associated to a DAG $G(N, E)$ *with nodes corresponding to random variables* X_i*,*

(2) for each X_i*, a classical channel* $P(X_i|P_a(X_i))$*, where* $P_a(X_i)$ *denotes the set of parents of* X_i*.*

Thus the joint probability distribution over $X_1, \ldots, X_n$ *is given by*

$$P(X_1, \ldots, X_n) = \prod_i P(X_i|Pa(X_i)). \tag{1.3}$$

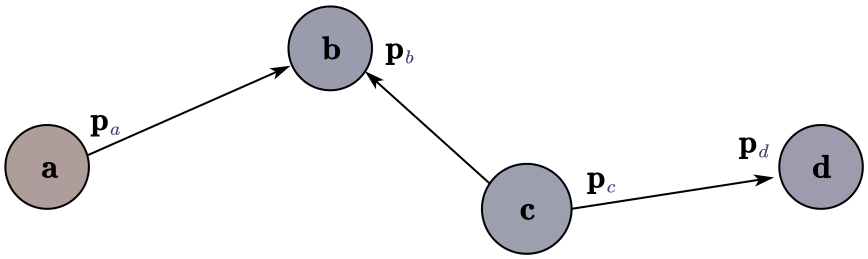


Figure 1.1: An example of classical causal model.

There are two categories of causal analysis: causal inference and causal discovery [6]. Causal inference relies on the ability to empirically control the variables under investigation and is widely employed in scientific laboratory settings. In contrast, causal discovery seeks to recover the underlying causal structure solely from observational data, where experimenters cannot directly control any variables — a situation commonly encountered in social and economic research.

Chapter 2

Introduction to Quantum Information Theory

The formalism of quantum theory provides an operational probabilistic framework that extends beyond Shannon's classical information theory. Each quantum system is associated with a Hilbert space—a complex normed linear vector space—in which linear operations and distances between states are well defined. A quantum state is represented by an element of the corresponding Hilbert space. The superposition p rinciple of quantum states enables a reduction in data-encoding resources owing to its inherently high parallelism. Moreover, quantum correlations such as entanglement offer an even more powerful resource for information processing compared with the classical counterpart.

2.1 State Space and Operators

Definition 2.1.1 (Hilbert space)**.** *A complete inner product space is called a Hilbert space, denoted by* $\mathcal{H}$*. In quantum mechanics we focus on complex Hilbert spaces. The dimension of* $\mathcal{H}$ *is the number of its basis vectors needed to span the space, denoted by* $\dim \mathcal{H}$*.*

In quantum information theory, we focus on the complex finite-dimensional Hilbert space and employ the *Dirac' notation* [4, 9]. Given $\mathcal{H} = \mathbb{C}^n$ with $n = dim\mathcal{H} < \infty$, an element in $\mathcal{H}$ can be represented as a columm vector, denoted by the "ket" $|\psi\rangle$. Its dual vector is represented using the "bra" notation $\langle\psi|$.

Definition 2.1.2 (Projector)**.** *The projector of a unit vector* $|\psi\rangle$ *on* $\mathcal{H}$ *is the outer*

product of itself:

$$P = |\psi\rangle \langle\psi| , \tag{2.1}$$

which maps from $\mathcal{H}$ onto the one-dimensional space $\mathbb{C}|\psi\rangle$.

Definition 2.1.3 (Quantum operator)**.** *The set of linear maps from Hilbert space $\mathcal{H}_A$ onto $\mathcal{H}_B$ is called the quantum operators, denoted by*

$$L(\mathcal{H}_A, \mathcal{H}_B) := \{O : \mathcal{H}_A \to \mathcal{H}_B | O \text{ is linear}\}. \tag{2.2}$$

Each $O \in L(\mathcal{H}_A, \mathcal{H}_B)$ has a adjoint operator, denoted by $O^\dagger$.

Definition 2.1.4 (Self-adjoint/Hermitian operator)**.** *An operator A is called self-adjoint (or Hermitian) if $A^\dagger = A$.*

Theorem 2.1.5. *The eigenvalues of a Hermitian operator are always real.*

Theorem 2.1.6 (Spectral Theorem)**.** *If an operator $O \in L(\mathcal{H})$ is Hermitian, then it can be decomposed in terms of the linear combination of one of its orthonormal basis:*

$$O = \sum_{i=0}^{d-1} \lambda_i |\psi_i\rangle \langle\psi_i| , \tag{2.3}$$

where $\lambda_1, \ldots, \lambda_d \in \mathbb{R}$ are real eigenvalues of O, $\{|\psi_i\rangle\}_{i=0}^{d-1}$ forms an orthormal eigenvalues of O, and $d = dim(\mathcal{H})$.

We summarize and clarify the main concepts of operators, observables, and transformations within the framework of quantum information theory. These notations and concepts constitute the foundation of our work.

- **Quantum operators**. The quantum operators from Hilbert space $\mathcal{H}_A$ to $\mathcal{H}_B$ are elements of the collection of linear maps from $\mathcal{H}_A$ onto $\mathcal{H}_B$.
- **Quantum observables**. In quantum mechanices, the observables manifest as *self-adjoint operators* on a complex Hilbert space $\mathcal{H}$.
- **Quantum channels (transformations)**. A quantum channel is a *complete positive trace preserving (CPTP)* superoperator that maps from $L(\mathcal{H}_A)$ to $L(\mathcal{H}_B)$. A quantum superoperator can be represented by the *Choi operator*.

Definition 2.1.7 (Positive semi-definite operator)**.** *An operator $A \in L(\mathcal{H})$ is positive semi-definite if A is Hermitian and all its eigenvalues are nonnegative. We define the set of positive semi-definite operators on $\mathcal{H}$ as*

$$PSD(\mathcal{H}) := \{A \in L(H) : A \text{ positive semidefinite}\}. \tag{2.4}$$

Throughout this thesis, we adopt a comprehensive perspective on quantum states from the viewpoint of operators. Unless otherwise specified, the term quantum state refers to a density operator.

Definition 2.1.8 (Density operator)**.** *The set of density operators $D(\mathcal{H})$ on $\mathcal{H}$ is defined as the following*

$$D(\mathcal{H}) := \{\rho \in PSD(\mathcal{H}) : tr\rho = 1\}. \tag{2.5}$$

$D(\mathcal{H})$ is called the state space on $\mathcal{H}$.

Proposition 2.1.9. *$D(\mathcal{H})$ is a convex set, i.e.*

$$\rho = \sum_i p_i\rho_i \in D(\mathcal{H}), \forall \rho_i \in D(\mathcal{H}), \forall p_i \in [0,1] \ s.t. \sum_i p_i = 1. \tag{2.6}$$

An operational interpretation meaning of the mixed states originates from the fact that, in practice, because of the decoherence one can not prepare a number of identical quantum states but an *ensemble* $\{p_i, \rho_i\}$ of quantum states obeying some probability distribution p_i. To describe the *average* of such an ensemble, we should take $\rho = \sum_i p_i\rho_i$. It can be readily verified that $\rho \in D(\mathcal{H})$. To describe the average behavior of such an ensemble, the state is represented as

$$\rho = \sum_i p_i\rho_i. \tag{2.7}$$

Definition 2.1.10 (Pure and mixed states)**.** *A quantum state ρ is called pure if it can be represented in the form of $\rho = |\psi\rangle\langle\psi|$ for some unit vector $|\psi\rangle \in \mathcal{H}$, equivalent to the criterion $\rho = \rho^2$ (or $tr\rho^2 = 1$).*
A quantum state ρ is called mixed if it's not pure.

2.2 Quantum State Distance

In this section, we provide a comprehensive introduction to the geometric measures in an (linear) operator space $L(\mathcal{H}_A, \mathcal{H}_B)$, and in a state space $D(\mathcal{H})$. A linear matric space naturally defines a distance function based on inner product.

Definition 2.2.1 (Distance Function and Metric Space)**.** *A metric space is a set M with a real-valued distance function (or matric) $d(\cdot,\cdot)$ on $M \times M$,*

$$d : M \times M \to \mathbb{R} \tag{2.8}$$

$$d : (x, y) \mapsto d(x, y), \tag{2.9}$$

which satisfies the following properties

1. $d(x, y) \geq 0$;
2. $d(x, y) = 0 \iff x = y$;
3. $d(x, y) = d(y, x)$;
4. *(triangle inequality)* $d(x, z) \leq d(x, y) + d(y, z)$.

Generally we denote the metric space by $\langle M, d \rangle$.

Throughout the following discussion, we use the term *norm* to refer generally to the distance function defined on a linear metric space, such as a Hilbert space, a state space, or an operator space.

Definition 2.2.2 (l^p-Norm)**.** *In linear algebra, the* l^p*-norm of a vector* $x = (x_1, \ldots, x_n)^T \in \mathbb{C}^n$ *is defined by*

$$\|x\|_p := \begin{cases} (\sum_{i=1}^n |x_i|^p)^{1/p}, & if p \in [1, \infty) \\ max_i\{|x_i|\}, & if p = \infty \end{cases} \tag{2.10}$$

Definition 2.2.3 (Norm of Operator)**.** *Given an operator* $M \in L(\mathcal{H}_A, \mathcal{H}_B)$ *whose singular values are* $s_1, \ldots, s_r > 0$*, we define*

- *the Schatten p-norm by* $\|M\|_p := \|s\|_p$*, where* $s = (s_1, \ldots, s_r)^T$;
- *the trace norm* $\|M\|_1 = tr(\sqrt{M^\dagger M})$ *as the Schatten* 1*-norm;*
- *the Frobenius norm* $\|M\|_2$ *as the Schatten* 2*-norm;*
- *the operator norm* $\|M\|_\infty$ *as the Schatten* ∞*-norm.*

In quantum information processing, we often encounter the task of comparing a received quantum state with a target state. This motivates us to define distance measures (or pseudo distance measures) between two density matrices, known as the *trace distance* and the *fidelity*.

Definition 2.2.4 (Trace Distance)**.** *The (normalized) trace distance* $T(\rho, \sigma)$ *between two state* $\rho, \sigma \in D(\mathcal{H})$ *is defined as a half of the trace norm of* $\rho - \sigma$*,*

$$T(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_1 \tag{2.11}$$

Clearly, the *trace distance* is a matric on $D(\mathcal{H})$.

Definition 2.2.5 (Fidelity)**.** *The fidelity between two states $\rho, \sigma \in D(\mathcal{H})$ is defined as*

$$F(\rho, \sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1 = tr\left(\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right) = tr\left(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right). \tag{2.12}$$

Fidelity dosen not fulfill the definition of distance measures. Indeed, it's maximized when $\rho = \sigma$.

2.3 Composite Systems

In this section, we introduce the mathematical formalism for a composite quantum system comprising multiple subsystems. This framework is essential in quantum information theory, as one often needs to analyze the global system from an information-theoretic perspective while having only partial information about its local constituents. A further motivation stems from the fact that any quantum system is inevitably coupled to its environment. Therefore, a detailed investigation of the system's evolution and decoherence requires us to consider the entire, joint system.

Axiom 2.3.1 (Composite Hilbert Space)**.** *The composite (or global) Hilbert space of a quantum system composed by n subsystems is given by*

$$\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n. \tag{2.13}$$

Definition 2.3.2 (Product States, Separable States, and Entangled States)**.** *The global state $\rho \in D(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n)$ is called a*

- *product state if $\rho = \rho_1 \otimes \cdots \otimes \rho_n$, $\rho_i \in D(\mathcal{H}_i)$;*
- *separable state if $\rho = \sum_k p_k \rho_{1k} \otimes \cdots \otimes \rho_{nk}$, where $\rho_{ik} \in D(\mathcal{H}_i)$, $p_k \in [0,1]$ and $\sum_k p_k = 1$;*
- *entangled state if ρ is not separable.*

One of the most fundamental distinctions between quantum and classical mechanics lies in the quantum superposition principle, commonly referred to as quantum coherence. This principle originates from the interference (or coherence) phenomena of light. Given a density matrix $\rho \in \mathcal{H}$, represented in some orthonormal basis

$\{|i\rangle\}$:

$$\rho = \begin{pmatrix} a_{00} & \cdots & a_{0n} \\ & a_{11} & \cdots & a_{1n} \\ & & \ddots & \vdots \\ & & & \ddots \\ a_{n0} & a_{n1} & \cdots & a_{nn} \end{pmatrix}, \tag{2.14}$$

the diagonal entries can be distinctly extracted by performing projective operators of basis vector $|i\rangle$s (also called projective measurement, as will be discussed soon). In contrast, the off-diagonal terms encode the coherence properites of ρ, which determime the interference between the outcomes of projective measurements corresponding to diffenrent orthonormal basis vectors.

As expected, orthogonal projective operators represent the most ubiquitous measurement scheme in quantum theory. This is especially true in quantum computing, where projective measurement is employed in the final step to sample an algorithm's output state. Moreover, the notion of projective measurement can be generalized to describe the foundational principles governing information extraction from a quantum system.

Definition 2.3.3 (Generalized Measurement (POVM))**.** *A generalized measurement, also called positive-operator valued measure, on a Hilbert space $\mathcal{H}$ assotiated to the outcome set Ω (generally real) is a function*

$$E : \Omega \to PSD(\mathcal{H}) s.t. \sum_{x\in\Omega} E(x) = I. \tag{2.15}$$

Sometimes we use the notation $E_x = E_x$. Every E_x takes the form of $E_x = M_x^\dagger M_x$ s.t. $\sum_x M_x^\dagger M_x = I$.
When E_x spans a orthognal set of operators, the function E reduces to the projective operator.

Axiom 2.3.4 (Born's Rule)**.** *Performing a POVM, denoted by E, on a quantum state $\rho \in D(\mathcal{H})$, then the probability of observing outcome $x \in \Omega$ is*

$$p_x = tr(E_x\rho). \tag{2.16}$$

The post-measured state of $E_x = M_x^\dagger M_x$ is

$$\rho^{pm} = \frac{1}{p_x} M_x \rho M_x^\dagger. \tag{2.17}$$

Definition 2.3.5 (Coherent State[10]). *A coherent state $|\psi\rangle$ in a Hilbert space $\mathcal{H}$ is the superposition of one basis $\{|0\rangle, \ldots, |d-1\rangle\}$*

$$|\psi\rangle = \sum_{i=0}^{d-1} p_i |i\rangle, \tag{2.18}$$

where $d = dim(\mathcal{H})$, $p_i \leq 0$ and $\sum_i p_i = 1$.

Definition 2.3.5 can be generalized to an arbitrary finite-dimensional multipartite system consisting of n subsystems $A_0, \ldots, A_{n-1}$ with respect to the reference product basis $\{|i_{0,\ldots,n-1}\rangle := |i_1\rangle \otimes \cdots \otimes |i_n\rangle\}_{i=0}^{d_0 \ldots d_{n-1}}$, where $d_0, \ldots, d_{n-1}$ are dimensions of subsystems [10, 11, 12].

Definition 2.3.6 (Incoherent States and Incoherent Set [10]). *For an n-party composite quantum system $A_0 \ldots A_{n-1}$, in the reference product basis $\{|i_{0,\ldots,n-1}\rangle := |i_1\rangle \otimes \cdots \otimes |i_n\rangle\}_{i=0}^{d_0 \ldots d_{n-1}}$, an incoherent state $\sigma_{A_0,\ldots,A_{n-1}}$ takes the form*

$$\sigma_{A_0,\ldots,A_{n-1}} = \sum_{i_{0,\ldots,n-1}} p_{i_{0,\ldots,n-1}} |i_{0,\ldots,n-1}\rangle \langle i_{0,\ldots,n-1}|, \tag{2.19}$$

where $p_{i_{0,\ldots,n-1}} \geq 0$ and $\sum_{i_{0,\ldots,n-1}} p_{i_{0,\ldots,n-1}} = 1$.
All the possible incoherent states of the quantum system $A_0 \ldots A_{n-1}$ form an incoherent set $\mathcal{I}_{A_0 \ldots A_{n-1}}$.

