QQIF: Quantum Quantitative Information Flow (invited paper)

Arthur Américo
EECS
Queen Mary University of London
London, United Kingdom
a.passosderezende@qmul.ac.uk

Pasquale Malacaria
EECS
Queen Mary University of London
London, United Kingdom
p.malacaria@qmul.ac.uk

***Abstract*—The field of Quantitative Information Flow (QIF) is concerned with quantifying information leakage in systems. This work generalises the QIF framework to the quantum setting, having as foundations the recently developed g-leakage framework. Next, some recent results of the field of quantum statistical decision are presented and translated in terms of the "quantum" g-leakage framework, including a recently developed quantum generalisation of the Coriasceous Theorem, an important result for QIF.**

1. Introduction

Recent advances on physical quantum computing, like Google announcement of "quantum supremacy" [1] seems to indicate that quantum computers are getting closer to mainstream real world applications. Security is a the heart of quantum algorithms: as well known the implications of a real-world implementation of the Shor algorithm would be devastating for current cryptography. The problem of analyzing security of quantum systems is hence an important and timely problem.

In this work we describe a mathematical framework to reason about leakage of confidential information in the presence of quantum systems. The main contribution of this paper is presenting a generalisation of the framework commonly used in the field of quantitative information flow capable of covering both classical and quantum computation.

The field of quantitative information flow (QIF) concerns itself with quantifying and minimising information leakage in security systems. The most popular framework in QIF describes systems as a probabilistic mapping that takes as input some *secret* value x from a set $\mathcal{X}$ and produces an *observable* y from a set $\mathcal{Y}$. The leakage of information is then taken w.r.t. an *adversary* whose knowledge about the secret value is modelled by a probability distribution over $\mathcal{X}$. Being aware of the inner workings of the system, he updates his prior probability to a posterior probability distribution, gaining information about the secret value. The specifics on how to measure this leakage of information (i.e., what "information-measuring" function to use) may depend both on the operational aspects of the system and on the objectives and capabilities of the adversary.

The simplest model for a system is that of a (discrete memoryless) *channel* K, which can be seen as a probabilistic transition matrix s.t. $K(y|x)$ is the conditional probability of the system yielding y given that the secret value is x.

In this paper, we aim to explore the possible connections between the field of Quantum Computation and Information Theory and of QIF. While there seems to be little overlap between the types of problems each field devotes itself to, this paper shows deep connections, based on statistical decision theory.

Statistical decision theory [2] can be introduced by looking at the following problem: A state of nature generates a probability distributions over a set of possible observations. An observer has to make a decision based on such observations. His decision is subject to a loss function and his overall aim is to make the decision minimising his loss.

The stochastic mechanism by which a state of nature generates a probability distributions over a set of possible observations is often called an *experiment* and can be presented as a row stochastic matrix. A fundamental problem in the field was the following: Are some experiments more informative than other? That is: are some experiments always allowing an observer a smaller loss than other experiments no matter what loss function is chosen?

The answer is yes and remarkably Sherman [3] and Blackwell [4] proved that this partial order of information is completely characterized by matrix multiplication of row stochastic matrices: that is given experiments (i.e. row stochastic matrices) E, E', experiment E is more informative than E' if and only if there exists another experiment E'' s.t. $E = E'E''$.

The theory of statistical decision and Blackwell theorem have been independently recently rediscovered in the field of Quantitative Information Flow and gain functions [5]. Here the observer is the attacker, the experiment is the secret dependent system the attacker observes and the loss

function is the gain function of the attacker which model his objectives and capabilities.

A simple example is the setting of an attacker observing the computational time of some crypto operations, with the attacker's objective being guessing the secret key.

The Blackwell theorem (rediscovered in QIF as the Coriaceus theorem [6]) has here a natural interpretation as characterizing the security ordering between systems in terms of "S leaks always more than S' for all gain functions".

In recent work [7], Buscemi proved a generalisation of the Blackwell Theorem to the quantum statistical decision theory framework. Motivated by the applications the original Blackwell Theorem finds in QIF, we provide a translation of Buscemi's result using the quantum framework we develop in this paper, stating Buscemi's result in terms of a quantum generalisation of the g-leakage framework [5], [8].

2. A short introduction to the formalism of Quantum Theory

In this section we introduce some of the mathematical tools and concepts used in quantum theory. We do not attempt here to give much of the physical interpretation for these concepts, to this end we refer to Nielsen and Chuang's elucidative book on the subject [9]. We also constrain our exposition only to finite-dimensional Hilbert spaces.

In what follows we will introduce the necessary definitions and results. Besides Nielsen and Chuang's book, this section also follows the succint treatment given by Holevo, Efroimsky and Gamberg [10].

2.1. Hilbert Spaces, Vectors and Linear Transformations

Definition 1. *A* Hilbert space $\mathscr{H}$ *is a complex vector space with finite dimension* d*, together with an inner product* $\langle \cdot | \cdot \rangle$ *such that, for all* $\phi, \psi, \sigma \in \mathscr{H}$*, and all* $a, b \in \mathbb{C}$

1) $\langle \psi | \psi \rangle \geq 0$,
2) $\langle \psi | \psi \rangle = 0 \Leftrightarrow \psi = 0$,
3) $\langle \phi | \psi \rangle = \overline{\langle \psi | \phi \rangle}$*, where* $\overline{z}$ *is the complex conjugate of* z,
4) $\langle \sigma | a\phi + b\psi \rangle = a \langle \sigma | \phi \rangle + b \langle \sigma | \psi \rangle$.

The norm *of a vector is defined as* $\|\psi\| = \sqrt{\langle \psi | \psi \rangle}$.

Bra-ket notation:. A vector $\phi \in \mathscr{H}$ will be denoted by $|\phi\rangle$, while $\langle \phi|$ represents the complex function $|\psi\rangle \mapsto \langle \phi | \psi \rangle$ over $\mathscr{H}$. That is, $\langle \phi|$ is the adjoint of $|\phi\rangle$ on the dual space $\mathscr{H}^*$.

In finite dimensions, it might be useful to think of Hilbert spaces in terms of matrix notation. For example, suppose we have a three dimensional Hilbert space with orthonormal basis $\{|0\rangle, |1\rangle, |2\rangle\}$. Let $|\phi\rangle = 3|0\rangle + 2i|1\rangle + (1+i)|2\rangle$. We may represent $|\phi\rangle$ by the column vector

$$\begin{bmatrix} 3 \\ 2i \\ 1+i \end{bmatrix}. \tag{1}$$

And its adjoint $\langle \phi|$ by the row vector

$$\begin{bmatrix} 3 & -2i & 1-i \end{bmatrix}.$$

Notice that in this representation, $\langle \phi|$ is the conjugate transpose of $|\phi\rangle$. We denote this by writing $\langle \phi| = |\phi\rangle^\dagger$.

Definition 2. *A vector* $|\psi\rangle$ *is called a* pure state *if* $\|\psi\| = 1$.

We are also interested in *linear operators* from one Hilbert space to another. Let $\mathscr{H}_A$, $\mathscr{H}_B$ be Hilbert spaces with basis $\{|i\rangle\}_{\{1,\dots,m\}}$, $\{|j\rangle\}_{\{1,\dots,n\}}$ respectively. A linear operator $A : \mathscr{H}_A \to \mathscr{H}_B$ can be completely defined by

$$A = \sum_{i,j} a_{j,i} |j\rangle \langle i|.$$

where each coefficient $a_{j,i}$ is a complex number. Henceforth, for brevity's sake, we use *operator* as a short-hand for linear operator.

Given the coefficients of A, the result of applying A to a vector can be easily calculated whenever we have the the basis is orthonormal (i.e., if $\langle i|j\rangle = 1$ if $i = j$ and 0 otherwise).

