Optimal Eavesdropping in Quantum Cryptography

A Thesis

Submitted in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

in Computer Science

Author:
Atanu Acharyya

Supervisor:
Dr. Goutam Paul

Submitted on: September 2021

Applied Statistics Unit
Indian Statistical Institute
203, B. T. Road, Kolkata - 700 108.

LIST OF PUBLICATIONS

[1] Atanu Acharyya and Goutam Paul, *"Revisiting optimal eavesdropping in quantum cryptography: Optimal interaction is unique up to rotation of the underlying basis,"* in Physical Review A (American Physical Society), article 022326, volume **95**, issue 2, February **2017**, doi:10.1103/PhysRevA.95.022326.

[2] Atanu Acharyya and Goutam Paul, *"A complete characterization of the optimal unitary attacks in quantum cryptography with a refined optimality criteria involving the attackers Hilbert space only,"* in European Physical Journal D (EPJ D) , volume **75**, issue 8, pages 215, July **2021**, doi:10.1140/epjd/s10053-021-00203-7. arXiv:1712.04475 as of 2019.

Paper-to-Chapter mapping:

- Reference [1] above, i.e., [AP17], is mainly used in Chap. 4. Moreover, some of its results are used in Chap. 7 and Chap. 3 as well.
- Reference [2] above, i.e., [AP21], is mainly used in Chap. 5, and Chap. 6. Moreover, some of its results are also used in Chap. 7 and Chap. 3.

Note: Reference [2] above actually combines two separate works:

- Finding a necessary and sufficient condition for optimality.
- Characterizing optimal unitary evolutions.

For the urge of publication within the tenure, we had to merge these two works into a single paper, as evident from the arxiv developments.

CHAPTERWISE CONTRIBUTIONS

Here we explicitly mark the chapters as fully or partially contributory. *Fully contributory chapters* represent a publication of our work. *Partially contributory chapters* either have some contents from the published works, or have contents that couldn't be fitted into any publication. Note that, Ref. [AP21] consists of two separate works.

Chap. 3	Existing works on Optimal Eavesdropping	Partly Contributory, , Ref. [AP17, AP21]
Chap. 4	Characterizing the optimal Interactions [AP17]	Fully Contributory, Ref. [AP17]
Chap. 5	Necessary and sufficient conditions [AP21]	Fully Contributory, Ref. [AP21]
Chap. 6	Characterizing the optimal unitary evolutions [AP21]	Fully Contributory, Ref. [AP21]
Chap. 7	Post-processing and Comparative study	Mostly Contributory, Ref. [AP17, AP21]
Chap. 8	Coherent eavesdropping	Semi-contributory, no publication.

Contributory Sections in partially contributed Chapters:

1. In Chap. 3, some of the contributory sections are as follows: Sec. 3.3.1.6 , Fig. 3.2, Sec. 3.3.11; Sec. 3.4.1,3.4.2, 3.4.6.3 on optimal POVM; Sec. 3.4.7, 3.4.8 and Eq. (3.4.9) on succ-prob; Sec. 3.4.10 to illustrate the proofs, etc. A part of these results can be found in [AP17].
2. The existing literature is found quite unorganized and sometimes contradictory to consolidate certain aspects which are discussed in Chap. 7. We have tried to address those things to complement out first work, and a small part of it can be found in [AP21].
3. In Chap. 8, Sec. 8.4.1 explains some possible alternative approaches. The chapter as such has explained a paper elaborately that is not found in the literature and it bothered a lot to achieve it. To prove each equation of that paper indeed was a challenge.

to

my belated father,

my mother,

my belated maternal uncle,

my Maths teacher,

… …

ACKNOWLEDGMENT

I would like to thank my supervisor to retain patience at the final stage. Some delay due to consistent obstacles (including accident or death of close ones, and some funded disturbances) started introducing some question marks regarding my final outcome. Thankfully, my mother didn't entirely loose faith on me regarding this.

I must thank Prof. Tapas Samanta for his elderly advice pertaining to RFAC (research fellow advisory committee) related matters. Even I had to approach Prof. Bimal Roy in some critical situations.

I would like to thank my colleagues for whatever way they could extend their help. Particularly, I had to take some dictation help while latexing some of my hand-written works since I had been in trouble due to partial blockages with my lateral vision.

Last but not the least, thanks must go to those funded disturbances around my accommodation, or workplaces, or at other whereabouts. It helped me to know partly the back-end factors better than I could know in its absence. That knowledge allowed me to excel in turbulent situations despite the blockages incurred.

ABSTRACT

Quantum key distribution (QKD) has raised some promise for more secured communication than its classical counterpart. It allows the legitimate parties to detect eavesdropping which introduces error in the channel. If disturbed, there are ways to distill a secure key within some threshold error-rate. The amount of information gained by an attacker is generally quantified by (Shannon) mutual information. Knowing the maximum amount of information that an intruder can gain is important for post-processing purposes, and we mainly focus on that side in the thesis. Rényi information is also useful especially when post-processing is considered.

The scope of this thesis is to first describe some relevant ingredients for QKD and then study some open-ended issues. We mostly focus on the BB84 protocol and some issues relating optimal eavesdropping on it when each information-carrying particles are attacked individually. However, our results and techniques can also be applied for other protocols and different eavesdropping strategies. We felt a few other eavesdropping techniques worthy to analyze in that line, despite limitations to achieve newer results.

First we study the optimal eavesdropping technique on the BB84 protocol and show that the optimal information can be achieved in infinitely many different ways to interact and measure the information-carriers. Although they are mathematically equivalent in some sense, that variety may help when designing the eavesdropping setup.

However, it was not clear whether more such optimal interactions exist or not. This has lead us to derive them through a chain of necessary and sufficient conditions (NSC), which are shown to be in a one-to-one correspondence with the earlier interactions. In this process we arrive at a new NSC restricting attackers particles to a specific orientation, establishing the geometry of the attack more explicitly than earlier. Some explicit connections are shown with other modes of gleaning information like cloning.

Nevertheless, for practical purposes all an attacker requires is the evolution that entangles her ancilla with the senders particle, and the corresponding measurement that will

lead her to optimal information gain. This is generally neglected in the literature as they exhibit a specific interaction. In our case, having infinitely many options to interact, we felt it better to address the issue of findings optimal evolutions.

Overall, we have added more mathematical structures in the framework of optimal eavesdropping. We wanted to analyze the more generalized ways to attack, where a whole chunk of information-carrying particles can be evolved and then measured at a go. The process becomes complex to tackle when the chunks go bigger. Yet, we have explained the mathematical details of some of the existing results to point out the difficulties.

Contents

List of Figures

List of Tables

ABBREVIATIONS-I

QKD	Quantum Key Distribution
cbit	classical bit
qubit	quantum bit
IP	inner product
p&m	prepare and measure
eb	entanglement based
MUB	Mutually Unbiased Bases
IR	Intercept-Resend
BB84	Bennett-Brassard-1984 (protocol)
E91	Ekert-1991 (protocol)
B92	Bennett-1992 (protocol)
QBER	Quantum Bit Error Rate
POVM	Positive Operator Valued Measure
POM	Probability Operator Measure
OW	One Way
CPP	Classical Post Processing

ABBREVIATIONS-II

IS	Initial State
PIJS	Post-Interaction Joint State
IV	Interaction Vector
KG	Knowledge Gain (IG, MI etc.)
IG	Information Gain
MI	Mutual Information
bpsp	bits per sifted-photon
NSC	Necessary and Sufficient Condition
C-NOT	controlled NOT
LOCC	Local Operation and Classical Communication
CHSH	Clauser-Horne-Shimony-Holt
EPR	Einstein-Podilsky-Rosen
iff	if and only if
w.r.t.	with respect to
s.t.	such that
t.p.t.	to prove that
$:=$	defined as

GREEK LETTERS WITH PRONUNCIATION

Character	Name	Pronounce	Character	Name	Pronounce
α	alpha	*AL-fuh*	ν	nu	*NEW*
β	beta	*BAY-tuh*	ξ, Ξ	xi	*KSIGH*
γ, Γ	gamma	*GAM-muh*	o	omicron	*OM-uh-CRON*
δ, Δ	delta	*DEL-tuh*	π, Π	pi	*PIE*
ε	epsilon	*EP-suh-lon*	ρ	rho	*ROW*
ζ	zeta	*ZAY-tuh*	σ, Σ	sigma	*SIG-muh*
η	eta	*AY-tuh*	τ	tau	*TOW (as in cow)*
θ, Θ	theta	*THAY-tuh*	υ, Υ	upsilon	*OOP-suh-LON*
ι	iota	*eye-OH-tuh*	ϕ, Φ	phi	*FEE, or FI (as in hi)*
κ	kappa	*KAP-uh*	χ	chi	*KI (as in hi)*
λ, Λ	lambda	*LAM-duh*	ψ, Ψ	psi	*SIGH, or PSIGH*
μ	mu	*MEW*	ω, Ω	omega	*oh-MAY-guh*

Capitals shown are the ones that differ from Roman capitals.

CHAPTER 1

INTRODUCTION

The thesis provides a structured mathematical model for optimal eavesdropping for quantum key distribution (QKD). The optimal attack is characterized completely in the form of optimal post-interaction states, optimal measurements, and the optimal unitary evolution. Some connections with other aspects of QKD became more explicit here.

The power of quantum theory [Eng13] is better manifested through the manipulation of quantum information for secure electronic communication [ZBB$^+$05]. There is a vast literature developed over time on QKD. For a first reader interested to get a broad survey on the applications of QKD in quantum cryptography can read [ABB$^+$14] and a the related ones. That quantum cryptography is possible in practice with a single photon source is given by [BBG$^+$02]. Leaving those aside, let's get directly into the thesis content.

Thesis organization

We have first discussed the preliminary ideas in Chap. 2 that appeared useful for our purpose, the basic QKD protocols and eavesdropping models are also introduced therein. An optimal attack due to [FGG$^+$97] is explained elaborately in Chap. 3.

In Chap. 4 we have shown that the optimal information can be attained in infinitely many ways. We proved it further as a necessary and sufficient condition in Chap. 5. The corresponding optimal unitary evolutions are characterized completely in Chap. 6. Some of the relevant ingredients of Chap. 4 are moved in Chap. 3 to retain the flow of ideas.

We provide a basic sketch of the ingredients for classical post-processing in Chap. 7, where we also compare the bipartite informations across protocols. Finally, we have explained the notion of the generalized attack as a coherent model in Chap. 8.

We conclude the thesis work in Chap. 9 by discussing the motivations behind choosing the problems while summarizing the results, and by mentioning some further scopes.

1.1 Characterizing the Optimal Interactions

An eavesdropper can perform the interaction in many ways, but not all of them are optimal. We are concerned about the optimal interactions that provide maximum possible (Shannon) information. Given an interaction, she needs a suitable measurement to extract that information. Again, not all measurements provide the maximum information of the signal. The ones providing maximum information consists the optimal measurements. Thus, characterizing the optimal interactions includes the specification of the optimal measurement along with the interaction. However, when an interaction is optimal, that information is extractable irrespective of whether we know the right measurement or not.

Once she knows the reconciled bases, a suitable measurement provides some outcome from which she can infer the identity of Alice's signal using some strategy. That inference is also an integral part of the characterization process.

The maximum information can be obtained by various interactions. We have done the characterization to the maximum extent [AP17, AP21]. However, the difficulty to derive such interaction lies in the fact that specification of each optimal IV should also accompany the specification of its optimal measurement and there could be infinitely many such optimal IVs. However, the trick to tackle the problem is not too difficult once get noticed. First, we noted that the measurement setup is integral part of the interactions. Thus, if we express the IVs w.r.t. measurement directions, then we can calculate the probability components which in turn defines the information. To start the derivation with a general expression of the IVs, It is helpful to note that Eve's IVs form two mutually orthogonal sets. The rest is merely calculations to compare the corresponding IGs with the optimal value and figure out the optimal IVs. They also satisfy the necessary and sufficient condition for optimality, as expected.

In an arbitrary measurement basis, all the optimal IVs look same. Once we specify a measurement basis, we get the corresponding optimal IVs. Various choices of measurement setup provides variety of optimal IVs. The two specific instances from [FGG$^+$97] are two such special cases of our generalized expression. For the IVs [FGG$^+$97] with equal rates, no optimal measurement was specified. Although one can calculate them by diagonalizing the corresponding observable, it follows directly once we compare those with our general form.

Unlike [FGG$^+$97], we do not need a separate analysis for equal or unequal error rate. We did the derivation for unequal error rates to cover the general scenario, while the equal error cases follow trivially from it. In practice, at Bob's end, the disturbance in the channel may not appear same across the bases unless all the experimental setups including the channel are perfect. That delta differences could be covered easily with our generalized approach.

1.2 A new necessary and sufficient condition for optimality and deriving optimal interaction vectors

Although the above-mentioned approach identify infinitely many optimal interactions, those didn't appear out of a necessary and sufficient condition which might certify the population to be exhaustive. With this dissatisfaction, we started further with the NSC in [FGG$^+$97] and derived a chain of NSCs for optimality which finally provided the expression of the optimal IVs. However, the expression didn't look same as those in [AP17], and raised a question whether newer IVs could be found. But, we could establish an one-to-one correspondence between the populations in [AP17] and those in [AP21].

In the above derivation, as a byproduct, we got a special NSC that mentions the restriction on the spatial orientation of Eve's IVs to be optimal. Note that the NSC in [FGG$^+$97] involves both Alice and Eve's Hilbert space, while our byproduct involves only Eve's Hilbert space. The two unequal overlaps should be same and equal to the difference between the fidelity and disturbance introduced in the Alice-Bob's channel due to eavesdropping. This optimal overlap is quite important as it has a deep connection with the amount of CHSH violation for an equivalent entanglement-based protocol. A similar connection is found with cloning mechanism as well.

1.3 Characterizing the optimal unitary evolutions

All we found so far in the name of optimal interactions are only Eve's optimal IVs. However, for practical purposes, all an attacker need to know is the unitary evolution on Alice-Eve's joint system that she needs to perform during the interaction. The associated optimal measurement is already there. Given a specific set of IVs, getting the corresponding unitary evolution may not be that difficult, one can always try numerical approach. However, once the expressions are parametric, or more general to include all of them in a generalized measurement basis, It is worthy to find the unitary as a generalized form which we have done indeed.

Getting the initial motivation to tackle that generalized approach was not so easy. We could come up with a hack to get the optimal unitary that works with a specific initial state (IS) of Eve's ancilla. Interestingly, this particular unitary appeared the simplest one out of all the others that we could comprehend. However, even for this particular IS, there could be infinitely many other optimal unitaries (up to some unitary factor) that produces the same IVs: the unitary factor was also specified then. This is due to an interesting observation that an optimal unitary is actually the product of two unitaries: one rotates the space of the IS, while the other rotates the space of the post-interaction joint states. The later (PIJSs) actually is described in terms of the measurement directions. Thus, it supports the intuition that the optimal unitary has two degrees of freedom: the choice of the IS (its orientation) and the choice of the measurement directions. That is to say, an unitary evolution automatically specifies the IS and the measurement, and vice versa (although this direction has various alternate choices of the unitary). For the sake of theoretical completeness, we have shown how the choice of a different IV or the choice of a different measurement can easily be tackled from a known case (e.g., the initial hack that we had).

The work may look merely a theoretical framework, but, getting an (Unitary evolution, IS, measurement) specification suitable to design for practical purposes is altogether a different challenge. The majority of its theoretical backbone is provided here.

1.4 Post-processing and Comparative study

Once the legitimate parties (Alice and Bob) have arrived with a sifted key, they look for the possibility to filtrate from it a shared secret on which Eve has virtually no knowledge. The possibility depends on a threshold disturbance. So far they are within the threshold, they can go further to post-process it with the help of a classical channel.

They can identify the disturbance level by considering a random subsequence of their bit-stream and publicly tally the bits. Without eavesdropping, they should always agree with their bits. But, if eavesdropped, some of the bits at Bob's end will flip, and are in disagreement with those of Alice. The fraction of mismatches will provide them a rough estimate of the error rate. Depending on that rate, they can design their post-processing methodology.

Let's think about the remnant sifted key having some possible disagreement without a knowledge of where could they be. The objective is to identify the locations and either remove those corrupted bits or correct them. If they publicly discuss the whole content, then Eve may listen to that as well. However, they can do a bit better. They may divide the string into blocks (possibly containing one error at most) and tally the XOR (parity) of the blocks publicly. If there is an error in a block, the XOR will disagree. In that case, they can perform a binary search to locate the error. Thus, with some sacrifice, they can reconcile into a common string from the disagreed strings.

Although their key-strings are now in sync, Eve might have gained some knowledge due to her measurements and from later public discussions. The legitimate parties should eliminate that knowledge before they call it a shared secret. In this case, they can simply replace various blocks of the sting by the XOR value (on which Eve has literally no knowledge). In practice, there are better efficient (may look complex) methods to do the task.

The key-rate (ratio of the final key to that of the sifted key) depends on the disturbance level and the method that they employed. However, there are theoretical limits that they can achieve at best within their capacity. The literature seems a bit cryptic to address those ends with clarity while some differences are observed. We have tried to address some of the parameters within our capacity.

1.5 Coherent eavesdropping

There are various varieties with the eavesdropping techniques. One of the advanced model is to attack a whole chunk of information carrying particles (qubits) at a go. She may evolve them jointly, measure them jointly if possible, and even may defer the measurement till she gets the full knowledge of all the public discussions. This coherent way of learning the key is discussed to some extent. Mostly, we addressed some derivations minutely where we could feel difficulties. The main challenge remains to classify the number of free parameters to describe the unitary, as those defined the success probability, mutual information, key-rate etc. The description of the unitary ultimately boils down to the description of different overlaps between Eve's post-interaction states. The complexity increases with the size of the chunk to be attacked. To have tried to exhibit the difficulty during our involved calculations. We have considered the 4s protocol throughout to allow the similarities and differences to be noticed. The analysis could easily be extended for the 6s protocol as well.

We have explained the attack on a chuck of size two. An unitary evolution entangles her ancilla with Alice's qubit pairs. Thus, an incoming chunk can get into a superposition of all the four possible states, and so does Eve's ancilla. The four different states with Alice can then produce sixteen (16) different post-interaction states for Eve and four (4) for Bob. Eve is thus left with 256 overlaps among her states (compared to only four for incoherent attack). However, mutual orthogonality favors the initial step to reduce the number of different overlaps. For a given chunk of Alice, all the 4 states with Eve are mutually orthogonal. Further, all the 16 states can be classified into four groups based on the number of errors introduced and their location. Within each group, some of the overlaps may be same. If the attack model is considered same error across the bases, the rules of symmetry help reduce the number of parameters further. While attacking a two-qubit chunk, the unitary is ultimately characterized by only five (5) real parameters. We have discussed that reduction process with minute details in the thesis.

A 2-qubit attack on the 6s protocol is characterized by only 2 real parameters due to more symmetry in the protocol itself. A 3-qubit attack doesn't improve Eve's Shannon information or her chance for correct guess except that she learns fully the states that Bob receives undisturbed.

CHAPTER 2

BACKGROUND KNOWLEDGE

Here we discuss some of the relevant basic ideas from quantum mechanics and quantum information alongwith the mathematical pre-requisites. One can consult the book [NC11], while *wikipedia* is always a good resource to get a first impression on such topics.

Quantum information is described by the states of a quantum system which in turn is represented by a wave-function in quantum mechanics. A quantum measurement allows an wave-function to collapse into a specific quantum state that we ultimately observe to study the system. Over time, depending on the nature of the environmental perturbations, the system is evolved towards a newer wave-function. Further measurement reveals a state of the evolved system.

The basics of quantum mechanics typically starts with a few postulates. However, to understand these postulates, one needs to have a minimal knowledge on the mathematical pre-requisites which we discuss first.

2.1 Elements of quantum information processing

The classical methods of electronic communication restricts to cbits (classical bits) as the information carrier. However, the advent in quantum mechanics has raised the communication prospect by introducing qubits (quantum bits) as the information carrier that can encode multiple cbits in superposition – a quantum measurement can then extract the classical information out of the transmission.

The role of quantum information processing is all about to store, process (evolve), and retrieve (measure) the quantum information.

2.1.1 Qubits, Unitary, Gates, Measurements

A qubit is a quantum system that can encode the classical bits 0, 1 as a pair of mutually orthogonal normalized quantum states $|0\rangle := (1,0)^T, |1\rangle := (0,1)^T$, respectively. These two quantum states form a computational basis in a two dimensional space. However, a quantum system can do more – it can conceal both the cbits simultaneously in a quantum state as a superposition of these two basis states.

$$|\psi\rangle \quad = \quad \alpha_0|0\rangle + \alpha_1|1\rangle.$$

The co-efficients α_0, α_1, which can consume a complex value in general, are the amplitudes of the two basis states satisfying $\sum_{b=0}^{1} |\alpha_b|^2 = 1$.

When a quantum measurement is performed, for instance in the computational basis itself, it collapses to one of these basis states, leading to the corresponding cbits. It is not known apriori which of the two outcomes it will produce. However, the frequency to get an outcome is already known if the description of the state is known. A measurement will reveal the state $|b\rangle$ with probability $|\alpha_b|^2$. Measurements in succession on the copy of such a state will reveal this statistics for large number of trials.

A qubit is physically the state of an elementary particle, like polarized photon, spin of electron etc. For communication purposes, for instance, optical fibers serve the purpose to carry the photons.

One can prepare a quantum register to store multiple such states. For instance, a 3-qubit register storing $|011\rangle$ which in turn is a tensor product $|0\rangle \otimes |1\rangle \otimes |1\rangle$ of three basis states, will encode the classical state 011 = 3. However, the power of quantum mechanics allows the register to store multiple such classical states simultaneously. For instance, the 3-qubit register in state $\frac{1}{\sqrt{2}}(|011\rangle + |101\rangle)$ will store both the classical states 3 and 5 simultaneously. Similarly, one can store all the classical digits from 0 to 7 as a superposition $\sum\limits_{b=0}^{7} |b\rangle$ (ignored scaling) of 8 different basis states $|0\rangle$ to $|7\rangle$ in a single register.