Note that the definitions of coherent and incoherent states are basis-dependent. Indeed, since any density matrix is Hermitian, it admits a spectral decomposition with respect to some orthonormal basis. This implies that any quantum state can, in principle, be represented as an incoherent state in its own eigenbasis.

Definition 2.3.7 (Incoherent Operation [10]). *An incoherent operation $\mathcal{O}_{IC}$ is an operation mapping incoherent states to incoherent states*

$$\mathcal{O}_{IC} : \mathcal{I}_{A_0 \ldots A_{n-1}} \to \mathcal{I}_{A_0 \ldots A_{n-1}}. \tag{2.20}$$

Definition 2.3.8 (Classical-quantum States). *Let $\mathcal{H}_X = \mathbb{C}^\Sigma$ and $\mathcal{H}_Q = \mathbb{C}^n$, a state $\rho_{XQ} \in D(\mathcal{H}_X \otimes \mathcal{H}_Q)$ is called a classical-quantum (cq) state, if it can be written in the form*

$$\rho_{XQ} = \sum_{x \in \Sigma} p(x) |x\rangle \langle x| \otimes \rho_{Q,x}, \tag{2.21}$$

corresponding to the ensemble $\{p(x), |x\rangle \langle x| \otimes \rho_{Q,x}\}$ on $D(\mathcal{H}_{XQ})$.

The classical-quantum (cq) state captures the most general quantum communication scenario: Alice encodes the classical message $X \sim p_x$, with $x \in \Sigma$, into a quantum state $\rho_{Q,x} \in D(\mathcal{H}_Q)$ which is then sent to Bob. Bob attempts to extract the information form the encoded ensemble $\mathcal{E} = \{p_x, \rho_{Q,x}\}$ by performing a POVM on the received system, yielding a measurement outcome $Z \sim p_z$. The overall communication channel is therefore described by a classical-quantum state.

Proposition 2.3.9. *For any bipartity system AB satisfying $d_A \geq d_B$, its incoherent states σ_{AB} can be represented w.r.t. $\{|i\rangle_A\}$*

$$\sigma_{AB} = \sum_i p_i \, |i\rangle \langle i|_A \otimes \rho_{B|i}, \tag{2.22}$$

where $\{|i\rangle_A\}$ is an orthonormal basis of system A and $\rho_{B|i}$ is the density matrix of B in $\{|i\rangle_A\}$. Incoherent states $\sigma_{AB} \in \mathcal{I}_{B|A}$ in such a form is called A-incoherent states. Maps $\Lambda_{IC}^{B|A} : \mathcal{I}_{B|A} \to \mathcal{I}_{B|A}$ are called A-incoherent operations.

A-incoherent states [10, 12] can be viewed as a generalization of classical-quantum states. The underlying idea stems from the fact that, in quantum information processing, we are often interested in how a quantum channel influences the correlations between a measured qubit and a prepared ancillary qubit, whose initial state can be engineered in a specific basis. After the evolution of the quantum channel, the measurement is likewise performed in the basis of the ancillary qubit. The final state of the ancilla thus encodes the information about both the quantum channel and the monitored qubit.

Definition 2.3.10 (Coherence Measure)**.** *A measure of coherence $f_C(\rho)$ is a nonnegative function which fulfills the following properties:*

- *$f_C(\rho) = 0$ if ρ is an incoherent state;*
- *$f_C(\rho)$ is a non-increasing monotone function under incoherent operations, i.e.*

$$\forall \Lambda_{IC}(\rho), f_C(\rho) \geq f_C(\Lambda_{IC}(\rho)). \tag{2.23}$$

2.4 Quatnum Entropy

Analogous to classical Shannon entropy (Definition 1.1.1), the information content of a quantum state is associated with the uncertainty in the outcomes of quantum measurements. To quantify the intrinsic information of a density operator, Von Neumann introduced a basis-independent entropy function [13], which depends solely on the eigenvalues of the density matrix.

Definition 2.4.1 (Von Neumann Entropy)**.** *The Von Neumann entropy $S(\rho)$ of a quantum state $\rho \in D(\mathcal{H})$ is defined by*

$$S(\rho) := -tr(\rho \log \rho) \tag{2.24}$$

$$= -\sum_{i}^{d} \lambda_i \log \lambda_i, \tag{2.25}$$

where $\{\lambda_i\}$ is the set of the normalized eigenvalues of ρ and $d = dim\mathcal{H}$.

Proposition 2.4.2 (Properties of Von Neumann Entropy)**.**

1. *Nonegativity. $S(\rho) \leq 0$. $S(\rho) = 0$ iff ρ is pure.*
2. *Upper bound. $S(\rho) \leq \log rank(\rho) \leq \log dim(\mathcal{H})$. $S(\rho) = \log dim(\mathcal{H})$ iff ρ is maximally mixed, i.e. $\rho = \frac{I}{dim\mathcal{H}}$.*
3. *Invariance under isometries: $S(V\rho V^\dagger = S(\rho))$ for any isometry V.*
4. *Contnuity. $S(\rho)$ is continuous.*
5. *Concavity. $S(\rho)$ is a strictly concave function of $\rho \in D(\mathcal{D})$, i.e. $S(\sum_i p_i\rho_i) \leq \sum_i p_i S(\rho_i)$ for any $\rho_i \in D(\mathcal{H})$ and a discrete probability distribution $\{p_i\}$.*

Definition 2.4.3 (Purity and Linear Entropy)**.** *The purity of a quantum state $\rho \in D(\mathcal{H})$ is defined as*

$$\gamma(\rho) := tr(\rho^2). \tag{2.26}$$

The linear entropy of ρ is defined as

$$S_l(\rho) := \frac{d}{d-1}\left(1 - \gamma(\rho)\right). \tag{2.27}$$

Proposition 2.4.4.

1. $\frac{1}{d} \leq \gamma \leq 1$.
2. $0 \leq S_l(\rho) \leq 1$.
3. *ρ is pure $\iff$ $\gamma(\rho) = 1$ $\iff$ $S_l(\rho) = 0$.*

Definition 2.4.5 (Entropy of Subsystems)**.** *Given a bipartite quantum system $\rho_{AB} \in D(\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B})$, the Von Neumann entropies of the global system and subsystems are defined by*

$$S(AB) := S(\rho_{AB}), S(A) := S(\rho_A), S(B) := S(\rho_B), \tag{2.28}$$

where $\rho_A = tr_B\rho_{AB}$ and $\rho_B = tr_A\rho_{AB}$ are reduced states of subsystems A and B respectively.

Proposition 2.4.6 (Properties of Entropy of Subsystems)**.**

1. Sunadditivity.

$$S(A) + S(B) \geq S(AB). \tag{2.29}$$

2. Non-monotinicity. Generally,

$$S(AB) \ngeq S(A), S(AB) \ngeq S(B). \tag{2.30}$$

3. Araki-Lieb (triangle) inequality.

$$S(AB) \geq |S(A) - S(B)|. \tag{2.31}$$

4. Strong subadditivity. For a tripartite system ρ_{ABC}, we have

$$S(AC) + S(BC) \geq S(ABC) + S(C). \tag{2.32}$$

5. Weak monotinicity. For a tripartite system ρ_{ABC}, we have

$$S(AC) + S(BC) \geq S(A) + S(B). \tag{2.33}$$

Definition 2.4.7 (Quantum Conditional Entropy)**.** *For a quantum system ρ_{AB}, the conditional entropies are defined as*

$$S(A|B) := S(AB) - S(B), \tag{2.34}$$
$$S(B|A) := S(AB) - S(A). \tag{2.35}$$

It is important to note that, in contrast to its classical counterpart, the quantum conditional entropy is not necessarily non-negative. A prominent example is the Bell state $|\Phi\rangle_{AB}^{+} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. For this state, the calculation $S(A|B) = S(AB) - S(B) = -S(B) \leq 0$ demonstrates this fact.

To measure the shared correlation between two quantum subsystems, we introduce the quantum mutual information.

Definition 2.4.8 (Quantum Mutual Information)**.** *The quantum mutual information of a bipartite state ρ_{AB} is defined as*

$$I(A:B) := S(A) + S(B) - S(AB) \tag{2.36}$$
$$= S(A) - S(A|B) \tag{2.37}$$
$$= S(B) - S(B|A). \tag{2.38}$$

Proposition 2.4.9 (Properties of Mutual Information)**.**

1. *Nonnegativity.* $I(A:B) \geq 0$. $I(A:B) = 0$ *iff.* $\rho_{AB} = \rho_A \otimes \rho_B$.
2. *Upper bounds.* $I(A:B) \leq 2\min\{S(A), S(B)\}$.
3. *Invariance under isometries. For any local isometries* $V_{A\to A'}$, $W_{B\to B'}$, *the mutual information dosen't change, i.e.* $I(A:B) = I(A':B')$.
4. *If* ρ_{AB} *is pure, then* $I(A:B) = 2S(A) = 2S(B)$.
5. *Monotonicity. For any tripartite quantum state* ρ_{ABC}, *it holds that* $I(A : BC) \leq I(A:B)$.

Theorem 2.4.10. *Given a bipartite quantum system* ρ_{AB}, *the correlations between observables* M_A, M_B *are upper bounded by the quantum mutual information between subsystems* A *and* B. *Formally,*

$$I(A:B) \geq \frac{C^2(M_A, M_B)}{2\|M_A\|^2\|M_B\|^2}, \tag{2.39}$$

where $C(M_A, C_B) = \langle M_A \otimes M_B\rangle - \langle M_A\rangle\langle M_B\rangle$ *is the correlation function of* M_A, M_B.

Theorem 2.4.10 [14] provides an upper bound for the total correlations between two local observables M_A and M_B. A crucial clarification is needed: despite its prevalence in protocols for quantifying transmitted information, quantum mutual information remains a measure of correlation, not causation.

Another fundamental question in quantum information theory concerns the influence of prior knowledge about a third subsystem C on the quantum mutual information between A and B. To formalize this, we introduce the concept of *conditional quantum mutual information*.

Definition 2.4.11 (Conditional Quantum Mutual Information)**.** *Given a tripartite quantum state* $\rho_{ABC} \in D(\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$, *we define the conditional quantum mutual information between* A *and* B *(conditioned by* C*) through the follwoing way:*

$$I(A:B|C) := S(A|C) + S(B|C) - S(AB|C) \tag{2.40}$$
$$= S(AC) + S(BC) - S(ABC) - S(C). \tag{2.41}$$

Proposition 2.4.12 (Properites of Conditional Mutual Information)**.**

1. *Nonnegativity.* $I(A:B|C) \leq 0$. *The nonegativity of conditional quantum information is equivalent to the strong subadditivity of entropies of subsystems.*

2. Upper bound (the dimension bound). For any tripartite quantum state $\rho_{ABC} \in D(\mathcal{H}_A \otimes \mathcal{H}_B) \otimes \mathcal{H}_C$, it holds that

$$I(A:B|C) \leq 2\log[\min\{dim(\mathcal{H}_A), dim(\mathcal{H}_B)\}]. \tag{2.42}$$

Theorem 2.4.13 (Carlen-Lieb Extension of Strong Subadditivity [15])**.** *It holds for any tripartite quantum state ρ_{ABC} that*

$$I(A:B|C) \geq 2\max\{-S(B|A), -S(A|B), 0\}. \tag{2.43}$$

Definition 2.4.14 (Quantum Relative Entropy)**.** *Given two states $\rho, \sigma \in D(\mathcal{H})$, the relative entropy of ρ with respect to σ is defined by*

$$S(\rho||\sigma) := \begin{cases} tr(\rho\log\rho) - tr(\rho\log\sigma), & if\ ker\ \sigma \subseteq\ ker\ \rho; \\ \infty, & otherwise. \end{cases} \tag{2.44}$$

Quantum relative entropy is a kind of pseudodistance in the state space, which doesn't fulfill all the properties of a distance function, but still provides a comparison quantity between two states. Becaus of its simplicity in calculation, there's a wide range of applications of quantum relative entropy. We will employ relative entropy to generalize the definition of quantum discord of multipartite systems in Chapter 3.

2.5 Holevo Bound

Definition 2.5.1 (Holevo Quantity)**.** *The Holevo quantity of an ensemble $\mathcal{E} = \{p_x, \rho_{Q,x}\}$ is defined as*

$$\chi(\mathcal{E}) := I(X:Q) = S(\sum_x p_x\rho_{Q,x}) - \sum_x p_x S(\rho_{Q,x}), \tag{2.45}$$

where the mutual information is computed in the cq-state $\rho_{XQ} = \sum_x p_x |x\rangle\langle x| \otimes \rho_{Q,x}$.

Proposition 2.5.2. *For any ensemble $\mathcal{E} = \{p_x, \rho_{Q,x}\}$ of $\rho_{Q,x} \in D(\mathcal{H}_{XQ})$, the follwoing holds*

$$0 \leq \chi(\mathcal{E}) \leq S(\sum_x p_x\rho_x) \leq \log\dim\mathcal{H}_Q. \tag{2.46}$$

A natural question that arises is: *what is the maximum amount of information Bob can obtain using an optimal measurement strategy?* To quantify thesis, we employ the notation of *accessible information*, denoted $I_{acc}(\mathcal{E})$, which measures the maximum mutual information Bob can gain by performing an optimal POVM μ on his received quantum system Q [4, 9]. It is defined as

$$I_{acc}(\mathcal{E}) := \max_{\mu} I(X : Z). \tag{2.47}$$

Theorem 2.5.3 (Holevo Bound)**.** *$I(X : Z) \leq I(X : Q) = \chi(\mathcal{E})$ for any ensemble $\{p_x, \rho_x\}$ and POVM μ.*

2.6 Entanglement and Quantum Discord

In this section, we first summarize the family of the measures of quantum entanglement for bipartite systems.

Table 2.1: Measures of entanglement for a bipartite state ρ_{AB}.

Name	Definition	State	Criterion of separability
Entanglement Entropy	$S_E := S(A) = S(B)$	pure	$S(A) = 0$
Linear Entropy	$S_l(\rho) := \frac{d}{d-1}(1 - tr(\rho^2))$	pure	$S_l(\rho_{AB}) = 0$
Concurrence	$C(\rho) := \max\{0, 2\lambda_{\max}(\hat{\rho} - \mathrm{tr}\hat{\rho})\}$	mixed	$C(\rho_{AB}) = 0$
Squashed Entanglement	$E_{sq}(\rho_{AB}) := \frac{1}{2}\inf_{\rho_{ABC}} I(A : B\|C)$	mixed	$E_{sq}(\rho_{AB}) = 0$
Entanglement of Formation	$E_f(\rho) := \min_{\{p_i, \|\psi_i\rangle\}} \sum_i p_i E(\|\psi_i\rangle)$	mixed	$E_f(\rho_{AB}) = 0$

Definition 2.6.1 (Entanglement Entropy[4])**.** *If ρ_{AB} is pure, then $S(A) = S(B)$ is called the entanglement entropy.*

The entanglement entropy quantifies the number of Bell pairs needed to create the entangled state via LOCC in the *asymptotic limit*. This procedure is known as *entanglement dilution*.

Definition 2.6.2 (Concurrence)**.** *The concurrence of a quantum state ρ is defined by*

$$C(\rho) := \max\{0, 2\lambda_{\max}(\hat{\rho} - \mathrm{tr}\hat{\rho})\}. \tag{2.48}$$

$\lambda_{\max}$ is the maximum eigenvalue of $\hat{\rho} = \sqrt{\sqrt{\rho}\tilde{\rho}\sqrt{\rho}}$ with $\tilde{\rho}$ denoting the complex conjugate of ρ.

The *squashed entanglement*, an intrinsic measure of entanglement for bipartite quantum systems, is defined using the conditional quantum mutual information [16, 17].