As an example, consider the vector $|\phi\rangle = 3|0\rangle + 2i|1\rangle + (1+i)|2\rangle$ in $\mathscr{H}_A$, suppose an orthonormal basis of $\mathscr{H}_B$ is $\{|0\rangle, |1\rangle\}$ and that A is given by the coefficients $a_{0,0} = a_{1,1} = a_{1,2} = 1$, $a_{1,0} = a_{0,1} = 1/2$ and $a_{0,2} = 0$. Then

$$\begin{aligned} A|\phi\rangle &= \sum_{j,k} a_{k,j} |k\rangle \langle j|\phi\rangle \\ &= \sum_k 3a_{k,0} |k\rangle \langle 0|0\rangle + \sum_k 2ia_{k,1} |k\rangle \langle 1|1\rangle + \\ &\quad \sum_k (1+i)a_{k,2} |k\rangle \langle 2|2\rangle \\ &= (3+i)|0\rangle + \left(\frac{5}{2} + 3i\right)|1\rangle. \end{aligned}$$

The operator A is perhaps more easily understood by thinking again in terms of matrices. The effect of A can be thought as multiplying the vector representation of $|\phi\rangle$ (1), by the matrix with coefficients $a_{k,j}$, i.e.,

$$\begin{bmatrix} 1 & 1/2 & 0 \\ 1/2 & 1 & 1 \end{bmatrix}.$$

We end this section with some important definitions and results. We say that an operator A is defined *on* a Hilbert space $\mathscr{H}$ if its both domain and codomain are $\mathscr{H}$.

Definition 3. *An operator* A *on* $\mathscr{H}$

- *is called* Hermitian *if* $A = A^\dagger$*, where* $A^\dagger$ *is the conjugate transpose of* A.

- *is called* positive, *if* $\forall \phi \in \mathscr{H}$, $\langle\phi| A |\phi\rangle \geq 0$. *In this case, we write* $A \geq 0$.

Proposition 1 ([9]). *Any positive operator on a Hilbert space is Hermitian.*

Theorem 1 ([10, Theorem 1.2]). *Given any Hermitian operator* A*, there is a orthonormal basis* $\{|i\rangle\}_{i\in\{0,\ldots,m\}}$ *and* real *numbers* $a_1, \ldots, a_m$ *s.t.*

$$A = \sum_{i=0}^{m} a_i |i\rangle \langle i| .$$

We say that A *is diagonal w.r.t. this basis, and that the numbers* a_i *are the* eigenvalues *of* A.

It is easy to verify that an Hermitian operator is positive if and only if all its eigenvalues are non-negative.

Definition 4. *Let* A *be an operator on* $\mathscr{H}$*. Its* trace *is defined as*

$$\mathbf{tr}(A) = \sum_i \langle i| A |i\rangle$$

where $|i\rangle$ *is any orthonormal basis of* $\mathscr{H}$.

Any choice of orthonormal basis will yield the same value for the trace of an operator, and thus the definition above is sound.

2.2. The Density Operator

Instead of vectors in $\mathscr{H}$, one can also use a subset of linear operators from $\mathscr{H}$ to $\mathscr{H}$ to describe quantum states. These operators are called *density operators*. This alternative formalism is the one mainly adopted in the majority of this paper.

In this formalism, a pure state $|\phi\rangle$ is represented by the operator $\rho = |\phi\rangle\langle\phi|$. For example, take the pure state

$$|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{i}{\sqrt{3}}|1\rangle - \frac{1}{\sqrt{6}}|2\rangle . \tag{2}$$

Its respective density operator is

$$\rho = \begin{bmatrix} 1/2 & -i/\sqrt{6} & -1/\sqrt{12} \\ i/\sqrt{6} & 1/3 & -i/\sqrt{18} \\ -1/\sqrt{12} & i/\sqrt{18} & 1/6 \end{bmatrix} . \tag{3}$$

Notice that the trace of ρ is 1. In fact, this property together with positiveness are sufficient to characterize all valid density operators.

Definition 5. *A* density operator ρ *is a positive operator such that* $\mathbf{tr}(\rho) = 1$.

Any quantum state is completely described by a density operator. Notice that this definition is more permissive than Definition 2. Consider, for example, the following operator

$$\rho_m = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} .$$

It is not hard to see that there is no pure state ϕ such that $\rho_m = |\phi\rangle\langle\phi|$. The state ρ_m is what we call a *mixed* state, i.e. a (convex) mixture of the pure states $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$.

Definition 6. *A state* ρ *is called* pure *if* $\mathbf{tr}(\rho^2) = 1$ *(equivalently, if it has rank 1). If this is not the case (i.e.,* $\mathbf{tr}(\rho^2) < 1$*) the state is called* mixed.

Notice that because all eigenvalues of a positive operator are non-negative, $\mathbf{tr}(\rho^2)$ can never be greater than 1.

2.3. Completely Positive Trace-preserving Maps

It is sometimes necessary to define an operation that takes density operators on one Hilbert space to density operators on another. This is often accomplished by the use of a *completely positive trace-preserving map* (CPTP) [9].

Definition 7. *A linear map* $\mathcal{E}$ *from density operators on* $\mathscr{H}_A$ *to density operators on* $\mathscr{H}_B$ *is a* completely positive trace-preserving map *(CPTP) if there is a family of linear operators* $\{A_i\}_i$ *from* $\mathscr{H}_A$ *to* $\mathscr{H}_B$ *such that*

- $\mathcal{E}(\rho) = \sum_i A_i \rho A_i^\dagger$,
- $\sum_i A_i^\dagger A_i = I$,

where I *is the identity operator on* $\mathscr{H}_A$ *(i,e.* $I = \sum_j |j\rangle\langle j|$ *for any orthonmormal basis* $\{|j\rangle\}_j$ *of* $\mathscr{H}_A$*).*

2.4. Measurement and POVMs

Whenever one makes a *measurement* of a quantum state ρ, the state "collapses" to one of possible observable events, which depend on the measurement being made. There is a myriad of different formalisms for measurements in Quantum Theory, but to our purposes we will present only the so-called *Positive Operator-Valued Measure* (POVM).

Definition 8. *A* POVM *is a collection of linear operators* $E = \{E_j\}$ *such that*

- *for all* j, $E_j \geq 0$,
- $\sum_j E_j = I$.

Each operator E_j represents an outcome (to which we might give the name E_j or j), and the probability of the quantum state ρ collapsing on this observable is given by $\mathbf{tr}(\rho E_j)$. Notice that, as the operators are positive and sum to 1, the values $\{\mathbf{tr}(\rho E_j)\}$ are indeed a valid probability distribution.

As an example, consider the quantum state (3), and the POVM given by the operators

$$E_0 = |0\rangle\langle 0| \quad E_1 = |1\rangle\langle 1| \quad E_2 = |2\rangle\langle 2| .$$

The probability of each observable would be:

$$\begin{aligned} \mathbf{Pr}(E_0) &= \mathbf{tr}(\rho |0\rangle\langle 0|) = \langle 0|\rho|0\rangle = 1/2, \\ \mathbf{Pr}(E_1) &= 1/3, \\ \mathbf{Pr}(E_2) &= 1/6. \end{aligned}$$

The state represented above by ρ is a pure state, with vector $|\psi\rangle$ given by (2). This means that we

may construct a POVM that contains $|\psi\rangle\langle\psi|$, for example:

$$\{|\psi\rangle\langle\psi|, I - |\psi\rangle\langle\psi|\}.$$

For which we have:

$$\mathbf{Pr}(|\psi\rangle\langle\psi|) = \mathbf{tr}(\rho\,|\psi\rangle\langle\psi|) = \langle\psi|\,\rho\,|\psi\rangle = 1,$$
$$\mathbf{Pr}(I - |\psi\rangle\langle\psi|) = 0.$$

As illustrated above, the probability distribution given by a state depends on the choice of POVM.

2.5. Tensor Product

Tensor products are an important concept both for linear algebra and quantum mechanics, but one which is tricky to give a very formal definition. Here we introduce the basic ideas and how they work.