The strength of quantum information processing lies in its ability to start from an elementary basis state and evolve it unitarily into a superposition of more and more basis states (orthonormal) that could later be used to perform quantum computations and measurements. The unitary evolution is done by elementary quantum logic gates. The most common such gate is the Hadamard gate that transforms the computational basis states $|0\rangle, |1\rangle$ into the Hadamard states $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, respectively. It creates superposition, which is better understood involving many qubits. For instance, one can start with $|00\rangle$ as a basis state and apply the Hadamard on both the qubits to generate a superposition $|0\rangle + |1\rangle + |2\rangle + |3\rangle$ of all the four basis states. In general, one may start with an n-cbit register and grow it to a state to store all the 2^n different cbits simultaneously in superposition. A measurement in the computational basis on each register will reveal one of these 2^n digits.

Creating entanglement: A powerful aspect of quantum systems (two or more) is entanglement. For instance, a 2-qubit state can be prepared into the following superposition $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ which cannot be written as a tensor product of two single-qubit states. That is, the information in the two registers are no more independent of each other.

Hadamard alone cannot serve the purpose, it only creates a superposition within a single qubit. To entangle the two qubits, one needs a 2-qubit quantum gate like $\mathsf{C\text{-}NOT}$, i.e., the controlled-NOT gate that flips the second bit (target) only when the first bit (control) is Boolean YES (=1).

$$\mathsf{C\text{-}NOT} \quad : \quad |c\rangle|t\rangle \quad \mapsto \quad |c\rangle|t+c\rangle.$$

Clearly, until the control bit is in superposition, we won't get the entanglement. The superposition is done by the Hadamard gate. Thus,

$$|0\rangle|0\rangle \quad \xmapsto{\mathbb{H}\otimes\mathbb{1}} \quad \tfrac{1}{\sqrt{2}}(|0\rangle+|1\rangle)|0\rangle \quad \xmapsto{\mathsf{C\text{-}NOT}} \quad \tfrac{1}{\sqrt{2}}(|00\rangle+|11\rangle).$$

It is tempting to conclude at the first place that a $\mathsf{C\text{-}NOT}$ gate can successfully make a carbon copy of a given quantum state owing to the following:

$$\mathsf{C\text{-}NOT} \quad : \quad |c\rangle|0\rangle \quad \mapsto \quad |c\rangle|c\rangle, \quad \text{for} \quad c = 0, 1.$$

However, it fails to copy an arbitrary state $\alpha|0\rangle + \beta|1\rangle$ in superposition to its replica, rather it entangles them. This is described in the well-known *no cloning theorem* [WZ82], that allows quantum communication to identify an attempt for eavesdropping, which is not possible in the classical domain.

2.2 Mathematical framework

2.2.1 Hilbert space

A *Hilbert space* is a *real* or *complex vector space* having an *inner product* and is *complete*.

For instance, the space $\mathbb{C}^n$ of n-dim Complex numbers. We'll mostly deal with $\mathbb{C}^2$, realizing the state of a qubit.

We do not need much to bother about the second criterion for our purposes, so we just address it informally. Completeness means that every *Cauchy sequence* of its elements converges to a unique element (*limit point*) of that space under consideration.

Completeness in general is defined on a set having some *norm* (*metric*) to measure distance between any two elements of the set. For instance, consider the closed interval

$[0,1] \in \mathbb{R}$ with *Euclidean distance*. The *sequence* $\{1-\frac{1}{r}\}_{r=2}^{\infty}$ is Cauchy that converges to the limit point 1 which indeed belongs to the set. It can be verified that any such Cauchy sequence defined in this interval indeed converges to an element of that set, certifying completeness.

Triangle inequality: For any three elements x,y,z in a set with a *metric* 'dist',

$$\mathsf{dist}(x,z) \quad \leq \quad \mathsf{dist}(x,y)+\mathsf{dist}(y,z).$$

Remember the usual triangle inequality for three 2-dim vectors in $\mathbb{R}^2$. One can extend it for usual *dot products* in $\mathbb{R}^n$, and for *complex inner product* in $\mathbb{C}^n$.

Cauchy-Schwarz inequality: Any two elements x,y of an *inner product space* (IPS) or a *Hilbert space* (HS) satisfy [Ste04, WW]

$$|\langle x|y\rangle| \quad \leq \quad ||x||\,||y||\,.$$

2.2.2 Quantum operators: Hermitian, Unitary, Normal, positive operators

An operator is a rule to transform a function to another function, e.g., derivative operator. In quantum mechanics, it transforms kets to kets, thereby is expressed as a matrix.

Observables are physical properties that can be measured, e.g., momentum. There is an operator associated with an observable.

For our purposes, operators are matrices. Some useful operators are defined here.

Hermitian matrix A Hermitian (or self-adjoint) matrix is a complex square matrix that is equal to its own conjugate transpose. Both $A^{\dagger}, A^{*}$ are used to denote conjugate transpose (i.e., $\overline{A^{\mathsf{T}}}$) of A.

Thus, in a Hermitian matrix, the ij-th element and the ji-th element are the complex conjugate of each other.

$$A \text{ is Hermitian} \quad \Longleftrightarrow \quad A=\overline{A^{\mathsf{T}}} \quad \Longleftrightarrow \quad a_{ij}=\overline{a}_{ji}.$$

Positive and positive semi-definite operators An operator A over some Hilbert space $\mathcal{H}$ is positive, if

$$\langle\psi|A|\psi\rangle \geq 0, \qquad \forall\ |\psi\rangle \in \mathcal{H}.$$

Eigenvalues of such operator are non-negative. They are useful to construct POVMs. Definite means strictly positive.

Unitary operator An operator $\mathcal{U}$ is unitary, if it produces the identity matrix ($\mathbb{1}$) whenever (pre and post) multiplied by its conjugate transpose, i.e.,

$$\mathcal{U}\mathcal{U}^{\dagger} \ = \mathbb{1} \ = \mathcal{U}^{\dagger}\mathcal{U}.$$

The rows (or column) are normalized and are mutually orthogonal. Thereby, they form an orthonormal basis for the associated Hilbert space.

It is useful to describe the time evolution of a quantum state.

Normal operator An operator A is normal, if

$$AA^{\dagger} \ = \ A^{\dagger}A.$$

Thus, a normal operator commutes with its adjoint. Hermitian and unitary operators are normal.

Eigenvalue and eigenvector of an operator A state (vector) $|\psi\rangle$ is an eigenstate (eigenvector) of an operator A, if $A|\psi\rangle = \lambda|\psi\rangle$, for some scalar λ aka eigenvalue.

For instance, σ_z has eigenstates $|0\rangle, |1\rangle$ with eigenvalues +1 and -1 respectively.

Spectral decomposition theorem A normal operator is diagonalizable (i.e., a diagonal matrix) in some basis of that Hilbert space. That is, a normal operator A can be written as [NC11]

$$A = \sum_{i=1}^{n} \lambda_i |e_i\rangle\langle e_i| = \sum_{i=1}^{n} \lambda_i E_i,$$

for eigenstates $|e_i\rangle$ with eigenvalues λ_i. The projectors $E_i := |e_i\rangle\langle e_i|$ satisfy the completeness relation $\sum_i E_i = \mathbb{1}$.

For instance, the Pauli operators are normal. $\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|$ has eigenstates $|0\rangle, |1\rangle$ with eigenvalues +1,-1 respectively.

Note that, quantum measurements deal with the Hermitian operators for which the above result is quite straightforward to prove since the eigenspace form an orthonormal basis for the underlying Hilbert space.

Projection operators An operator P of the form $|\psi\rangle\langle\psi|$ is called projection operator. Any such operator is Hermitian (self-adjoint). For a normalized state $|\psi\rangle$, we have $P^2 = P$.

For instance, the operator $P = \sum\limits_{s=1}^{n} |s\rangle\langle s|$ projects onto the subspace spanned by the kets $|1\rangle, |2\rangle, \cdots, |n\rangle$.

Expected value of an operator The expectation of an operator A is the mean or average value $\langle\psi|A|\psi\rangle$ for a given quantum state $|\psi\rangle$. It indicates the average value of the outcomes of a measurement A when applied many times on copies of a given quantum state.

2.2.3 Pauli Matrices

Pauli matrices are a set of 2 × 2 complex Hermitian and unitary matrices.

$$\begin{aligned}\sigma_1 = \sigma_x &:= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \sigma_2 = \sigma_y &:= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\ \sigma_3 = \sigma_z &:= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.\end{aligned}$$

Together with the identity matrix $\mathbb{1}$ (or, σ_0), the Pauli matrices form an orthogonal basis, in the complex Hilbert space of all 2×2 matrices.

It is easy to find the following algebraic properties of these matrices.

- They are self-inverse:

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \mathbb{1} \quad = \quad -i\sigma_1\sigma_2\sigma_3.$$

  They anti-commute:

$$\sigma_i\sigma_j \quad = \quad -\sigma_j\sigma_i.$$

- Their determinant is -1, and are traceless.

$$\det\sigma_i = -1, \qquad \mathrm{tr}\ \sigma_i = 0, \quad \forall i \in \{1,2,3\}.$$

Therefore (together with $\sigma_i^2 = \mathbb{1}$), they have eigenvalues ±1 with the eigenvectors

$$\psi_{x+} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} = |0\rangle_x \quad , \quad \psi_{x-} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix} = |1\rangle_x,$$
$$\psi_{y+} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\i\end{bmatrix} = |0\rangle_y \quad , \quad \psi_{y-} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-i\end{bmatrix} = |1\rangle_y,$$
$$\psi_{z+} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\end{bmatrix} = |0\rangle_z \quad , \quad \psi_{z-} = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\end{bmatrix} = |1\rangle_z.$$

Thus, any of them can be written as

$$\sigma_s = |0_s\rangle\langle 0_s| - |1_s\rangle\langle 1_s|,$$

while the eigenprojectors span the 2-dimensional Hilbert space.

$$|0_s\rangle\langle 0_s| + |1_s\rangle\langle 1_s| = \mathbb{1}.$$

Therefore, each Pauli matrix represents an observable with two outcomes +1 and -1. In quantum mechanics, such an observable depicts the spin of a spin-$\frac{1}{2}$ particle in the three spatial directions.

Result 1. *The* 4 *Hermitian matrices constitute a linear vector space. A basis in this space can be chosen using the tensor products* $\sigma_i \otimes \sigma_j$ *involving the 2-dim Pauli matrices. Therefore, all operators involving two-qubit observables can be expanded over this basis. One can similarly extend the idea for higher dimensions.*

2.2.4 Bloch vector representation of a quantum state

Given an orthonormal basis, any pure state $|\psi\rangle$ of a two-level quantum system can be written as a superposition of the basis vectors $|0\rangle$ and $|1\rangle$, where the coefficient or amount of each of the two basis vector is a complex number.

$$|\psi\rangle := \alpha|0\rangle + \beta|1\rangle, \quad \text{for } \alpha, \beta \in \mathbb{C} \quad \text{with } |\alpha|^2 + |\beta|^2 = 1.$$

This means that the state is described by four real numbers. However only the relative phase between the coefficients of the two basis vectors has any physical meaning, so that there is redundancy in this description. We can take the coefficient of $|0\rangle$ to be real and non-negative. This allows the state to be described (up to a global phase) by only three real numbers, giving rise to the three dimensions of the Bloch sphere. The logic is developed as follows.

Up to a global phase, the state can also be written as [1]

$$\begin{aligned}|\psi\rangle &= \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle,\\ &\quad \text{where,}\ \ \theta\in[0,\pi],\ \ \phi\in[0,2\pi)\\ &\equiv e^{-i\phi/2}\cos\frac{\theta}{2}|0\rangle + e^{+i\phi/2}\sin\frac{\theta}{2}|1\rangle.\end{aligned}$$

However, a state can be identified by a suitable observable. One can verify that the above state is an eigenstate of the following observable σ_r corresponding to the eigenvalue $+1$.

$$\sigma_r \ = \hat{r}\cdot\hat{\sigma} = \ x\sigma_x + y\sigma_y + z\sigma_z,$$

with the unit vector

$$\hat{r} \ = \ (x,y,z) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta).$$

The other eigenstate for eigenvalue -1 is

$$|\psi^{\perp}\rangle \ = \ -e^{-i\phi/2}\sin\frac{\theta}{2}|0\rangle + e^{+i\phi/2}\cos\frac{\theta}{2}|1\rangle.$$

The direction $\hat{r}$ can be mapped to a point on the surface of a sphere. That sphere is known as the *Bloch sphere*, while the direction $\hat{r}$ is called the *Bloch vector*. Thus, any 2-d quantum state corresponds to a Bloch vector defined by three Cartesian variables and alternately by two angles θ,ϕ on the Bloch sphere. Here, θ is the angle that the direction $\hat{r}$ makes with the z-axis, while its projection on the xy plane creates the azimuthal angle ϕ with the x-axis. One can orient the measurement apparatus in a Stern-Gerlac experiment towards the direction $\hat{r}$ and let the state collapse to one of the eigenstates with outcome as the eigenvalue ± 1.

For mixed states, one considers the density operator. Any two-dimensional density operator ρ can be expanded using the identity $\mathbb{1}$ and the Hermitian, traceless Pauli matrices $\vec{\sigma}$, which together defines a orthogonal basis of a 2-dimensional Hilbert space.

$$\rho \ := \ \frac{1}{2}(\mathbb{1} + \vec{r}\cdot\vec{\sigma}), \quad ||\vec{r}|| \leq 1.$$

While $\vec{r}$ having unit norm represents pure states, other cases cover the mixed states living inside the sphere.

A projective measurement $P^{\pm} := \frac{1}{2}(\mathbb{1} + \hat{m}\cdot\vec{\sigma})$ results in an outcome ± 1 with probability $\mathsf{Tr}(\rho P^{\pm}) = \frac{1}{2}(1 + \hat{m}\cdot\vec{r})$ leaving the post-measurement state $\dfrac{P^{\pm}\rho P^{\pm}}{\mathsf{Tr}(\rho P^{\pm})}$.

[1] Unlike a classical bit that can take one of two states at a time, a qubit can in principle have infinitely many possible states parametrized by the continuous variables α,β, or equivalently, by the angles θ,ϕ.

Result 2. *Denoting the two eigenstates $|\psi\rangle$ and $|\psi^{\perp}\rangle$ as $|+\rangle_r$ and $|-\rangle_r$, respectively, one can write*

$$|+\rangle_r|-\rangle_r \; - \; |-\rangle_r|+\rangle_r \;\; = \;\; \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) =: |\psi^{-}\rangle.$$

for all possible Bloch vectors $\hat{r}$.

Result 3. *When Alice and Bob share a singlet $|\psi^{-}\rangle$ and measures their particles in the directions $\hat{a}$ and $\hat{b}$ respectively, one can show that*

$$E(\hat{a},\hat{b}) \; := \; \langle\psi^{-}|\sigma_a^A \otimes \sigma_b^B|\psi^{-}\rangle \;\; = \;\; -\hat{a}\cdot\hat{b}.$$

This is the expected value that Alice and Bob measure σ_a and σ_b respectively and gets the eigenstates $\hat{a}$ and $\hat{b}$ respectively.

Proof. First, note that

$$\begin{aligned} E(\hat{a},\hat{b}) &= \frac{1}{2}\langle 01-10|\sigma_a^{\mathtt{A}} \otimes \sigma_b^{\mathtt{B}}|01-10\rangle \\ &= \frac{1}{2}[\langle 01|\sigma_a^{\mathtt{A}} \otimes \sigma_b^{\mathtt{B}}|01\rangle - \langle 01|\sigma_a^{\mathtt{A}} \otimes \sigma_b^{\mathtt{B}}|10\rangle \\ &\quad - \langle 10|\sigma_a^{\mathtt{A}} \otimes \sigma_b^{\mathtt{B}}|01\rangle + \langle 10|\sigma_a^{\mathtt{A}} \otimes \sigma_b^{\mathtt{B}}|10\rangle]. \end{aligned}$$

Since $\langle ij|\sigma_a^{\mathtt{A}} \otimes \sigma_b^{\mathtt{B}}|kl\rangle = \langle i|\sigma_a^{\mathtt{A}}|k\rangle\langle j|\sigma_b^{\mathtt{B}}|l\rangle$, we use the following observations

$$\begin{aligned} \langle 0|\sigma_r|0\rangle &= z, & \langle 1|\sigma_r|1\rangle &= -z, \\ \langle 0|\sigma_r|1\rangle &= x-iy, & \langle 1|\sigma_r|0\rangle &= x+iy. \end{aligned}$$

to get the desired result. □

2.2.5 Mixed states

A pure quantum state can be described by a single ket. A mixed quantum state is a statistical ensemble (a probability distribution of states that some particles can be found in) of pure states. Mixed states inevitably arise from pure states when, for a composite quantum system $H_1 \otimes H_2$ with an entangled state on it, the part H_2 is inaccessible to the observer. The state of the part H_1 is expressed then as the partial trace over H_2.

A mixed state cannot be described with a single ket vector. Instead, it is described by its associated density matrix, usually denoted ρ. Moreover, a mixed quantum state on a given quantum system described by a Hilbert space H can be always represented as the partial trace of a pure quantum state on a larger bipartite system $H \otimes K$ for a sufficiently large Hilbert space K.

The **density matrix** describing a mixed state is defined to be an operator of the form

$$\rho = \sum_s p_s |\psi_s\rangle\langle\psi_s|$$

where p_s is the fraction of the ensemble in each pure state $|\psi_s\rangle$.

A simple criterion to check whether a given density matrix describes a pure or mixed state is that the trace of ρ^2 must be 1 if the state is pure, and less than 1 if mixed. Another equivalent criterion is that the von Neumann entropy is 0 for a pure state, and strictly positive for a mixed state.

The rules for measurement in quantum mechanics becomes simple when stated in terms of density matrices. For instance, the ensemble average (expectation value) of a measurement corresponding to some observable A is given by

$$\langle A\rangle = \sum_s p_s \langle\psi_s|A|\psi_s\rangle = \sum_s\sum_i p_s a_i |\langle\alpha_i|\psi_s\rangle|^2 = \mathrm{tr}(\rho A)$$

where $|\alpha_i\rangle$, a_i are eigenkets and eigenvalues, respectively, for the operator A.

2.2.6 Density matrix

A density matrix is a matrix that describes the statistical state, whether pure or mixed, of a system in quantum mechanics. The probability for any outcome of any well-defined measurement upon a system can be calculated from the density matrix for that system. The extreme points in the set of density matrices are the pure states, which can also be written as state vectors or wavefunctions. Density matrices that are not pure states are mixed states. Any mixed state can be represented as a convex combination of pure states.

Describing a quantum state by its density matrix is a fully general alternative formalism to describing a quantum state by its state vector (its "ket") or by a statistical ensemble of kets. However, in practice, it is often most convenient to use density matrices for calculations involving mixed states, and to use kets for calculations involving only pure states. Mixed states arise in situations where the experimenter does not know which pure state the system is in. For instance, in an entangled system, each subsystem must be treated as a mixed state even if the complete system is in a pure state.

A density matrix is self-adjoint (i.e., Hermitian), positive semi-definite, and of trace one. Thereby, its eigenvalues are non-negative and sum to one, defining a probability distribution.

A density operator describes a pure state if it is a rank one projection. A necessary and sufficient condition for a density matrix ρ describes a pure state if and only if $\rho = \rho^2$.

2.2.7 Partial Trace and reduced density matrices

When two parties share a pure entangled state, the state of the individual subsystems is no more a pure state. It is possible though to trace out one of the subsystem and get the density of the other one as a mixed state. Thus, for a pure state $|\psi\rangle_{\mathtt{AB}}$, shared between Alice and Bob, considering $\rho_{\mathtt{AB}}$ as the joint density operator, $\rho_{\mathtt{A}} = \mathsf{Tr}_{\mathtt{B}}(\rho_{\mathtt{AB}})$ and $\rho_{\mathtt{B}} = \mathsf{Tr}_{\mathtt{A}}(\rho_{\mathtt{AB}})$ are the density operators with Alice and Bob, respectively. Mathematically speaking, for a tensor product $A \otimes B$ of two matrices, $\mathsf{Tr}_{\mathtt{B}}(A \otimes B) := A \cdot \mathsf{Tr}(B)$ traces out Bob's subsystem.

For example, consider the Bell state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)_{\mathtt{AB}}$, which has the following density

$$\begin{aligned}\rho_{\mathtt{AB}} &= \frac{1}{2}\big(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|\big)\\ &= \frac{1}{2}\sum_{(a,b)\in\{0,1\}\times\{0,1\}} |a\rangle\langle b| \otimes |a\rangle\langle b|.\end{aligned}$$

Now, use the property that $\mathsf{Tr}(|a\rangle\langle b|) = \langle a|b\rangle = 1,0$, depending on whether the two states are same, or, are orthogonal, respectively. Then,

$$\rho_{\mathtt{A}} = \frac{1}{2}\sum_{a=b\,\in\{0,1\}} |a\rangle\langle b|,$$

which is a mixed state. Bob's local state is same as that of Alice in this case.

However, the calculations go difficult when the state with one or both the parties are themselves in superposition disrupting orthogonality. In such situations, Schmidt decomposition becomes useful to express a bipartite state in orthogonal bases with each of the parties.

2.2.8 Schmidt Decomposition

For any bipartite pure state $|\psi\rangle_{\mathtt{AB}} \in \mathcal{H}_{\mathtt{A}} \otimes \mathcal{H}_{\mathtt{B}}$ with $d = \min\{dim\mathcal{H}_{\mathtt{A}}, dim\mathcal{H}_{\mathtt{B}}\}$, there are orthonormal bases $\{|\alpha_i\rangle_{\mathtt{A}}\}_{i=1}^{d} \in \mathcal{H}_{\mathtt{A}}$ and $\{|\beta_i\rangle_{\mathtt{B}}\}_{i=1}^{d} \in \mathcal{H}_{\mathtt{B}}$, such that

$$|\psi\rangle_{\mathtt{AB}} = \sum_{i=1}^{d} \sqrt{\nu_i}\,|\alpha_i\rangle_{\mathtt{A}}|\beta_i\rangle_{\mathtt{B}}, \quad \text{with} \quad \nu_i \geq 0,\ \sum_{i=1}^{d} \nu_i = 1.$$

The frequencies $\sqrt{\nu_i}$ are called the Schmidt co-efficients, while the number of non-zero ν_i's is called the *Schmidt rank* of the bipartite state. For example, the Bell pair under consideration is already Schmidt decomposed, with Schmidt co-efficients $\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}$ and Schmidt rank 2.

A pure state is separable iff it has Schmidt rank 1, and is maximally entangled if

$\lambda_j = 1/d, \forall j$.