Definition 2.6.3 (Squashed Entanglement)**.** *The squashed entanglement of a bipartite quantum state ρ_{AB}, is defined by*

$$E_{sq}(\rho_{AB}) := \frac{1}{2} \inf_{\rho_{ABC}} I(A:B|C), \tag{2.49}$$

where the infimum is over all extensions ρ_{ABC} s.t. ρ_{AB} is the reduced state of ρ_{ABC}, i.e. $\rho_{AB} = tr_C(\rho_{ABC})$.

Definition 2.6.4 (Entanglement of Formation)**.** *If a quantum state ρ admits a decomposition into pure states:*

$$\rho = \sum_i p_i \left|\psi_i\right\rangle \left\langle\psi_i\right|, \tag{2.50}$$

then the entanglement of formation of ρ is defined as the minimum (normalized) average entanglement taken all such pure-state ensembles:

$$E_f(\rho) := \min_{\{p_i, |\psi_i\rangle\}} \sum_i p_i E(|\psi_i\rangle). \tag{2.51}$$

For any mixed bipartite state with no more than two nonzero eigenvalues, S. Hill and W. K. Wootters [18, 19] proved a functional relation between the entanglement of formation and the concurrence.

Theorem 2.6.5. *Let ρ be any density matrix of a bipartite quantum system having no more than two nonzero eigenvalues. Then the entanglement of formation of ρ is the following function of concurrence of ρ:*

$$E_f(\rho) = \mathcal{E}(C(\rho)), \tag{2.52}$$

where the function $\mathcal{E}(x)$ is given by

$$\mathcal{E} := H(\frac{1}{2} + \frac{1}{2}\sqrt{1-x^2}), \tag{2.53}$$

with $H(\cdot)$ the entropy function.

There exist other measures of bipartite entanglement—such as the distillable entanglement—which are useful in different operational contexts. Quantifying multipartite entanglement with a single measure is substantially more difficult. Indeed, detecting and characterizing global entanglement remains a challenging problem in quantum information theory, with many open questions still unresolved.

In the following, we turn to quantum discord, which characterizes quantum correlations beyond entanglement.

Tracing back to the definition of quantum mutual information, we can identify two natural interpretations. In the following, we illustrate these interpretations for a bipartite quantum system AB:

- $I(AB) = S(\rho_{AB}||\rho_A \otimes \rho_B) = S(A) + S(B) - S(AB)$ that quantifies the statistical distance between the joint state and the product of its marginals (independence assumption), representing their *total* correlations.
- $J_A(\rho_{AB}) = S(B) - S(\rho_{B|A})$ that is the mutual information left after a prior complete knowledge on A (in practice through state tomography).

Defining a quantum version of J is not straightforward. The primary difficulty lies in interpreting the prior complete knowledge of subsystem A. In practice, we generally do not know the exact density matrix of a quantum system; we can only gain partial information about it through measurement. Therefore, a useful and natural approach is to extend the conditional entropy $S(B|A)$ as the linear combination of the entropy of B conditioned by a *projective measurement* on A. This gives

$$S(B|\{\Pi_i^A\}) = \sum_i p_i S(\rho_{B|\Pi_i^A}), \tag{2.54}$$

where $p_i = Tr[(\Pi_i^A \otimes \mathbb{I}^B)\rho_{AB}]$ is the probability for obtaining the ith outcome and $\rho_i^A = Tr_B[(\Pi_i^A \otimes \mathbb{I}^B)\rho_{AB}(\Pi_i^A \otimes \mathbb{I}^B)/p_i]$ is the correponding post-measurement state of system A. (Note that the state of B after measurement Π_i^A is $\rho_{B|\Pi_i^A} = tr_A(\Pi_i^A \otimes \mathbb{I}^B)\rho_{AB}(\Pi_i^A \otimes \mathbb{I}^B)/p_i^A$, and $S(\rho_{B|\Pi_i^A}) \equiv S(B|\rho_{iA})$) [20, 21].

So the expression of J can now be extended as

$$J_{\{\Pi_i^A\}}(\rho_{AB}) := S(\rho_B) - S(B|\{\Pi_i^A\}). \tag{2.55}$$

$J_A(\rho_{AB})$ depends on both ρ_{AB} and the measurement $\{\Pi_i^A\}_{i=0}^{d_A-1}$. Usually, we are interested in the *maximum* over the set of all possible pojective measrements $\{\Pi_i^A\}$ to keep the value of J measurement independent:

$$\max_{\{\Pi_i^A\}\in PM(\rho_A)} J_{\{\Pi_i^A\}}(\rho_{AB}) = S(\rho_{AB}) - \min_{\{\Pi_i^A\}\in PM(\rho_A)} S(\rho_B|\{\Pi_i^A\}), \tag{2.56}$$

where $PM(\rho_A)$ denotes the set of all possible projective measurements on system A.

We can thus define the *quantum discord* to quantify the discrepancy between these two nonequivalent expressions of the quantum mutual information.

Definition 2.6.6 (Quantum Discord [20]). *For a bipartite quantum system ρ_{AB}, its quantum discord $\mathcal{D}_A$ w.r.t. subsystem A is defined as*

$$\begin{aligned}\mathcal{D}_A &:= I(\rho_{AB}) - \max_{\{\Pi_i^A\}} J_{\{\Pi_i^A\}}(\rho_{AB}) \\ &= \min_{\{\Pi_i^A\}} [I(\rho_{AB}) - J_{\{\Pi_i^A\}}(\rho_{AB})] \\ &= S(\rho_A) - S(\rho_{AB}) + \min_{\{\Pi_i^A\}} S(\rho_B|\{\Pi_A^i\}) \\ &= \min_{\{\Pi_i^A\}} S(\rho_B|\{\Pi_A^i\}) - S(\rho_B|\rho_A). \end{aligned} \tag{2.57}$$

Notations

- Quantum discord is *asymmetric*, meaning that $\mathcal{D}_A$ can differ from $\mathcal{D}_B$.
- In practice, minimizing the difference between $I(\rho_{AB})$ and $J_{\Pi_i^A}(\rho_{AB})$ is challenging because it requires optimizing over projective measurements defined by a specific reference basis. This implies that the chosen reference basis must be specified when discussing quantum discord.
- Zero discord states are equivalent to *pure classical correlated states* whose joint statistics can be completely explained by a local hidden variable model.

Definition 2.6.7 (Discord Measure with Pseudodistance). *The coherence measure $D^\delta(\rho)$ with a pseudodistance δ of a state ρ to the zero-discord state set $\mathcal{C}$ is defined as*

$$D^\delta(\rho) := \min_{\sigma \in \mathcal{C}} \delta(\rho, \delta). \tag{2.58}$$

The relative entropy $S(\rho||\sigma)$ is one of the most widely used pseudo-distances in quantum information theory. The coherence measure, defined via this pseudo-distance, is given by $C(\rho) = S(\rho||\Phi^i(\rho))$, where $\Phi^i(\rho) = \sum_i p_i |i\rangle \langle i|$ and $p_i = \langle i| \rho |i\rangle$ dephases the state in the reference basis $\{|i\rangle\}$. By analogy, we define the discord measure in terms of relative entropy as $D(\rho) = \min_{\sigma \in \mathcal{C}} S(\rho||\sigma)$, where $\mathcal{C}$ denotes the set of classical states. Thus, the quantity $D^\delta(\rho)$ follows the same formal definition as $C^\delta(\rho)$.

Theorem 2.6.8 (Entanglement Monogamy [22, 23]). *For a pure quantum state ρ_{ABC}, the monogamy of entanglement holds in the form of*

$$E_f(\rho_{AB}) + J(A : \check{C}) = S(A), \tag{2.59}$$

where $E_f(\rho_{AB})$ is the entanglement of formation of AB, and $J(A : \check{C})$ is the maximum Holevo quantity corresponding to the projective measurement on C.

2.7 Quantum Channels Theory

In classical communication, a channel is a function that maps one random variable to another. Analogously, in quantum information theory, a quantum channel is a superoperator that transforms one density matrix into another. Note that in both classical and quantum contexts, the dimensions of input and output states are not necessarily equal.

Definition 2.7.1 (Quantum Channel)**.** *A quantum channel from Hilbert $L(\mathcal{H}_A)$ to $L(\mathcal{H}_B)$ is a CPTP linear superoperator*

$$\Phi_{A\to B} : L(\mathcal{H}_A) \to L(\mathcal{H}_B), \tag{2.60}$$

$$\rho_A \mapsto \rho_B. \tag{2.61}$$

Definition 2.7.2 (Choi Operator)**.** *The Choi operator associated with a superoperator $\Phi_{A\to B}$ is defined by*

$$J^{\Phi}_{AB} := \sum_{i,j} |i\rangle \langle j| \otimes \Phi_{A\to B}(|x\rangle \langle y|) \in L(\mathcal{H}_A, \mathcal{H}_B), \tag{2.62}$$

where $\{|i\rangle\}$ is an arbitrary orthonormal basis of $\mathcal{H}_A$.

A quantum channel is completely characterized by its Choi operator. This leads to a one-to-one correspondence between quantum channels and Choi states, a concept generally referred to as *Channel-State Duality*.

Theorem 2.7.3 (Choi-Jamiolkowski Isomorphism)**.** *The following map is an isomorphism,*

$$L(L(\mathcal{H}_A), L(\mathcal{H}_B)) \to L(\mathcal{H}_A \otimes \mathcal{H}_B), \Phi_{A\to B} \mapsto J^{\Phi}_{AB}, \tag{2.63}$$

whose inverse is

$$\Phi_{A\to B}[M_A] = tr_A[(M_A^T \otimes I_B) J^{\Phi}_{AB}], \forall M_A \in L(\mathcal{H}_A). \tag{2.64}$$

Theorem 2.7.4. *When a superoperator $\Phi_{A\to B}$ is complete positive, its Choi operator J^{Φ}_{AB} is positive semidefinite.*
$\Phi_{AB}[M],\ \forall M \in L(\mathcal{H}_A)$ can be represented by

$$\Phi_{A\to B}[M] = \sum_{i=1}^{r} X_i M X_i^{\dagger}, \tag{2.65}$$

$X_1, \ldots, X_r \in L(\mathcal{H}_A, \mathcal{H}_B)$. This representation is called the Kraus representation.

The preceding discussion provides a mathematical description of quantum channels. Crucially, these representations correspond to physical processes, extending beyond mere formalism. Furthermore, a complete description of a quantum channel is essential for modeling quantum causal structures. This is because the functional relation between two nodes in such a structure is equivalent to a Choi matrix, which ultimately governs the probabilities of measurement outcomes at quantum nodes.

Chapter 3

Measure of One-way Quantum Information

One might be interested in identifying the quantum counterpart of classical transfer entropy. However, due to the no-cloning theorem, a generic quantum state cannot be copied without disturbing the original. A natural way to "capture" the past state of a quantum system is to introduce an ancilla qubit that becomes entangled with the original qubit. This approach effectively maps the temporal correlations of a single qubit at different times onto the spatial entanglement between two qubits.

3.1 Quantum Transfer Entropy

Formally, we introduce the following protocol to define quantum transfer entropy [24], as depicted in Figure 3.1.

Protocol 3.1.1.

- *Initial state* $\rho_A^{in} \otimes \rho_B^{in}$ *with two alcilla qubits* $\rho_{A'}^{in} = \rho_{B'}^{in} = |0\rangle\langle 0|$, *where* $\rho_A^{in} = \sum_{ik} \rho_{ik} |i\rangle\langle k|$ *and* $\rho_B^{in} = \sum_{jl} \sigma_{jl} |j\rangle\langle l|$.
- *STEP 1 Apply CNOT on* AA' *and* BB' *respectively, let* A *and* B *control qubits and* A' *and* B' *target qubits. We obtain* $\rho^1_{A'ABB'} = \rho^1_{A'A} \otimes \rho^1_{BB'}$ *with* $\rho^1_{A'A} = \sum_{ik} \rho_{ik} |ii\rangle\langle kk|_{A'A}$ *and* $\rho^1_{BB'} = \sum_{jl} \sigma_{jl} |jj\rangle\langle ll|_{BB'}$.
- *STEP 2 System* AB *evolves with the unknown channel* U_{AB}, *after the evolution the quantum state is*

$$\rho^2_{A'ABB'} = (I_{A'} \otimes U_{AB} \otimes I_{B'}) \rho^1_{A'ABB'} (I_{A'} \otimes U^\dagger_{AB} \otimes I_{B'}) \tag{3.1}$$

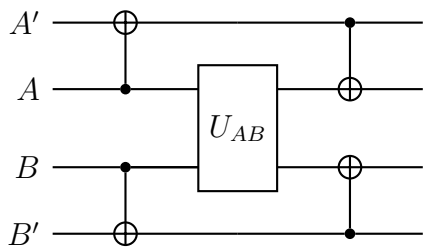


Figure 3.1: Protocol for defining quantum transfer entropy.

- *STEP 3 Apply $CNOT_{A'A}$ and $CNOT_{B'B}$ independently, the final state of the global system is*

$$\begin{aligned}\rho^f_{A'ABB'} &= \rho^3_{A'ABB'} \\ &= (CNOT_{A'A} \otimes CNOT_{B'B})\rho^2_{A'ABB'}(CNOT_{A'A} \otimes CNOT_{B'B}). \quad (3.2)\end{aligned}$$

Definition 3.1.2 (Quantum Transfer Entropy)**.** *Let the Protocol 3.1.1 yields the final global state $\rho^f_{A'B'AB}$, the quantum transfer entropy (QTE), denoted by $\mathcal{C}(A \to B)$ is defined as:*

$$\begin{aligned}\mathcal{C}(A \to B) &= I(B : A'A|B') && (3.3)\\ &= S(B|B') + S(A'A|B') - S(A'AB|B') && (3.4)\\ &= S(BB') + S(A'AB') - S(A'ABB') - S(B') && (3.5)\\ &\geq S(BB') + S(A'AB') - (S(A'AB') + S(BB') - S(B')) - S(B')\\ &= 0.\end{aligned}$$

Given an unknow quantum channel U_{AB}, one can detect the one-way information flow between two nodes connected by the channel. The quantum transfer entropy (QTE), which is fully determined by the output state of Protocol 3.1.1, can be regarded as a characterization of U without performing the full channel tomography.

A nonzero QTE indicates a directional causal influence between the two quantum nodes, and its magnitude quantifies the strength of this causal relation. It's important to note that QTE is basis-dependent. This arises from the fact that the roles of the control and target qubits may be interchanged under a change to a maximally biased basis. For example, the gate $C^{0,1\to 0,1}_{A\to B}$ in computational basis $\{|0\rangle, |1\rangle\}$ is equivalent to $C^{+,-\to +,-}_{A\to B}$ in the Hardmard basis $\{|+\rangle, |-\rangle\}$. One may argue that the generation of quantum entanglement is a symmetric process; however, it remains essential to examine the behavior of information flow in a specified basis. We also remark that, as in the classical case, the causal structure

(DAG) associated with a given set of random variables is *not* unique. Nevertheless, the causal structures obtained under different choices of basis must be isomorphic.

3.2 Quantum-Classical Information Splitting of Controlled Rotation Channels

Lemma 3.2.1. *In the computational basis, the value of* $\mathcal{C}(A \to B)$ *for a CNOT channel* ***only*** *depends on the input state of system A, and* $\mathcal{C}(A, B) = I(A' : B)$.

Proof. The input quantum states of system A and B are $\rho_A^{in} = \sum_{ik} \rho_{ik} |i\rangle \langle k|$ and $\rho_B^{in} = \sum_{jl} \sigma_{jl} |j\rangle \langle l|$.