Let $\mathscr{H}_A$, $\mathscr{H}_B$ be vector spaces of dimensions d_A, d_B. Given $|\psi\rangle \in \mathscr{H}_A$ and $|\phi\rangle \in \mathscr{H}_B$, we denote by $|\psi\rangle \otimes |\phi\rangle$ what we call their *tensor product*.

The Hilbert space $\mathscr{H}_A \otimes \mathscr{H}_B$ is called the tensor product of $\mathscr{H}_A$, $\mathscr{H}_B$, and it contains all elements $|\psi\rangle \otimes |\phi\rangle$ for $|\psi\rangle \in \mathscr{H}_A$ and $|\phi\rangle \in \mathscr{H}_B$, and their linear combinations. If $\{|a_0\rangle, \ldots, |a_m\rangle\}$ is a basis for $\mathscr{H}_A$ and $\{|b_0\rangle, \ldots, |b_n\rangle\}$ is a basis for $\mathscr{H}_B$, then the set $\{|a_i\rangle \otimes |b_j\rangle \mid i \leq m, j \leq n\}$ is a basis for $\mathscr{H}_A \otimes \mathscr{H}_B$.

For simplicity's sake, we write $|\psi\rangle|\phi\rangle$ or $|\psi\phi\rangle$ to mean $|\psi\rangle \otimes |\phi\rangle$. Also, we may refer to $\mathscr{H}_A$, $\mathscr{H}_B$ as the *systems* A, B, and to $\mathscr{H}_A \otimes \mathscr{H}_B$ as the *composite system* AB.

2.5.1. Manipulation Rules for Tensor Products. Now we introduce the rules to algebraically manipulate tensor products and operators in these spaces. These rules are as described in [9, Section 2.17].

For $z \in \mathbb{C}$, $|\psi_1\rangle, |\psi_2\rangle \in \mathscr{H}_A$ and $|\phi_1\rangle, |\phi_2\rangle \in \mathscr{H}_B$.

1) $z(|\psi_1\rangle|\phi_1\rangle) = (z\,|\psi_1\rangle)|\phi_1\rangle = |\psi_1\rangle(z\,|\phi_1\rangle)$,
2) $(|\psi_1\rangle + |\psi_2\rangle)|\phi_1\rangle = |\psi_1\rangle|\phi_1\rangle + |\psi_2\rangle|\phi_1\rangle$,
3) $|\psi_1\rangle(|\phi_1\rangle + |\phi_2\rangle) = |\psi_1\rangle|\phi_1\rangle + |\psi_1\rangle|\phi_2\rangle$.

Let $\{A_i\}$, $\{B_i\}$ be collections of operators over $\mathscr{H}_A \otimes \mathscr{H}_B$. Then, given a collection of complex numbers $\{c_i\}$, the action of the operator

$$C = \sum_i c_i A_i \otimes B_i$$

on vectors $|\psi\rangle|\phi\rangle$, where $|\psi\rangle \in \mathscr{H}_1$, $|\phi\rangle \in \mathscr{H}_2$, is defined as

$$\left(\sum_i c_i A_i \otimes B_i\right)|\psi\rangle|\phi\rangle = \sum_i c_i (A_i\,|\psi\rangle) \otimes (B_i\,|\phi\rangle),$$

and extended by linearity to $\mathscr{H}_A \otimes \mathscr{H}_B$. Finally, given two vectors $|\sigma\rangle = \sum_i z_i\,|\psi_i\rangle|\phi_i\rangle$, $|\sigma'\rangle = \sum_j z'_j\,|\psi'_j\rangle|\phi'_j\rangle$ belonging to $\mathscr{H}_A \otimes \mathscr{H}_B$, their inner product is defined by

$$\langle\sigma|\sigma'\rangle = \sum_{i,j} \overline{z_i} z'_j \langle\psi_i|\psi'_j\rangle\langle\phi_i|\phi'_j\rangle.$$

2.5.2. Partial trace. The operator algebra from the previous section gives a natural way to think about density operators in tensor product spaces. An important notion is that of a *partial trace*. Given an operator ρ^{AB} in a composite system AB, it is natural to ask what ρ^{AB} means as a description of state A. This is obtained by the partial trace operator

$$\rho^A = tr_B(\rho^{AB}),$$

where tr_B is defined as

$$tr_B(|\psi_1\rangle\langle\psi_2| \otimes |\phi_1\rangle\langle\phi_2|) = |\psi_1\rangle\langle\psi_2|\,\langle\phi_1|\phi_2\rangle \qquad (4)$$

if ρ^{AB} can be expressed as the tensor product of two pure states. If that is not the case, tr_B is completely defined as the linear operator that respects (4).

For interpretations of the partial trace, we refer to Section 2.4.3 of [9].

3. Quantum QIF: the basic structure

In this section we introduce the basic framework of Quantum Quantitative Information Flow. This construction is similar to *Quantum Statistical Models* in [7], but in this work we are interested on limiting the set of feasible POVMs, as a way to modelling possible attackers.

The setting is as follows. Consider a (nonempty, finite) set of secret values $\mathcal{X} = \{x_1, \ldots, x_n\}$. Some system takes a secret value $x \in \mathcal{X}$ as input and performs a computation, producing a quantum state ρ^x. A system is, thus, just a collection of states $\rho^{\mathcal{X}} = \{\rho^x\}_{x\in\mathcal{X}}$ indexed by $\mathcal{X}$ (with possible repetitions), that are density operators on some Hilbert space $\mathscr{H}$.

An adversary then makes a measurement on ρ^x, selecting a POVM $E = \{E_y\}_{y\in\mathcal{Y}}$ from a set of "allowed" POVMs $\mathcal{P}$. In doing so, he is able to obtain information about the secret value. Notice that each POVM is indexed by a (finite, nonempty) set $\mathcal{Y} = \{y_1, \ldots, y_m\}$, which is akin to the *output set* in classical QIF. This connection will be justified later.

Remark 1. *Albeit not used in this paper, another way to succinctly represent the pair p_X, $\rho^{\mathcal{X}}$ in our framework is as a density operator that is* classical *w.r.t. a Hilbert space with basis $\{|x\rangle\}_{x\in\mathcal{X}}$ (see e.g. [11, Section 2.1.3] and [7])*

$$\rho = \sum_x p_X(x)\,|x\rangle\langle x| \otimes \rho^x.$$

3.1. Quantifying Information on Quantum QIF

The quantification of information in QQIF is similar to the classical case, with the caveat that the adversary has a choice between different POVMs.

Our framework is based on g-vulnerabilities [5], in which a *gain function* g is used to model the interests and capabilities of the adversary.

The adversary has some prior knowledge about the secret, modelled by a probability distribution p_X. He also has at his disposal a set of possible actions $\mathcal{W}$. Whenever the adversary takes action $w \in \mathcal{W}$ and the secret value is $x \in \mathcal{X}$, he obtains a gain equal to $g(w,x)$. The *prior g-vulnerability* is then simply the expected gain of an optimal action.

$$V_g(p_X) = \max_{w\in\mathcal{W}} \sum_{x\in\mathcal{X}} p_X(x) g(w,x).$$

After the execution of the system, the attacker chooses a POVM $\{E_y\}_{y\in\mathcal{Y}}$ to perform a measurement on the resulting quantum state, and then chooses the action $w \in \mathcal{W}$ that maximises his gain.

The *posterior g-vulnerability* in the quantum setting can then be expressed by

$$V_{g,\mathcal{P}}(p_X, \rho^{\mathcal{X}}) = \max_{E\in\mathcal{P}} \sum_{y\in\mathcal{Y}} \max_{w\in\mathcal{W}} \sum_{x\in\mathcal{X}} p(x) g(w,x) \mathbf{tr}(\rho^x E_y). \tag{5}$$

3.2. Recovering classical QIF

The first important point to be made regarding QQIF is that it properly generalises the classical QIF scenario.