2.2.9 Information content: Entropy, Mutual information

Entropy is a way to quantify the amount of information in a signal. Roughly speaking, entropy indicates how much is unknown about the signal, while information means the left-over unknown due to some learning on the signal [2]. We elaborate it mathematically in the following.

Entropy: Let a random variable (r.v.) can take finitely many values x with probability $p(x)$. Then, the Shannon entropy of X is defined as

$$H(X) \quad := \quad -\sum_x p(x) \log_2 p(x),$$

expressed in bits. It is a concave function. When the r.v. takes n different values, the maximum value of the entropy is $\log_2 n$ that occurs for uniform distribution $p(x) = {}^1\!/_n \, \forall x$. Increasing entropy indicates decreasing knowledge about the random variable. Thus, it quantifies the uncertainty of the r.v.

A frequently useful entropy function is the binary entropy function

$$H_2(X) \quad = -p(x) \log_2 p(x) - (1 - p(x)) \log_2 (1 - p(x)) \quad =: h(p(x)).$$

It can reach the maximum (=1 bit) for $x = \frac{1}{2}$. The concavity can be visualized from the figure 3.1.

In quantum communications, when one measures a quantum state, the outcomes define a random variable. One can estimate the information gained out of the measurement by computing the entropy of the r.v.

Similarly, for two random variables X, Y, one can get a distribution $p(x,y)$ of the sample points (x,y), for which the joint Shannon entropy is defined as

$$H(X,Y) \quad := \quad -\sum_{(x,y)} p(x,y) \log p(x,y).$$

The conditional entropy of X given the knowledge on the r.v. Y is defined as any of the following

$$\begin{aligned} H(X|Y) \quad &:= \quad H(X,Y) - H(Y) = -\sum_{(x,y)} p(x,y) \log p(x|y) \\ &= \quad \sum_y p(y) H(X|y) = -\sum_y p(y) \log p(x|y). \end{aligned}$$

[2] The quantification is done in the log scale. Generally, logarithms are considered to the base 2, and these information quantifiers are thus measured in bits.

It reflects the remaining uncertainty about X due to the knowledge of Y.

von Neumann entropy It determines the amount of entropy in a quantum state. For a quantum-mechanical system described by a density matrix ρ, the von Neumann entropy is defined as

$$S := -\mathrm{tr}(\rho \ln \rho).$$

ρ being Hermitian is diagonalizable in its eigenbasis $|1\rangle, |2\rangle, |3\rangle \ldots$:

$$\rho = \sum_j \eta_j |j\rangle\langle j|.$$

Thereby, a measurement in the eigenbasis gives rise to the classical eigenvalues as outcomes and the von Neumann entropy is merely [3]

$$S = -\sum_j \eta_j \ln \eta_j.$$

In this form, S can be seen as the information theoretic Shannon entropy.

Mutual information The mutual information between two random variables X, Y is defined as any of the following

$$\begin{aligned} I(X,Y) &:= H(X) + H(Y) - H(X,Y) \\ &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X). \end{aligned}$$

It is thus symmetric in X, Y. It reflects the amount of knowledge common in the two random variables. The last two expressions read as the reduction in entropy (uncertainty) of a r.v. due to the knowledge on the other r.v.

Rényi entropy and Rényi information It is an one-parametric extension of the Shannon entropy.

Rényi entropy of order α ($> 0, \neq 1$) is defined as

$$R_\alpha(X) = \frac{1}{1-\alpha} \log_2 \sum_a (p(x))^\alpha.$$

It is maximized ($= \log N$) for the uniform distribution $p(x) = 1/N \; \forall x$.

[3] The catch here is that for logarithm of diagonal matrices, the logarithm transcends to the diagonal itself.

There is no generally accepted definition of conditional Rényi entropy, but it can be considered similar to the conditional Shannon entropy. Following it, one can consider Rényi (mutual) information of order α between the two random variables $\mathcal{A},\mathcal{E}$ to be defined as follows.

$$\begin{aligned} I^R_\alpha(X,Y) &= R_\alpha(X) - R_\alpha(X|Y) \\ &= \frac{1}{1-\alpha}\log_2 \sum_x (p(x))^\alpha - \frac{1}{1-\alpha}\sum_y p(y)\log_2 \sum_x (p(x|y))^\alpha. \end{aligned}$$

However, unlike Shannon information, it's not symmetric, and a proper choice of α is often not clear. Moreover, conditional Rényi entropy doesn't follow the chain rule, and thereby, the Rényi MI cannot be thought in general as reduction in uncertainty due to the knowledge of the other.

Rényi information of order $\alpha > 1$ is an upper bound on the Shannon mutual information. Sometimes, $\alpha = 2$ is of special interest and is used here.

2.2.10 Expected value

The expected value of a random variable A is defined as

$$E(A) \; := \; \sum_a a P_A(a).$$

The expected value of the product of two random variables are defined as

$$E(AB) \; := \; \sum_{a,b} ab P_{AB}(a,b).$$

Variance The variance of a random variable is defined as

$$Var(A) \; := \; E(A - E(A))^2 = E(A^2) - (E(A))^2.$$

S.D. The *standard deviation* of a random variable is defined as the square root of variance.

$$\sigma(A) = \sqrt{Var(A)} \; := \; \sqrt{E(A^2) - (E(A))^2}.$$

Covariance The *covariance* of two random variables is defined as

$$Covar(A,B) \; := \; E[(A - E(A))(B - E(B))] = E(AB) - E(A)E(B).$$

Correlation coefficient The *correlation coefficient* between two random variables is defined as

$$\rho_{(A,B)} \quad := \quad Covar(A,B)/\sigma(A)\sigma(B).$$

When two random variables take values from +1 and -1, then

$$E(AB) = \sum_{a,b=\pm 1} P_{AB}(a=b) - P_{AB}(a \neq b),$$

$$E(A) = 0 = E(B),\ \ E(A^2) = 1 = E(B^2),\ \ \sigma_A = \sigma_B = 1,\ \ \rho_{(A,B)} = E(AB).$$

This result will be useful in CHSH violation.

2.3 Quantum Mechanics

2.3.1 Interpretations of Quantum Mechanics

There are various schools of interpretations to explain quantum mechanics. The standing one that is used in quantum communication is the Copenhagen interpretation. Therefore, a few relevant facts on it are added here.

Copenhagen interpretation: The wave-function has no reality. Nature is only probabilistic and only a measurement forces it to choose a state, before this there is no realism. The process of measurement causes a collapse of the wave function and the result corresponds with the eigenvalue of the measurement operator: mapping the operator to a real value. Furthermore Heisenberg's uncertainty principle prevents us of knowing all parameters of a system at once. Nowadays, most physicists prefer the Copenhagen interpretation of quantum mechanics.

Some of the other schools of interpretations are: i) hidden-variable theories, e.g., Bohmian mechanics. ii) many-worlds interpretation, etc.

2.3.2 Quantum operators do not commute

Properties in the quantum world correspond to operators that do not commute. This basic feature of the formalism is at the root of the observation that orders of measurements matter. It is also the key to the understanding that it does not make sense in quantum theory to think of two different quantities corresponding to non-commuting observables to "take specific values". They do not. This does not mean, of course, that no two observables necessarily commute.

Two observables are compatible if they commute. Then it is possible to know precisely the value of both observables at the same time.

For non-commuting observables, measuring one randomizes the other.

2.3.3 Uncertainty principle

Uncertainty is a statistical measure (standard deviation) of the spread of measurements about the mean. For some measurement operator A, it is given by

$$\Delta A \;:=\; \sqrt{\langle A^2\rangle - \langle A\rangle^2},$$

where, $\langle A\rangle := \langle\psi|A|\psi\rangle$ denotes the mean value of the outcomes to measure the state $|\psi\rangle$, while $\langle A^2\rangle := \langle\psi|A^2|\psi\rangle$ is the 2nd order moment.

Then, for two measurement operators A,B, the product of the uncertainties satisfies the following inequality

$$\Delta A\Delta B \;\geq\; \frac{1}{2}|\langle AB - BA\rangle|.$$

For incompatible observables (i.e., non-commutative: $AB \neq BA$), if one of the uncertainties go smaller, the other one go larger – both cannot be measured simultaneously with high precision [NC11].

1. One cannot "know the values of two non-commuting observables at once".
2. measurement of one observable makes the outcome of another non-commuting observable less certain.
3. know the value of A precisely, then the measurement of B will be a lot disturbed.

For instance, for the state $|0\rangle$ and observables σ_z, σ_x, we get $\Delta\sigma_z\Delta\sigma_x \geq 1$.

2.3.4 Postulates and quantum measurements

Postulates of quantum mechanics [NC11] broadly describes the following:

1. how the states of a physical system are described.
2. how measurements work.
3. how the evolution of a physical system is described.

Postulate 1: Quantum states The state of a quantum system is a vector $|\psi\rangle$ in a Hilbert space. A qubit is a 2-dimensional state. We consider generally normalized states. Linear combination (superposition) of two states is another state.

Postulate 2: Quantum measurements A quantum measurement corresponds to an observable which is Hermitian and thereby has a spectral decomposition in an eigenbasis. The eigenvalues are the outcomes, while the eigenstates are the post-measurement state of the system.

Let the observable is an Hermitian operators A having spectral decomposition $A = \sum_a a\Pi_a$ with orthonormal eigenprojectors $\Pi_a = |a\rangle\langle a|$ and eigenvalues a. As the eigenbasis spans the Hilbert space, it satisfies the completeness relation $\sum_a \Pi_a = \mathbb{1}$.

Given a state $|\psi\rangle$, the measurement outcome a occurs with probability (Born's rule)

$$p_a = \langle\psi|\Pi_a|\psi\rangle = \mathsf{Tr}(\Pi_a\rho_\psi) = |\langle\psi|a\rangle|^2.$$

The state of the system after the measurement is $\frac{1}{\sqrt{p_a}}\Pi_a|\psi\rangle = |a\rangle$.

The expected value (average) of the observable A w.r.t. $|\psi\rangle$ is then

$$\langle A\rangle = \sum_a a\langle\psi|\Pi_a|\psi\rangle = \sum a p_a.$$

The observable in such cases is a collection of projection operators $\{\Pi_a\}$ satisfying the completeness relation, known as **(von Neumann) measurement**. There are generalized measurements like POVMs that we'll discuss shortly.

The state of the system may not be pure. For an arbitrary density ρ, the Born's rule and the post-measurement state needs be upgraded as follows.

$$p_a = \mathsf{Tr}(\Pi_a\rho\Pi_a) = \mathsf{Tr}(\Pi_a^2\rho) = \mathsf{Tr}(\Pi_a\rho).$$

and the post-measurement state becomes

$$\rho_a = \frac{1}{\sqrt{p_a}}\Pi_a\rho\Pi_a.$$

The projectors Π_a in this case called *detection operators*.

Postulate 3: Evolution Dynamical evolution of a closed system corresponds to an unitary operator that transform a quantum state to another state. It preserves the length and the overlap between two states.

POVM For projective measurements, the number of outcome is limited by the dimension of the Hilbert space due to orthogonality restriction on the projectors. However, in circumstances, it is often desirable that the number of outcomes exceed the dimension of the Hilbert space while keeping positivity and normalization of the probability distribution, which is possible by relaxing the orthogonality restriction.

The trick to play is with the Born's rule. So far P_a^2 is a positive (semi-definite) operator, it generates non-negative probabilities. It need not be a projector then.

So, we can introduce a set of positive operators $P_a \geq 0$ generating the probabilities $p_a = \mathsf{Tr}(\rho P_a)$. To make it a distribution, it should satisfy normalization $\sum_a P_a = \mathbb{1}$. This collection $\{P_a\}$ is known as *positive operator valued measure* (POVM).

The detection operators need not be the projectors. Denoting them as M_a, the Born's rule can be written as

$$p_a \quad = \quad \mathsf{Tr}(M_a \rho M_a^\dagger) = \mathsf{Tr}(\rho P_a).$$

Thus, the POVM elements can be considered as $P_a = M_a^\dagger M_a$, which is a positive operator.

A good example could be the POVMs used in [BBM92] to distinguish two non-orthogonal states $|a_0\rangle, |a_1\rangle$. The POVM $\{B_0, B_1, B_?\}$ is defined as follows.

$$\begin{aligned} B_0 &:= (\mathbb{1} - |a_1\rangle\langle a_1|)/(1 + \langle a_0|a_1\rangle), \\ B_1 &:= (\mathbb{1} - |a_0\rangle\langle a_0|)/(1 + \langle a_0|a_1\rangle), \\ B_? &:= (\mathbb{1} - B_0 - B_1). \end{aligned}$$

Note that the operators are Hermitian, but mutually non-orthogonal, and thereby are not projection operators. However, they are positive (only eigenstate $|a^\perp\rangle$ with eigenvalue +1), and sum to identity.

If Alice sends some $|a_s\rangle$, Bob gets either $|a_s\rangle$, or inconclusive result (i.e., $B_?$ clicked), but never $|a_{s+1}\rangle$. Any B_s detects $s \in \{0, 1, ?\}$ w.p. $\frac{1-|\langle a_0|a_1\rangle|^2}{1+\langle a_0|a_1\rangle} = 1 - \langle a_0|a_1\rangle$.

To summarize, a projective measurement is a collection of mutually orthogonal projection operators $\{\Pi_m\}$ satisfying the completeness relation $\sum_m \Pi_m = \mathbb{1}$.

And a POVM is a collection of positive operators $\{E_\lambda\}$ that sum to the identity. Note that positivity makes it Hermite.

2.3.5 Entanglement and non-locality

An entangled state cannot be written as a product of separate states. For instance, consider the famous EPR-pair

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

Suppose such a pair of particles are distributed between two parties, say, Alice and Bob. If one of them measures and get an outcome $x \in \{0, 1\}$, then the state automatically collapses to $|xx\rangle$. Thereby, whenever the other party measures, both of their results agree. It seems that the action of the measurement by the first party instantaneously effect the outcome of the other party. However, the EPR pair doesn't violate the so-called local-realism. In 1960s, John Bell's entanglement-based experiment exhibited *quantum non-locality* that cannot be reproduced by any *local realistic theory*.

2.3.5.1 Bell states

For a 2-qubits (4-dimensional) Hilbert space, one can define an orthonormal basis consisting of four Bell states each of which is a maximally entangled state. The basis states map two bits $a, b \in \{0, 1\}$ into two entangled qubits as follows.

$$|\beta_{ab}\rangle \quad := \quad \frac{1}{\sqrt{2}}\left(|0, b\rangle + (-1)^a |1, b+1\rangle\right).$$

These two-particle maximally entangled states are often symbolized as follows:

$$\begin{aligned}
|\phi^+\rangle &:= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) &=:& \ |\beta_{00}\rangle, \\
|\psi^+\rangle &:= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) &=:& \ |\beta_{01}\rangle, \\
|\phi^-\rangle &:= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) &=:& \ |\beta_{10}\rangle, \\
|\psi^-\rangle &:= \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) &=:& \ |\beta_{11}\rangle.
\end{aligned}$$

The preparation of the Bell states are given in Fig. 2.3.

The *singlet* $|\psi^-\rangle$, being a spin-0 particle, exhibits an interesting property: that it is invariant w.r.t. rotation of the measurement apparatus. Up to a global phase, they are all equivalent.

$$\begin{aligned}
|\psi^-\rangle &= \frac{1}{\sqrt{2}}(|0_z 1_z\rangle - |1_z 0_z\rangle) \\
&= -\frac{1}{\sqrt{2}}(|0_x 1_x\rangle - |1_x 0_x\rangle) \\
&= -i\,\frac{1}{\sqrt{2}}(|0_y 1_y\rangle - |1_y 0_y\rangle).
\end{aligned}$$

Therefore, measuring the state in any arbitrary direction provides the same measurement outcomes. This is useful to establish the Bell-CHSH inequality.

Note that, $|\phi^+\rangle$ also looks same along Z and X directions.

2.3.5.2 Locality and realism

For a shared system of particles A, B, *locality* means that measuring one particle shouldn't disturb the state of the other particle. *Realism* means that the values of measurable properties of each subsystem are objectively real: they have definite values even before the measurement. Thus local realism represent classical view of the world. However, in quantum world, the wave-function assumes reality only after measurement. In experiments with bipartite systems, the measurements are local and a classical communication is allowed to tabulate the measurement statistics and do some computations. Such computations are known as *local operations and classical communication* (LOCC).

2.3.5.3 Bell-CHSH inequality [CHSH69]

John Bell (1964) shown that no local realistic (i.e., classical) model can explain all quantum predictions, pointing to the difference between quantum and classical world. Some conditions necessary for the local realistic models are given by CHSH inequality which is defined on the local measurement statistics on a bipartite system.

Consider a joint bipartite system of particles in state $|\psi\rangle_{AB}$ is pre-shared between Alice and Bob. Each of them are allowed to perform one of the two possible measurements on their respective subsystem: $\mathbb{A}_0, \mathbb{A}_1$ with Alice and $\mathbb{B}_0, \mathbb{B}_1$ with Bob respectively. Let the corresponding measurement outcomes are $A_0, A_1, B_0, B_1 \in \{+1, -1\}$, respectively.

Let them repeat the bipartite measurements $\mathbb{A}_i \otimes \mathbb{B}_j$ on various identical states $|\psi\rangle_{AB}$ and tabulate the outcomes (A_i, B_j), and their products A_iB_j. They can compute the average $\{\langle A_iB_j\rangle\}_{i,j\in\{0,1\}}$ of these four different products and compute the following correlation co-efficient, known as CHSH sum

$$\langle S_{\mathsf{CHSH}}\rangle \quad := \quad \langle \mathbb{A}_0 \otimes \mathbb{B}_0\rangle + \langle \mathbb{A}_0 \otimes \mathbb{B}_1\rangle + \langle \mathbb{A}_1 \otimes \mathbb{B}_0\rangle - \langle \mathbb{A}_1 \otimes \mathbb{B}_1\rangle.$$

If the shared system is a separable state, they will find $|\langle S_{\mathsf{CHSH}}\rangle| \leq 2$. But, if it is entangled, they'll find (for some suitable measurement) $2 < |\langle S_{\mathsf{CHSH}}\rangle| \leq 2\sqrt{2}$. This agrees both in theory and in practice. Following is a theoretical justification. The two cases are considered separately. The first one is known as the CHSH inequality and is explained below.

Note that the joint measurement will generate a pair of outcomes following some probability distribution $\{p(A_iB_j|\mathbb{A}_i \otimes \mathbb{B}_j)\}$. So far the system is entangled, the outcomes exhibit some correlation and thus, won't factorize. But, as they go separable, the outcomes are independent and the probabilities factorize as $p(A_iB_j|\mathbb{A}_i \otimes \mathbb{B}_j) = p(A_i|\mathbb{A}_i)p(B_j|\mathbb{B}_j)$. This is the reason why we find the difference in the CHSH sum.

The expected value of the joint observable is defined as

$$\langle \mathbb{A}_i \otimes \mathbb{B}_j\rangle \quad = \quad \langle \psi_{AB}|\mathbb{A}_i \otimes \mathbb{B}_j|\psi_{AB}\rangle = \sum_{A_i=\pm 1, B_j=\pm 1} A_iB_jp(A_i, B_j).$$

For a separable state, it reduces to

$$\begin{aligned}\langle \mathbb{A}_i \otimes \mathbb{B}_j \rangle &= \sum_{A_i=\pm1, B_j=\pm1} A_i B_j p(A_i) p(B_j) \\ &= \sum_{A_i=\pm1} A_i p(A_i) \sum_{B_j=\pm1} B_j p(B_j) \quad = \langle \mathbb{A}_i \rangle \langle \mathbb{B}_j \rangle.\end{aligned}$$

Note that, $\langle \mathbb{A}_i \rangle = \sum_{A_i=\pm1} A_i Pr(A_i) = Pr(A_i = +1) - Pr(A_i = -1) \in [-1, +1]$.

Thereby, the CHSH sum becomes

$$\begin{aligned}\langle S_{\mathsf{CHSH}} \rangle &:= \langle \mathbb{A}_0 \rangle \langle \mathbb{B}_0 \rangle + \langle \mathbb{A}_0 \rangle \langle \mathbb{B}_1 \rangle + \langle \mathbb{A}_1 \rangle \langle \mathbb{B}_0 \rangle - \langle \mathbb{A}_1 \rangle \langle \mathbb{B}_1 \rangle \\ &= \langle \mathbb{A}_0 \rangle [\langle \mathbb{B}_0 \rangle + \langle \mathbb{B}_1 \rangle] + \langle \mathbb{A}_1 \rangle [\langle \mathbb{B}_0 \rangle - \langle \mathbb{B}_1 \rangle] \\ &= \langle A_0 [B_0 + B_1] + A_1 [B_0 - B_1] \rangle\end{aligned}$$

The last equality follows due to the linearity of the expected values and the independence of the random variables. Note that, one of $B_0 \pm B_1$ is 2, while the other one is 0. Thus, for a specific bi-partite measurement, $S_{\mathsf{CHSH}} = \pm 2$. Then, for a finitely many such measurements, the outcome statistics satisfy the following inequality

$$-2 \leq \quad \langle S_{\mathsf{CHSH}} \rangle \quad \leq 2.$$

The extreme ends are achieved when all the experiments provide the same value of S: either +2, or, -2.

However, this classical view of the outcomes statistics in not true in entanglement driven quantum mechanics as the joint distribution exhibits some dependence.

2.3.5.4 Bell-violation by Quantum mechanical entanglement

Let Alice and Bob share some entangled state state $|\psi\rangle_{AB}$. Their outcomes are mutually dependent based on the quality of entanglement, e.g., with maximally entangled states, the outcomes are completely correlated. Since the distribution is no more factored, the above approach doesn't help to get the inequality.

Let, Alice's observables corresponds to the Bloch vectors $\hat{a}_0, \hat{a}_1$, while Bob's directions are $\hat{b}_0, \hat{b}_1$. Then,

$$\langle \mathbb{A}_i \otimes \mathbb{B}_j \rangle \quad = \quad \langle \psi_{AB} | \mathbb{A}_i \otimes \mathbb{B}_j | \psi_{AB} \rangle = \langle \psi_{AB} | \hat{a}_i \cdot \vec{\sigma} \otimes \hat{b}_j \cdot \vec{\sigma} | \psi_{AB} \rangle.$$

When they share the EPR pair $\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$, the expected value becomes

$$\langle \mathbb{A}_i \otimes \mathbb{B}_j \rangle \quad = \quad -\hat{a}_i \cdot \hat{b}_j = -\cos \theta_{ij}.$$

Here, θ_{ij} is the angle between the two directions $\hat{a}_i, \hat{b}_j$.