After step 1, the global state is

$$\begin{aligned} \rho^1_{A'ABB'} = & \sum_{jkjl} \rho_{ik}\sigma_{jl} |ii\rangle \langle kk|_{A'A} \otimes |jj\rangle \langle ll|_{BB'} \\ & \sum_{ikjl} \rho_{ik}\sigma_{jl} |iijj\rangle \langle kkll|_{A'ABB'} . \end{aligned} \tag{3.6}$$

If in step 2 $U_{AB} = CNOT_{AB}$, then after the evolution of the quantum channel and step 3, we have

$$\begin{aligned} \rho^1_{A'ABB'} \overset{step2}{\to} \rho^2_{A'ABB'} &= \sum_{ikjl} \rho_{ik}\sigma_{jl} |i, i, i \oplus j, j\rangle \langle k, k, k \oplus l, l| \\ \overset{step3}{\to} \rho^f_{A'ABB'} &= \sum_{ikjl} \rho_{ik}\sigma_{jl} |i0ij\rangle \langle k0kl| \\ &= \sum_{ik} \rho_{ik} |ii\rangle \langle kk|_{A'B} \otimes |0\rangle \langle 0|_A \otimes \sum_{jl} \sigma_{jl} |j\rangle \langle l| . \end{aligned} \tag{3.7}$$

The final state of $A'ABB'$ is the *product state* of subsystems $A'B$, A and B', so

$$\mathcal{C}(A \to B) = S(B) + S(A') - S(A'B) = I(A' : B) \tag{3.8}$$

Moreover $\rho_{A'B} = \sum_{ik} \rho_{ik} |ii\rangle \langle kk|$, $\rho_B = Tr_{A'}\rho_{A'B} = \sum_i \rho_{ii} |i\rangle \langle i|$, $\rho_{A'} = Tr_B\rho_{A'B} = \sum_i \rho_{ii} |i\rangle \langle i|$. Thus the value of $I(A' : B)$ only depends on the matrix elements ρ_{ik} of initial state of A. □

Theorem 3.2.2 (Classical-Quantum Splitting of Information Flow)**.** *For a CNOT quantum channel, the transfer entropy* $\mathcal{C}(A \to B)$ *can be split into the classical and quantum information flows, i.e.*

$$\mathcal{C}(A \to B) = h + c, \tag{3.9}$$

where $h = S(\Phi(\rho_A^{in}))$ *is the entropy of the diagonalized input state of* A *and* $c = h - S(\rho_A^{in})$ *is the coherence of the input state of* A.

Proof.

$$\begin{aligned}
\mathcal{C}(A \to B) = I(A' : B) &= S(A') + S(B) - S(A'B) \\
&= H(A') - D(A') + H(B) - D(B) - H(A'B) + D(A'B) \\
&= (H(A') + H(B) - H(A'B)) - (D(A') + D(B) - D(A'B)) \\
&= h + h - h + D(AB) \\
&= h + h - S(\rho_A^{in}) \qquad (3.10) \\
&= h + c. \qquad (3.11)
\end{aligned}$$

□

Comments For a controlled gate $CR(\theta)$, if $\theta \neq \pi$, i.e. $CR(\theta) \neq CNOT$, $D(A)$ and $D(B)$ may not equal to 0, the quantum information flow is $D(A'B) - D(A) - D(B)$.

In the following, we demonstrate the equivalence of transfer entropy splitting based on the quantum discord defined by the measurement.

According to the expression of $\rho_{A'B}^f$, to define the quantum discord based on measurement, we just need measurements on A'. In the following, we consider the projective measurement $\Pi_{A'} = \{|0\rangle\langle 0|, |1\rangle\langle 1|\}$ in computational basis on A'.

After implementing $\Pi_{A'}$, the post-measurement global state is

$$\rho_{A'ABB'}^{pm} = \sum_i \rho_{ii} |ii\rangle\langle ii|_{A'B} \otimes |0\rangle\langle 0| \otimes \sum_{jl} \sigma_{jl} |j\rangle\langle l|, \qquad (3.12)$$

and $\rho_{A'B}^{pm} = \sum_i \rho_{ii} |ii\rangle\langle ii|$, $\rho_{A'}^{pm} = \rho_B^{pm} = \sum_i \rho_{ii} |i\rangle\langle i|$.

Then we can compute the quantum discord from $\rho_{A'B}^f$ and $\rho_{A'B}^{pm}$.

$$\begin{aligned}
I(A' : B) &= S(\rho_{A'}^f) + S(\rho_B^f) - S(\rho_{A'B}^f) \\
&= 2h - S(\rho_A^{in}). \qquad (3.13)
\end{aligned}$$

$$\begin{aligned}
J(A' : B) &= S(\rho_{A'}^{pm}) + S(\rho_B^{pm}) - S(\rho_{A'B}^{pm}) \\
&= h + h - h \\
&= h. \qquad (3.14)
\end{aligned}$$

Thus the quantum discord is

$$D(A' : B) = I(A' : B) - J(A' : B) = h - S(\rho_A^{in}) = C(\rho_A^{in}), \tag{3.15}$$

which coincides with the definition based on distance.

The simulation results, as shown in Figure 3.2 shows that a CNOT channel transfers the full information, i.e. both the incoherence and coherence of the controled qubit A to the target qubit B. Particularly, a pure control qubit possesses only the relative enrtopy of coherence and a classical control qubit possesses only the entropy of the diagonal entries.

More generally, there exists the linear relation between the input information and the transfer entropy for the controlled X-gates.

To examine the linear relationship between the input information and the resulting information flow, we performed a linear regression analysis. The results are summarized in Tables 3.1-3.3. We note that, except for the CNOT channel, the transfer entropy is not an exact functional dependence on the input information for the other controlled gates; rather, the two quantities exhibit a strong linear correlation.

Comments

- We did not minimize the quantum discord in the preceding computations, which implies that the decomposition of the QTE remains basis-dependent. This observation is consistent with our intuition: the reconstruction of an unknown quantum channel must be described with respect to a chosen basis—namely, the preferred reference frame of the observer (or observers).

- One may argue that if two observers choose different reference bases, they may arrive at different interpretations of the information flow. Hence, a prior agreement on the reference basis among observers must be ensured. This can be accomplished within the LOCC framework, where the measurement setup can be coordinated through classical communication.

3.3 Bounds for Quantum Transfer Entropy

The quantum transfer entropy is defined in terms of the *quantum conditional mutual information*, which is associated with several subsystem-dependent bounds. In this

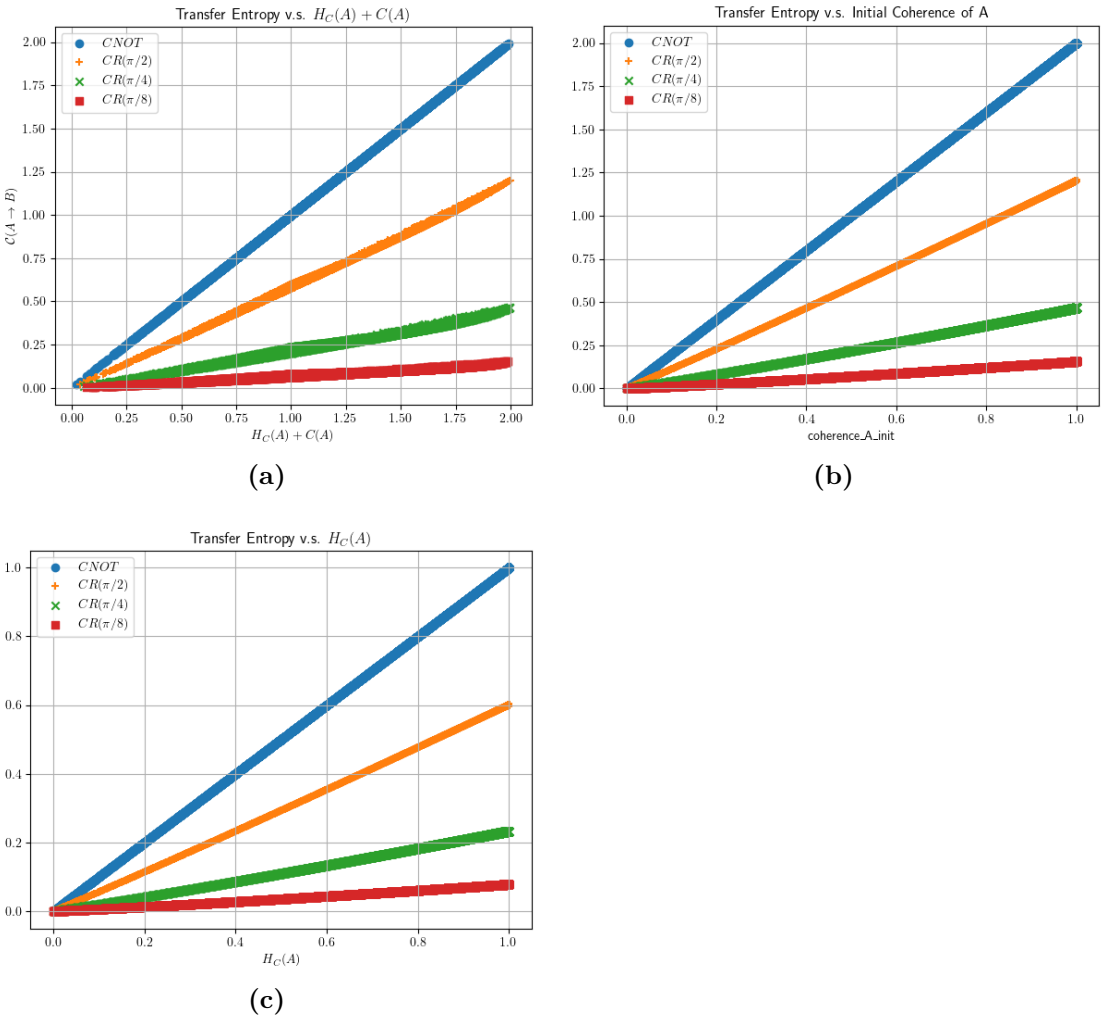


Figure 3.2: Information flow of arbitrary input states, pure states, and classical states of A.

Table 3.1: Linear Regression Results of Arbitrary input states of A

Channel	Slope	r value	p value
$CNOT$	1.0	1.0	0
$CR\left(\frac{\pi}{2}\right)$	0.588	0.9993	0
$CR\left(\frac{\pi}{4}\right)$	0.217	0.9944	0
$CR\left(\frac{\pi}{8}\right)$	0.067	0.9852	0

Table 3.2: Linear Regression Results of pure input states of A

Channel	Slope	r value	p value
$CNOT$	2.0	1.0	0
$CR\left(\frac{\pi}{2}\right)$	1.216	0.99995	0
$CR\left(\frac{\pi}{4}\right)$	0.481	0.99966	0
$CR\left(\frac{\pi}{8}\right)$	0.163	0.99921	0

work, we adopt the result presented in [17].

Theorem 3.3.1. *For every tripartite finite-dimensional state ρ_{ABE},*

$$I(A:B|E) \geq \frac{1}{2\ln 2}\|\rho_{AB} - S_{A:B}\|^2_{1-LOCC}, \tag{3.16}$$

where $\|\rho_{AB} - S_{A:B}\| := \min_{\sigma\in S_{A;B}} \|\rho_{AB} - \sigma\|$ is the distance between ρ_{AB} and the set $S_{A:B}$ of separable states on $A:B$ [17].

The computation of $\|\rho_{AB} - S_{A:B}\|$ is not straightforward. In fact, evaluating this trace distance requires solving an optimization problem over positive semidefinite operators, which naturally takes the form of a semidefinite program (SDP). Consequently, obtaining the exact value typically involves numerical optimization rather than a closed-form analytical expression.

Theorem 3.3.2. *For a tripartite quantum system ρ_{ABC}, there exists an explicit lower bound of conditional mutual information $I(A:B|C)$ represented only in terms of ρ_A and ρ_B*

$$I(A:B|C) \geq \frac{C^2(M_A, M_B)}{2\|M_A\|^2\|M_B\|^2} + f(d,\gamma). \tag{3.17}$$

Here M_A and M_B are local observables on A and B respectively. $f(d,\gamma)$ is the function of the dimension $d = dim\mathcal{H}_{AB}$ and the purity γ of ρ_{AB}, specifically.

Table 3.3: Linear Regression Results of classical input states of A

Channel	Slope	r value	p value
$CNOT$	1.0	1.0	0
$CR\left(\frac{\pi}{2}\right)$	0.608	0.99995	0
$CR\left(\frac{\pi}{4}\right)$	0.240	0.99965	0
$CR\left(\frac{\pi}{8}\right)$	0.082	0.99921	0

$$f(d,\gamma) = \left[\left(1-\frac{1}{d}-\sqrt{(1-\frac{1}{d})(\gamma-\frac{1}{d})}\right)\log\left(1-\frac{1}{d}-\sqrt{(1-\frac{1}{d})(\gamma-\frac{1}{d})}\right)\right.$$
$$\left.+\left(\frac{1}{d}+\sqrt{(1-\frac{1}{d})(\gamma-\frac{1}{d})}\right)\log\left(\frac{1}{d}+\sqrt{(1-\frac{1}{d})(\gamma-\frac{1}{d})}\right)\right] \tag{3.18}$$

3.4 Causality Interpretation

Is causality epistemological or realistic? We leave this fundamental question open, as resolving it lies beyond the scope of this work. Instead, we adopt an operational perspective to examine the definition of QTE and the corresponding mechanisms of information transmission. This approach allows us to highlight how quantum transfer entropy differs from its classical counterpart in both interpretation and behavior.

In analogy with the classical case, a quantum causal structure can also be represented by a directed acyclic graph (DAG), which specifies the allowed directions of influence among the relevant quantum systems. In this framework, each node of the DAG is associated with a density matrix describing the state of the corresponding quantum subsystem. When a measurement is performed on a given subsystem, the quantum state at that node gives rise to a classical random variable, thereby linking the quantum causal structure to observable statistical data.

Definition 3.4.1 (Quatnum Causal Structure [8, 7, 25, 26])**.** *A quantum causal structure C^Q, is a DAG where each observed node has a corresponding (classical) random variable, and each unobserved node is associated with a quantum system.*

Considering two correlated binary time series X_t and Y_t, where the state of Y_t is a deterministic function of the previous states X_{t-1} and Y_{t-1}. In this setting, the classical transfer entropy from X_{t-1} to Y_t is upper bounded by $H(X_{t-1})$, formally

$$\mathcal{TE}(X_{t-1}\to Y_t) := I(X_{t-1}:Y_t|Y_{t-1}) \le H(X_{t-1}) \le 1. \tag{3.19}$$

Theorem 3.2.2 provides a novel measure beyond the classical causal model. The quantum entanglement between two interacting qubits can violate the upper bound of the classical transfer entropy. It's crucial to point out that this violation depends on the operations of observers. Again, the choice of reference basis determines how an observer "reads" or "represents" the quantum states, in turn decides the quantities of classical-quantum information splitting.

We use the example of input state $|+0\rangle$ to illustrate how the choice of reference bases influences the value of $\mathcal{C}(A \to B)$:

1. The entire protocol is implemented in computational basis $\{|0\rangle, |1\rangle\}$. One can directly verify that $\mathcal{C}(A \to B) = 2$ and $\mathcal{C}(B \to A) = 0$.

2. The entire protocol is implemented in Hardmard basis $\{|+\rangle, |-\rangle\}$. The value of quantum transfer entropy of different directions is inverse, i.e. $\mathcal{C}(A \to B) = 0$ and $\mathcal{C}(B \to A) = 2$.

3. System $A'A$ is operated and measured in the Hardmard basis, while system BB' is in the computational basis. One can calculate that $\mathcal{C}(A \to B) = \mathcal{C}(B \to A)$.

This example illustrates that the interpretation of quantum transfer entropy can vary depending on the choice of reference bases. In other words, what an observer infers as "information flow" is influenced by how the underlying quantum states are represented. The basis dependence revealed here suggests that such an interpretation is not absolute but may depend on the operational framework adopted by the observers.

Chapter 4

Quantum Causal Discovery Algorithm

Although many quantum causal discovery and quantum causal inference algorithms have emerged recently [27, 28, 29, 30], there's stll a lack of model-free scenario with the prespective of quantifying the information flows in a quantum network.