In classical QIF, systems are modelled by information theoretical *channels* K, with inputs in $\mathcal{X}$ and outputs in $\mathcal{Y}$. The probability of the system outputting y when the secret value is x is $K(y|x)$. The classical posterior vulnerability is given by

$$V_g^{\mathbf{c}}(p_X, K) = \sum_{y\in\mathcal{Y}} \max_{w} \sum_{x\in\mathcal{X}} p(x) g(w,x) K(y|x).$$

The system K above can be easily translated to the quantum setting by

- Letting $\{|y\rangle\}_{y\in\mathcal{Y}}$ be an orthonormal basis of a $|\mathcal{Y}|$-dimensional Hilbert space $\mathscr{H}$,
- Defining the quantum states as $\rho_K^x = \sum_y K(y|x) |y\rangle \langle y|$,
- Defining the set of allowed POVMs to be the singleton $\mathcal{P} = \{E\}$, such that $E_y = |y\rangle\langle y|$.

Then, (5) reduces to

$$\begin{aligned}
&V_{g,\mathcal{P}}(p_X, \rho_K^{\mathcal{X}}) \\
=&\max_{E\in\mathcal{P}} \sum_{y\in\mathcal{Y}} \max_{w\in\mathcal{W}} \sum_{x\in\mathcal{X}} p(x) g(w,x) \mathbf{tr}(\rho_K^x E_y) \\
=&\sum_{y\in\mathcal{Y}} \max_{w\in\mathcal{W}} \sum_{x\in\mathcal{X}} p(x) g(w,x) \mathbf{tr}(\rho_K^x |y\rangle\langle y|) \\
=&\sum_{y\in\mathcal{Y}} \max_{w\in\mathcal{W}} \sum_{x\in\mathcal{X}} p(x) g(w,x) \sum_{y'} K(y'|x) \langle y|y'\rangle \langle y'|y\rangle \\
=&\sum_{y\in\mathcal{Y}} \max_{w\in\mathcal{W}} \sum_{x\in\mathcal{X}} p(x) g(w,x) K(y|x) \\
=&V_g^{\mathbf{c}}(p_X, K).
\end{aligned}$$

Another way to see this connection it to interpret ρ_K^x as a description of the probability distribution on the corresponding row of K. Thus, for the classical case, the system in QQIF is a mapping from secrets to probability distributions on outputs.

3.3. Quantum Framework as a Collection of Classical Channels

In the last section it was shown how the quantum setting generalises classical QIF. In making this connection, it was seen that the chosen states and POVM characterised the channel K.

This is in fact a general property: given a collection $\rho^{\mathcal{X}}$ and a POVM E, there will be a channel associated with $\rho^{\mathcal{X}}$,E.

To see this, recall from 2.4 that $\mathbf{tr}(\rho^x E_y)$ is the probability of outcome y when measuring state ρ^x. Thus, letting $K_E(y|x) = \mathbf{tr}(\rho^x E_y)$, we have

$$\begin{aligned}
&\sum_{y\in\mathcal{Y}} \max_{w\in\mathcal{W}} \sum_{x\in\mathcal{X}} p(x) g(w,x) \mathbf{tr}(E_y \rho^x) \\
&= \sum_{y\in\mathcal{Y}} \max_{w\in\mathcal{W}} \sum_{x\in\mathcal{X}} p(x) g(w,x) K_E(y|x) \\
&= V_g^{\mathbf{c}}(p_X, K_E).
\end{aligned}$$

One possible interpretation of QQIF is thus that the quantum states together with $\mathcal{P}$ generate a space of possible channels from which the adversary is able to choose from by selecting a POVM.

The quantum version of posterior g-vulnerability is thus the maximum value obtainable from the set of channels induced by $\rho^{\mathcal{X}}$ and $\mathcal{P}$.

$$V_{g,\mathcal{P}}(p_X, \rho^{\mathcal{X}}) = \max_{E\in\mathcal{P}} V_g^{\mathbf{c}}(p_X, K_E). \tag{6}$$

As an example, let $\mathcal{X} = \{x_1, x_2\}$, $\mathcal{Y} = \{y_1, y_2\}$ and consider the system given by $\rho^{x_1} = |0\rangle\langle 0|$ and $\rho^{x_2} = |+\rangle\langle +|$, where $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Finally, let $\mathcal{P} = \{E, E'\}$, with

- $E_{y_1} = |0\rangle\langle 0|$, $E_{y_2} = |1\rangle\langle 1|$,
- $E'_{y_1} = |+\rangle\langle +|$, $E'_{y_2} = |-\rangle\langle -|$, where $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$.

The choice of E, E' induce the following channels.

K_E	y_1	y_2
x_1	1	0
x_2	$1/2$	$1/2$

$K_{E'}$	y_1	y_2
x_1	$1/2$	$1/2$
x_2	1	0

In general, the optimal choice of POVM will be dependent both on p_X and on g. For example, let $g_{id}(x,x') = \delta_{x,x'}$, where $\delta_{x,x'}$ is the Kronecker delta, and let $p_1 = (0.6, 0.4)$, $p_2 = (0.4, 0.6)$. Then, $V_{g_{id}}^{\mathbf{c}}(p_1, K_E) > V_{g_{id}}^{\mathbf{c}}(p_1, K_{E'})$, whereas $V_{g_{id}}^{\mathbf{c}}(p_2, K_E) < V_{g_{id}}^{\mathbf{c}}(p_2, K_{E'})$. An optimal adversary will thus choose E for p_1 and E' for p_2.

Fix now $p = (1/2, 1/2)$ and let g_1, g_2 be given by

$$g_1(x,x') = \begin{cases} 2, & \text{if } x = x' = x_1, \\ 1, & \text{if } x = x' = x_2, \\ 0, & \text{otherwise.} \end{cases}$$

$$g_2(x,x') = \begin{cases} 2, & \text{if } x = x' = x_2, \\ 1, & \text{if } x = x' = x_1, \\ 0, & \text{otherwise.} \end{cases}$$

In this case, it is easy to check that $V_{g_1}^{\mathbf{c}}(p, K_E) > V_{g_1}^{\mathbf{c}}(p, K_{E'})$ and $V_{g_2}^{\mathbf{c}}(p, K_E) < V_{g_2}^{\mathbf{c}}(p, K_{E'})$.

3.4. Using the QQIF Framework to Model Classical Scenarios

3.4.1. Example: POVMs as different attacks. One interesting application of the framework is modelling situations in which multiple attacks are possible.

Let $\mathcal{X}$ be the set of three-bit strings, perhaps signifying the value of a cryptographic key. Define the states of $\rho^{\mathcal{X}}$ by $\rho^x = |x\rangle\langle x|$ — i.e., $\rho^{000} = |000\rangle\langle 000|$,$\rho^{001} = |001\rangle\langle 001|$ and so on, where the hilbert space is the tensor product of three qubits.

Now, consider three possible attacks on the key: observing the first bit, observing the XOR of the first two bits of the string, and observing its hamming weight. These can be modelled, respectively, by the following POVMs.

- Observing the first bit:
  - $E_0^{1^{\text{st}}} = |000\rangle\langle 000| + |001\rangle\langle 001| + |010\rangle\langle 010| + |011\rangle\langle 011|$,
  - $E_1^{1^{\text{st}}} = |100\rangle\langle 100| + |101\rangle\langle 101| + |110\rangle\langle 110| + |111\rangle\langle 111|$.
- Observing the XOR of the first two bits:
  - $E_0^{\oplus} = |000\rangle\langle 000| + |001\rangle\langle 001| + |110\rangle\langle 110| + |111\rangle\langle 111|$,
  - $E_1^{\oplus} = |010\rangle\langle 010| + |011\rangle\langle 011| + |100\rangle\langle 100| + |101\rangle\langle 101|$.
- Observing the hamming weight:
  - $E_0^H = |000\rangle\langle 000|$,
  - $E_1^H = |001\rangle\langle 001| + |010\rangle\langle 010| + |100\rangle\langle 100|$,
  - $E_2^H = |011\rangle\langle 011| + |110\rangle\langle 110| + |101\rangle\langle 101|$,
  - $E_3^H = |111\rangle\langle 111|$.

These different POVMs give rise, respectively, to the following channels.