Thus, the CHSH correlation coefficient becomes

$$\begin{aligned} |\langle S_{\mathsf{CHSH}} \rangle| &:= |\langle \mathbb{A}_0 \otimes \mathbb{B}_0 \rangle + \langle \mathbb{A}_0 \otimes \mathbb{B}_1 \rangle + \langle \mathbb{A}_1 \otimes \mathbb{B}_0 \rangle - \langle \mathbb{A}_1 \otimes \mathbb{B}_1 \rangle| \\ &= \cos\theta_{00} + \cos\theta_{01} + \cos\theta_{10} - \cos\theta_{11}. \end{aligned}$$

One can choose the angles properly to violate the CHSH inequality.

For instance, consider Alice using $\mathbb{A}_0 = \sigma_z, \mathbb{A}_1 = \sigma_x$, and Bob using $\mathbb{B}_0 = -\frac{\sigma_z+\sigma_x}{\sqrt{2}}, \mathbb{B}_1 = \frac{\sigma_z-\sigma_x}{\sqrt{2}}$. Then, the CHSH expected values of the joint measurements $\mathbb{A}_i \otimes \mathbb{B}_j$ become

$$\langle \mathbb{A}_i \otimes \mathbb{B}_j \rangle = (-1)^{ij} \frac{1}{\sqrt{2}}.$$

Thereby, the average value of the CHSH sum becomes $2\sqrt{2} > 2$. Thus, quantum mechanical entanglement violates the Bell-CHSH inequality, which is not possible in its classical counterpart due to factored states. If the shared state is not maximally entangled (possibly due to eavesdropping), the CHSH sum drops a bit, but remain above 2 so far the joint state is not separable.

Thereby, Bell-CHSH violation is an way to demarcate classical and quantum world. In quantum communication with EPR pairs, one can check whether the shared state is indeed entangled or factored. In the later case, the system is distinguishable perfectly by an eavesdropper which is equivalent to a classical communication. The amount of Bell violation indicates the extent of entanglement degradation, from which they can decide whether further filtration is possible.

2.3.5.5 Tsirelson's inequality [Cir80]

In the above case, we have considered a particular quantum system – a maximally entangled EPR pair. However, choice of the joint system could be many. The question is, what is the maximum CHSH-violation in the quantum domain? The answer is $2\sqrt{2}$.

2.4 Quantum gates

In classical computation, any sequence of elementary operations (e.g., NAND and COPY) allows one to build up any complex computation. Similarly, in quantum computation, any unitary operation in the Hilbert space of n qubits can be decomposed into one-qubit and two-qubit elementary gates.

Here we discuss some useful and relevant quantum gates. Some common 1-qubit gates are Hadamard, bit-flip, phase-flip etc. Some relevant 2-qubit gates are C-NOT, SWAP etc.

2.4.1 Hadamard gate

It is useful to create superposition. It corresponds to the following unitary transformation

$$\mathbb{H} \;:=\; \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}.$$

It transforms the computational basis states $|0\rangle, |1\rangle$ into the Hadamard states $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, respectively. In general, an arbitrary state $|s\rangle$ can be transformed as follows:

$$\mathbb{H} \;:\; |s\rangle \;\mapsto\; |s\rangle + (-1)^s|1-s\rangle.$$

Hadamard transforms the computational basis into the Hadamard basis and vice versa. The later part is true because Hadamard matrix is self-invertible owing to $\mathbb{H}^2 = \mathbb{1}$. It is Hermitian as well, because $\mathbb{H}^\dagger = \mathbb{H}$, i.e., the conjugate transpose is same as the transformation matrix as well. Thereby, it can be considered both for unitary evolution, and for measurement purposes. It is useful to note that $\mathbb{H} = \frac{1}{\sqrt{2}}(\sigma_x + \sigma_z)$.

2.4.2 C-NOT gate

It is a 2-qubit gate that flips the second bit only when the first bit is Boolean YES. It is often useful to create entanglement between two qubits.

A NOT-gate (i.e., σ_x) simply flips the input bit: $0/1 \mapsto 1/0$. A C-NOT gate flips (or not) the target-bit (here, second) only when the control-bit (here, first) is Boolean YES (or NO), i.e.,

$$\begin{aligned}\textsf{C-NOT} \;:\; |0\rangle|x\rangle &\;\mapsto\; |0\rangle|x\rangle\\ |1\rangle|x\rangle &\;\mapsto\; |1\rangle|x+1\rangle.\end{aligned}$$

and, in general,

$$\textsf{C-NOT} \;:\; |c\rangle|t\rangle \;\mapsto\; |c\rangle|t+c\rangle.$$

Clearly, until the control bit is in superposition, we won't get the entanglement. The superposition is done by the Hadamard, for instance. Thus,

$$|0\rangle|0\rangle \;\xmapsto{\mathbb{H}\otimes\mathbb{1}}\; \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|0\rangle \;\xmapsto{\textsf{C-NOT}}\; \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

The matrix representation of the C-NOT gate and a circuit diagram is as follows.

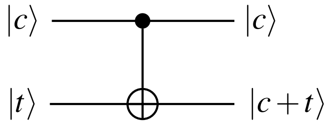
Since, $\textsf{C-NOT}^2 = \mathbb{1}$, it is self-invertible.

One can generalize the concept by considering a controlled-Unitary, where the target qubit is evolved by the unitary only when the control-bit is Boolean YES (=1).

Figure 2.1 | C-NOT gate: matrix representation and circuit.

The C-NOT operator corresponds to $\mathbb{1}_2 \otimes \sigma_x$. The circuit has in input lines a control qubit $|c\rangle$ and a target qubit $|t\rangle$.

$$C \ := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$



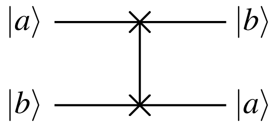
2.4.3 SWAP gate

Another useful 2-qubit gate is the SWAP-gate that swaps the 2-qubits states $|x\rangle|y\rangle \mapsto |y\rangle|x\rangle$. The transformation matrix is as follows.

Figure 2.2 | SWAP gate: matrix representation and circuit.

The SWAP operator exchanges the input lines $|a\rangle$ and $|b\rangle$.

$$S_w \ := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$



Mathematically, it corresponds to the permutation Π_{1324}.

A *Toffoli gate* is a C^2NOT gate, which applies a NOT operation to the target qubit only when the two control qubits are set to 1.

2.4.4 Preparation of the initial state

Preparation of a general state in general is not efficient in quantum domain, as the number of gates grow exponentially in the number of qubits.

However, in special cases, an wave function can be prepared efficiently, i.e., number of gates required is polynomial in the number of qubits. For instance, an equal superposition of all the states in some n-dim computational basis is obtained by applying n Hadamard gates to the state $|0\rangle^{\otimes n}$.

For 3-qubit unitary in translucent attack model, there seems to be a trade-off between preparing the 2-qubit initial state and the 3-qubit unitary.

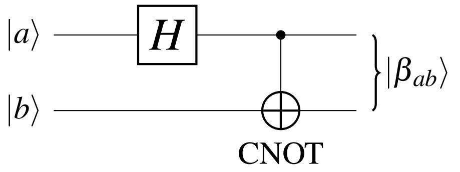
2.4.5 Preparation of Bell states

The preparation of the Bell states are given in Fig. 2.3. Mathematically, the mapping is as follows

$$\text{C-NOT}(\mathbb{H} \otimes \mathbb{1}) \quad : \quad |ab\rangle_{a,b\in\{0,1\}} \quad \mapsto \quad |\beta_{ab}\rangle.$$

Figure 2.3 | A circuit to prepare Bell states.

For inputs $a, b \in \{0, 1\}$, Hadamard on the first line creates a superposition, and the CNOT creates the entanglement.



2.5 What secures quantum communication?

The following features are the integral part of the security of quantum communication.

- No cloning theorem forbids an eavesdropper to copy non-orthogonal states perfectly.

- Uncertainty principle forbids her to measure two incompatible properties precisely at a go. For instance, measuring both the z-spin and the x-spin are not feasible with certainty, i.e., measuring σ_z with certainty means σ_x randomizes the outcomes. Mainly the *p&m* schemes get the security following this principle.

- For *eb* schemes, Bell-violation certifies security. On the other hand, *p&m* schemes typically depend on the estimated error-rate without a Bell test.

2.5.1 No cloning theorem [WZ82]

Unlike classical world, where copying classical bits is feasible without being caught, in quantum world, copying an arbitrary state is not possible [WZ82] other than an inferior copy.

Let $|\psi\rangle$ is the state to be copied, $|b\rangle$ is the blank copy. Then, assuming copying is possible by some unitary $\mathcal{U}$, we have

$$\mathcal{U}|\psi\rangle \otimes |b\rangle = |\psi\rangle \otimes |\psi\rangle.$$

Now, one can come up with a CNOT gate that can copy two mutually orthogonal states perfectly as follows.

$$\begin{array}{rrcl}\mathsf{C\text{-}NOT} : & |0\rangle|0\rangle & \mapsto & |0\rangle|0\rangle \\ & |1\rangle|0\rangle & \mapsto & |1\rangle|1\rangle.\end{array}$$

However, it cannot copy a linear combination of these two states

$$\begin{array}{rcl}\mathsf{C\text{-}NOT} \ : & (\alpha|0\rangle+\beta|1\rangle)|0\rangle & =\alpha|0\rangle|0\rangle+\beta|1\rangle|0\rangle \\ \mapsto & \alpha|0\rangle|0\rangle+\beta|1\rangle|1\rangle. & \end{array}$$

The later one is an entangled state, and not a copy of the input state.

One can apply the same logic as the above and gets

$$\mathcal{U}|0\rangle|b\rangle = |0\rangle|0\rangle,$$
$$\mathcal{U}|1\rangle|b\rangle = |1\rangle|1\rangle.$$

But,

$$\begin{aligned}\mathcal{U}(|0\rangle+|1\rangle)\otimes|b\rangle &= \mathcal{U}|0\rangle|b\rangle+\mathcal{U}|1\rangle|b\rangle \\ &= |0\rangle|0\rangle+|1\rangle|1\rangle.\end{aligned}$$

Clearly, although it can copy two mutually orthogonal states perfectly, it cannot copy an arbitrary state.

The inability to perfectly copy an unknown state restricts the capacity of an eavesdropper in quantum domain than in classical cases, assuring some security. However, there are cloning machines indeed that can copy an arbitrary state up to some fidelity, if not perfect. We'll discuss it later.

2.6 Some Frequently Useful Results

We often use $|a\rangle|b\rangle$ in place of $|a\rangle\otimes|b\rangle$.

Result 2.6.1. $|a\rangle|b\rangle\ \langle c|\langle d| = |a\rangle\langle c|\otimes|b\rangle\langle d|$

Result 2.6.2. $\left(\hat{O}_1\otimes\hat{O}_2\right)(|c\rangle\otimes|d\rangle) = \left(\hat{O}_1|c\rangle\right)\otimes\left(\hat{O}_2|d\rangle\right)$

Result 2.6.3. $\mathsf{Tr}\left(|a\rangle\langle b|\right) = \langle b|a\rangle$

Result 2.6.4. *For two density operators ρ_A,ρ_B with two parties Alice and Bob,*

$$\begin{aligned}\mathsf{Tr}_A\left(\rho_A\otimes\rho_B\right) &= \mathsf{Tr}\left(\rho_A\right)\rho_B, \\ \mathsf{Tr}_B\left(\rho_A\otimes\rho_B\right) &= \rho_A\mathsf{Tr}\left(\rho_B\right)\end{aligned}$$

Result 2.6.5. *For elements* $|z_1\rangle, |z_2\rangle \in \mathbb{C} \times \mathbb{C}$,

$$\begin{aligned}\langle z_1|z_2\rangle + \langle z_2|z_1\rangle &= 2\,\mathbf{Re}\langle z_1|z_2\rangle,\\ \langle z_1|z_2\rangle - \langle z_2|z_1\rangle &= i \cdot 2\,\mathbf{Re}\langle z_1|z_2\rangle.\end{aligned}$$

2.7 Quantum Key Distribution

The existing classical way of communicating secret information depends mainly on the RSA cryptosystem where the security is based on the hardness of factorization etc. However, Shor's algorithm can efficiently break the RSA encryption. As soon as a quantum computer is realized, the current way of communication is no more secure. On the other hand, QKD has shown the promise to establish secret keys among two parties, albeit a bit inefficient in terms of key-rate etc.

From its inception, this particular discipline has come up with some protocols to allow two legitimate parties share some secret key bit-stream. It started with BB84 [BB84, BB14] which is a prepare-and-measure type scheme. It uses two conjugate bases to encode a classical bit into a quantum bit. Later, Bruß [Bru98, 6s] extended it with a third basis for encoding spanning the whole Bloch sphere. Bennett made a simplification to his 4s protocol to come up with a 2s protocol using only two non-orthogonal bases [Ben92, 2s]. Ekert provided an entanglement based protocol [Eke91, eb] that certifies security via Bell violation. Bennett *et al.* further (being a bit critic) connected [BBM92] this eb protocol to BB84 and shown that Bell violation is not essential to provide security in QKD.

Subsec. 2.7.1 provides a brief overview of the BB84 protocol and the encoding model. The basic structure for classical post-processing to filter a secret key from a partially secret and erroneous key is also briefed therein. Subsec. 2.7.2 discusses the other protocols as mentioned above. Subsec. 2.7.3 describes a broad overview of the eavesdropping models with some illustrations.

2.7.1 The BB84 protocol [BB84, p&m]

The BB84 protocol [BB84] can establish an information-theoretically secure secret key between two distant parties. Alice encodes a stream of classical bits (*cbits*) into an ensemble of quantum bits (*qubits*) using two *mutually unbiased bases* (MUBs). She then transmits the qubits one-by-one over a quantum channel. Bob, at the receiving end, measures individually in one of the encoding bases, chosen randomly. Later they reconcile bases publicly over an authenticated classical channel to filtrate a *sifted key*.

However, the quantum channel may introduce errors in the flying qubits. Or, an eavesdropper may try to listen to the channel. She may either measure the qubit in some basis (that may not match that of Alice) and resend the resulted qubit to Bob. Or, she may attach some ancilla qubit with it, evolve the joint system unitarily, and release the carrier qubit towards Bob, and measures her qubit later. In any case, her attempt to learn the state of the qubit introduces an error, which is detectable by the legitimate parties by estimating the error rate. Here comes the advantage of quantum mechanics over classical cryptography. Masking a classical bit by a basis state does the job difficult for Eve to choose a correct basis. A wrong choice randomizes the state where Bob may get the error.

Alice uses two bases, indexed 0,1: basis 0 is the computational basis $\{|0\rangle, |1\rangle\}$, while basis 1 is the Hadamard basis $\{|+\rangle, |-\rangle\}$. Fig. 2.4 captures the encoding process. A more details of the encoding is given at the end of this section.

Table 2.1 | BB84 exemplified.

Alice encodes cbits 0,1 in randomly chosen bases z,x. Bob measures in randomly chosen bases z,x. For matching bases, he gets the outcome in sync that defines the raw key. For mismatching bases, his outcomes are randomized (?) and are disregarded. This is the scenario with no eavesdropping.

Alice	cbits	0	1	0	0	1	0	1	1	1	0	0
	bases	z	x	x	x	z	x	z	z	x	z	z
	qubits	$\|0\rangle_z$	$\|1\rangle_x$	$\|0\rangle_x$	$\|0\rangle_x$	$\|1\rangle_z$	$\|0\rangle_x$	$\|1\rangle_z$	$\|1\rangle_z$	$\|1\rangle_x$	$\|0\rangle_z$	$\|0\rangle_z$
Bob	bases	z	z	x	z	x	x	z	x	z	z	x
	cbits	0	?	0	?	?	0	1	?	?	0	?

Let's describe the protocol briefly. An example in Table 2.1 may help to better understand the steps of the protocol.

1. Alice wants to share a bit-string $a_1a_2\ldots a_n$ with Bob, where each bit is chosen randomly. She chooses another random bit-string $\beta_1\beta_2\ldots\beta_n$, where each bit stands for her choice of the encoding basis. She then prepares the qubit-stream $|a_1\rangle_{\beta_1}|a_2\rangle_{\beta_2}\ldots|a_n\rangle_{\beta_n}$ and sends the qubits one-by-one over the quantum channel to Bob. For instance, if $a_i = 1 = b_i$, she sends $|-\rangle$.

2. Bob chooses a base-string $\beta_1'\beta_2'\ldots\beta_n'$ randomly and measures the received qubits one-by-one in those bases. The resulting bit-string $a_1'a_2'\ldots a_n'$ is called the *raw key*.

3. They publicly compare their base-strings. For the mis-matched positions, they discard the bits in their own bit-string. The remaining bit-string is called the *sifted key*. Since Bob may choose roughly half of the bases wrong, the length of the sifted key is half of the raw key.

4. The sifted keys match only in error-free channel. For an erroneous channel, they need to estimate the error-rate from their sifted key. They choose half of the positions randomly and compare their respective bits publicly. The fraction of the mismatches provides a good estimate of the error-rate for a large key. If the error-rate is crosses some threshold, they abort the protocol, as no further filtering is possible beyond that. Otherwise, the continue further with a refined string where the compared bits are discarded. Again, the key-length is halved.

5. The remaining bit-string contains errors, but at unknown locations. They now perform classical post-processing of their respective bit-strings. *Information reconciliation* allows them to discard or remove the errors and filtrate an identical shared string. Although error-free, Eve may have partial knowledge on this string which could be essentially eliminated using some *privacy amplification* procedure.

Classical post-processing

Eavesdropping introduces disturbance in the quantum channel. Thus, the sifted key with Bob may not match that of Alice. They can estimate the error-rate by publicly comparing some part of it. If it is within a threshold value, they can perform the classical post-processing. First, they should remove or rectify the errors to come up with a common string. But, Eve still may have some knowledge on it, that could be eliminated by shortening the string intelligently. Some more details can be found in Chap. 7.

2.7.1.1 More about the transmission

The MUBs and the states Alice and Bob want to share a secret key using BB84 protocol. To encode a classical bit into a quantum bit, Alice randomly chooses a basis from $\mathfrak{B}_{xy} = \{|x\rangle, |y\rangle\}$ and $\mathfrak{B}_{uv} = \{|u\rangle, |v\rangle\}$, where

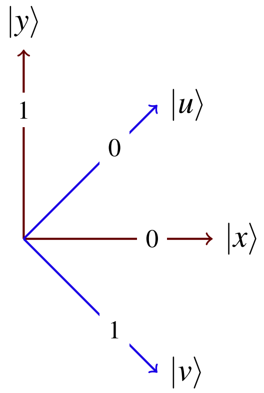
$$|x\rangle = \frac{1}{\sqrt{2}}\left(|u\rangle + |v\rangle\right), \qquad |y\rangle = \frac{1}{\sqrt{2}}\left(|u\rangle - |v\rangle\right), \tag{2.1}$$

i.e., the bases are conjugate to each other. She encodes 0 into either $|x\rangle$, or, $|u\rangle$ depending on whether she has randomly chosen a basis xy, or, uv, respectively. Similarly, she encodes 1 into $|y\rangle$, or, $|v\rangle$. The encoding bases are better understood from Fig. 2.4.

Alice encodes her key-bits, each as a polarized photon, and sends it to Bob.

Figure 2.4 | Two encoding bases of BB84.

Two encoding bases for the sender, conjugate to each other: the computational basis $\{|x\rangle, |y\rangle\}$, and the Hadamard basis $\{|u\rangle, |v\rangle\}$. 0 is encoded by $|x\rangle$, or $|u\rangle$; 1 is encoded by $|y\rangle$, or $|v\rangle$.



Receiving end Upon receiving a signal, Bob measures it randomly in one of the bases xy, or, uv and registers the outcome cbit. Once done with all the transmitted signals, he is left with a bit-stream, called the *raw key*.

Then, the legitimate parties use the public classical channel for basis reconciliation. Both the legitimate parties makes their choice of bases public. Wherever they agree, the corresponding cbits are retained and the rest of the bits due to mismatched bases are thrown out. Approximately half of the bit-stream gets wasted. The remaining bit-stream is called the *sifted key*.

Then, they use the classical channel to estimate the error-rate of the quantum channel. For that, they choose a stipulated sequence of bits from the sifted key and publicly tally their bits for those positions. The amount of mismatches are noted down, and the fraction of mismatches is denoted the *quantum bit error rate* (QBER).

A classical post-processing is done on the sifted key, if the error rate is within a tolerable limit.

Alice's encoding For encoding, Alice uses two orthonormal bases conjugate to each other: the *computational basis*, and the *Hadamard basis*. The basis states correspond to the eigenstates of the phase-flip operator σ_z and bit-flip operator σ_x, respectively. The following notations for the bases and their states are used interchangeably throughout the thesis.

Table 2.2 | Different symbols to denote Alice's encoding.

Alice uses two conjugate bases each having two basis states to encode the key-bits.

Computational basis		Hadamard basis	
Basis	States	Basis	States
$+$	$\{\lvert 0\rangle, \lvert 1\rangle\}$	$\times$	$\{\frac{\lvert 0\rangle+\lvert 1\rangle}{\sqrt{2}}, \frac{\lvert 0\rangle-\lvert 1\rangle}{\sqrt{2}}\}$
Various labellings used			
xy	$\{\lvert x\rangle, \lvert y\rangle\}$	uv	$\{\lvert u\rangle, \lvert v\rangle\}$
0	$\{\lvert 0\rangle^0, \lvert 1\rangle^0\}$	1	$\{\lvert 0\rangle^1, \lvert 1\rangle^1\}$
Z	$\{\lvert +z\rangle, \lvert -z\rangle\}$ $= \{\lvert 0_z\rangle, \lvert 1_z\rangle\}$	X	$\{\lvert +x\rangle, \lvert -x\rangle\}$ $= \{\lvert 0_x\rangle, \lvert 1_x\rangle\}$

$\bar{\beta}$ denotes the conjugate of a basis β. The Hadamard transform $\mathbb{H} := \frac{1}{\sqrt{2}}\left(\sigma_z + \sigma_x\right)$ flips the bases ($\mathbb{H} : \beta \mapsto \bar{\beta}$) while the basis states can be written with respect to the computational basis elements as $\lvert a\rangle^{\beta} = \mathbb{H}^{\beta}\lvert a\rangle$ for $a = 0, 1$. The orthogonal counterpart of a state $\lvert a\rangle$ is denoted by $\lvert a \oplus 1\rangle$ or $\lvert \bar{a}\rangle$. Alice encodes the cbit 0 into a qubit in state $\lvert x\rangle$ or $\lvert u\rangle$, and encodes 1 into $\lvert y\rangle$ or $\lvert v\rangle$.