In this chapter, we introduce a causal discovery algorithm that is formulated on the basis of the quantum transfer entropy defined in the previous sections. This algorithm provides an operational procedure for inferring directional influence between quantum subsystems, extending classical causal discovery methods into the quantum domain.

4.1 Bipartite Systems

Assume that two observers, Alice and Bob, each possess a spatially separated qubit, which may in general be entangled. Alice and Bob independently perform measurements on their respective qubits in chosen local bases. The measurement outcomes are represented by two binary random variables $X, Y \in \{0, 1\}$ which obey the marginal distributions p_X and p_Y respectively. These outcomes form the observable data from which the underlying causal or informational relationships between the two quantum systems may be inferred.

We study the case that p_X and p_Y are encoded to two *pure* quantum states that serve as the input to the QTE protocol described in Protocol 3.1.1. Given an interaction channel U_{AB}, we can explicitly compute the quantum transfer entropy induced by this channel. This allows us to analyze how the chosen unitary

interaction mediates information flow between the two subsystems within the QTE framework. The procedure is outlined as following, assuming all steps are preformed in the computational basis:

1. The probabilities are encoded by pure quantum states

$$\begin{aligned} p_x \to |\psi_A^{in}\rangle &= \sum_x \sqrt{p_x}\,|x\rangle\,, \\ p_y \to |\psi_B^{in}\rangle &= \sum_y \sqrt{p_y}\,|y\rangle\,. \end{aligned} \tag{4.1}$$

   Employ ancilla qubits $|\psi_{A'}\rangle = |\psi_{B'}\rangle = |0\rangle$. The global state is then $\rho^1_{A'ABB'} = |\psi^1\rangle\langle\psi^1|$, $|\psi^1\rangle_{A'ABB'} = \sum_{xy}\sqrt{p_x p_y}\,|0xy0\rangle$.

2. Apply $C^{0,1\to 0.1}_{A\to A'}$ and $C^{0,1\to 0.1}_{B\to B'}$, obtaining $|\psi^2\rangle_{A'ABB'} = \sum_{xy}\sqrt{p_x p_y}\,|xxyy\rangle$.

3. Perform the unitary evolution U_{AB} on the system AB. The density matrix after evolution is $\rho^3_{A'ABB'} = U_{AB}\rho^2_{A'ABB'}U^\dagger_{AB}$.

4. Disentangle $A'A$, BB' with $C^{0,1\to 0.1}_{A'\to A}$ and $C^{0,1\to 0.1}_{B'\to B}$, resulting in $\rho^4_{A'ABB'}$.

5. Dephase the state $\rho^4_{A'ABB'}$, denote the final state by $\rho^f_{A'ABB'} = \Phi(\rho^4_{A'ABB'})$.

Lemma 4.1.1. *For a bipartite product state $\rho_{AB} = \rho_A \otimes \rho_B$, the dephasing on ρ_{AB} is equivalent to the local independent dephasing on the subsystems*

$$\Phi(\rho_{AB}) \equiv \Phi(\rho_A) \otimes \Phi(\rho_B). \tag{4.2}$$

Proof. Let $\{|ij\rangle\}_{i,j=0}^1$ be a basis on Hilbert space $\mathbb{C}^4$.

$$\Phi(\rho) = \sum_{ij} \langle ij|\,\rho_{AB}\,|ij\rangle\,|ij\rangle\langle ij|_{AB} \tag{4.3}$$

$$= \sum_{ij} \langle ij|\,\rho_A \otimes \rho_B\,|ij\rangle\,|ij\rangle\langle ij|_{AB} \tag{4.4}$$

$$= \sum_{ij} \left(\langle i|\,\rho_A\,|i\rangle \otimes \langle j|\,\rho_B\,|j\rangle\right)|ij\rangle\langle ij| \tag{4.5}$$

$$= \left(\sum_i \langle i|\,\rho_A\,|i\rangle\,|i\rangle\langle i|\right) \otimes \left(\sum_j \langle j|\,\rho_B\,|j\rangle\,|j\rangle\langle j|\right) \tag{4.6}$$

$$= \Phi(\rho_A) \otimes \Phi(\rho_B). \tag{4.7}$$

□

By applying Lemma 4.1.1, one can verify that if in step 3 of the protocol, the channel satisfies $U_{AB} = U_A \otimes U_B$-that is, the operation is strictly local-then the resulting density matrix factorizes into a product state across the bipartition $A'A$ and BB'

$$\rho^f_{A'ABB'} = \rho^f_{A'A} \otimes \rho^f_{BB'}. \tag{4.8}$$

Hence the quantum transfer entropy

$$\begin{aligned}
\mathcal{C}(A \to B) &= I(A'A : B|B') && (4.9)\\
&= S(A'AB') + S(BB') - S(A'ABB') - S(B') && (4.10)\\
&= S(A'A) + S(B') + S(BB') - S(A'A) - S(BB') - S(B') && (4.11)\\
&= 0. && (4.12)
\end{aligned}$$

Analogously, we also have $\mathcal{C}(B \to A) = 0$.

When taking the interaction channel to be the controlled-NOT gate $U_{AB} = CNOT_{AB}$ in step 3, we can explicitly compute the quantum transfer entropy associated with the update rule $Y_t = X_{t-1} \otimes Y_{t-1}$. From Proposition 3.2.1 we obtain the global state after step 4 takes the form of the following product state:

$$\rho^4_{A'ABB'} = \rho^4_{A'B} \otimes |0\rangle\langle 0|_A \otimes \rho^f_{B'}. \tag{4.13}$$

Therefore, the final global state after the dephasing operation is given by

$$\begin{aligned}
\rho^f_{A'ABB'} &= \Phi(\rho^4_{A'ABB'}) && (4.14)\\
&= \Phi(\rho^4_{A'B}) \otimes |0\rangle\langle 0|_A \otimes \Phi(\rho^f_{B'}) && (4.15)\\
&= \sum_x p_x |xx\rangle\langle xx|_{A'B} \otimes |0\rangle\langle 0|_A \otimes \rho^f_{B'}. && (4.16)
\end{aligned}$$

Notice $\rho^f_{A'B} = \sum_x p_x |xx\rangle\langle xx|_{A'B}$ and $\rho_{A'} = \rho_B = \sum_x p_x |x\rangle\langle x|$, the transfer entropy (from A to B) is thus

$$\begin{aligned}
\mathcal{C}(A \to B) &= I(\rho^f_{A'} : \rho^f_B) && (4.17)\\
&= S(\rho^f_{A'}) + S(\rho^f_B) - S(\rho^f_{A'B}) && (4.18)\\
&= H(X), && (4.19)
\end{aligned}$$

where $H(X) = \sum_x p_x \log p_x$ is the Shannon entropy of random variable X. And $\mathcal{C}(B \to A) = 0$.

The case is similar for the inverse $U_{AB} = CNOT_{B\to A}$:

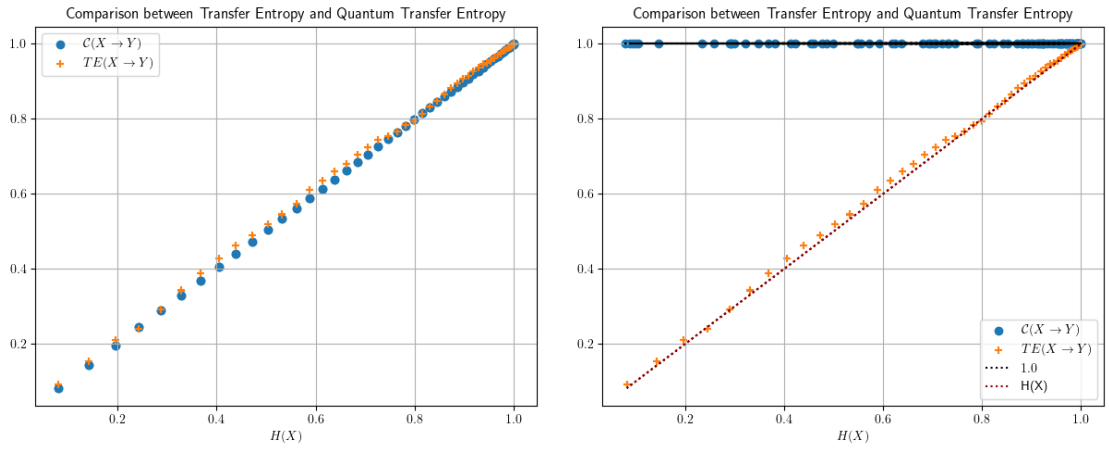
$$\begin{aligned}
\mathcal{C}(B \to A) &= H(Y), && (4.20)\\
\mathcal{C}(A \to B) &= 0. && (4.21)
\end{aligned}$$

Figure 4.1 illustrates the comparison between the classical transfer entropy and quantum transfer entropy for binary, stationary, first-order Markov process, where the update rule for the random variable Y is given by

$$Y_t = X_{t-1} \oplus Y_{t-1}. \tag{4.22}$$

As shown in Figure 4.1a, when all the input data are encoded in the computational basis, QTE $\mathcal{C}(X \to Y)$ and CTE $TE(X \to Y)$ exhibit nearly identical behavior. However, when the probability distributions p_x and p_y are instead encoded in the maximally biased bases, as illustrated in Figure 4.1b, the presence of additional quantum coherence produces a significantly more robust information-flow measure. In particular, we observe that $\mathcal{C}(A \to B)$ stabilizes at the constant value 1, effectively elevating the quantum transfer entropy for any input distribution p_x to match that of the case with maximal Shannon entropy.

This enhanced robustness, arising from coherence-assisted information flow, suggests a potentially valuable feature for future applications in quantum causal discovery, where stable and basis-sensitive indicators of directional influence are of practical importance.



(a) **(b)**

Figure 4.1: Comparison between quantum and classical transfer entropy for a binary stable 1st-order Markov time series, whose generator is $Y_t = Y_{t-1} \oplus X_{t-1}$. From left to right, the result for data encoded in the coincident basis and in the maximum biased basis ($|\psi_A^{in}\rangle$ and $|\psi_{A'}^{in}\rangle$ in the Hardmard basis, ($|\psi_B^{in}\rangle$ and $|\psi_{B'}^{in}\rangle$ in the computational basis).

Finally, we examine the causal influence exhibited by other double-qubit channels, summarized in Table 4.1. This comparison highlights how different choices of controlled operations modify the magnitude of information flow between the subsystems.

Channel	$C(A \to B)$	$C(B \to A)$
$CX(0)$	0	0
$CX(\pi/8)$	0.156	0
$CX(\pi/4)$	0.467	0
$CX(3\pi/8)$	0.834	0
$CX(\pi/2)$	1.201	0
$CX_{B\to A}(\pi/2)$	0	1.201
$I_A \otimes I_B$	0	0
$I_A \otimes X_B$	0	0

Table 4.1: The numerical results of the causal influence $C(A \to B)$, $C(B \to A)$ for several two-qubit channels. Both of the input states of A and B are $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.

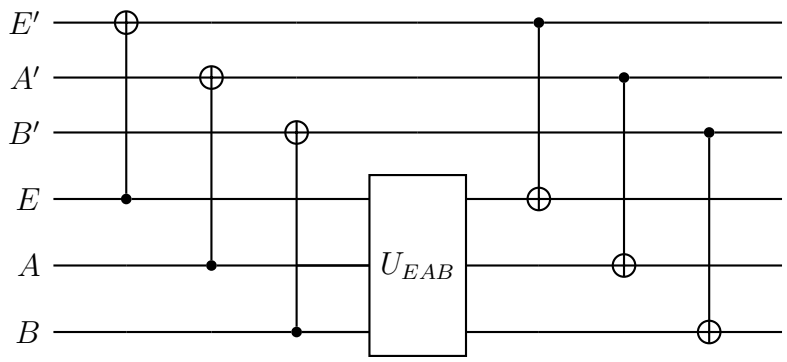


Figure 4.2: QTE protocol when considering the influence of environment E with ancilla environmental qubit E'.

4.2 Tripartite Systems

In practice, when attempting to diagnose the information flow between two quantum systems A and B, one must inevitably take into account the presence of noise arising from their interaction with an external environment E. The degree of decoherence in the joint state ρ_{AB} is directly related to the strength of the coupling between the composite system AB and its environment. Various theoretical frameworks have been developed to model the dynamics of open quantum systems, such as linear response theory for weak environmental perturbations.

In this section, we introduce an information-theoretic method based on quantum transfer entropy to quantitatively analyze the information flow between two quantum systems while explicitly incorporating environmental effects. By the purification theorem, any mixed state ρ_{AB} can be represented as the reduced state of a pure tripartite state $|\psi\rangle_{ABE}$. Motivated by this perspective, we consider a model in which the evolution of the tripartite system EAB is governed by a Toffoli gate, $CCNOT_{EA\to B}$. In this setup, the monitored qubit A and the environment E exert symmetric and equivalent control over the target qubit B.

Table 4.2: Truth Table of $CCNOT_{zx\to y}$

$z \cdot x$	y	output
0	0	0
0	1	1
1	0	1
1	1	0

Analogous to the procedure described in the previous section, independent local measurements performed on systems A, B, and E yield binary random vairables X, Y and Z, respectively. These variables are associated with the corresponding marginal probability distributions p_x, p_y and p_z. These measurement outcomes constitute the observable data from which we analyze how information is distributed and transferred among the three subsystems in the presence of environmental interactions. For simplicity, we can model the environment as a single qubit, which allows us to clearly illustrate the fundamental mechanisms by which environmental interactions influence the observed information flow.

If one can completely control the third subsystem E, the quantum transfer entropy can be defined similarly in an operational way. Employ the following protocol:

Protocol 4.2.1.

1. *We encode the distributions through pure states:*

$$p_x \to |\psi_A^{in}\rangle = \sum_x \sqrt{p_x}\,|x\rangle\,, \tag{4.23}$$

$$p_y \to |\psi_B^{in}\rangle = \sum_y \sqrt{p_y}\,|y\rangle\,, \tag{4.24}$$

$$p_z \to |\psi_E^{in}\rangle = \sum_z \sqrt{p_z}\,|z\rangle\,. \tag{4.25}$$

Each qubit is associated to an ancilla qubit at state $|0\rangle$, the initial global state is

$$\rho^1_{E'A'B'EAB} = |\psi^1\rangle\langle\psi^1|_{E'A'B'EAB}\,, \tag{4.26}$$

where $|\psi^1\rangle_{E'A'B'EAB} = \sum_{zxy}\sqrt{p_z p_x p_y}\,|000zxy\rangle$.

2. *Entangle each qubit to its own ancilla qubit, we obtain*

$$|\psi^2\rangle_{E'A'B'EAB} = \sum_{zxy}\sqrt{p_z p_x p_y}\,|zxyzxy\rangle_{E'A'B'EAB}\,. \tag{4.27}$$

3. *Apply U_{EAB}, the global state turns to*

$$|\psi^3\rangle_{E'A'B'EAB} = I_{E'A'B'}\otimes U_{EAB}\,|\psi^2\rangle_{E'A'B'EAB} \tag{4.28}$$

.