$K_{E^{1^{\text{st}}}}$	0	1
000	1	0
001	1	0
010	1	0
011	1	0
100	0	1
101	0	1
110	0	1
111	0	1

$K_{E^{\oplus}}$	0	1
000	1	0
001	1	0
010	0	1
011	0	1
100	0	1
101	0	1
110	1	0
111	1	0

K_{E^H}	0	1	2	3
000	1	0	0	0
001	0	1	0	0
010	0	1	0	0
011	0	0	1	0
100	0	1	0	0
101	0	0	1	0
110	0	0	1	0
111	0	0	0	1

If the adversary can choose between any of these three attacks, then the vulnerability of the system is given by $V_{g,\mathcal{P}}(p_X, \rho^{\mathcal{X}})$, where $\mathcal{P} = \{E^{1^{\text{st}}}, E^{\oplus}, E^H\}$. Notice that, because each of the channels above cannot be obtained from postprocessing another, then for each POVM there is a gain-function g and a p_X such that said POVM is strictly better than the others.

3.4.2. Using Quantum Notation to Represent Programs. Consider the program P in Figure 1, that takes as a secret input a value $\mathtt{h} \in \{0, 1, 2, 3\}$, and alters a three digit binary string `s`.

```
1.    s=000
2.    if h<2 and coin(0.5) then
3.        s=s|001
4.    if h%2==0  then
5.        s=s|010
6.    if s>000 and coin(0.5) then
7.        s=s|100
```

Figure 1. A probabilistic program that alters the value of `s`. The Boolean function `coin(p)` returns `True` with probability equal to `p`, and `False` otherwise.

The framework developed from Section 3 is useful to model the leakage of and adversary that can make a measurement on the variable `s` at some point of the execution. This can be done by associating, for each `h` and each line of the execution, a state $\rho_i^{\mathtt{h}}$ composed of three qubits, modelling the value of `s` after the execution of line i.

For example, consider the case $\mathtt{h} = 0$. At line 1 we have $\rho_1^0 = |000\rangle\langle 000|$. At line three, there is a fifty percent chance the bit is flipped, and we obtain $\rho_3^0 = \frac{1}{2}|000\rangle\langle 000| + \frac{1}{2}|001\rangle\langle 001|$. By proceeding in similar fashion, one obtains $\rho_5^0 = \frac{1}{2}(|010\rangle\langle 010| + |011\rangle\langle 011|)$ and $\rho_7^0 = \frac{1}{4}(|010\rangle\langle 010| + |110\rangle\langle 110| + |011\rangle\langle 011| + |111\rangle\langle 111|)$.

If, at any point of the execution, an adversary measures the value of `s`, obtain the respective channel can be obtained simply by applying the appropriate POVM on the quantum states. For example, suppose an attacker measures the hamming weight of `s` after the execution of line 7, by using the POVM E^H described in Section 3.4.1. The states at line 7 and the resulting channel are depicted in Figure 2.

4. Blackwell Theorem for QQIF

One important notion in QIF is that of *postprocessing* a program — i.e., executing a probabilistic remapping on the outputs of a system. In the language of information theoretical noisy channels, this is captured by the *degradedness* relation [12], which is a preorder over channels.[1]

Definition 9. *Given channels $K_1 : \mathcal{X} \to \mathcal{Y}$, $K_2 : \mathcal{X} \to \mathcal{Z}$, we say that K_2 is* degraded from K_1, *and write $K_1 \geq_{\mathrm{d}} K_2$ if there is a channel $W : \mathcal{Y} \to \mathcal{Z}$ such that*

$$\forall x \in \mathcal{X}, \forall z \in \mathcal{Z}, \quad K_2(z|x) = \sum_y W(z|y) K_1(y|x).$$

1. In the QIF literature, this relation is usually called the *refinement relation* .

$$\rho_7^0 = {}^1/_4\,|010\rangle\langle 010| + {}^1/_4\,|110\rangle\langle 110| + {}^1/_4\,|011\rangle\langle 011| + {}^1/_4\,|111\rangle\langle 111|$$

$$\rho_7^1 = {}^1/_4\,|000\rangle\langle 000| + {}^1/_4\,|100\rangle\langle 100| + {}^1/_4\,|001\rangle\langle 001| + {}^1/_4\,|101\rangle\langle 101|$$

$$\rho_7^2 = {}^1/_2\,|010\rangle\langle 010| + {}^1/_2\,|110\rangle\langle 110|$$

$$\rho_7^3 = |000\rangle\langle 000|$$

$P^7_{E^H}$	0	1	2	3
0	0	$^1/_4$	$^1/_2$	$^1/_4$
1	$^1/_4$	$^1/_2$	$^1/_4$	0
2	0	$^1/_2$	$^1/_2$	0
3	1	0	0	0

Figure 2. (top) Quantum states representing s after execution of line 7 of program P.s (bottom) The channel $P^7_{E^H}$ obtained by the POVM E^H after the execution of the 7th line of program P.

It is usual to write $K_2 = K_1 W$ to signify that K_2 is obtained by postprocessing K_1 by W. This notation is justified as the channel matrix of K_2 is equal to the product of the channel matrices of K_1 and W.

An important result both for QIF and for the study of comparison of experiments is the *Blackwell-Sherman-Stein* Theorem [3], [4], [13]. This theorem is also known as the *Coriaceous Theorem* in the QIF literature.

In our setting, the Blackwell Theorem may be stated as follows.

Theorem 2. *$K_1 \geq_{\mathrm{d}} K_2$ if, and only if, for all p_X and all g,*

$$V_g^{\mathbf{c}}(p_X, K_1) \geq V_g^{\mathbf{c}}(p_X, K_2).$$

In [7], Buscemi provided a result generalising the Blackwell theorem for quantum statistical models, which are very similar to the QQIF framework from Section 3. In this Section, we present the results by Buscemi, and translate them to the notation we have developed so far.

We first introduce some necessary definitions, adapted from [7], and show how they translate to our framework.

4.1. Quantum Statistical Models

Definition 10. *A* quantum statistical model (QSM) *is a triple $\mathbf{R} = (\mathcal{X}, \mathscr{H}, \rho^{\mathcal{X}})$, where $\mathscr{H}$ is a Hilbert space and $\mathcal{X}, \rho^{\mathcal{X}}$ are as defined in Section 3.*

Definition 11. *Given a QSM $\mathbf{R}$, an action set $\mathcal{W}$ and a gain function g, we define the* maximum expected payoff *as*

$$\$_g(\mathbf{R}) = \max_E \frac{1}{|\mathcal{X}|} \sum_{x\in\mathcal{X}} \sum_{w\in\mathcal{W}} g(w,x)\mathbf{tr}(\rho^x E_w),$$

where the maximum is taken over all possible POVMs indexed by elements in $\mathcal{W}$.

The similarity between the framework developed by Buscemi and the one we introduced in Section 3 is quite clear. The next proposition shows how much these frameworks are related.