2.7.2 Other protocols

2.7.2.1 6s protocol [Bru98, p&m]

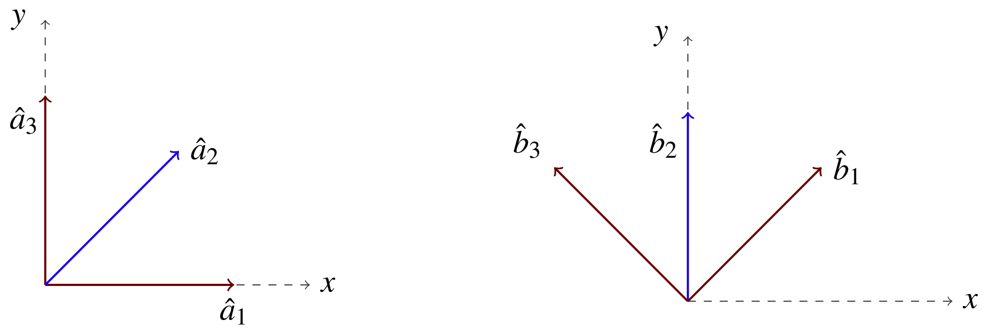
It considers 3 mutually unbiased bases for encoding two bits.

$$\begin{array}{llll}
\lvert 0_z\rangle & := \lvert 0\rangle, & \lvert 1_z\rangle & := \lvert 1\rangle, \\
\lvert 0_x\rangle & := \frac{1}{\sqrt{2}}(\lvert 0\rangle + \lvert 1\rangle), & \lvert 1_x\rangle & := \frac{1}{\sqrt{2}}(\lvert 0\rangle - \lvert 1\rangle), \\
\lvert 0_y\rangle & := \frac{1}{\sqrt{2}}(\lvert 0\rangle + i\lvert 1\rangle), & \lvert 1_y\rangle & := \frac{1}{\sqrt{2}}(\lvert 0\rangle - i\lvert 1\rangle).
\end{array}$$

The bases are also denoted in short as Z, Y, X, respectively. In the QKD protocol, the prior probability to randomly choose one such basis is $\frac{1}{3}$.

Figure 2.5 | Measurement bases for E-91 protocol.

Alice and Bob uses two measurement setups, as in left and right, respectively. The measurement directions with Alice are $\hat{a}_1,\hat{a}_2,\hat{a}_3$, corresponding to azimuthal angles $\phi_a = 0, \frac{\pi}{4}, \frac{\pi}{2}$. Bob uses directions $\hat{b}_1,\hat{b}_2,\hat{b}_3$, with angles $\phi_b = \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$. The matching directions lead to the key, while the mismatched directions allow to test CHSH violation.



The 6s protocol provides more symmetry than its 4s counterpart when the Bloch sphere representation is considered. The three bases describe the three mutually orthogonal directions that span the entire Bloch sphere. While, for the BB84 protocol, the two bases covers only two orthogonal directions and span a great circle on the Bloch sphere. That symmetry for the 6s protocol often found to provide less number of parameters to describe an optimal attack, reduces Eve's maximum information on the transmitted states, and greatly simplifies the security analysis. For an IR-attack, the 6s protocol can tolerate a QBER of 33%, compare to 25% of the 4s protocol.

2.7.2.2 E-91 protocol [Eke91, eb]

It is an *eb* scheme that uses an EPR-pair to encode a classical bit and distribute the two particles to the two legitimate parties.

1. Let's choose the EPR-pair (either distributed by a source, or by Alice) as the following maximally entangled Bell state.

$$|\psi^-\rangle_{\mathtt{AB}} \;:=\; \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)_\beta .$$

   It is worthy to note that the state looks same in any orthonormal basis β.

2. As in Fig. 2.5, each of Alice and Bob use three measurement directions $\hat{a}_i$, and $\hat{b}_j$ $(i,j = 1,2,3)$, respectively, with outcomes ± 1 . They publicly tally their measurement choices. The outcomes can be divided into two groups as follows:

   (a) For mismatched directions, they make the outcomes public, which allows them to test the CHSH inequality. So far the inequality is violated, there is some entanglement remained and they can perform classical post-processing on their strings to come up with a shared secret key. The inequality is described below.

Let $p_{\lambda,\mu}(\hat{a}_i,\hat{b}_j)$ denotes the probability that Alice-Bob chooses measurement direction $\hat{a}_i,\hat{b}_j$ and get outcome $\lambda,\mu \in \{+1,-1\}$. Define the correlation coefficients (i.e., expected value of the outcomes in this case)

$$E(\hat{a}_i,\hat{b}_j) \quad := \quad p_{+,+}(\hat{a}_i,\hat{b}_j)+p_{-,-}(\hat{a}_i,\hat{b}_j)-p_{+,-}(\hat{a}_i,\hat{b}_j)-p_{-,+}(\hat{a}_i,\hat{b}_j).$$

Then, the CHSH correlation coefficient is defined as [4]

$$S \quad := \quad E(\hat{a}_1,\hat{b}_1)-E(\hat{a}_1,\hat{b}_3)+E(\hat{a}_3,\hat{b}_1)+E(\hat{a}_3,\hat{b}_3).$$

In absence of any disturbance (eavesdropping), the quantum measurement will find the value $S=-2\sqrt{2}$, which violates the CHSH inequality $|S|\leq 2$. Presence of noise leaves it in the range $2<|S|<2\sqrt{2}$. If $|S|<2$, the shared state is no more entangled.

(b) For matched directions, their measurements along the same axis are anticorrelated, i.e.,

$$E(\hat{a}_2,\hat{b}_1)=E(\hat{a}_3,\hat{b}_2)=-1.$$

The outcomes for these measurements are the raw key, once Bob flips his bits. They can perform classical post-processing on their strings to come up with a shared secret key.

An attempt to eavesdropping actually introduces *elements of physical reality* to perturb the orientation of the particle and reduces the quantum value of S.

Compare with the BB84 protocol: E-91 protocol provides less throughput than BB84, as the fraction of reconciled bases is lower than that of BB84. But, eavesdropping causes higher error rate for E91. For instance, IR attack introduces 33% error than 25% for BB84. The maximum information that Eve learns is also less than that for BB84. The symmetry of the qubit states simplifies the security analysis.

Note 2.1. *A lot of similar results are observed among the E-91 protocol and the BB84 protocol, that points to an interesting connection between these two protocols. Note that, while E-91 protocol deals with two qubits for distributing data, the BB84 protocol uses only one. Thereby, any unitary evolution (that realizes the noise in the channel) that may evolve the two qubits in the earlier scheme, have a counterpart unitary evolution to work*

[4] The only negative sign corresponds to the two directions making 145°angle, all other angles are 45°.

on only one of the qubits as desired by the later scheme.

$$\mathcal{U}_1 \otimes \mathcal{U}_2 |\phi^-\rangle = \mathbb{1} \otimes \mathcal{U}_2 \mathcal{U}_1^\dagger |\phi^-\rangle.$$

However, Bell test is not essential to certify security and a simpler measurement can serve the purpose as follows.

2.7.2.3 BBM92 protocol [BBM92, connects p&m to eb]

It is a *p&m* protocol to complement *eb* E91. A source distributes legitimate parties halves of EPR-pairs. The legitimate parties measure only two MUBs (Z, X). It makes BBM92 more efficient than E91 which uses three bases. During basis reconciliation, their raw key consists of the bits for matched bases where they obtain correlated measurement outcomes. A part of it is then sacrificed to publicly estimate the QBER.

BBM92 actually connects E91 to BB84, while being a critic of the former. If Alice is to distribute the data, her measurement (Z/X) would collapse the EPR state to a product state, as if she prepared the state for Bob and sent to him. That's precisely the scenario for BB84 which doesn't need a Bell test to certify security. This observation is the base for *eb* version of the *p&m* protocols, useful to prove the security.

The security argument is as follows. An eavesdropper cannot gain any information without introducing disturbance which is detectable. Suppose, Eve interacts via a probe, which was in state $|E\rangle$ before interaction. For Alice's arbitrary two states $|s\rangle, |t\rangle$, let the unitary interaction go as follows:

$$\begin{aligned} |s\rangle|E\rangle &\mapsto |s\rangle|E_s\rangle, \\ |t\rangle|E\rangle &\mapsto |t\rangle|E_t\rangle. \end{aligned}$$

Unitarity preserves the inner product:

$$\langle s|t\rangle\langle E|E\rangle = \langle s|t\rangle\langle E_s|E_t\rangle.$$

If the Alice's two states are non-orthogonal, they have a non-zero overlap. Then, the overlap between Eve's two post-interaction states must be 1, leaving them indistinguishable.

In order to distinguish them, she must introduce an error. Then, the PIJS cannot be a product state, it goes entangled. The legitimate parties can detect it by comparing a subsequence of the raw key.

2.7.2.4 2s protocol [Ben92, p&m]

In 1992, Bennett suggested a simplification on his celebrated BB84 protocol by incorporating only two non-orthogonal states for encoding. It serves the purpose that an eavesdropper cannot distinguish them unambiguously. Although it allowed the experiments to become easier, it is performance is quite poor than BB84, e.g., the noise tolerance level. The protocol is as follows, illustrated for a special case.

1. Alice encodes cbits 0, 1 into qubits $|a_0\rangle, |a_1\rangle$, respectively. The qubits are non-orthogonal to forbid Eve getting full information. For simplicity, let's consider

$$|a_0\rangle = |0_z\rangle \qquad |a_1\rangle = |0_x\rangle.$$

2. Bob randomly measures σ_z or σ_x.
   Note that, σ_z detects $|a_0\rangle$ perfectly, while randomizes $|a_1\rangle$. On the other hand, σ_x detects $|a_1\rangle$ perfectly, while randomizes $|a_0\rangle$. The distribution of Bob's outcome is given below.

$\mathtt{B}$	σ_z		σ_x	
$\mathtt{A}$	0_z	1_z	0_x	1_x
0_z	1	0	$1/2$	$1/2$
0_x	$1/2$	$1/2$	1	0

Table 2.3 The probabilities $P(\mathtt{B} = b|\mathtt{A} = a)$.

Comes to Bob's strategy to interpret, consider the probabilities $P(\mathtt{A} = a|\mathtt{B} = b)$.

$$P(\mathtt{A} = 0_x|\mathtt{B} = 1_z) \quad = \quad P(\mathtt{A} = 0_z|\mathtt{B} = 1_x) = 1.$$

Thus, whenever Bob gets 1_z, he interprets Alice's bit as 1, and whenever Bob gets 1_x, he interprets Alice's bit as 0. The remaining outcomes are labeled *inconclusive*.

3. Bob intimates Alice publicly the inconclusive positions and individually discard those bits. The conclusive remnant should agree with probability 1 in absence of eavesdropping, which consists the *raw key*.

4. To detect any eavesdropping that might have altered the states, the legitimate parties can publicly tally a fraction of their *supposed-to-be identical* bits (and later discard) to estimate the error rate, that may follow a classical post-processing.

2.7.3 Eavesdropping strategies

As Alice's qubits moves through the quantum channel, Eve has the liberty to attack the qubits in various ways, and measure till she can maintain her register. Some categorization can be considered as follows.

- *Individual attack:* Eve can interact as well measure each qubit separately.
- *Coherent attack:* Eve can interact as well measure a whole chunk of qubits at a go.
- *Collective attack:* Eve can interact individually, but measures a whole chunk of qubits. It leads to more information than individual attack.
  The measurement timing for Eve is also a factor. She can measure either immediately, or after post-processing, or anywhere in between depending on her capacity to store her ancilla system.

The first kind of attack model itself may be considered with different varieties.

- *Intercept-Resend:* Eve measures each qubit on the fly. She may measure in the same bases as that of the legitimate parties.
- *Measure in the Intermediate basis:* Eve measures each qubit on the fly. She may choose her own orthonormal bases for measurement. An optimal attack corresponds to the measurement in the *Breidbart basis* [BBBW82].
- *Translucent attack:* Eve can attach an ancilla with Alice's qubit, evolve the joint system unitarily to glean some information from Alice's qubit into her ancilla, and measures later until bases reconciliation is done by the legitimate parties. For instance, [FGG$^+$97] falls into that category.

In any of such category or sub-category of eavesdropping under consideration, one of the important objective remains is to figure out those attacks which leads to maximum information gain among all others. Our objective in this thesis is to study such attacks in the name of *optimal eavesdropping*. Any such optimal attack corresponds to a strategy that interprets Eve's measurement outcomes as the best guess on Alice's signal: such a strategy is called *optimal strategy*.

For individual attacks, all the three parties are left with a classical random variable out of their measurement results, and the joint distribution can directly be analyzed for the classical post-processing.

We discuss in this section the first two of the individual attacks. The translucent attack will be discussed in detail in the subsequent chapters. We'll also discuss coherent attacks later in a dedicated chapter as well. The protocol that we'll consider as BB84 for consistency and comparison of results. But, the results for the corresponding six-state protocol will also be briefed for completeness.

2.7.3.1 Intercept-Resend attack

In this attack model, Eve introduces a fixed QBER of 25% in the channel between Alice and Bob. It is so because, Eve measures in the matching basis 50% of the time, and randomizes the state in other 50% of the cases. In the former case, (after basis reconciliation, Bob's base is same as that of Alice,) Bob's base matches both of Alice and Eve leading Bob get the correct result always. In the later case, as Eve has altered the state, Bob gets wrong result in half the cases, leading to the QBER.

We'll show here that Eve gains an information that amounts to 1/2 *bpsp* (bits per send photon), with success probability 3/4.

The MI among Alice and Eve is defined as follows

$$I_{\mathtt{AE}} \;=\; H(\mathtt{A}) - H(\mathtt{A}|\mathtt{E}).$$

It indicates the drop in entropy in Alice's random variable due to measurement knowledge of Eve. The above two entropies are defined on the prior and the posterior probability distribution, respectively. Since Alice's distribution is uniform, $H(\mathtt{A}) = 1$. The later entropy is defined as follows

$$\begin{aligned} H(\mathtt{A}|\mathtt{E}) &= \sum_e P(\mathtt{E}=e)H(\mathtt{A}|\mathtt{E}=e), \\ H(\mathtt{A}|\mathtt{E}=e) &= -\sum_a P(a|e)\log_2 P(a|e). \end{aligned}$$

The posterior probability $P(a|e)$ that Eve actually gets Alice's bit a, given that Eve's outcome is e, can be given by the Baye's rule

$$\begin{aligned} P(a|e) &= \frac{P(e|a)P(a)}{P(e)}, \\ P(e) &= \sum_a P(e|a)P(a). \end{aligned}$$

The distribution of Alice-Eve is given as below For any state from Alice, Eve can get one of the four states $0_z, 1_z, 0_x, 1_x$. Let, due to basis reconciliation, (*w.l.o.g.*) Alice and Bob

E A	0_z	1_z	0_x	1_x
0_z	1	0	1/2	1/2
1_z	0	1	1/2	1/2
0_x	1/2	1/2	1	0
1_x	1/2	1/2	0	1
	1/2	1/2	1/2	1/2

Table 2.4 The probabilities $P(\mathtt{E}=e|\mathtt{A}=a)$. Last row indicates the probabilities $P(\mathtt{E}=e)$.

agrees with the basis Z. Then, we get

$$P(\mathtt{A}=0_z|\mathtt{E}=0_z)=1,\quad P(\mathtt{A}=0_z|\mathtt{E}=0_x)=1/2,\quad P(\mathtt{E}=e)=1/2.$$

Therefore,

$$I_{\mathtt{AE}} \;=\; 1-\frac{1}{2}\,h(1)-\frac{1}{2}\,h\left(\frac{1}{2}\right)=\frac{1}{2}.$$

However,

$$I_{\mathtt{AB}} \;=\; 1-h\left(\frac{1}{4}\right)\approx 0.19 \;<\; 0.5=I_{\mathtt{AE}}.$$

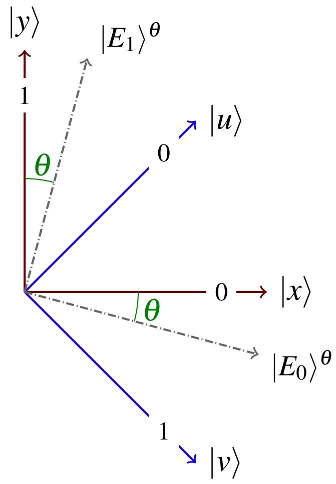
Since Bob's information is less than Eve's information on Alice's bit, an OWCPP is unfaithful.

2.7.3.2 Measure in the Intermediate basis [BBBW82]

In this case, Eve measures in an intermediate basis than what Alice uses. The optimal attack stands for the Breidbart basis that is the clockwise $\pi/8$ rotation of the computational basis.

Figure 2.6 | Attacking BB84 with intermediate basis: optimal for Breidbart basis [BBBW82].

Two MUBs for Alice to encode the bitstream: the computational basis $\{|x\rangle,|y\rangle\}$, and the Hadamard basis $\{|u\rangle,|v\rangle\}$. 0 is encoded by $|x\rangle$, or $|u\rangle$; 1 is encoded by $|y\rangle$, or $|v\rangle$. Eve uses an intermediate basis $|E_0\rangle^\theta,|E_1\rangle^\theta$ which is the clockwise θ rotation of the computational basis. The optimal attack stands for the Breidbart basis which corresponds to clockwise $\pi/8=22.5^0$ rotation.



Eve's probability of correct guess is

$$P_c \quad = \quad \cos^2(\pi/8) = \frac{1}{2}\left(1+\frac{1}{\sqrt{2}}\right) = 0.854,$$

corresponding to

$$\text{QBER} \quad = \quad 2P_c(1-P_c) = 25\%,$$

and Shannon information gain

$$I_{\text{E}} \quad = \quad 1 - H(P_c) = 0.399.$$

Here we provide a brief outline of the proof.

Eve measures with the projectors

$$E_0 := |E_0\rangle\langle E_0|, \quad E_1 := |E_1\rangle\langle E_1| \quad \text{with } E_0 + E_1 = \mathbb{1},$$

where the measurement directions correspond to anti-clockwise θ rotation of the computational basis, defined as follows

$$\begin{aligned} |E_0\rangle \quad &= \quad \cos\theta|x\rangle - \sin\theta|y\rangle, \\ |E_1\rangle \quad &= \quad \sin\theta|x\rangle + \cos\theta|y\rangle. \end{aligned} \tag{2.2}$$

As visible in Fig. 2.6, $|E_0\rangle$ is close to $|x\rangle,|v\rangle$, while $|E_1\rangle$ is close to $|y\rangle,|u\rangle$. Thus, Eve's *strategy* will be to interpret outcome 0 as $|x\rangle,|v\rangle$ and outcome 1 as $|y\rangle,|u\rangle$. Her probability of success is then

$$\begin{aligned} F_{\text{E}} = \Pr(\mathsf{success}) \quad &= \quad \frac{1}{4}\big(\langle x|E_0|x\rangle + \langle v|E_0|v\rangle + \langle y|E_1|y\rangle + \langle u|E_1|u\rangle\big) \\ &= \quad \frac{1}{4}\left(|\langle x|E_0\rangle|^2 + |\langle v|E_0\rangle|^2 + |\langle y|E_1\rangle|^2 + |\langle u|E_1\rangle|^2\right). \end{aligned}$$

For the defined measurement directions, we get

$$F_{\text{E}} \quad = \quad \frac{1}{2} + \frac{1}{4}(\cos 2\theta + \sin 2\theta).$$

The maximum takes place (by equating its first derivative to zero) when $\tan 2\theta = 1$, i.e., at $\theta = \pi/8 = 22.5^0$. The corresponding basis is called the Breidbart basis. In that case, the QBER with Bob is also minimized, as Eve measures and passes the states $|E_0\rangle,|E_1\rangle$ to Bob.

Bob commits an error, when say, Alice sends x, Bob gets y etc. (other cases are symmetric), irrespective of whatever state (E_0,E_1) is send by Eve. Due to Eve's strategy, when Alice sends $|x\rangle$, Eve gets $|E_0\rangle$ with proportion $F_{\mathtt{E}}$, and she gets $|E_1\rangle$ with proportion $1-F_{\mathtt{E}}$. Thus, Bob experiences a QBER

$$\begin{aligned} D &= F_{\mathtt{E}}\langle E_0|B_y|E_0\rangle+(1-F_{\mathtt{E}})\langle E_1|B_y|E_1\rangle \\ &= F_{\mathtt{E}}|\langle y|E_0\rangle|^2+(1-F_{\mathtt{E}})|\langle y|E_1\rangle|^2 \\ &= F_{\mathtt{E}}\sin^2\theta+(1-F_{\mathtt{E}})\cos^2\theta, \end{aligned}$$

when Bob's projector $B_y=|y\rangle\langle y|$ clicks. It is enough to consider this case, as the other prior probability weightings of $|x\rangle,|y\rangle$ are all equal. It turns out that

$$\begin{aligned} D &= 2F_{\mathtt{E}}(1-F_{\mathtt{E}}), \\ D &= \frac{1}{4}. \end{aligned}$$

So, the legitimate parties can detect Eve if they find a quarter of the data incorrect.

However,

$$I_{\mathtt{AB}} = 1-h\left(\frac{1}{4}\right)\approx 0.19 \; < \; 0.39 = I_{\mathtt{AE}}.$$

Since Bob's information is less than Eve's information on Alice's bit, an OWCPP is unfaithful.

CHAPTER 3

EXISTING WORKS ON OPTIMAL EAVESDROPPING

The framework of optimal eavesdropping using ancillary probe on the BB84 protocol [BB84] was addressed in [FGG$^+$97]. The attacker applies a suitable unitary evolution to entangle her probe with the senders signal, and the joint system is later measured by a specific POVM. The amount of information gathered by Eve is quantified by two functions: IG, and MI. The maximum amount of information is then calculated for both the quantifiers for each of the signal basis. An optimal interaction is suggested there that can achieve the maximum of information. The associated optimal measurement is specified as well. Verification of optimality is subjected to satisfying a set of necessary and sufficient conditions.