4. *Disentangle $A'A$, $B'B$ and $E'E$,*

$$|\psi^4\rangle_{E'A'B'EAB} = C^{0,1\to0,1}_{E'\to E}\otimes C^{0,1\to0,1}_{A'\to A}\otimes C^{0,1\to0,1}_{B'\to B}\,|\psi^3\rangle_{E'A'B'EAB}\,. \tag{4.29}$$

Proposition 4.2.2.

$$\mathcal{C}(A \to B|E) = \mathcal{C}(EA \to B) - \mathcal{C}(E \to B). \tag{4.30}$$

Proof.

$$\begin{aligned}
\mathcal{C}(A \to B|E) &= I(A'A : B|E'B'E) && (4.31)\\
&= S(E'A'B'EA) + S(E'B'EB) - S(E'B'E) - S(E'A'B'EAB)\\
&= [S(E'A'B'EA) + S(B'B) - S(B') - S(E'A'B'EAB)]\\
&\quad - [S(E'B'E) + S(B'B) - S(B') - S(E'B'EB)]\\
&= I(E'A'EA : B|B') - I(E'E : B|B'). && (4.32)
\end{aligned}$$

□

Proposition 4.2.3 (Lieb Lower Bound)**.** *If $U_{EAB} = CCNOT_{EA\to B}$, the following inequalities hold*

$$2\max\{0, -S^f(B|A'A), -S^f(A'A|B)\} \leq \mathcal{C}(A \to B|E), \tag{4.33}$$

$$2\max\{0, -S^f(B|E'E), -S^f(E'E|B)\} \leq \mathcal{C}(E \to B|A). \tag{4.34}$$

Proof. These inequalities hold directly form the Carlen-Lieb inequality 2.4.13.

□

Lemma 4.2.4. *For $CCNOT_{EA\to B}$ channel,*

$$\mathcal{C}(A \to E) \equiv 0, \tag{4.35}$$

$$\mathcal{C}(A \to B|E) \equiv \mathcal{C}(A \to BE). \tag{4.36}$$

Proof. If $U_{EAB} = CCNOT_{EA\to B}$, the output state in Protocol 4.2.1 is

$$\begin{aligned}
|\psi^4\rangle_{E'A'B'EAB} &= \sum_{zxy} \sqrt{p_z p_x p_y}\, |z, x, y, 0, 0, zx\rangle_{E'A'B'EAB}\\
&= \left(\sum_{zx} \sqrt{p_z p_x}\, |z, x, zx\rangle_{E'A'B}\right) \otimes \left(\sum_{y} \sqrt{p_y}\, |y\rangle_{B'}\right) \otimes |00\rangle_{EA} && (4.37)\\
&= |\psi_{E'A'B}\rangle \otimes |\psi_{B'}\rangle \otimes |\psi_E\rangle \otimes |\psi_A\rangle\,. && (4.38)
\end{aligned}$$

Thus we have

$$\begin{aligned}
\mathcal{C}(A \to E) &= I(A'A : E|E') && (4.39)\\
&= S(E'A'A) + S(E'E) - S(E') - S(E'A'EA)\\
&= S(E'A') + S(A) + S(E') + S(E)\\
&\quad - S(E') - S(E'A') - S(E) - S(A)\\
&= 0. && (4.40)
\end{aligned}$$

$$\begin{aligned}\mathcal{C}(A \to B|E) &= I(A'A : B|E'B'E) && (4.41)\\ &= S(E'A'B'EA) + S(E'B'EB)\\ &- S(E'B'E) - S(E'A'B'EAB)\\ &= S(E'A'B'A) + S(E'B'EB)\\ &- S(E'B') - 0 - S(E'A'B'EAB)\\ &= I(A'A : EB : E'B')\\ &= \mathcal{C}(A \to BE). && (4.42)\end{aligned}$$

□

Proposition 4.2.5 (Observable Lower Bound)**.** *If $U_{EAB} = CCNOT_{EA\to B}$, the observable lower bound of $\mathcal{C}(A \to B \mid E)$ reduces to*

$$\mathcal{C}(A \to B \mid E) = I(A'B \mid E') \geq \frac{C^2(M_{A'}, M_B)}{2\|M_{A'}\|^2\|M_B\|^2} + f(d,\gamma). \tag{4.43}$$

Here $M_{A'}$ and M_B are local observables on A' and B respectively. $f(d,\gamma)$ is the function of the dimension $d = dim\mathcal{H}_{A'B}$ and the purity γ of $\rho_{A'B}$, specifically.

$$\begin{aligned}f(d,\gamma) = &\left[\left(1 - \frac{1}{d} - \sqrt{(1-\frac{1}{d})(\gamma - \frac{1}{d})}\right)\log\left(1 - \frac{1}{d} - \sqrt{(1-\frac{1}{d})(\gamma - \frac{1}{d})}\right)\right.\\ &\left.+\left(\frac{1}{d} + \sqrt{(1-\frac{1}{d})(\gamma - \frac{1}{d})}\right)\log\left(\frac{1}{d} + \sqrt{(1-\frac{1}{d})(\gamma - \frac{1}{d})}\right)\right] \end{aligned} \tag{4.44}$$

Proposition 4.2.6 (Upper Bound)**.** *If $U_{EAB} = CCNOT_{EA\to B}$, the following inequalities hold*

$$\mathcal{C}(A \to B|E) \leq h_A + c_A, \tag{4.45}$$
$$\mathcal{C}(E \to B|A) \leq h_E + c_E. \tag{4.46}$$

Proof. We illustrate the proof of the first inequality.
By observabing the final state corresponding to $CCNOT_{EA\to B}$ chennel, we have

$$\begin{aligned}\mathcal{C}(A \to B|E) &= I(A'A : B|E'B'E) && (4.47)\\ &= S(E'A'B'EA) + S(E'B'EB)\\ &- S(E'B'E) - S(E'A'B'EAB)\\ &= S(E'A') + S(E'B) - S(E') && (4.48)\\ &= S(A'|E') + S(B|E') && (4.49)\\ &\leq h_A + c_A. && (4.50)\end{aligned}$$

The proof of $\mathcal{C}(E \to B|A)$ follwos the same procedure. □

However, in realistic scenarios it is impossible to fully control the environment. For this reason, we do not employ an ancilla qubit E', since in practice one cannot coherently entangle such an ancilla with the actual environment. Furthermore, to ensure that the output of the protocol becomes effectively classical, we introduce a dephasing operation. This step is operationally equivalent to performing a projective measurement on the system and serves to eliminate residual quantum coherence that would otherwise obscure the classical information extracted from the process.

To simplify the discussion, we focus on the specific channel

$$U_{EAB} = CCNOT_{EA\to B}$$

in the following.

Protocol 4.2.7 (QTE protocol considering single third party qubit and projective measurement)**.**

1. *We encode the distributions through pure states:*

$$p_x \to |\psi_A^{in}\rangle = \sum_x \sqrt{p_x}\,|x\rangle\,, \tag{4.51}$$

$$p_y \to |\psi_B^{in}\rangle = \sum_y \sqrt{p_y}\,|y\rangle\,, \tag{4.52}$$

$$p_z \to |\psi_E^{in}\rangle = \sum_z \sqrt{p_z}\,|z\rangle\,. \tag{4.53}$$

Each monitored qubit is associated to an ancilla qubit at state $|0\rangle$, the initial global state is

$$\rho^1_{A'B'EAB} = |\psi^1\rangle\langle\psi^1|_{A'B'EAB}\,, \tag{4.54}$$

where $|\psi^1\rangle_{A'B'EAB} = \sum_{zxy}\sqrt{p_z p_x p_y}\,|00zxy\rangle$.

2. *Entangle two monitored qubits to their own ancilla qubit, we obtain*

$$|\psi^2\rangle_{A'B'EAB} = \sum_{zxy}\sqrt{p_z p_x p_y}\,|xyzxy\rangle_{A'B'EAB}\,. \tag{4.55}$$

3. *Apply $U_{EAB} = CCNOT_{EA\to B}$, from the truth table of $CCNOT$ gate (Table 4.2) the global state turns to*

$$|\psi^3\rangle_{A'B'EAB} = \sum_{zxy}|x,y,z,x,zx\otimes y\rangle_{A'B'EAB}\,. \tag{4.56}$$

.

4. Disentangle $A'A$, $B'B$, resulting

$$|\psi^4\rangle_{A'B'EAB} = \sum_{zxy} \sqrt{p_z p_x p_y}\, |x, y, z, 0, zx\rangle_{A'B'EAB}\,. \tag{4.57}$$

5. Dephase $\rho^4_{A'B'EAB} = |\psi^4\rangle\langle\psi^4|_{A'B'EAB}$, we obtain the final density matrix (a classical state)

$$\begin{aligned}
\rho^f_{A'B'EAB} &= \Phi(\rho^4_{A'B'EAB}) \\
&= \sum_{zxy} p_z p_x p_y\, |x, y, z, 0, zx\rangle\langle x, y, z, 0, zx|_{A'B'EAB} \\
&= \left(\sum_{zx} p_z p_x\, |x, z, zx\rangle\langle x, z, zx|_{A'EB}\right) \otimes |0\rangle\langle 0|_A \\
&\otimes \left(\sum_{y} p_y\, |y\rangle\langle y|_{B'}\right) \\
&= \rho^f_{A'E'B} \otimes \rho^f_E \otimes \rho^f_A \otimes \rho^f_{B'}.
\end{aligned} \tag{4.58}$$

In the following, we illustrate some results of conditional quantum transfer entropy $\mathcal{C}(A \to B|E)$. These examples demonstrate how the presence of the environment modifies the effective information flow from A to B once the influence of E id completely taken into account.

Proposition 4.2.8.

$$\mathcal{C}(A \to B|E) = \alpha - h_z, \tag{4.59}$$

where $h_z = -\sum_z p_x \log p_z$ denotes the classical Shannon entropy of the environmental input state $|\psi^{in}_E\rangle$ and $\alpha = H(X'Z')$ represents the joint Shannon entropy of the distribution summarized in Table 4.3. Consequently, the conditional quantum transfer entropy reduces to the conditional Shannon entropy associated with the distribution given in Table 4.3:

$$\mathcal{C}(A \to B|E) \equiv H(X'|Z'). \tag{4.60}$$

Similarly, we obtain

$$\mathcal{C}(E \to B|A) = H(Z''|X''), \tag{4.61}$$

where the conditional Shannon entropy $H(Z'' : X'')$ is computed with respect to the probability distribution summarized in Table 4.4. This expression quantifies the effective information flow from the environment E to the target qubit B once the influence of control qubit A has been accounted for.

Table 4.3: Probability Distribution for $\mathcal{C}(A \to B|E)$

X' \ Z'	0	1
0	$p_{z=0}$	$p_{x=0}p_{z=1}$
1	0	$p_{x=1}p_{z=1}$

Table 4.4: Probability Distribution for $\mathcal{C}(E \to B|A)$

X'' \ Z''	0	1
0	$p_{x=0}$	0
1	$p_{x=1}p_{z=0}$	$p_{x=1}p_{z=1}$

Proposition 4.2.9. *For the dephased final state the following inequalities hold*

$$0 = 2\max\{0, -S^f(B|A'A), -S^f(A'A|B)\} \leq \mathcal{C}(A \to B|E) \leq h_x, \tag{4.62}$$
$$0 = 2\max\{0, -S^f(B|E'E), -S^f(E'E|B)\} \leq \mathcal{C}(E \to B|A) \leq h_z. \tag{4.63}$$

Proof. We give the proof of $\mathcal{C}(A \to B|E)$ in the following. For $\mathcal{C}(E \to B|A)$, it takes the same procedure.
According to Proposition 4.2.6, the left inequality holds.
By observing the final state of the Protocol 4.2.7, we obtain:

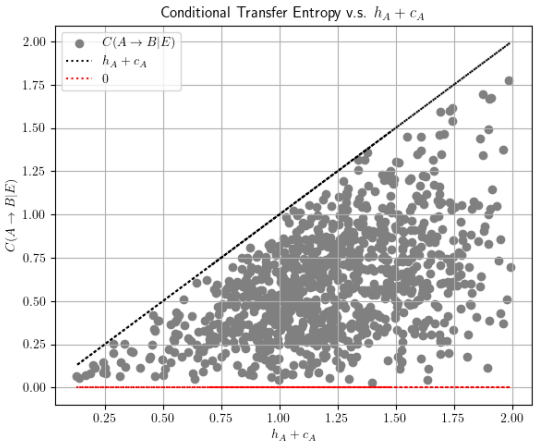
$$\begin{aligned} S^f(B|A'A) &= S(A'AB) - S(A'A) && (4.64)\\ &= S(A'B) + S(A) - S(A') - S(A) && (4.65)\\ &= S(A'B) - S(A') && (4.66)\\ &\geq 0, && (4.67)\\ S^f(A'A|B) &\geq 0. && (4.68) \end{aligned}$$

In addition, after the dephasing, only the classical information left, i.e. $h_A = 0$ and $c_A = h_x$, yielding the right inequality. □

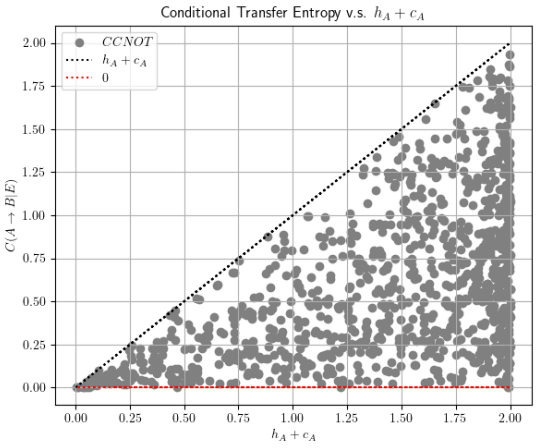
Figure 4.3 shows the range of $\mathcal{C}(A \to B|E)$ when the final diagonalization step is omitted. By comparing Figures 4.3a and 4.3b, we observe that the choice of data-encoding basis again does not alter the attainable range of the conditional transfer entropy. Furthermore, Figure 4.3c indicates that when the input state of the environment E is fixed to $|1\rangle$, the channel $CCNOT_{EA\to B}$ becomes operationally equivalent to $CNOT_{A\to B}$. In this situation, $\mathcal{C}(A \to B|E)$ saturates its

upper bound, as the environmental subsystem no longer introduces additional influence or ambiguity in the flow of information from A to B.

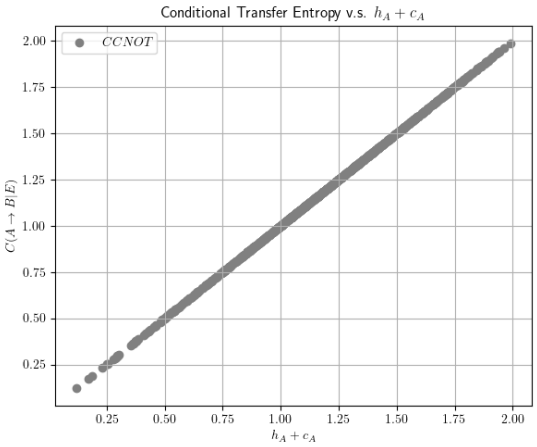
The numerical results corresponding to Proposition 4.2.9 are presented in Figure 4.4, where the Leib lower bound of the conditional QTE is observed to be exactly zero.