Proposition 2. *Define a QSM $\mathbf{R} = (\mathcal{X}, \mathscr{H}, \rho^{\mathcal{X}})$, an action set $\mathcal{W}$ and a gain function g. Let p_u be the uniform distribution over $\mathcal{X}$, and $\mathcal{P}$ be all POVMs in $\mathscr{H}$. We have*

$$\$_g(\mathbf{R}) = V_{g,\mathcal{P}}(p_u, \rho^{\mathcal{X}}).$$

Proof: First, we prove that $\$_g(\mathbf{R}) \geq V_{g,\mathcal{P}}(p_u, \rho^{\mathcal{X}})$. We have

$$\begin{aligned} V_{g,\mathcal{P}}(p_u, \rho^{\mathcal{X}}) &= \max_{E\in\mathcal{P}} \sum_{y\in\mathcal{Y}} \max_{w\in\mathcal{W}} \sum_{x\in\mathcal{X}} p(x) g(w,x) \mathbf{tr}(\rho^x E_y) \\ &= \frac{1}{|\mathcal{X}|} \max_{E\in\mathcal{P}} \sum_{y\in\mathcal{Y}} \max_{w\in\mathcal{W}} \sum_{x\in\mathcal{X}} g(w,x)\mathbf{tr}(\rho^x E_y). \end{aligned}$$

Now, let w_y be an action that maximizes $\sum_{x\in\mathcal{X}} p(x)g(w,x)\mathbf{tr}(\rho^x E_y)$ for each y, and define $\mathcal{Y}_w = \{y \in \mathcal{Y} \mid w_y = w\}$. We have

$$\begin{aligned} & V_{g,\mathcal{P}}(p_u, \rho^{\mathcal{X}}) \\ =& \frac{1}{|\mathcal{X}|} \max_{E\in\mathcal{P}} \sum_{y\in\mathcal{Y}} \max_{w\in\mathcal{W}} \sum_{x\in\mathcal{X}} g(w,x)\mathbf{tr}(\rho^x E_y) \\ =& \frac{1}{|\mathcal{X}|} \max_{E\in\mathcal{P}} \sum_{w'\in\mathcal{W}} \sum_{y\in\mathcal{Y}_{w'}} \max_{w\in\mathcal{W}} \sum_{x\in\mathcal{X}} g(w,x)\mathbf{tr}(\rho^x E_y) \\ =& \frac{1}{|\mathcal{X}|} \max_{E\in\mathcal{P}} \sum_{w\in\mathcal{W}} \sum_{y\in\mathcal{Y}_w} \sum_{x\in\mathcal{X}} g(w,x)\mathbf{tr}(\rho^x E_y) \\ =& \frac{1}{|\mathcal{X}|} \max_{E\in\mathcal{P}} \sum_{x\in\mathcal{X}} \sum_{w\in\mathcal{W}} \sum_{y\in\mathcal{Y}_w} g(w,x)\mathbf{tr}(\rho^x E_y) \\ =& \frac{1}{|\mathcal{X}|} \max_{E\in\mathcal{P}} \sum_{x\in\mathcal{X}} \sum_{w\in\mathcal{W}} g(w,x)\mathbf{tr}\left(\rho^x \sum_{y\in\mathcal{Y}_w} E_y\right). \end{aligned}$$

Now, let E' maximize the double sum above, and define $E''_w = \sum_{y\in\mathcal{Y}_w} E'_y$ for each $w \in \mathcal{W}$ (if $\mathcal{Y}_w = \varnothing$, then $E''_w = 0$). Then, $E'' = \{E''_w\}_{w\in\mathcal{W}}$ is a POVM, and

$$\begin{aligned} V_{g,\mathcal{P}}(p_u, \rho^{\mathcal{X}}) &= \frac{1}{|\mathcal{X}|} \sum_{x\in\mathcal{X}} \sum_{w\in\mathcal{W}} g(w,x)\mathbf{tr}\left(\rho^x \sum_{y\in\mathcal{Y}_w} E'_y\right) \\ &= \frac{1}{|\mathcal{X}|} \sum_{x\in\mathcal{X}} \sum_{w\in\mathcal{W}} g(w,x)\mathbf{tr}\left(\rho^x E''_w\right) \\ &\leq \frac{1}{|\mathcal{X}|} \max_E \sum_{x\in\mathcal{X}} \sum_{w\in\mathcal{W}} g(w,x)\mathbf{tr}\left(\rho^x E_w\right) \\ &= \$_g(\mathbf{R}). \end{aligned}$$

To see that $\$_g(\mathbf{R}) \leq V_{g,\mathcal{P}}(p_u, \rho^{\mathcal{X}})$, just take $\mathcal{Y} = \mathcal{W}$. Then

$$\sum_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}} g(y,x)\mathbf{tr}\left(\rho^x E_y\right) \leq \sum_{y\in\mathcal{Y}} \max_{w\in\mathcal{W}} \sum_{x\in\mathcal{X}} g(w,x)\mathbf{tr}\left(\rho^x E_y\right).$$

And therefore,

$$\begin{aligned}\$_g(\mathbf{R}) &= \frac{1}{|\mathcal{X}|}\max_E \sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}} g(y,x)\mathbf{tr}\left(\rho^x E_y\right)\\ &\leq \frac{1}{|\mathcal{X}|}\max_{E\in\mathcal{P}} \sum_{y\in\mathcal{Y}}\max_{w\in\mathcal{W}}\sum_{x\in\mathcal{X}} g(w,x)\mathbf{tr}\left(\rho^x E_y\right)\\ &= V_{g,\mathcal{P}}(p_u,\rho^{\mathcal{X}}).\end{aligned}$$

□

4.2. Statistical Models and the Quantum Blackwell Theorem

Introduced by Buscemi [7], *statistical morphisms* play, in the quantum version of the Blackwell Theorem, the role played by postprocessing in the classical version. Before defining Statistical morphisms and enuciating the quantum version of Blackwell Theorem given by [7], we introduce some necessary definitions. Let $\mathcal{G}(\mathscr{H})$ be the set of density operators in $\mathscr{H}$, and $\mathcal{L}(\mathscr{H})$ the set of linear operators in $\mathscr{H}$.

Definition 12. *A family $\{F_w\}_{w\in\mathcal{W}}$ of operators over H is called a $\mathcal{W}$-test on a subset $\mathcal{G}\subset\mathcal{G}(\mathscr{H})$ iff there is a POVM $E=\{E_w\}_{w\in\mathcal{W}}$ indexed by $\mathcal{W}$ such that for all $w\in\mathcal{W}$, $\rho\in\mathcal{G}$, we have $\mathbf{tr}(\rho F_w)=\mathbf{tr}(\rho E_w)$.*

Definition 13 ([7]). *Let $\mathcal{G}\subset\mathcal{G}(\mathscr{H})$, $\mathcal{G}'\subset\mathcal{G}(\mathscr{H}')$. A linear map $L:\mathcal{L}(\mathscr{H})\to\mathcal{L}(\mathscr{H}')$ induces a statistical morphism $L:\mathcal{G}\to\mathcal{G}'$ if*

1) *for all $\rho\in\mathcal{G}$, $L(\rho)\in\mathcal{G}'$,*
2) *The dual transformation $L^*:\mathcal{L}(\mathscr{H}')\to\mathcal{L}(\mathscr{H})$ defined by trace duality (that is, $\forall A\in\mathcal{L}(\mathscr{H}'), B\in\mathcal{L}(\mathscr{H})$, $\mathbf{tr}(L^*(A)B)=\mathbf{tr}(AL(B))$) maps $\mathcal{W}$-tests on $\mathcal{G}'$ to $\mathcal{W}$-tests in $\mathcal{G}$.*

In particular, every linear CPTP map (as defined in Section 2.3) is a statistical morphism [7, Remark 8].

Given a collection of states $\rho^{\mathcal{X}}$, define the set $\mathcal{G}(\rho^{\mathcal{X}})=\{\rho^x \mid x\in\mathcal{X}\}$. We are now in position to state the Quantum version of the Blackwell Theorem.

Theorem 3 ([7, Theorem 3]). *Let $\mathbf{R}=(\mathcal{X},\mathscr{H},\rho^{\mathcal{X}})$, $\mathbf{S}=(\mathcal{X},\mathscr{H}',\sigma^{\mathcal{X}})$ be QSMs. Then,*

$$\$_g(\mathbf{R})\geq\$_g(\mathbf{S})$$

for all gain functions g if, and only if, there is a statistical morphism $L:\mathcal{G}(\rho^{\mathcal{X}})\to\mathcal{G}(\sigma^{\mathcal{X}})$ such that $\forall x\in\mathcal{X},\ L(\rho^x)=\sigma^x$.