An attackers presence gets detected by estimating the amount of error introduced and is calculated for each of the bases. The above analysis considered the general scenario with asymmetric error rate for the two bases. Generally, for practical purposes, the error rate becomes equal across the two bases. Such a symmetric attack is further discussed. Another optimal interaction is specified therein.

Such an attack can withstand some level of disturbance in the channel, beyond which a key distillation is not possible. That critical value is calculated following a logic that the receivers information must dominate that of the attacker. The key-rate can be calculated from the difference between these two informations, and we have plotted it within the tolerable error limit.

Some generalizations are done in [AP17], and we have placed some of that discussion here only to retain the flow of thoughts.

3.1 Brief overview

3.1.1 Eavesdropping

A third party (Eve) is allowed to tamper the quantum channel. However, any approach to learn the state of the qubit introduces an error which is further detectable by the recipient. The legitimate parties can estimate the *quantum bit error rate* (QBER) by discussing over the public channel on a part of the sifted key. Within a threshold value QBER*, a classical post-processing (CPP) is faithful to filter a shared secret on which Eve has virtually no information.

3.1.2 The attack model by Fuchs *et al.* [FGG$^+$97]

An advanced eavesdropping model [FGG$^+$97] is to extract the information of a transmitted qubit via an ancilla qubit by interacting unitarily. Given that the attacker is allowed to defer her measurement until after basis reconciliation, an one-way (OW) CPP is faithful if the estimated QBER remains below the critical value 0.1464 where the secret key-rate becomes zero. The authors could estimate the maximum *knowledge gain* (KG) by an attacker that eventually appeared a tight bound due to an witness interaction. Nonetheless, there could be infinitely many such saturating candidates (interactions) which are unitarily equivalent [AP17]. In that attack model, a candidate interaction must pass a formal verification of optimality, viz., a *necessary and sufficient condition*(NSC) [FGG$^+$97] involving the joint Hilbert space of the sender and the attacker.

Chapter organization

The section wise work-flow is as follows. Sec. 3.3 is dedicated to discuss the attack model provided by [FGG$^+$97]. Sec. 3.3.9 presents the excerpts of the attack: a circuit diagram, some practical sides of the eavesdropping like key-rate etc. All the detailed calculations and ideas are deferred to Sec. 3.4. The chapter ends with a brief conclusion and by mentioning the further scopes to explore.

The main results are briefly described in Sec. 3.3. It includes the mathematical framework of optimal eavesdropping, the functions to be optimized, a necessary and sufficient condition for optimality, an optimal interaction and its optimal POVM. Some technical details are provided in Sec. 3.3.11. The main results are however proved in Sec. 3.4 with the proofs in between the lines placed at the end.

3.2 Notations and some results in use

Notation: For a function parametrized by an index i, $\arg\max$ of the function is that particular index for which the functional value is maximum among all other indices. Mathematically,

$$\arg\max_{i \in I} f(i) = j \in I \ : \ f(j) = \max_{i} f(i).$$

3.2.1 Some useful results

Theorem 3.1 (Jensen's inequality). *If $f(x)$ is a concave function of x, then a convex combination of its functional values never exceeds the functional value of the same convex combination of its arguments, i.e.,*

$$\sum q_x f(x) \leq f\left(\sum q_x x\right).$$

Equality occurs when all the x's are equal.

In the domain $x \in [0,1]$, we consider the concavity of some functions as below.

Lemma 3.1. *$z(x) := \sqrt{x(1-x)}$ is a concave function of x.*

Hint: Observe that its double derivative is negative:

$$z''(x) = -\frac{1}{4z^3(x)} < 0.$$

Lemma 3.2. *The following function*

$$\phi(x) = (1+x)\ln(1+x) + (1-x)\ln(1-x)$$

is m.i., and non-concave.

Hint: The reason is that its first derivative is non-negative:

$$\phi'(x) \quad = \quad \ln\left(\frac{1+x}{1-x}\right) \ \geq \ \ln 1 = 0.$$

Result 4. *It is useful to define* $0 \ln 0 := 0$*, i.e.,*

$$\lim_{x \to 1} x \ln x = 0.$$

3.2.2 Concavity-diagram: Entropy, IG-max, MI-max, phiFunc

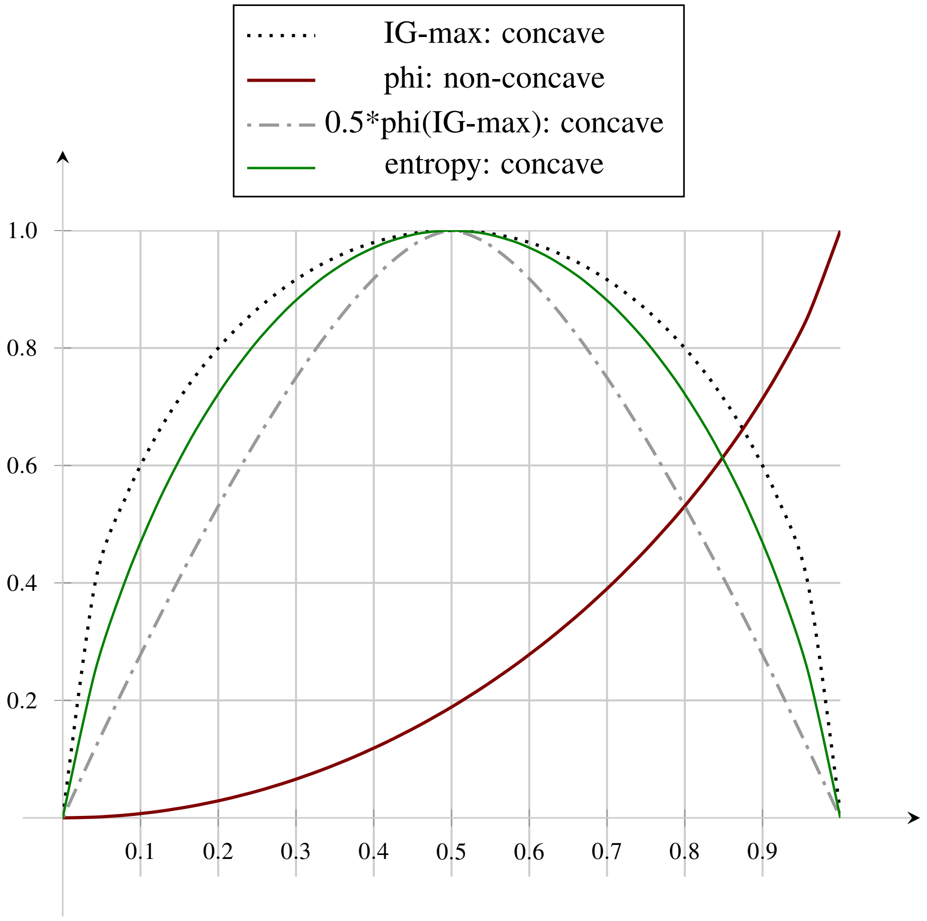
We plot here the following four functions to visualize their concavity in the domain $x \in [0,1]$.

$$
\begin{aligned}
z(x) &= \sqrt{x(1-x)} \\
\phi(x) &= (1+x)\ln(1+x) + (1-x)\ln(1-x) \\
\frac{1}{2}\cdot\phi(2z(x)) & \\
H(x) &= -x\log(x) - (1-x)\log(1-x).
\end{aligned}
$$

Note that, while $\phi(x)$ is m.i., but, non-concave, $\phi(z(x))$ is concave. We'll see later that $2z(x)$ corresponds to IG-max, while $\frac{1}{2}\phi(2z(x))$ corresponds to MI-max for error-rate x in the range of $[0,1]$.

Figure 3.1 | Plotted: four functions for concavity: $2z(x)$, $\phi(x)$, $\frac{1}{2}\phi(2z(x))$, $H(x)$**.**

All are concave except $\phi(x)$ which is m.i.



3.3 Broad overview: Our reformulation of Optimal incoherent Eavesdropping

Optimal eavesdropping means that an eavesdropper performs the interaction and the measurement in such a way that she can extract maximum information about the signal sent by Alice, ensuring that the disturbance at Bob's end remains bounded by a suitable threshold. In the QKD literature, it is interpreted as maximizing the information gain by Eve or mutual information between Alice and Eve. For BB84 protocol, considering the interaction to be unitary and restricted to equal prior ($p_x = \frac{1}{2} = p_y$), Fuchs *et al.* [FGG$^+$97] provided an upper bound on information gain and mutual information over all possible interaction-POVM pairs. A criterion to achieve the bounds was provided there. To show that these bounds are attainable, an interaction-POVM pair for unequal error rates and another for equal error rates were provided therein. These results are discussed briefly in this section. Since these results hold for equal prior, the subsequent sections follow the same assumption unless explicitly mentioned.

The analysis done here is kept close to [FGG$^+$97] for easy comparison. However, we have incorporated our own approach to tackle some of the problems and some results of [AP17] are placed here to allow the overall concept become more clearer.

3.3.1 Basic ingredients

3.3.1.1 Mathematical modeling of eavesdropping

Suppose, an eavesdropper Eve interferes the communication and involve a probe to interact unitarily with the qubit that was transmitted by Alice.

Suppose Alice has picked up a signal, say, $|x\rangle$ (corresponding density operator being $\rho_x^{\mathtt{A}} = |x\rangle\langle x|$), in the basis $\mathfrak{B}_{xy}$. Eve's probe was initially in state $|\psi_0\rangle$ (corresponding density operator $\rho_0^{\mathtt{E}} = |\psi_0\rangle\langle\psi_0|$). It interacts unitarily (the unitary operator $\mathcal{U}$) with the qubit sent by Alice. The post-interaction joint state $|X\rangle$ between Alice and Eve's system is an entangled state [1] realized by

$$|x\rangle \otimes |\psi_0\rangle \xrightarrow{\mathcal{U}} |X\rangle.$$

Thus, the state received by Bob is no more pure, but a simple mixture of the two basis states (here $\mathfrak{B}_{xy}$) chosen by Alice. So, Bob's density matrix is diagonal in the basis chosen by Alice. Thus, the Schmidt decomposition of the post-interaction joint state $|X\rangle$ can be written as

$$|X\rangle = \sqrt{\alpha}\; |x\rangle|\xi_x\rangle + \sqrt{1-\alpha}\; |y\rangle|\zeta_x\rangle,$$

[1] If order to gain some information of Alice's state, Eve must disturb it, allowing the joint system go entangled.

such that

$$|\xi_x\rangle \perp |\zeta_x\rangle, \tag{3.1}$$

where $|\xi_x\rangle, |\zeta_x\rangle$ are component of Eve's part of the joint state post interaction.

Similarly, when Alice sends $|y\rangle$, the post-interaction state $|Y\rangle$ must be of the form

$$|Y\rangle = \sqrt{\beta}\,|y\rangle|\xi_y\rangle + \sqrt{1-\beta}\,|x\rangle|\zeta_y\rangle,$$

such that

$$|\xi_y\rangle \perp |\zeta_y\rangle. \tag{3.2}$$

The density operator for the post-interaction state $|X\rangle$ is given by

$$\rho_x^{\mathtt{AE}} = |X\rangle\langle X| = \mathcal{U}\left(\rho_x^{\mathtt{A}} \otimes \rho_0^{\mathtt{E}}\right)\mathcal{U}^\dagger. \tag{3.3}$$

Eve's description of her system will be

$$\rho_x \;:=\; \rho_x^{\mathtt{E}} \;=\; \mathtt{tr}_{\mathtt{A}}\left(\rho_x^{\mathtt{AE}}\right) \;=\; \mathtt{tr}_{\mathtt{A}}\left(|X\rangle\langle X|\right), \tag{3.4}$$

where $\mathtt{tr}_{\mathtt{A}}$ represents the partial trace over Alice's qubit.

Since the interaction is done unitarily, it follows from Eqs. (2.1, 3.3) that

$$|X\rangle = \frac{1}{\sqrt{2}}\left(|U\rangle + |V\rangle\right), \qquad |Y\rangle = \frac{1}{\sqrt{2}}\left(|U\rangle - |V\rangle\right). \tag{3.5}$$

3.3.1.2 Eve's measurement

Before performing any measurement on her ancilla, Eve waits until Alice's declaration of her choice of basis.

Eve's measurement is considered to be a POVM [Fuc96, NC11]: $\{E_\lambda\}$ or $\{F_\lambda\}$ depending on whether Alice's choice is *xy* or *uv* basis. Denote them commonly as $\{M_\lambda\}^\beta$.

For the BB84 protocol, for the [FGG+97] attack model, and for a reconciled basis, Eve need to distinguish four states after an interaction. She needs to incorporate a generalized measurement with four outcomes labeled by $\lambda \in \{0,1,2,3\}$.

She then interprets the outcome following a *strategy* which is a rule for Eve to assign a guess for the state of the signal sent by Alice.

3.3.1.3 The probability space

Assume, Alice sends a signal in *xy* (or, *uv*) basis with the prior probabilities p_x, p_y (or, p_u, p_v) respectively. As Alice reveals her basis to be *xy*, Eve uses a POVM $\{E_\lambda\}$ to measure her probe. Considering $\mathcal{A}, \mathcal{B}, \mathcal{E}$ as the random variables corresponding to the signal sent by Alice, the signal received by Bob, and, the measurement outcome of Eve, the conditional probability of occurrence of various outcomes λ of that measurement for an input state is given as follows.

$$\begin{aligned} P_{\lambda x} &:= \Pr[\mathcal{E}=\lambda|\mathcal{A}=x] = \mathrm{tr}\left(\rho_x E_\lambda\right) = \langle X|\mathbb{1}\otimes E_\lambda|X\rangle, & (3.6)\\ P_{\lambda y} &:= \Pr[\mathcal{E}=\lambda|\mathcal{A}=y] = \mathrm{tr}\left(\rho_y E_\lambda\right) = \langle Y|\mathbb{1}\otimes E_\lambda|Y\rangle. & (3.7)\end{aligned}$$

The probability that Eve gets an outcome λ, while Alice uses xy basis is therefore

$$q_{xy}(\lambda) \ := \ \Pr[\mathcal{E}=\lambda] = P_{\lambda x}p_x + P_{\lambda y}p_y.$$

She needs to interpret the signal. Looking at outcome λ, Eve assigns a guess on the signal sent by Alice guided by some strategy. The posterior probability $Q_{x\lambda}$ (or $Q_{y\lambda}$) of the event that Alice had sent the signal x (or y) given that Eve has observed an outcome λ is given by Bayes' theorem.

$$\begin{aligned} Q_{x\lambda} &:= \Pr[\mathcal{A}=x|\mathcal{E}=\lambda] = \frac{P_{\lambda x}p_x}{q_{xy}(\lambda)},\\ Q_{y\lambda} &:= \Pr[\mathcal{A}=y|\mathcal{E}=\lambda] = \frac{P_{\lambda y}p_y}{q_{xy}(\lambda)}.\end{aligned}$$

3.3.1.4 Quantifying Eve's information gain

A simple approach that Eve can utilize these likelihoods in order to to perform a guess realized by the following function.

$$\arg\max\{Q_{x\lambda}, Q_{y\lambda}\} = \begin{cases} x, & \text{if } Q_{x\lambda} > Q_{y\lambda},\\ y, & \text{if } Q_{y\lambda} > Q_{x\lambda}.\end{cases}$$

A convenient measure of Eve's **information gain** for an outcome λ, as proposed in [FGG+97], is

$$G_{xy}(\lambda) \ := \ \left|Q_{x\lambda} - Q_{y\lambda}\right|.$$

On average, Eve's information gain over all outcomes is

$$G_{xy} \ := \ \sum_\lambda q_{xy}(\lambda)G_{xy}(\lambda) = \sum_\lambda \left|P_{\lambda x}p_x - P_{\lambda y}p_y\right|.$$

In particular, for equiprobable signals,

$$G_{xy} \ = \ \frac{1}{2}\sum_\lambda \left|P_{\lambda x} - P_{\lambda y}\right|.$$

A more sophisticated way to process her outcome-statistics is **mutual information** [CT06] that keeps track of all the q_λ and $Q_{i\lambda}$ of her observations. For equal prior probabilities of the input signals, this is given by

$$I_{xy} := \ln 2 + \sum_\lambda q_{xy}(\lambda)\left(Q_{x\lambda}\ln Q_{x\lambda} + Q_{y\lambda}\ln Q_{y\lambda}\right).$$

Similarly, one can define G_{uv}, I_{uv}, D_{uv} while considering Alice's signal was prepared in uv basis. For simplicity, we drop the subscripts xy and uv, and use G, I, when both the bases to be considered in discussion.

3.3.1.5 Eve's strategy and probability of success $\mathsf{Pr(success)}$ on a correct guess

Definition 3.1. *A **strategy** S of Eve is a function $S(\lambda)$ which assigns a unique guess of the signal sent by Alice given the measurement outcome λ of Eve.*

Definition 3.2. *Given an observation λ, if Eve's guess matches the signal sent by Alice, i.e., $S(\lambda) = A$, we call the event a **success**.*

Definition 3.3. *The **conditional success probability** of Eve is given by*

$$\Pr(\mathsf{success}|\mathcal{E} = \lambda) := \Pr\left[S(\lambda) = A|\mathcal{E} = \lambda\right]$$

*and the **success probability** of Eve is given by*

$$\Pr(\mathsf{success}) := \sum_\lambda \Pr\left[\mathcal{E} = \lambda\right].\Pr\left[S(\lambda) = A|\mathcal{E} = \lambda\right]$$

Definition 3.4. *Among all possible strategies, the one providing the maximum success probability is called*[2] *the **optimal strategy** S_{opt} and the corresponding success probability is called the **optimal success probability** $\Pr_{opt}(\mathsf{success})$.*

Eve's probability of success and failure to recognize Alice's states correctly are denoted as $F^{\mathcal{E}}_{xy}$, and $D^{\mathcal{E}}_{xy} := 1 - F^{\mathcal{E}}_{xy}$ respectively. Those depends actually on the overlap of her non-orthogonal states.

Eve's objective is to achieve maximum of these parameters: IG, MI, PS, whenever possible. We'll show a relation between the optimal values of these parameters.

[2] We use 'opt' and $\star$ interchangeably to demarcate optimal quantities.

3.3.1.6 Disturbance at Bob's end

Eve's unitary interaction creates entanglement, and thereby, introduces **disturbance** to the signal sent by Alice which is detectable by Bob. Considering the signal sent in *xy* basis, the disturbance introduced by Eve could be described by

$$D_{xy} \;:=\; \sum_{\lambda} q_{xy}(\lambda) d_{xy}(\lambda),$$

where, $d_{xy}(\lambda)$ is the average error for Bob to read the signal that was sent by Alice while Eve finds an outcome λ. For equal prior,

$$d_{xy}(\lambda) \;:=\; \frac{1}{2}\left(d_{\lambda x} + d_{\lambda y}\right),$$

where, $d_{\lambda x}$ is the error for Bob as Alice sends *x* and Eve detects λ (i.e., Bob reads *y*), i.e.,

$$d_{\lambda x} \;:=\; \Pr[\mathcal{B} = y | (\mathcal{A} = x, \mathcal{E} = \lambda)],$$

and $d_{\lambda y}$ is the error for Bob as Alice sends *y* and Eve detects λ (i.e., Bob reads *x*), i.e.,

$$d_{\lambda y} \;:=\; \Pr[\mathcal{B} = x | (\mathcal{A} = y, \mathcal{E} = \lambda)].$$

Clearly, D_{xy} is the observable error rate that Bob experiences in order to read the signal sent by Alice prepared in *xy* basis. $F_{xy} = 1 - D_{xy}$ is the fraction of correctly received states.

The presence of Eve introduces an error D_{xy} in the channel between the two legitimate parties. The channel works like a binary symmetric channel with bit-flip rate D_{xy}. On the other hand, Eve creates for herself another binary symmetric channel with bit-flip rate $D^{\mathcal{E}}_{xy}$.

3.3.2 Eve's max. information: Maximum IG and MI

For equal prior, Fuchs *et al.* [FGG$^+$97] deduced an upper bound on the information gain (*G*). It was then used to provide another upper bound on the mutual information (*I*). A necessary and sufficient condition to achieve the maximum values was also provided therein. We recollect these results here.

Proposition 3.1. (*An upper bound on information gain (G)*) [FGG$^+$97, Eqs. (23,24)]

$$G_{xy} \;\leq\; 2\sqrt{D_{uv}(1-D_{uv})}, \tag{3.3.2.1}$$

$$G_{uv} \;\leq\; 2\sqrt{D_{xy}\left(1-D_{xy}\right)}. \tag{3.3.2.2}$$

To be more individual, for a measurement outcome λ with Eve, the bound on information gain [FGG+97, Eq. (20)] is expressed by the following inequality

$$G_{xy}(\lambda) \leq 2\sqrt{d_{uv}(\lambda)\left[1-d_{uv}(\lambda)\right]}. \tag{3.3.2.3}$$

Interestingly, while Eve's information gain corresponds to signals sent in the *xy* basis, Bob's error rate corresponds to signals sent in the *uv* basis and vice versa.

Proposition 3.2. (*An Upper Bound on Mutual Information (I)*) [FGG+97, Eqs. (31,32)]

$$I_{xy} \leq \frac{1}{2}\,\phi\left[2\sqrt{D_{uv}(1-D_{uv})}\right], \tag{3.3.2.4}$$

$$I_{uv} \leq \frac{1}{2}\,\phi\left[2\sqrt{D_{xy}\left(1-D_{xy}\right)}\right], \tag{3.3.2.5}$$

where $\phi(z) = (1+z)\ln(1+z) + (1-z)\ln(1-z)$.

The basis-subscripts in the inequations emphasize that the mutual information and the error rate in the upper bound refer to signals sent in two different bases. We'll use *bpsp* (bits per send photon) to scale mutual information among two parties.