(a)

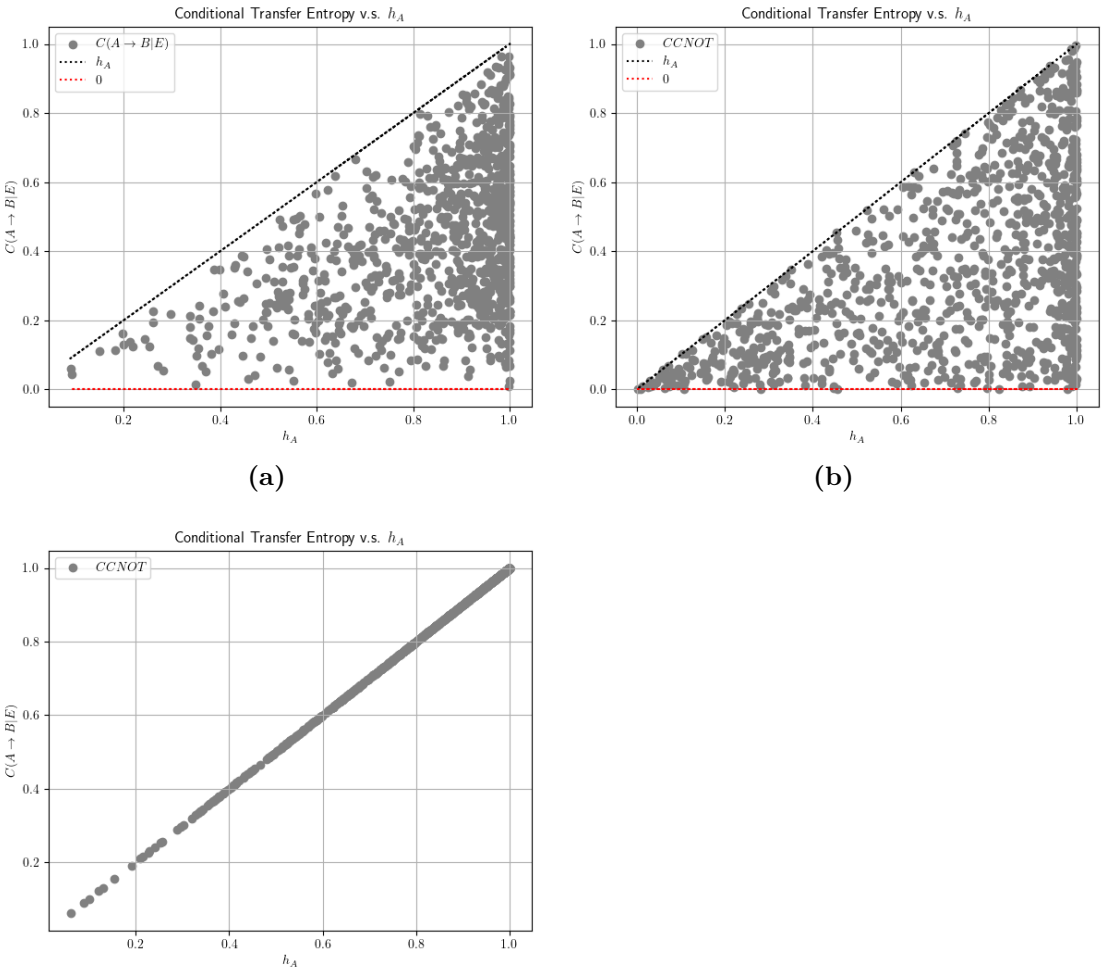


(b)



(c)

Figure 4.3: $\mathcal{C}(A \to B|E)$ without diagonalization. 4.3a The result of random input density matrices. 4.3b The result of pure input states. 4.3c The results of random input ρ_A^{in} and $\rho_E^{in} = |1\rangle\langle 1|$.



(a) **(b)**

(c)

Figure 4.4: $\mathcal{C}(A \to B|E)$ with diagonalization. 4.4a The result of random input density matrices. 4.4b The result of pure input states. 4.4c The results of random input ρ_A^{in} and $\rho_E^{in} = |1\rangle\langle 1|$.

In contrast to the behavior of $CNOT_{A\to B}$ acting on a closed bipartite quantum system, acting on a closed bipartite quantum system, once the strong influence of the environment is taken into account, the information flow from A to B is no longer an explicit linear (or even quasi-linear) function of the quantum information initially encoded in subsystem A. Instead, the maximum amount of definite knowledge that an observer can reliably infer is quantified by the transfer entropy *conditioned on the total information available about the environmental states* ρ_E^{in} and ρ_E^f.

This conditioned transfer entropy lies within the same causal cone determined by the tripartite dynamics, as illustrated in Figures 4.3 and 4.4.

The observable lower bound of $\mathcal{C}(A \to B|E)$ for a $CCNOT$ channel is always nonpositive when the final state is diagonalized. A representative example is depicted in Figure 4.5, where we consider a classical joint state

$$\rho_{AB} = \sum_{ij} p_{ij} \left|ij\right\rangle \left\langle ij\right| \tag{4.69}$$

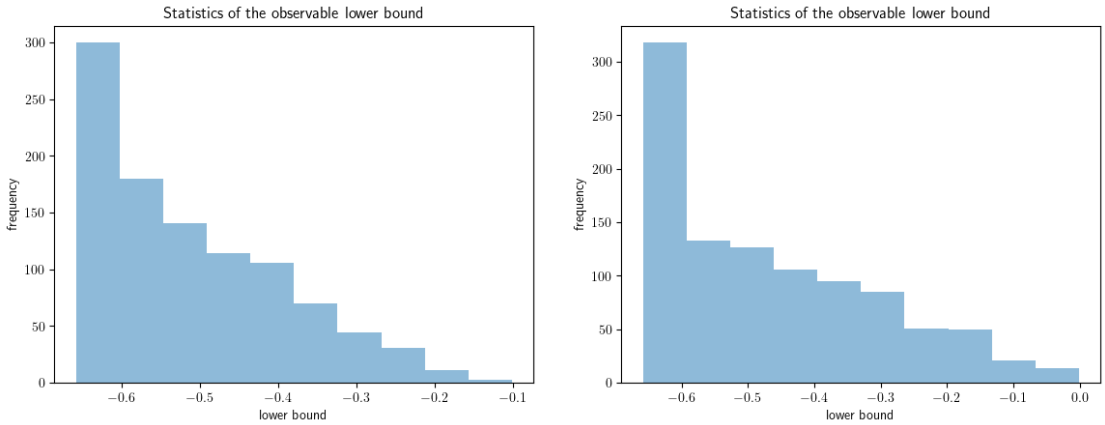
in the Hilbert space $\mathbb{C}^2 \otimes \mathbb{C}^2$, together with local observables $M_A = M_B = \sigma_Z$. One can verify in this observable configuration, the observable lower bound is alway *negative*.

Without applying dephasing to the final states, it becomes possible to obtain a positive observable lower bound, as illustrated in Figure 4.6. In this scenario, the Lieb lower bound is always non-negative, reflecting the fact that coherence is preserved throughout the protocol. As summarized in Tables 4.5 and 4.6, both the Lieb bound and the observable lower bound attain their maximal values when the environmental input state is $\left|E\right\rangle^{in} = \left|1\right\rangle$, for which the action of $CCNOT_{EA\to B}$ effectively reduces to $CNOT_{A\to B}$.

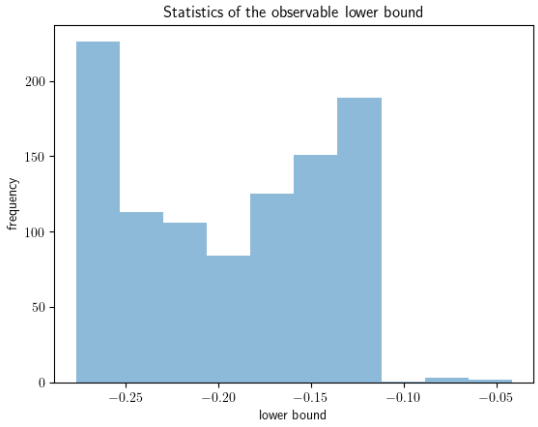
Moreover, the use of pure input states in Protocol 4.2.1 significantly increases the frequency with which the observable lower bound takes positive values. This highlights the role of input-state coherence in enhancing the detectability of conditional quantum transfer entropy in open-system settings.

Table 4.5: Frequencies of positive Lieb lower bounds for non-diag final states.

random input	pure input	$\rho_E^{in} = \left\|1\right\rangle \left\langle 1\right\|$
0.157	0.633	1



(a) **(b)**



(c)

Figure 4.5: Observable lower bounds of $\mathcal{C}(A \to B|E)$ for the $CCNOT$ channel, when the final state is diagonalized. The observables are chosen to be $M_A = M_B = \sigma_Z$. From 4.5a to 4.5c: lower bounds of random input states, pure input states, and $\rho_E^{in} = |1\rangle\langle 1|$.

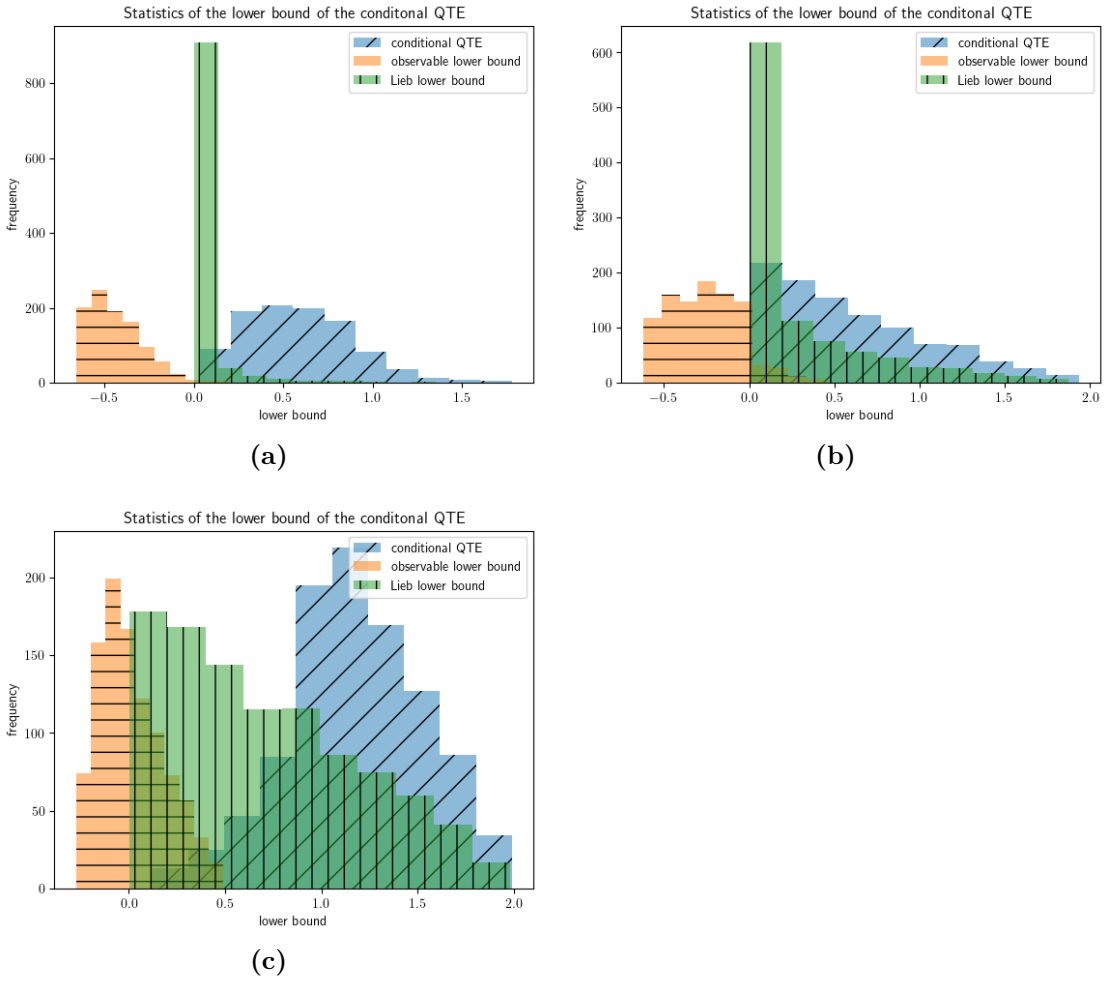


Figure 4.6: Leib and observable lower bounds of $\mathcal{C}(A \to B|E)$ for the $CCNOT$ channel, and non-diagonalized final state. The observables are chosen to be $M_A = M_B = \sigma_Z$. From 4.6a to 4.6c: lower bounds of random input states, pure input states, and $\rho_E^{in} = |1\rangle\langle 1|$.

Table 4.6: Frequencies of positive observable lower bounds for non-diag final states, with $M_A = M_B = \sigma_Z$.

random input	pure input	$\rho_E^{in} = \|1\rangle\langle 1\|$
0.014	0.098	0.476

Indeed, for a tripartite quantum system ρ_{ABC}, we can split the conditional mutual information in the following way:

$$I(A:B|C) = I(A:BC) - I(A:C) \tag{4.70}$$
$$= D(\check{A}:\check{B}\check{C}) - D(\check{A};\check{C}) + J(\check{A}:\check{B}\check{C}) - J(\check{A}:\check{C}). \tag{4.71}$$

Since we know that the quantum part is upper bounded by

$$D(\check{A}:\check{B}\check{C}), D(\check{A};\check{C}) \leq S(A), \tag{4.72}$$

then $I(A:B|C) > S(A)$ implies

$$I(\check{A}:\check{B}|\check{C}) = J(\check{A}:\check{B}\check{C}) - J(\check{A}:\check{C}) > 0. \tag{4.73}$$

This fact means it's sufficient to conclude a positive classical part of conditional mutual information when condition $I(A:B|C) > S(A)$ holds. So in order to select the state corresponding to positive Lieb lower bounds of $\mathcal{C}(A \to B|E)$, we just take those fulfilling the following inequality

$$2\max\{0, -S(B|A'A), -S(A'A|B)\} - \min\{S(A'A), S(B)\} > 0. \tag{4.74}$$

Definition 4.2.10 (Classical Causation Inference Quantity)**.** *The classical csusation inference quantity based the local (total) correlation of subsystems is defined as*

$$\mathcal{I}_C := 2\max\{0, -S(B|A'A), -S(A'A|B)\} - \min\{S(A'A), S(B)\}. \tag{4.75}$$

Theorem 4.2.11.

$$\mathcal{I}_C > 0 \implies \mathcal{C}^{\Phi}(A \to B|E) > 0, \tag{4.76}$$

where $\mathcal{C}^{\Phi}(A \to B|E)$ is the conditional quantum transfer entropy computed from the dephased final state.

Theorem 4.2.11 provides a method for identifying the purely classical causal influence between subsystems A and B induced by U_{EAB}, without requiring any prior knowledge of the environment E. This makes the criterion particularly useful

Table 4.7: Frequencies of positive $\mathcal{I}_C$ for non-diag final state.

random input	pure input	$\rho_E^{in} = \|1\rangle\langle 1\|$
0.011	0.187	0.396

for open-system scenarios, where direct access to environmental degrees of freedom is typically unavailable. The frequencies with which $\mathcal{I}_C$ takes positive values under different input settings are summarized in Table 4.7.

We conclude this chapter by comparing several representative tripartite channels. As summarized in Table 4.8, both $C(A \to B \mid E)$ and its Lieb lower bound exhibit a clear increase as the rotation angle of the controlled operation becomes larger. This trend reflects the enhanced ability of stronger controlled rotations to mediate directional information flow from A to B when considering the environment E. Furthermore, when the dynamics consist solely of independent local operations on the three subsystems, the conditional quantum transfer entropy $C(A \to B \mid E)$ necessarily vanishes. This confirms that no causal influence can be generated without genuine interactions coupling A, B and E.

Channel	$C(EA \to B)$	$C(E \to B)$	$C(A \to B \mid E)$	Lieb bound	obs. bound	$\mathcal{I}_C$
$CCX(0)$	0	0	0	0	0	0
$CCX(\pi/8)$	0.123	0.062	0.062	0.035	-0.03	-0.009
$CCX(\pi/4)$	0.371	0.186	0.186	0.099	-0.091	-0.037
$CCX(3\pi/8)$	0.668	0.334	0.334	0.164	-0.161	-0.088
$CCX(\pi/2)$	0.968	0.484	0.484	0.211	-0.226	-0.166
$I_E \otimes I_A \otimes I_B$	0	0	0	0	0	0
$I_E \otimes I_A \otimes X_B$	0	0	0	0	0	0

Table 4.8: The numerical results of quantum conditional transfer entropy $C(A \to B \mid E)$ for several three-qubit channels. All the input states of A, B and E are $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.

Chapter 5

Conclusion and Outlook

In this work, we develop an operational measure of one-way information flow induced by quantum channels, which we refer to as quantum transfer entropy (QTE). QTE is defined in terms of conditional mutual information, where temporal correlations in a quantum process are mapped to spatial correlations by introducing ancilla qubits that record past system states through entanglement. This construction allows us to treat the history of a quantum system as an explicit quantum memory. Analogous to its classical counterpart, QTE is a model-free quantity and is therefore capable of capturing non-linear causal relationships in quantum processes with a finite time lag. A nonzero value of QTE indicates the presence of directed quantum causation, while its magnitude provides a quantitative measure of the corresponding causal strength.