4.2.1. The Blackwell Theorem for QQIF, and its Limitations.

In QQIF terms, statistical morphisms are transformations L with the following property. Suppose two family of states $\rho^{\mathcal{X}}$, $\sigma^{\mathcal{X}}$ such that $\forall x\in\mathcal{X},\sigma^x=L(\rho^x)$. Then, property 2 in Definition 13 guarantees that any channel obtainable by a POVM in the system given by $\sigma^{\mathcal{X}}$ is also obtainable by a suitable choice of POVM in the system $\rho^{\mathcal{X}}$. In particular, bearing in mind (6), this implies that if the adversary is able to choose any POVM, the vulnerability of $\sigma^{\mathcal{X}}$ is always going to be lower than that of $\rho^{\mathcal{X}}$.

We are now in position to state the the following corollary, which is a rewording of Theorem 3 using the quantum g-leakage framework, developed in Section 3.

Corollary 1. *Let $\mathcal{P}$ be the set of all possible POVMs. Then, there is a statistical morphism $L:\mathcal{G}(\rho^{\mathcal{X}})\to\mathcal{G}(\sigma^{\mathcal{X}})$ such that $\forall x\in\mathcal{X},\ L(\rho^x)=\sigma^x$ if, and only if, for all gain functions g and all p_X,*

$$V_{g,\mathcal{P}}(p_X,\rho^{\mathcal{X}})\geq V_{g,\mathcal{P}}(p_X,\sigma^{\mathcal{X}}).$$

Proof. The result for uniform distributions follows from Theorem 3 and Proposition 2. To see that it holds for an arbitrary distribution p_X, first eliminate states ρ^x,σ^x s.t. $p_X(x)=0$, which does not alter the value of posterior g-vulnerability. Then, the result follows by noticing that for each g, we can define $g'(w,x)=\frac{1}{p_X(x)|\mathcal{X}|}g(w,x)$, for which $\forall\rho, V_{g,\mathcal{P}}(p_X,\rho)=V_{g',\mathcal{P}}(p_u,\rho^{\mathcal{X}})$. □

One of the interesting parts of the model developed in Section 3 is that, by limiting the set $\mathcal{P}$, one can model different channels, expliciting different attacks and interests from adversaries. As presented here, many scenarios of interest can be modelled by taking a finite $\mathcal{P}$. Such a restriction means, however, that the result in Corollary 1 is not immediately applicable to the QQIF, as the completeness of $\mathcal{P}$ — that is, the fact that $\mathcal{P}$ contains all possible POVMs — is essential to the result. Without it, it is possible for a statistical morphism to increase information leakage.

To see why, consider the system with states $\rho^{x_1}=|+\rangle\langle+|$, $\rho^{x_2}=|-\rangle\langle-|$, and let $\mathcal{P}$ contain only one POVM, given by $E_{y_1}=|0\rangle\langle0|$, $E_{y_2}=|1\rangle\langle1|$. Thus, for any g, (6) yields

$$V_{g,\mathcal{P}}(p_X,\rho^{\mathcal{X}})=V_g^{\mathbf{c}}(p_x,\overline{0})=V_g(p_X),$$

where $\overline{0}(y|x)=1/2$ for all x,y (that is, $\overline{0}$ is a null channel). Therefore, this system leaks no information.

Now, consider the following CPTP

$$L(\rho)=H\rho H,\quad\text{where}\quad H=\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\1&-1\end{bmatrix}$$

i.e., the Hadamard gate [9]. By applying this transformation to the system above, we obtain a new system $\sigma^{\mathcal{X}}$ in which

- $\sigma^{x_1}=L(|+\rangle\langle+|)=|0\rangle\langle0|$,
- $\sigma^{x_2}=L(|-\rangle\langle-|)=|1\rangle\langle1|$.

Suppose that the set $\mathcal{P}$ of allowed POVMs remain unchanged. Then, we have

$$V_{g,\mathcal{P}}(p_X,\sigma^{\mathcal{X}})=V_g^{\mathbf{c}}(p_x,\overline{I}),$$

where $I(y_i|x_j)=\delta_{i,j}$, and the system has maximum leakage. In particular, if $g=g_{id}$, we have

$$V_{g_{id},\mathcal{P}}(p_X,\rho^{\mathcal{X}})=1/2\quad\text{and}\quad V_{g_{id},\mathcal{P}}(p_X,\sigma^{\mathcal{X}})=1$$

even though $\sigma^{\mathcal{X}}$ is obtained from $\rho^{\mathcal{X}}$ by a statistical morphism.

This phenomenon of a statistical morphism "increasing information leakage" is due to the limitations of the observer, not an actual increase in the amount of information in any fundamental sense. In fact, $\rho^{\mathcal{X}}$ and $\sigma^{\mathcal{X}}$ are equivalent, up to a change of basis. However, because we limit $\mathcal{P}$ to only one POVM, the adversary is only able to obtain any information on the latter system.

4.2.2. The Blackwell Theorem for QQIF is a generalisation of the Coriaceous Theorem. To see how the Coriaceous Theorem is a particular case of Corollary 1, consider now a construction of the quantum system from a classical channel similar to the one given in Section 3.2. Given a channel $K : \mathcal{X} \to \mathcal{Y}$, we let $\{|y\rangle\}_{y\in\mathcal{Y}}$ be an orthonormal basis for $\mathscr{H}$, and $\rho^x_K = \sum_y K(y|x)\,|y\rangle\langle y|$. This time, however, instead of restricting only to the POVM $E_y = |y\rangle\langle y|$, we let $\mathcal{P}$ be the set of all possible POVMs. We claim that, for all p_X and g

$$V_{g,\mathcal{P}}(p_X, \rho^{\mathcal{X}}_K) = V^{\mathbf{c}}_g(p_X, K).$$

In fact, let $\{E'_z\}_{z\in\mathcal{Z}}$ be any POVM, indexed by some finite set $\mathcal{Z}$. Then, by the discussion in Section 3.3, the channel induced by ρ^x_K and E' is given by

$$\begin{aligned}
K_{E'}(z|x) &= \mathbf{tr}(\rho^x_K E'_z)\\
&= \sum_y \langle y| \left(\sum_{y'} K(y'|x)\,|y'\rangle\langle y'|\right) E'_z\,|y\rangle\\
&= \sum_y \sum_{y'} K(y'|x)\,\langle y|y'\rangle\langle y'|\,E'_z\,|y\rangle\\
&= \sum_y K(y|x)\,\langle y|\,E'_z\,|y\rangle\\
&= \sum_y K(y|x)\mathbf{tr}(|y\rangle\langle y|\,E'_z)\\
&= \sum_y K(y|x) R_{E'}(z|y)
\end{aligned}$$

where $R_{E'}$ is a channel, given by $R_{E'}(z|y) = \mathbf{tr}(|y\rangle\langle y|\,E'_z)$ (i.e., the channel that gives the probability of output z given that the state is $|y\rangle\langle y|$). Thus, $K \geq_{\mathrm{d}} K_{E'}$ for any POVM $E' \in \mathcal{P}$, and (6) yields

$$V_g(p_X, \rho^{\mathcal{X}}_K) = \max_{E\in\mathcal{P}} V^{\mathbf{c}}_g(p_X, K_E) = V^{\mathbf{c}}_g(p_X, K). \quad (7)$$

As hinted by (7), whenever the states are constructed as above, the QQIF system behaves as in classical QIF. This result is congruent to Remark 3 and Postulate 1 in [7], which states that abelian QSMs are equivalent to classical models — which thus shows that Theorem 3 is a generalisation of the classical Blackwell Theorem.

For clarity, in Proposition 3 below, illustrating how the classical Blackwell Theorem can be obtained from its quantum version.