3.3.3 Certify optimality: a necessary and sufficient condition

Proposition 3.3. (*Necessary and Sufficient Conditions to Achieve* $G^{\star}$)[3] [FGG+97, Eqs. (38,39)]
The necessary and sufficient conditions for equality in Eq. (3.3.2.1) *are*

$$|V_{\lambda u}\rangle = \varepsilon_{\lambda}\sqrt{\frac{D_{uv}}{1-D_{uv}}}\;|U_{\lambda u}\rangle \tag{3.3.3.1}$$

and

$$|U_{\lambda v}\rangle = \varepsilon_{\lambda}\sqrt{\frac{D_{uv}}{1-D_{uv}}}\;|V_{\lambda v}\rangle, \tag{3.3.3.2}$$

where

$$\varepsilon_{\lambda} = \pm 1 = \mathsf{sgn}\left(Q_{x\lambda} - Q_{y\lambda}\right) \tag{3.3.3.3}$$

[3] $q^{\star}$ denotes optimal (maximum) value for any quantity q.

and

$$\begin{aligned}
|U_{\lambda u}\rangle &= B_u \otimes \sqrt{E_\lambda}\,|U\rangle, \qquad |V_{\lambda u}\rangle = B_u \otimes \sqrt{E_\lambda}\,|V\rangle,\\
|U_{\lambda v}\rangle &= B_v \otimes \sqrt{E_\lambda}\,|U\rangle, \qquad |V_{\lambda v}\rangle = B_v \otimes \sqrt{E_\lambda}\,|V\rangle,\\
B_u &= |u\rangle\langle u|,\;\; B_v = |v\rangle\langle v|, \quad \textit{with}\;\; B_u + B_v = \mathbb{1}. \qquad (3.3.3.4)
\end{aligned}$$

Similar conditions hold for a signal prepared in uv basis to attain the equality in Eq. (3.3.2.2).

It is quite worthy to note that the set of conditions that optimize the info-gain G also optimizes the mutual info I. Therefore, the necessary and sufficient conditions for equality in Eqs. (3.3.2.4, 3.3.2.5) becomes the same as those in Proposition 3.3. That is to say, for a signal sent in xy basis, an interaction-POVM tuple that attains the bound in Eq. (3.3.2.1) does the same in Eq. (3.3.2.4) and vice versa. Moreover, for the other basis, similar statement holds for Eqs. (3.3.2.2) and (3.3.2.5).

3.3.4 The postinteraction joint states

Eve's objective is to maximize the functions G or I, irrespective of what MUB was used by Alice for encoding. Both the bounds (3.3.2.4, 3.3.2.5) [and therefore the bounds (3.3.2.1, 3.3.2.2)] could be achieved simultaneously while fixing D_{xy}, D_{uv} independently. One of the conditions that must hold to achieve the bounds in xy basis is the following [FGG$^+$97, Eq. (33)]:

$$d_{\lambda u} = d_{\lambda v} = d_{uv}(\lambda) = D_{uv}, \;\; \forall \lambda.$$

An analogous condition holds good for signals sent in the uv basis.

Thus, for a signal sent in xy basis, the Schmidt decomposition of the postinteraction states are

$$\begin{aligned}
|X\rangle &= \sqrt{1-D_{xy}}\,|x\rangle|\xi_x\rangle + \sqrt{D_{xy}}\,|y\rangle|\zeta_x\rangle,\\
|Y\rangle &= \sqrt{1-D_{xy}}\,|y\rangle|\xi_y\rangle + \sqrt{D_{xy}}\,|x\rangle|\zeta_y\rangle. \qquad (3.3.4.1)
\end{aligned}$$

Assuming that all inner products $\langle \xi_i|\zeta_j\rangle$ are real, the restrictions (3.1, 3.2) on $|\xi_i\rangle, |\zeta_j\rangle$ becomes more restricted as

$$\{|\xi_x\rangle, |\xi_y\rangle\} \perp \{|\zeta_x\rangle, |\zeta_y\rangle\}. \qquad (3.3.4.2)$$

Similarly, for a signal sent in *uv* basis, the post-interaction states are

$$\begin{aligned} |U\rangle &= \sqrt{1-D_{uv}}\,|u\rangle|\xi_u\rangle + \sqrt{D_{uv}}\,|v\rangle|\zeta_u\rangle, \\ |V\rangle &= \sqrt{1-D_{uv}}\,|v\rangle|\xi_v\rangle + \sqrt{D_{uv}}\,|u\rangle|\zeta_v\rangle. \end{aligned} \tag{3.3.4.3}$$

From the orthogonality relation (3.3.4.2), one can conclude that Eve's probe lives in a Hilbert space having dimension at most four, and thereby is safe to consider 2 qubits (4 states). It is thus convenient to introduce the same bases (*xy* and *uv*, used by Alice) for each of Eve's qubits.

3.3.4.1 Interrelation between Eve's post-interaction states across the two bases

Since the bases $\mathfrak{B}_{xy}$ and $\mathfrak{B}_{uv}$ are conjugate to each other, one can expect a relationship between $|\xi_i\rangle, |\zeta_j\rangle$ in *uv* basis and those in *xy* basis which is described below.

$$\begin{aligned} 2\sqrt{1-D_{uv}}|\xi_u\rangle &= \sqrt{1-D_{xy}}(|\xi_x\rangle+|\xi_y\rangle)+\sqrt{D_{xy}}(|\zeta_x\rangle+|\zeta_y\rangle), \\ 2\sqrt{D_{uv}}|\zeta_u\rangle &= \sqrt{1-D_{xy}}(|\xi_x\rangle-|\xi_y\rangle)+\sqrt{D_{xy}}(|\zeta_y\rangle-|\zeta_x\rangle). \end{aligned} \tag{3.3.4.4}$$

Similarly,

$$\begin{aligned} 2\sqrt{1-D_{uv}}|\xi_v\rangle &= \sqrt{1-D_{xy}}(|\xi_x\rangle+|\xi_y\rangle)-\sqrt{D_{xy}}(|\zeta_x\rangle+|\zeta_y\rangle), \\ 2\sqrt{D_{uv}}|\zeta_v\rangle &= \sqrt{1-D_{xy}}(|\xi_x\rangle-|\xi_y\rangle)-\sqrt{D_{xy}}(|\zeta_y\rangle-|\zeta_x\rangle). \end{aligned} \tag{3.3.4.5}$$

3.3.5 Optimal interaction, optimal POVM

The main interest now will be focused to understand the nature of the interaction vectors ξ, ζ, the associated measurements, and the amount of information that Eve can achieve out of that.

An interaction is optimal if the IVs leads to maximum possible information with all possible POVMs. Needless to say, not all interactions are optimal. It is interesting to identify an (preferably all) *optimal interaction*(s). Unfortunately, in a general setting, this is a difficult problem to tackle with. Here we concentrate only on finding the *optimal interaction vectors*. But, finding the associated unitary is important for practical purposes – we'll address that issue in a separate chapter to find *optimal unitary evolutions*.

Given an interaction, one can extract varying amount of information out of various POVMs applied for measurement. One (or, some) of them lead to maximum amount of information out of all possible POVMs, and is called an *optimal POVM*. This is rather an easier problem to tackle with and is addressed below.

3.3.6 Optimal measurement (POVM): maximizes both IG and MI for a given interaction

Finding an optimal POVM for such IVs correspond to a rather easier optimization problem: maximize IG over all POVMs [Fuc96]. An upper bound exists and is achievable in each of the encoding bases. In xy basis, the maximum IG is attained by the orthonormal eigenprojectors $\{E_\lambda := |E_\lambda\rangle\langle E_\lambda|\}$ of the Hermitian $\frac{1}{2}\left(\rho_x - \rho_y\right)$. For equal prior (and not necessarily for unequal prior), the same measurement optimizes both IG and MI for an optimal interaction.

Relevant details to derive the optimal POVM and the Hermitian can be found in Sec. 3.4.2.

3.3.7 Optimal interaction: A specific choice

We need an interaction that achieves optimal information (i.e., attains $G^\star$ or $I^\star$). In [FGG$^+$97, Sec. III: Eqs. (50,51)], one such specific choice was provided for unequal error rates, which was shown to be a valid candidate (as it leads to optimality). Similarly, for equal error rates, another specific instance was introduced in [FGG$^+$97, Sec. IV, Eq. (69)]. Whether there are alternate candidates or not was remained open-ended issue.

3.3.7.1 For unequal error rates, i.e., $D_{xy} \neq D_{uv}$

Equations (50, 51) of [FGG$^+$97, Sec. III] are restated here. Consider a canonical basis for Eve's probe as $\{|\mathcal{E}_0\rangle, |\mathcal{E}_1\rangle, |\mathcal{E}_2\rangle, |\mathcal{E}_3\rangle\}$. Without loss of generality,

$$|\mathcal{E}_0\rangle = |x\rangle|x\rangle, |\mathcal{E}_1\rangle = |y\rangle|x\rangle, |\mathcal{E}_2\rangle = |x\rangle|y\rangle, |\mathcal{E}_3\rangle = |y\rangle|y\rangle. \tag{3.3.7.1}$$

The states $|\xi_i\rangle, |\zeta_j\rangle$ with Eve were described in the Bell Basis (w.r.t. Alice's encoding basis xy) which are defined as follows.

$$\begin{aligned} |\Phi^{\pm}_{xy}\rangle &:= \frac{1}{\sqrt{2}}\left(|x\rangle|x\rangle \pm |y\rangle|y\rangle\right) = \frac{1}{\sqrt{2}}\left(|\mathcal{E}_0\rangle \pm |\mathcal{E}_3\rangle\right), \\ |\Psi^{\pm}_{xy}\rangle &:= \frac{1}{\sqrt{2}}\left(|x\rangle|y\rangle \pm |y\rangle|x\rangle\right) = \frac{1}{\sqrt{2}}\left(|\mathcal{E}_2\rangle \pm |\mathcal{E}_1\rangle\right). \end{aligned} \tag{3.3.7.2}$$

In terms of the Bell-basis states to describe Eve's probe, the interaction was chosen as

$$\begin{aligned} |\xi_x\rangle &= \sqrt{1-D_{uv}}\,|\Phi^+_{xy}\rangle + \sqrt{D_{uv}}\,|\Phi^-_{xy}\rangle, \\ |\xi_y\rangle &= \sqrt{1-D_{uv}}\,|\Phi^+_{xy}\rangle - \sqrt{D_{uv}}\,|\Phi^-_{xy}\rangle, \\ |\zeta_x\rangle &= \sqrt{1-D_{uv}}\,|\Psi^+_{xy}\rangle - \sqrt{D_{uv}}\,|\Psi^-_{xy}\rangle, \\ |\zeta_y\rangle &= \sqrt{1-D_{uv}}\,|\Psi^+_{xy}\rangle + \sqrt{D_{uv}}\,|\Psi^-_{xy}\rangle. \end{aligned} \tag{3.3.7.3}$$

The corresponding optimal POVM [FGG$^+$97, Eqs. (55,56)] are the eigenprojectors

$$E_\lambda = |E_\lambda\rangle\langle E_\lambda|,$$

where

$$|E_0\rangle = |\mathcal{E}_0\rangle,\ |E_1\rangle = |\mathcal{E}_1\rangle,\ |E_2\rangle = |\mathcal{E}_2\rangle,\ |E_3\rangle = |\mathcal{E}_3\rangle. \tag{3.3.7.4}$$

The above analysis works for a signal chosen in *xy* basis. A similar analysis holds for the *uv* basis as well.

3.3.7.2 For equal error rates, i.e., $D_{xy} = D_{uv} = D$

For equal error rates, [FGG$^+$97, Sec. IV, Eq. (69)] comes up with another choice of $|\xi_i\rangle, |\zeta_j\rangle$. We describe it as below.

$$\begin{aligned} |\xi_x\rangle &= |x\rangle|x\rangle, \\ |\xi_y\rangle &= \big(\cos\alpha|x\rangle + \sin\alpha|y\rangle\big)\,|x\rangle \\ |\zeta_x\rangle &= |x\rangle|y\rangle, \\ |\zeta_y\rangle &= \big(\cos\beta|x\rangle + \sin\beta|y\rangle\big)\,|y\rangle. \end{aligned} \tag{3.3.7.5}$$

Optimality of G (or I) is reached when

$$\alpha = \beta \quad \text{and} \quad \sin\alpha = 2\sqrt{D(1-D)}.$$

The corresponding optimal POVM can be calculated by diagonalizing the observable $\rho_x - \rho_y$. We have mentioned it in our work as in the following chapter.

Although, both interactions (3.3.7.3, 3.4.8) lead to optimality, the way they were proposed in [FGG$^+$97] seems to be an intelligent guesswork. A derivation of the optimal interactions from the first principle is done in our work as described in the following chapter [Chap. 4].

3.3.8 Optimal strategy

Once the interaction and measurements are over, Eve's task remains to interpret her measurement outcomes to assign a guess on the signal sent by Alice. The optimal strategy of Eve can be described as follows.

As Alice declares her basis to be $\beta \in \{0,1\}$, Eve measures her ancilla in basis $\{|M_\lambda\rangle^\beta\}_{\lambda\in\{0,1,2,3\}}$ and interprets her measurement outcome in terms of a guess on Alice's bit. For +ve outcome, which occurs for $\lambda = 0,2$, she bets on 0, whereas, for −ve outcome, which occurs for $\lambda = 1,3$, she bets on 1.

To mount an optimal attack, Eve performs a suitable interaction unitarily, measures accordingly after basis reconciliation, and finally guesses the signal applying her strategy. Her KG is maximum *iff* her IVs are optimal and her measurement is also optimal.

Fig. 3.2 provides a schematic view of the attack model.

3.3.9 Optimal Eavesdropping in a nutshell

Alice encodes each *cbit* $a \in \{0,1\}$ by a randomly chosen basis $\beta \in \{0,1\}$ into a qubit in state $|a^{\beta}\rangle$. Eve attaches a two-qubit probe having state $|e\rangle$ with each of Alice's qubit qubit. She evolves the joint system unitarily ($\mathcal{U}$) from the pre-interaction joint state $|a^{\beta}\rangle|e\rangle$ to the post-interaction joint state $|S_a^{\beta}\rangle = \mathcal{U}\, |a^{\beta}\rangle|e\rangle$.

After basis reconciliation, whenever the legitimate parties agree on a basis $\beta \in \{0,1\}$, Eve measures her ancilla with a suitable POVM $\{M_{\lambda}\}^{\beta}_{\lambda \in \{0,1,2,3\}}$. Eve interprets the outcome following a *strategy* which is a rule to assign a guess for the state of the signal sent by Alice. For +ve outcome, which occurs for $\lambda = 0,2$, she bets on 0, whereas, for $-$ve outcome, which occurs for $\lambda = 1,3$, she bets on 1. Her guess a_{λ} is thus the following function

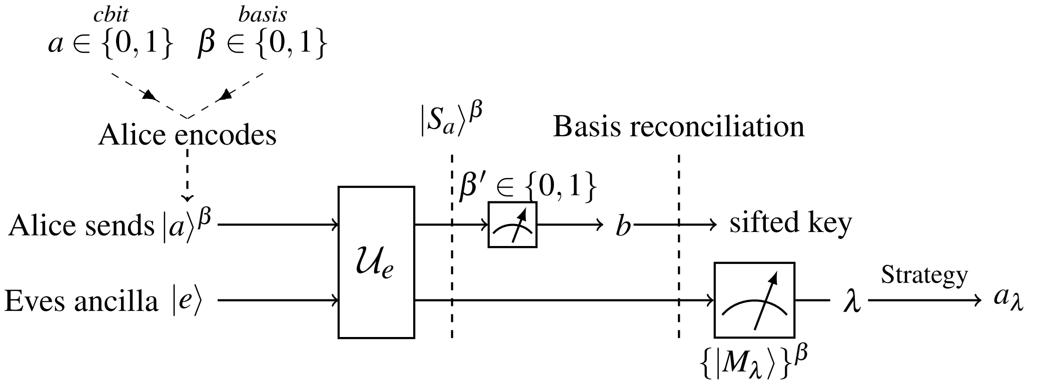
$$a_{\lambda} \;=\; \begin{cases} 0, & \text{if } \lambda = 0,2 \\ 1, & \text{if } \lambda = 1,3. \end{cases}$$

3.3.9.1 A schematic view of the eavesdropping model

Following is a circuit diagram illustrating the eavesdropping model in a nutshell.

Figure 3.2 | A circuit diagram for an optimal eavesdropping on BB84 protocol.

Alice uses one of the two MUBs, β, to encode a *cbit* 'a' into a *qubit* $|a\rangle^{\beta}$. Eve attaches an ancilla $|e\rangle$ and evolves the joint system unitarily ($\mathcal{U}_e$) that creates an entangled state $|S_a\rangle^{\beta}$. Bob measures the received qubit in basis β' to get the *cbit* b, and keeps it if the bases are matched. After basis reconciliation, Eve measures her ancilla in the POVM basis $\{|M_\lambda\rangle^{\beta}\}$. She interprets her outcome λ by a strategy and bet for $a_\lambda \in \{0,1\}$ to guess Alice's *cbit*. When Eves choices for the unitary and the measurement are optimal, she guesses the key best while not forcing to abort the protocol.



3.3.10 Practical eavesdropping

A practical eavesdropping should ideally leave the error rate symmetric across the two basses, i.e., $D_{xy} = D_{uv} = D$. Otherwise, the legitimate parties can detect the difference during the error-estimation phase, and thereby detect the presence of a malevolent party. For a QBER $= D$, the maximum amount of the IG in both the bases reaches $2\sqrt{D(1-D)}$, and is achievable [FGG+97].

Due to symmetric eavesdropping, the quantum channel between Alice-Bob and that between Alice-Eve can be interpreted as a binary symmetric channel with data-flipping rate D and $D_E = \frac{1}{2} - \sqrt{D(1-D)}$, respectively. Thus, at error-rate D, the respective bipartite mutual informations become

$$\begin{aligned} MI_{\mathtt{AB}} &= 1 - H(D) = \frac{1}{2}\,\phi\,(1-2D)\,, \\ MI_{\mathtt{AE}} &= 1 - H(D_E) = \frac{1}{2}\,\phi\left(2\sqrt{D(1-D)}\right), \end{aligned}$$

when expressed in *bits per sifted-photon* (*bpsp*).

Following the optimal strategy, Eve can glean $(1 - H(D_E))$ bits per sifted-photon of the transmission with fidelity $1 - D_E$ in lieu of introducing an error-rate D at Bob's end. The distinguishing advantage for an optimal attack is $\sqrt{D(1-D)}$.

3.3.10.1 The secure zone and key-rate

These are the two most relevant parameters for any attack. The amount of QBER characterizes the severity of the attack.

Definition 3.5 (Secure zone)**.** *The window of disturbance beyond which no (one-way) classical post-processing can distill a secret key for the legitimate parties.*

For the BB84 protocol, it is $D \in [0, D^\star = 0.1464)$ for the attack model considered in [FGG+97].

Definition 3.6 (Sifted key-rate)**.** *It is the ratio of the length of the sifted key and that of the raw key.*

It decreases as disturbance increases through the secure zone, as plotted in Fig. 3.3. From disturbance beyond 14.64%, it becomes zero, leaving no scope to retrieve a secret key with OW-CPP.

The *sifted key-rate K_{sif}* is bounded below by the difference $MI_{\mathtt{AB}} - MI_{\mathtt{AE}}$. An optimal attack achieves the maximum: for a QBER D, it amounts to $K_{sif}(D) = H(D_E) - H(D)$ *bpsp*. It decreases with growing QBER, and vanishes when the two MIs coincide which happens at the threshold [Fig. 3.3]

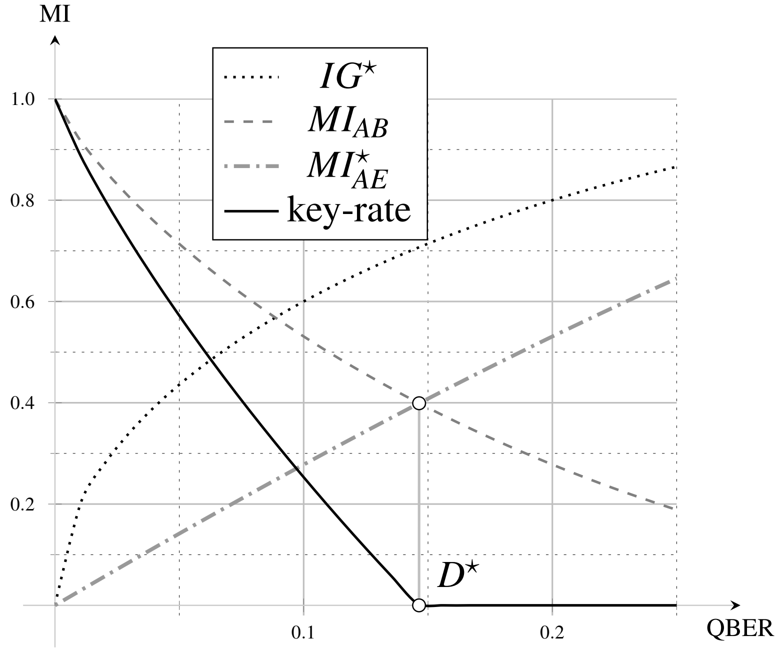
$$D^\star = \frac{1}{2}\left(1-\frac{1}{\sqrt{2}}\right) \approx 0.1464. \qquad (3.3.10.1)$$

Beyond this tolerable rate, an OW-CPP may not guarantee to filtrate a secure key. Within the *secure zone* $D \in [0, D^\star)$, key-filtration is guaranteed because Bob possess more information on Alice's bit than Eve does.

3.3.10.2 Sifted Key-rate, IG, MI(AB,AE)

Figure 3.3 | Sifted Key-rate for one-way classical post-processing.

Plotted: optimal Information Gain, bipartite Mutual Informations, and the secret key-rate. The graph of $MI_{\mathtt{AE}}$ reveals the information-disturbance trade-off. For QBER $D^\star = 0.1464$, $MI_{\mathtt{AB}}$ and $MI_{\mathtt{AE}}$ coincides, and the key-rate drops to zero. Below this error rate, an OW-CPP is faithful.



3.3.10.3 Connecting Bell violation and cloning

There is a very deep connection between a prepare-and-measure (*p&m*) scheme and its entanglement-based (*eb*) counterpart as well with cloning mechanisms to glean information. Particularly, for an optimal attack, the connections between *p&m* scheme, its *eb* counterpart, and optimal cloning mechanisms are quite clear. However, more deeper analysis may actually reveal more intricate connections in general.

In the *eb* protocol, the legitimate parties observe a Bell violation so far the estimated QBER remains in the secure zone of the *p&m* scheme. An optimal attack with QBER D reduces the CHSH correlation co-efficient to $\eta_D 2\sqrt{2}$ for $\eta_D := 1-2D$. An optimal attack also leaves Bob with the Bloch vectors contracted by a factor of η_D.