An important feature of QTE is its basis dependence, which reflects the necessity for observers to agree upon a reference basis prior to any causal analysis. Although different choices of reference basis may lead to different numerical values of QTE, the resulting quantum causal structures remain isomorphic, thereby preserving the essential pattern of directional influence among the underlying quantum systems.

We investigate the quantum-classical splitting of QTE in bipartite channels. Our analysis shows that a $CNOT_{A\to B}$ quantum channel transfers the entire information content of the control qubit - including both its classical information and quantum coherence - to the target qubit. In particular, we obtain

$$\mathcal{C}(A \to B) = h_A + c_A, \tag{5.1}$$

where h_A denotes the classical Shannon entropy of the control qubit A and c_A characterizes its coherence.

For a general controlled-Z channel, the quantum transfer entropy is highly linearly correlated with the total information of the control qubit. The slope of this linear relation is determined by the rotation angle of the controlled-Z operation, indicating that the extent of phase rotation directly modulates the strength of the induced information flow.

We futhre investigate the exact QTE and the associated entropy bounds in tripartite systems.When the influence of a third subsystem (or an external environment) is taken into account, the purely directional causation between qubits A and B is captured by the conditional mutual information

$$\mathcal{C}(A \to B \mid E). \tag{5.2}$$

We analyze the extreme case of $CCNOT_{EA\to B}$ channel, and derive two lower bounds of $\mathcal{C}(A \to B \mid E)$, in terms of only local subsystems. The first one, called Lieb lower bound, is

$$2\max\{0, -S^f(B|A'A), -S^f(A'A|B)\} \leq \mathcal{C}(A \to B|E). \tag{5.3}$$

For $CCNOT_{EA\to B}$ channel, we have another the lower bound in terms of local observables $M_{A'}$ and M_B on subsystems A and B respectively:

$$\mathcal{C}(A \to B \mid E) = I(A'B \mid E') \geq \frac{C^2(M_{A'}, M_B)}{2\|M_{A'}\|^2\|M_B\|^2} + f(d, \gamma). \tag{5.4}$$

The Lieb lower bound is always non-negative for diagonalized final state of QTE defining protocol, reflecting the fact that coherence is preserved throughout the protocol.

Finally, we introduce a quantity that serves as a criterion for classical causation inference, defined on the local composite system $A'AB$. The classical causation inference quantity is defined by

$$\mathcal{I}_C := 2\max\{0, -S(B|A'A), -S(A'A|B)\} - \min\{S(A'A), S(B)\}. \tag{5.5}$$

A positive $\mathcal{I}_C$ indicates a positive pure classical correlation between A and B after the evolution of a three-qubit channel U_{EAB}, in the presence of environmental qubit E.

Although quantum transfer entropy performs well for a variety of controlled channels, significant challenges arise when attempting to extend it to more complex quantum processes due to its nonadditivity. As a consequence, the current technique is applicable only to relatively small channel fragments, where the informational

contributions can still be meaningfully isolated.

Furthermore, because our framework relies on entanglement and coherence resources to enhance the capability of causal discovery in quantum processes, quantum channels that strongly disturb or extensively reshape coherence may lead to spurious or misleading causal structures. In such cases, the apparent information flow inferred from QTE may not faithfully represent the underlying dynamical causation but instead reflect coherence manipulation introduced by the channel itself.

Beyond the above discussion, further development of quantum causal discovery algorithms within this framework is both natural and desirable. An important direction for future work is to systematically characterize the class of quantum causal structures that can be reliably resolved by the present method. Such a classification would clarify the scope of applicability of QTE-based inference and help identify which structural features of quantum processes remain accessible or become ambiguous under this approach.

Bibliography

[1] Thomas Schreiber. «Measuring Information Transfer». In: *Physical Review Letters* 85.2 (July 2000). Publisher: American Physical Society, pp. 461–464. DOI: 10.1103/PhysRevLett.85.461. URL: https://link.aps.org/doi/10.1103/PhysRevLett.85.461 (visited on 05/12/2025) (cit. on p. 1).

[2] Ryan G. James, Nix Barnett, and James P. Crutchfield. «Information Flows? A Critique of Transfer Entropies». In: *Physical Review Letters* 116.23 (June 2016). Publisher: American Physical Society, p. 238701. DOI: 10.1103/PhysRevLett.116.238701. URL: https://link.aps.org/doi/10.1103/PhysRevLett.116.238701 (visited on 05/12/2025) (cit. on p. 1).

[3] Raul Vicente, Michael Wibral, Michael Lindner, and Gordon Pipa. «Transfer entropy—a model-free measure of effective connectivity for the neurosciences». en. In: *Journal of Computational Neuroscience* 30.1 (Feb. 2011), pp. 45–67. ISSN: 1573-6873. DOI: 10.1007/s10827-010-0262-3. URL: https://doi.org/10.1007/s10827-010-0262-3 (visited on 06/16/2025) (cit. on p. 1).

[4] Mark M. Wilde. *Quantum Information Theory*. 2nd ed. Cambridge: Cambridge University Press, 2017. ISBN: 978-1-107-17616-4. DOI: 10.1017/9781316809976. URL: https://www.cambridge.org/core/books/quantum-information-theory/247A740E156416531AA8CB97DFDAE438 (visited on 11/24/2025) (cit. on pp. 1, 5, 17).

[5] Leland Gerson Neuberg. «CAUSALITY: MODELS, REASONING, AND INFERENCE, by Judea Pearl, Cambridge University Press, 2000». en. In: *Econometric Theory* 19.4 (Aug. 2003), pp. 675–685. ISSN: 1469-4360, 0266-4666. DOI: 10.1017/S0266466603004109. URL: https://www.cambridge.org/core/journals/econometric-theory/article/abs/causality-models-reasoning-and-inference-by-judea-pearl-cambridge-university-press-2000/DA2D9ABB0AD3DAC95AE7B3081FCDF139 (visited on 11/25/2025) (cit. on p. 2).

[6] Wenjin Niu, Zijun Gao, Liyan Song, and Lingbo Li. *Comprehensive Review and Empirical Evaluation of Causal Discovery Algorithms for Numerical Data*. arXiv:2407.13054 [cs]. Sept. 2024. DOI: 10.48550/arXiv.2407.13054. URL: http://arxiv.org/abs/2407.13054 (visited on 05/05/2025) (cit. on pp. 2, 4).

[7] Jonathan Barrett, Robin Lorenz, and Ognyan Oreshkov. *Quantum Causal Models*. arXiv:1906.10726 [quant-ph]. Nov. 2020. DOI: 10.48550/arXiv.1906.10726. URL: http://arxiv.org/abs/1906.10726 (visited on 04/04/2025) (cit. on pp. 2, 30).

[8] Ding Jia. *Quantifying Causality in Quantum and General Models*. en. Jan. 2018. URL: https://arxiv.org/abs/1801.06293v1 (visited on 05/27/2025) (cit. on pp. 2, 30).

[9] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. en. ISBN: 9780511976667 Publisher: Cambridge University Press. Dec. 2010. DOI: 10.1017/CBO9780511976667. URL: https://www.cambridge.org/highereducation/books/quantum-computation-and-quantum-information/01E10196D0A682A6AEFFEA52D53BE9AE (visited on 11/25/2025) (cit. on pp. 5, 17).

[10] Jiajun Ma, Benjamin Yadin, Davide Girolami, Vlatko Vedral, and Mile Gu. «Converting Coherence to Quantum Correlations». In: *Physical Review Letters* 116.16 (Apr. 2016). Publisher: American Physical Society, p. 160407. DOI: 10.1103/PhysRevLett.116.160407. URL: https://link.aps.org/doi/10.1103/PhysRevLett.116.160407 (visited on 04/10/2025) (cit. on pp. 11, 12).

[11] Kavan Modi, Tomasz Paterek, Wonmin Son, Vlatko Vedral, and Mark Williamson. «Unified View of Quantum and Classical Correlations». In: *Physical Review Letters* 104.8 (Feb. 2010). Publisher: American Physical Society, p. 080501. DOI: 10.1103/PhysRevLett.104.080501. URL: https://link.aps.org/doi/10.1103/PhysRevLett.104.080501 (visited on 04/16/2025) (cit. on p. 11).

[12] Thomas R. Bromley, Marco Cianciaruso, and Gerardo Adesso. «Frozen Quantum Coherence». In: *Physical Review Letters* 114.21 (May 2015). Publisher: American Physical Society, p. 210401. DOI: 10.1103/PhysRevLett.114.210401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.114.210401 (visited on 04/14/2025) (cit. on pp. 11, 12).

[13] John von Neumann. *Mathematical Foundations of Quantum Mechanics: New Edition*. en. Princeton University Press, Feb. 2018. ISBN: 978-1-4008-8992-1. DOI: 10.1515/9781400889921. URL: https://www.degruyterbrill.com/document/doi/10.1515/9781400889921/html (visited on 11/24/2025) (cit. on p. 12).

[14] Michael M. Wolf, Frank Verstraete, Matthew B. Hastings, and J. Ignacio Cirac. «Area Laws in Quantum Systems: Mutual Information and Correlations». In: *Physical Review Letters* 100.7 (Feb. 2008). Publisher: American Physical Society, p. 070502. DOI: 10.1103/PhysRevLett.100.070502. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.070502 (visited on 05/11/2025) (cit. on p. 15).

[15] Eric A. Carlen and Elliott H. Lieb. «Bounds for Entanglement via an Extension of Strong Subadditivity of Entropy». en. In: *Letters in Mathematical Physics* 101.1 (July 2012), pp. 1–11. ISSN: 1573-0530. DOI: 10.1007/s11005-012-0565-6. URL: https://doi.org/10.1007/s11005-012-0565-6 (visited on 08/07/2025) (cit. on p. 16).

[16] Matthias Christandl and Andreas Winter. «"Squashed Entanglement" - An Additive Entanglement Measure». In: *Journal of Mathematical Physics* 45.3 (Mar. 2004). arXiv:quant-ph/0308088, pp. 829–840. ISSN: 0022-2488, 1089-7658. DOI: 10.1063/1.1643788. URL: http://arxiv.org/abs/quant-ph/0308088 (visited on 08/07/2025) (cit. on p. 17).

[17] Fernando G. S. L. Brandão, Matthias Christandl, and Jon Yard. «Faithful Squashed Entanglement». en. In: *Communications in Mathematical Physics* 306.3 (Sept. 2011), pp. 805–830. ISSN: 1432-0916. DOI: 10.1007/s00220-011-1302-1. URL: https://doi.org/10.1007/s00220-011-1302-1 (visited on 08/27/2025) (cit. on pp. 17, 29).

[18] Sam A. Hill and William K. Wootters. «Entanglement of a Pair of Quantum Bits». In: *Physical Review Letters* 78.26 (June 1997). Publisher: American Physical Society, pp. 5022–5025. DOI: 10.1103/PhysRevLett.78.5022. URL: https://link.aps.org/doi/10.1103/PhysRevLett.78.5022 (visited on 11/24/2025) (cit. on p. 18).

[19] William K. Wootters. «Entanglement of Formation of an Arbitrary State of Two Qubits». In: *Physical Review Letters* 80.10 (1998), pp. 2245–2248. DOI: 10.1103/PhysRevLett.80.2245 (cit. on p. 18).

[20] Harold Ollivier and Wojciech H. Zurek. «Quantum Discord: A Measure of the Quantumness of Correlations». In: *Physical Review Letters* 88.1 (Dec. 2001). Publisher: American Physical Society, p. 017901. DOI: 10.1103/PhysRevLett.88.017901. URL: https://link.aps.org/doi/10.1103/PhysRevLett.88.017901 (visited on 04/14/2025) (cit. on pp. 19, 20).

[21] Chandrashekar Radhakrishnan, Mathieu Laurière, and Tim Byrnes. «Multipartite Generalization of Quantum Discord». In: *Physical Review Letters* 124.11 (Mar. 2020). Publisher: American Physical Society, p. 110401. DOI: 10.1103/PhysRevLett.124.110401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.124.110401 (visited on 05/11/2025) (cit. on p. 19).

[22] Masato Koashi. «Monogamy of quantum entanglement and other correlations». In: *Physical Review A* 69.2 (2004). DOI: 10.1103/PhysRevA.69.022309 (cit. on p. 20).

[23] Paul Appel, Marcus Huber, and Claude Klöckl. «Monogamy of correlations and entropy inequalities in the Bloch picture». en. In: *Journal of Physics Communications* 4.2 (Feb. 2020). Publisher: IOP Publishing, p. 025009. ISSN: 2399-6528. DOI: 10.1088/2399-6528/ab6fb4. URL: https://dx.doi.org/10.1088/2399-6528/ab6fb4 (visited on 03/25/2025) (cit. on p. 20).

[24] Davide Girolami and Michele Minervini. «Quantitative bounds to propagation of quantum correlations in many-body systems». In: *Physics Letters A* 496 (Feb. 2024), p. 129315. ISSN: 0375-9601. DOI: 10.1016/j.physleta.2024.129315. URL: https://www.sciencedirect.com/science/article/pii/S0375960124000100 (visited on 09/28/2025) (cit. on p. 23).

[25] John-Mark A. Allen, Jonathan Barrett, Dominic C. Horsman, Ciarán M. Lee, and Robert W. Spekkens. «Quantum Common Causes and Quantum Causal Models». In: *Physical Review X* 7.3 (July 2017). Publisher: American Physical Society, p. 031021. DOI: 10.1103/PhysRevX.7.031021. URL: https://link.aps.org/doi/10.1103/PhysRevX.7.031021 (visited on 03/31/2025) (cit. on p. 30).

[26] Rafael Chaves, Christian Majenz, and David Gross. «Information–theoretic implications of quantum causal structures». en. In: *Nature Communications* 6.1 (Jan. 2015). Publisher: Nature Publishing Group, p. 5766. ISSN: 2041-1723. DOI: 10.1038/ncomms6766. URL: https://www.nature.com/articles/ncomms6766 (visited on 04/01/2025) (cit. on p. 30).

[27] Ge Bai, Ya-Dong Wu, Yan Zhu, Masahito Hayashi, and Giulio Chiribella. «Quantum causal unravelling». en. In: *npj Quantum Information* 8.1 (June 2022). Publisher: Nature Publishing Group, pp. 1–9. ISSN: 2056-6387. DOI: 10.1038/s41534-022-00578-4. URL: https://www.nature.com/articles/s41534-022-00578-4 (visited on 05/26/2025) (cit. on p. 32).

[28] Christina Giarmatzi and Fabio Costa. «A quantum causal discovery algorithm». en. In: *npj Quantum Information* 4.1 (Mar. 2018). Publisher: Nature Publishing Group, p. 17. ISSN: 2056-6387. DOI: 10.1038/s41534-018-0062-6. URL: https://www.nature.com/articles/s41534-018-0062-6 (visited on 07/08/2025) (cit. on p. 32).

[29] Rishi Goel, Casey R. Myers, and Sally Shrapnel. *Quantum Algorithms for Causal Estimands*. arXiv:2505.12873 [quant-ph]. May 2025. DOI: 10.48550/arXiv.2505.12873. URL: http://arxiv.org/abs/2505.12873 (visited on 07/08/2025) (cit. on p. 32).

[30] Yu Terada, Ken Arai, Yu Tanaka, Yota Maeda, Hiroshi Ueno, and Hiroyuki Tezuka. *Quantum-enhanced causal discovery for a small number of samples*. arXiv:2501.05007 [quant-ph]. July 2025. DOI: 10.48550/arXiv.2501.05007. URL: http://arxiv.org/abs/2501.05007 (visited on 07/08/2025) (cit. on p. 32).