Proposition 3. *Let $K : \mathcal{X} \to \mathcal{Y}$, $W : \mathcal{X} \to \mathcal{Z}$ be classical channels, $\{|y\rangle\}_{y\in\mathcal{Y}}$, $\{|z\rangle\}_{z\in\mathcal{Z}}$ be orthonormal basis of two hilbert spaces and let $\rho^{\mathcal{X}}_K$, $\sigma^{\mathcal{X}}_W$ be defined, for all $x \in \mathcal{X}$, as*

$$\begin{aligned}
\rho^x_K &= \sum_y K(y|x)\,|y\rangle\langle y|\,,\\
\sigma^x_W &= \sum_z W(z|x)\,|z\rangle\langle z|\,.
\end{aligned}$$

Then, the following statements are equivalent

1) $K \geq_{\mathrm{d}} W$,
2) *There is a statistical morphism L such that $\forall x \in \mathcal{X},\ L(\rho^x_K) = \sigma^x_W$,*
3) *For all p_X and g, $V_g(p_X, \rho^{\mathcal{X}}_K) \geq V_g(p_X, \sigma^{\mathcal{X}}_W)$,*
4) *For all p_X and g, $V^{\mathbf{c}}_g(p_X, K) \geq V^{\mathbf{c}}_g(p_X, W)$.*

Proof: (1⇒2) Suppose $KR = W$ for some channel $R : \mathcal{Y} \to \mathcal{Z}$. Let $A_{yz} = \sqrt{R(z|y)}\,|z\rangle\langle y|$, and define L as

$$L(\rho) = \sum_{y,z} A_{yz}\rho A^\dagger_{yz}.$$

Notice that $A^\dagger_{yz} = \sqrt{R(z|y)}\,|y\rangle\langle z|$. Thus, for all $x \in \mathcal{X}$

$$\begin{aligned}
L(\rho^x_K) &= \sum_{y,z} A_{yz}\rho^x_K A^\dagger_{yz}\\
&= \sum_{y,z} A_{yz} \left(\sum_{y'} K(y'|x)\,|y'\rangle\langle y'|\right) A^\dagger_{yz}\\
&= \sum_z |z\rangle\langle z| \sum_y R(z|y)K(y|x)\\
&= \sum_z W(z|x)\,|z\rangle\langle z| = \sigma^x_W
\end{aligned}$$

where the third equality follows from $\langle y|y'\rangle = \delta_{y,y'}$. Moreover,

$$\begin{aligned}
\sum_{y,z} A^\dagger_{y,z}A_{y,z} &= \sum_{y,z} R(z|y)\,|y\rangle\langle z|z\rangle\langle y|\\
&= \sum_y |y\rangle\langle y| \sum_z R(z|y)\\
&= \sum_y |y\rangle\langle y| = I
\end{aligned}$$

where the penultimate equality follows from R being a channel. Thus, L is a CPTP, and therefore a statistical morphism.

(2⇒3) Follows from Theorem 3.

(3⇒4) From (7), the inequality

$$V_{g,\mathcal{P}}(p_X, \rho^{\mathcal{X}}_K) \geq V_{g,\mathcal{P}}(p_X, \sigma^{\mathcal{X}}_W)$$

is equivalent to the inequality

$$V^{\mathbf{c}}_g(p_X, K) \geq V^{\mathbf{c}}_g(p_X, W).$$

(4⇒1) Follows from Theorem 2. □

5. Related Literature

In most of its early works, QIF made use of information-theoretic measures, such as Shannon entropy [14], [15], min-entropy [16], [17] and guessing entropy [18]. Introduced in [5], the g-leakage framework proved itself to be a versatile manner to compute leakage of information in a myriad of scenarios. The posterior development of the framework made it clear the connection between QIF and the theory of statistical decisions, as shown by the independent proof in [6] of the Blackwell-Sherman-Stein Theorem.

Beyond Buscemi's work [7] the problem of finding an optimal measurement that minimises a quantum statistical decision theory problem was already discussed in [19], [20]. More recently, Konig et al [21] proved that conditional quantum min-entropy [11] over a classical-quantum state (as in Remark 1) is equal to the negative logarithm of the probability of guessing with an optimal POVM — that is, to $-\log V_{g_{id},\mathcal{P}}$ where $\mathcal{P}$ is the set of all POVMs.

5.1. Conclusions and further work

The main objective of this work was to give a first generalisation of quantitative information flow framework to the quantum setting. Having established a connection with quantum statistical decision theory allows leveraging of recent results and tools from quantum information theory.

This connection is a first attempt into exploring a possible generalisation of the g-leakage framework, which has been widely adopted in QIF, to quantum settings. The general framework here presented aims to analyze and quantify confidentiality in quantum systems.

Future work might explore such applications, and study essential quantum phenomena like entanglement of states, non locality, teleportation within the developed framework.

References

[1] F. Arute *et al.*, "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, pp. 505–510, 2019.

[2] G. Dahl, "Matrix majorization," *Linear Algebra and its Applications*, vol. 288, pp. 53 – 73, 1999.

[3] S. Sherman, "On a theorem of Hardy, Littlewood, Polya, and Blackwell," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 37, no. 12, pp. 826–831, 1951.

[4] D. Blackwell, "Equivalent comparisons of experiments," *The Annals of Mathematical Statistics*, vol. 24, no. 2, pp. 265–272, 1953.

[5] M. S. Alvim, K. Chatzikokolakis, C. Palamidessi, and G. Smith, "Measuring information leakage using generalized gain functions," in *Proc. IEEE 25th Computer Security Foundations Symposium (CSF)*, 2012, pp. 265–279.

[6] A. McIver, C. Morgan, G. Smith, B. Espinoza, and L. Meinicke, "Abstract channels and their robust information-leakage ordering," in *Proc. 3rd Int. Conf. Principles of Security and Trust (POST)*, ser. LNCS, vol. 8414. Springer, 2014, pp. 83–102.

[7] F. Buscemi, "Comparison of quantum statistical models: Equivalent conditions for sufficiency," *Communications in Mathematical Physics*, vol. 310, no. 3, pp. 625–647, 2012.

[8] M. S. Alvim, K. Chatzikokolakis, A. McIver, C. Morgan, C. Palamidessi, and G. Smith, "An axiomatization of information flow measures," *Theoretical Computer Science*, vol. 777, pp. 32 – 54, 2019, in memory of Maurice Nivat, a founding father of Theoretical Computer Science - Part I.

[9] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.

[10] A. S. Holevo, *Quantum Systems, Channels, Information*. Berlin, Boston: De Gruyter, 2012.

[11] R. Renner, "Security of quantum key distribution," Ph.D. dissertation, ETH Zurich, Zürich, 2005.

[12] T. M. Cover, "Broadcast channels," *IEEE Transactions on Information Theory*, vol. 18, no. 1, pp. 2–14, 1972.

[13] D. Blackwell, "Comparison of experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, J. Neyman, Ed. Berkeley: Univ. of California Press, 1951, pp. 93–102.

[14] D. Clark, S. Hunt, and P. Malacaria, "Quantitative information flow, relations and polymorphic types," *J. Log. and Comput.*, vol. 15, no. 2, pp. 181–199, Apr. 2005.

[15] K. Chatzikokolakis, C. Palamidessi, and P. Panangaden, "Anonymity protocols as noisy channels," *Information and Computation*, vol. 206, no. 2, pp. 378 – 401, 2008.

[16] G. Smith, "On the foundations of quantitative information flow," in *Proc. 12th Int. Conf. Foundations of Software Science and Computational Structures (FOSSACS)*, ser. LNCS, vol. 5504. Springer, 2009, pp. 288–302.

[17] M. Boreale, F. Pampaloni, and M. Paolini, "Asymptotic information leakage under one-try attacks," *Mathematical Structures in Computer Science*, vol. 25, no. 2, pp. 292–319, 2015.

[18] B. Köpf and D. A. Basin, "An information-theoretic model for adaptive side-channel attacks," in *Proc. of CCS*. ACM, 2007, pp. 286–296.

[19] A. Holevo, "Statistical decision theory for quantum systems," *Journal of Multivariate Analysis*, vol. 3, no. 4, pp. 337 – 394, 1973.

[20] H. Yuen, R. Kennedy, and M. Lax, "Optimum testing of multiple hypotheses in quantum detection theory," *IEEE Transactions on Information Theory*, vol. 21, no. 2, pp. 125–134, March 1975.

[21] R. Konig, R. Renner, and C. Schaffner, "The operational meaning of min- and max-entropy," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 4337–4347, Sep. 2009.