An optimal attack on the *p&m* scheme can also be achieved via an optimal phase-covariant cloner [BCMDM00][4]. The cloner is asymmetric since it creates two clones of the senders state: a degraded copy for her own with fidelity $(\frac{1}{2}+\sqrt{D(1-D)})$, and a superior copy for Bob with fidelity $1-D$. At the threshold QBER, both the fidelity for Bob and Eve reaches the maximum of $1-D^\star = \frac{1}{2}\left(1+\frac{1}{\sqrt{2}}\right)$ i.e., 85.36%, both in cloning and in *p&m* scheme. Moreover, for the choice of the permuted measurement basis $\{|E_0\rangle, |E_1\rangle, |E_2\rangle, |E_3\rangle\} = \{|00\rangle, |11\rangle, |01\rangle, |10\rangle\}$, the optimal PIJSs are in sync with the outputs of an optimal pc-cloner [BCMDM00, Eq. (36)].

3.3.11 Practical eavesdropping related details

Here is some detail on a few aspects discussed in the earlier section.

3.3.11.1 Success probability of Eve's state discrimination

An optimal attack on the *p&m* scheme leaves Eve with an optimal state-discriminate problem. For a specific encoding basis, the four different post-interaction states of Eve's ancilla can be grouped into two mutually orthogonal sets: one with the two fidelity states, and the other with the two disturbed states. Since Eve can discriminate these orthogonal sets (whether disturbed or not), all she is left with is to distinguish the two states in a set, *e.g.*, distinguishing $|\xi_a\rangle$ from $|\xi_{\bar{a}}\rangle$, or, distinguishing $|\zeta_a\rangle$ from $|\zeta_{\bar{a}}\rangle$. Following the optimal strategy, Eve can distinguish the two such parity states (fidelity or disturbed) with probability [Hel69]

$$\begin{aligned} F_{\mathrm{E}}^{\beta} &= \frac{1}{2}+\frac{1}{2}\sqrt{1-|\langle\xi_a^{\beta}|\xi_{\bar{a}}^{\beta}\rangle|^2} \\ &= \frac{1}{2}+\frac{1}{2}\sqrt{1-(1-2D_{\bar{\beta}})^2} \\ &= \frac{1}{2}+\sqrt{D_{\bar{\beta}}(1-D_{\bar{\beta}})}. \end{aligned}$$

Now, some results used in practical eavesdropping are elaborated here for better understanding.

[4] A pc-cloner can copy all the states $|\psi\rangle = |0\rangle + e^{i\phi}|1\rangle$ with equal fidelity $\langle\psi|\rho_\psi|\psi\rangle$ for output clone ρ_ψ.

3.3.11.2 Secret-key rate

The *secrecy capacity* C_s of the quantum channel between Alice and Bob is defined [CK78] as the optimum rate at which Alice can reliably send information to Bob leaving Eve's information on that data arbitrarily small. A necessary and sufficient condition for a positive secret-key rate is not known, but a lower bound is known [CK78]. For a more general scenario, considering the knowledge gain of Eve over Bob's data ($I_{\mathtt{BE}}$) due to public discussion over the supplementary classical channel, one can lower bound the secrecy capacity [EHPP94] by the following formula

$$C_s \quad \geq \quad \max\{I_{\mathtt{AB}} - I_{\mathtt{AE}},\, I_{\mathtt{AB}} - I_{\mathtt{EB}}\}.$$

Thus the legitimate parties should consider the channel unsafe and abort the transmission whenever

$$I_{\mathtt{AB}} \quad \leq \quad \min\{I_{\mathtt{AE}}, I_{\mathtt{EB}}\}.$$

On the other hand, the legitimate parties can establish a secret key following some one-way CPP, *iff* $I_{\mathtt{AB}} > I_{\mathtt{AE}}$ or $I_{\mathtt{AB}} > I_{\mathtt{EB}}$. For an optimal symmetric attack, $I_{\mathtt{AE}} = I_{\mathtt{EB}}$. Therefore, Alice and Bob lives in the *secure zone* whenever $I_{\mathtt{AB}} > I_{\mathtt{AE}}$. The difference $I_{\mathtt{AB}} - I_{\mathtt{AE}}$, that captures the *secret-key rate*, remains same during the error correction and privacy amplification. Thus, the condition transcends in order to establish a shared secret between the two legitimate parties.

3.3.11.3 An optimal attack contracts the Bloch vectors

The state $|a\rangle^{\beta}$ of a two-level quantum system (qubit) corresponds to a Bloch vector $\vec{a}_{\beta}$ on the surface of the Poincaré sphere. Alices' density operator $\rho_{\mathtt{A}} = |a\rangle^{\beta}\langle a|$ is a convex combination $\frac{1}{2}\left(\mathbb{1} + \vec{a}_{\beta}\cdot\vec{\sigma}\right)$ of the Pauli operators. For the BB84 protocol, the states in the Z and the X bases correspond to the Bloch vectors $(0,0,\pm 1)$ and $(\pm 1, 0, 0)$, respectively. Therefore, Alice sends the density operators $\frac{1}{2}(\mathbb{1} \pm \sigma_s)$ (for, $s \in \{z,x\}$) to Bob. But, due to eavesdropping, Bob receives the density

$$\begin{aligned}
\rho_{\mathtt{B}} &= F|a\rangle^{\beta}\langle a| + D|\bar{a}\rangle^{\beta}\langle\bar{a}| \\
&= F\cdot\frac{1}{2}\left(\mathbb{1}_2 + \vec{a}_{\beta}\cdot\vec{\sigma}\right) + D\cdot\frac{1}{2}\left(\mathbb{1}_2 - \vec{a}_{\beta}\cdot\vec{\sigma}\right) \\
&= \frac{1}{2}\left(\mathbb{1}_2 + (F-D)\vec{a}_{\beta}\cdot\vec{\sigma}\right)
\end{aligned}$$

While Alice sends the density $\frac{1}{2}\left(\mathbb{1} + \vec{a}\cdot\vec{\sigma}\right)$, Bob receives $\frac{1}{2}\left(\mathbb{1} + \eta_D\vec{a}\cdot\vec{\sigma}\right)$ with $\eta_D = 1 - 2D$. To be specific, the density operators $\frac{1}{2}(\mathbb{1} \pm \sigma_s)$ (for, $s \in \{z,x\}$) from Alice get perturbed to $\frac{1}{2}(\mathbb{1} \pm \eta_D\sigma_s)$ when it reaches Bob. Thus, eavesdropping shrinks the Bloch vectors by a factor of $\eta_D = 1 - 2D$.

3.3.11.4 Optimal state-discrimination vs Bell-violation

An optimal state-discrimination based attack on a *p&m* scheme has some intriguing connection with Bell-violation in an equivalent *eb* scheme and is discussed here.

The *p&m* scheme has its equivalent *eb* counterpart where Alice prepares a maximally entangled state $\frac{|aa\rangle + |\bar{a}\bar{a}\rangle}{\sqrt{2}}$ and send one of the particles to Bob. Both the parties measure the observables σ_z, σ_x, chosen randomly.

The security of the *eb* scheme is linked to the tests of quantum nonlocality [FGG$^+$97]. Presence of non-locality is a certificate for OW-CPP. The degree of non-locality depends on the estimated value of the CHSH polynomial for which the legitimate parties sacrifice a subset of their particles. Alice measures one of the observables σ_z, σ_x chosen randomly, while Bob measures one of the observables $\frac{\sigma_z + \sigma_x}{\sqrt{2}}, \frac{\sigma_z - \sigma_x}{\sqrt{2}}$ chosen randomly. The binary measurement outcomes $a_i, b_j \in \{-1, +1\}$ are used to estimate the CHSH correlation-coefficient which in turn is the expected value of the product of the outcomes.

$$S \;:=\; E(a_1, b_1) + E(a_1, b_2) + E(a_2, b_1) - E(a_2, b_2).$$

Due to some channel error D, each of the correlations $E(a_i, b_j|D)$ get reduced from its error-free counterpart $E(a_i, b_j)$ by a factor of $1 - 2D$:

$$\begin{aligned} E(a_i, b_j|D) &= F \cdot E(a_i, b_j) - D \cdot E(a_i, b_j) \\ &= (1 - 2D) \cdot E(a_i, b_j). \end{aligned}$$

Consequently, $S_D = (1 - 2D)S_0$.

The CHSH inequality forbids the correlation coefficient S to exceed 2 for *local operations and classical communication* (LOCC). However, for an error-free quantum channel, this inequality is violated and the correlation amount reaches the maximum of $2\sqrt{2}$. Then, in a quantum channel with error D, the maximum amount of violation becomes $S_D^\star = (1 - 2D)2\sqrt{2}$. In order to maintain quantum non-locality, this reduced sum must exceed 2, which happens precisely for $D < D^\star$ as in Eq. (3.3.10.1).

3.4 Illustrated derivations on Optimal eavesdropping

3.4.1 The optimization problem

To learn a transmitted signal Eve performs two tasks – an interaction and a measurement. Her objective is to maximize the information (IG, MI) on the the signal while not surpassing a threshold error. For a signal prepared in xy basis, each of G_{xy} or I_{xy} is a function of five parameters $p_x, p_y, \rho_x, \rho_y, E_\lambda$. The prior probabilities p_x, p_y should be fixed for a given communication. Then, the optimization should be performed over remaining three parameters $\rho_x, \rho_y, E_\lambda$.

A simpler version to get started with could be the following: given an interaction (i.e., the parameters ρ_x, ρ_y get fixed), maximize information gain G_{xy} over all measurements E_λ. Under this restriction, it was shown in [Fuc96] that information gain is bounded above, while the optimal bound $G^\star_{xy}(\rho_x, \rho_y)$ has an analytical expression that could be achieved by a POVM E_λ consisting of the eigenprojectors onto the orthonormal eigenbasis of a fairly simple Hermitian operator. The derivation is shown in Subsec. 3.4.2.

Unfortunately, under the same restriction on the parameter set, unlike optimal information gain, optimal mutual information $I^\star_{xy}(\rho_x, \rho_y)$ over all measurements doesn't have any analytical expression [Fuc96]. For BB84 protocol, due to equal prior, a betterment was done in [FGG$^+$97] while optimizing over the entire parameter set $\rho_x, \rho_y, E_\lambda$. We discuss it in the following section.

They derived an upper bound for G_{xy} and I_{xy} over the parameter set $\rho_x, \rho_y, E_\lambda$. An upper bound for the information gain was achieved as

$$G^\star_{xy} = \max_{\{(\rho_x, \rho_y, E_\lambda)\}} G_{xy}(\rho_x, \rho_y, E_\lambda) \;=\; 2\sqrt{D_{uv}(1-D_{uv})}.$$

The same set of conditions also optimize the mutual information.

$$\begin{aligned} I^\star_{xy} &= \max_{\{(\rho_x, \rho_y, E_\lambda)\}} I_{xy}(\rho_x, \rho_y, E_\lambda) \\ &= \frac{1}{2}\,\phi\left[2\sqrt{D_{uv}(1-D_{uv})}\right] \;=\; \frac{1}{2}\,\phi[G^\star_{xy}]. \end{aligned}$$

Then combining all the conditions that optimize G_{xy} (or I_{xy}), a necessary and sufficient condition to achieve the bounds was derived there.

3.4.2 Optimal measurement (POVM) to maximize information gain (G) for a given interaction

Let's consider the problem below: given an interaction,

$$\text{maximize } G_{xy} \;=\; \sum_{\lambda} \left| P_{\lambda x} p_x - P_{\lambda y} p_y \right|$$

over all POVMs $\{E_\lambda\}$.

In [Fuc96], an optimal observable (describing the optimal measurement) for this maximization was derived. The maximization was done on *Kolmogorov Variational Distance* [Fuc96, Eq. (130)]. The calculation can be found in [Fuc96, Appendix (Sec. 7)]. It shows that an optimal measurement corresponds to a Hermitian operator (observable) given by [Fuc96, Eq. (21)] and the optimal POVM consists of the orthonormal eigenprojectors of that operator. We describe the result in our own way with a proof in terms of maximizing G.

Lemma 3.3. *Given an interaction, an optimal POVM that achieves the maximum information gain consists of the eigenprojectors $\{E_\lambda\}$ onto the orthonormal eigenbasis $\{|E_\lambda\rangle\}$ that diagonalizes the Hermitian operator*

$$\widetilde{\Gamma}_{xy} := p_x \rho_x - p_y \rho_y, \tag{3.4.2.1}$$

where ρ_x, as defined in Eq. (3.4.6.1)*, is the density with Eve, and is the partial trace (over Alice's qubit) of the post-interaction state $|X\rangle$. The maximum achievable information gain is* $\mathsf{tr}\left|\widetilde{\Gamma}_{xy}\right|$.

Proof. Given an interaction (i.e., the density operators ρ_x, ρ_y get fixed), the associated $\widetilde{\Gamma}_{xy}$ being Hermitian is diagonalizable by an orthonormal eigenbasis $\{|\gamma_i\rangle\}$. Let the corresponding eigenvalues (all real) are $\{\gamma_i\}$. Then, over all POVMs $\{E_\lambda\}$,

$$\begin{aligned}
G_{xy} &= \sum_\lambda \left| P_{\lambda x} p_x - P_{\lambda y} p_y \right| \\
&= \sum_\lambda \left| p_x \mathrm{tr}\left(\rho_x E_\lambda\right) - p_y \mathrm{tr}\left(\rho_y E_\lambda\right) \right|, \text{ using Eqs. (3.6, 3.7)} \\
&= \sum_\lambda \left| \mathrm{tr}\left(\widetilde{\Gamma}_{xy} E_\lambda\right) \right|, \text{ using Eq. (3.4.2.1)} \\
&= \sum_\lambda \left| \sum_i \gamma_i \, \langle \gamma_i | E_\lambda | \gamma_i \rangle \right|, \text{with the observable } \widetilde{\Gamma}_{xy} = \sum \gamma_i | \gamma_i \rangle \\
&\leq \sum_\lambda \sum_i |\gamma_i| \, \langle \gamma_i | E_\lambda | \gamma_i \rangle \\
&= \sum_i |\gamma_i| \, \langle \gamma_i | \sum_\lambda E_\lambda | \gamma_i \rangle \\
&= \sum_i |\gamma_i| = \mathrm{tr} \left| \widetilde{\Gamma}_{xy} \right|.
\end{aligned}$$

The upper bound is achievable by some POVM $\{E_\lambda\}$ consisting of the projectors onto an orthonormal eigenbasis of the observable $\widetilde{\Gamma}_{xy}$. □

Remark 3.1. *For BB84, due to equal prior, the Hermitian of* (3.4.2.1) *becomes*

$$\Gamma_{xy} := \frac{1}{2}\left(\rho_x - \rho_y\right) \tag{3.4.2.2}$$

and use it for our purposes in the contributory chapters.

Remark 3.2. *Given an interaction, a POVM that is optimal for G_{xy} may not necessarily be optimal for I_{xy} [FGG$^+$97, Fuc96]. However, for equal prior probabilities, once the bound* $\mathrm{tr}\left|\Gamma_{xy}\right|$ *of G_{xy} in Lemma 3.3 becomes equal to the upper bound $\mathscr{D}^2_{uv} - \overline{\mathscr{D}}^2_{uv}$ of G_{xy} in Eq.* (3.3.2.1)*, the interaction is called optimal. In such situations, the interaction-POVM pair also attains the upper bound* (3.3.2.4) *of I_{xy}.*

3.4.3 Maximizing Information Gain G_{xy}

Here we provide a sketch of proving the upper bound on IG in xy basis, i.e., we prove Eqs. (3.3.2.3, 3.3.2.1) saying $G_{xy}(\lambda) \leq \sqrt{d_{uv}(\lambda)\left(1-d_{uv}(\lambda)\right)}$, and $G_{xy} \leq \sqrt{D_{uv}(1-D_{uv})}$. The details in between the lines are deferred to Sec. 3.4.10.1. The basis information is dropped for simplicity.

Proof. For a given POVM $\{E_\lambda\}$,

$$\begin{aligned}
q(\lambda)G(\lambda) \quad &= \quad \frac{1}{2}\left\{\left|\langle X|\mathbb{1}\otimes E_\lambda|X\rangle - \langle Y|\mathbb{1}\otimes E_\lambda|Y\rangle\right|\right\} \\
\texttt{Use:} \quad & \qquad \sqrt{\mathbf{2}}|\mathbf{X}\rangle{=}|\mathbf{U}\rangle{+}|\mathbf{V}\rangle,\ \ \sqrt{\mathbf{2}}|\mathbf{Y}\rangle{=}|\mathbf{U}\rangle{-}|\mathbf{V}\rangle \qquad [\ \texttt{Calc in Sec } \mathbf{3.4.10.1}\ ] \\
&= \quad \frac{1}{2}\left\{\left|\langle U|\mathbb{1}\otimes E_\lambda|V\rangle + \langle V|\mathbb{1}\otimes E_\lambda|U\rangle\right|\right\} \\
\texttt{Use:} \quad & \qquad \mathbb{1}{=}\mathbf{B_u}{+}\mathbf{B_v}\text{:}\ \ \mathbf{B_u}{=}|\mathbf{u}\rangle\langle\mathbf{u}|,\ \ \mathbf{B_v}{=}|\mathbf{v}\rangle\langle\mathbf{v}|;\quad \mathbf{B_u^2}{=}\mathbf{B_u} \qquad [\ \texttt{Calc in Sec } \mathbf{3.4.10.1}\ ] \\
&= \quad \left|\mathbf{Re}\langle U|B_u\otimes E_\lambda|V\rangle + \mathbf{Re}\langle V|B_v\otimes E_\lambda|U\rangle\right| \\
\texttt{Use:} \quad & \qquad \begin{matrix} U_{\lambda u} = B_u\otimes\sqrt{E_\lambda}\ \ |U\rangle & V_{\lambda u} = B_u\otimes\sqrt{E_\lambda}\ \ |V\rangle \\ U_{\lambda v} = B_v\otimes\sqrt{E_\lambda}\ \ |U\rangle & V_{\lambda v} = B_v\otimes\sqrt{E_\lambda}\ \ |V\rangle \end{matrix} \quad [\ \texttt{Calc in Sec } \mathbf{3.4.10.1}\ ] \\
[\text{ Ineq. 1 }] \quad &\leq \quad \left|\langle U_{\lambda u}|V_{\lambda u}\rangle\right| + \left|\langle U_{\lambda v}|V_{\lambda v}\rangle\right| \\
\texttt{Use:} \quad & \qquad \texttt{Schwartz ineq:}\ \ ||\,\vec{\mathbf{a}}\cdot\vec{\mathbf{c}}\,|| \leq ||\,\vec{\mathbf{a}}\,||\,||\,\vec{\mathbf{c}}\,|| \\
[\text{ Ineq. 2 }] \quad &\leq \quad \sqrt{\langle U_{\lambda u}|U_{\lambda u}\rangle\langle V_{\lambda u}|V_{\lambda u}\rangle} + \sqrt{\langle U_{\lambda v}|U_{\lambda v}\rangle\langle V_{\lambda v}|V_{\lambda v}\rangle} \\
\texttt{Use:} \quad & \qquad \langle\mathbf{V}_{\lambda\mathbf{u}}|\mathbf{V}_{\lambda\mathbf{u}}\rangle{=}\langle\mathbf{V}|\mathcal{B}_\mathbf{u}{\otimes}\mathbf{E}_\lambda|\mathbf{V}\rangle{=}\Pr(\mathcal{A}{=}\mathbf{v},\mathcal{E}{=}\lambda,\mathcal{B}{=}\mathbf{u}){=}\mathbf{P}_{\lambda\mathbf{v}}\ \ \mathbf{d}_{\lambda\mathbf{v}} \\
&= \quad \underbrace{\sqrt{P_{\lambda u}P_{\lambda v}}}\left(\sqrt{(1-d_{\lambda u})d_{\lambda v}} + \sqrt{d_{\lambda u}(1-d_{\lambda v})}\right) \\
\texttt{Use:} \quad & \qquad \texttt{GM } \leq\texttt{AM} \\
[\text{ Ineq. 3 }] \quad &\leq \quad \underbrace{1/2\cdot\left(P_{\lambda u}+P_{\lambda v}\right)}_{=\mathbf{q}_\lambda}\left(\sqrt{(1-d_{\lambda u})d_{\lambda v}} + \sqrt{d_{\lambda u}(1-d_{\lambda v})}\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
G_\lambda \quad &\leq \quad \sqrt{d_{\lambda u}(1-d_{\lambda v})} + \sqrt{(1-d_{\lambda u})d_{\lambda v}} \\
\texttt{Use:} \quad & \qquad \mathbf{d}_{\lambda\mathbf{u}}{=}\mathbf{d}_\lambda{+}\mathbf{w},\ \ \mathbf{d}_{\lambda\mathbf{v}}{=}\mathbf{d}_\lambda{-}\mathbf{w}\ \ :\ \ \mathbf{d}_\lambda{=}(\mathbf{d}_{\lambda\mathbf{u}}{+}\mathbf{d}_{\lambda\mathbf{v}})/\mathbf{2} \\
&= \quad \underbrace{\sqrt{\left(d_\lambda+w\right)\left(1-d_\lambda+w\right)} + \sqrt{\left(d_\lambda-w\right)\left(1-d_\lambda-w\right)}}_{\texttt{Even fn, Max at } \mathbf{w{=}0} \qquad [\ \texttt{Calc in Sec } \mathbf{3.4.10.1}\ ]} \\
[\text{ Ineq. 4 }] \quad &\leq \quad 2\sqrt{d_\lambda\left(1{-}d_\lambda\right)}
\end{aligned}$$

So, the bound on G_{xy} is as follows.

$$\begin{aligned}
G_{xy} \quad &= \quad \sum_\lambda q_\lambda G_\lambda \ \leq\ 2\sum q_\lambda \underbrace{\sqrt{d_\lambda\left(1{-}d_\lambda\right)}} \\
[\text{ Ineq. 5 }] \quad &\leq \quad 2\sqrt{\sum q_\lambda d_\lambda\left(1-\sum q_\lambda d_\lambda\right)} = 2\sqrt{D\left(1-D\right)}
\end{aligned}$$

□

3.4.4 Maximizing Mutual Information I_{xy}

The bound on G_{xy} is used to bound I_{xy}. A proof sketch for Eq. (3.3.2.4) is as following.

Proof.

$$\begin{aligned}
I_{xy} &= \ln 2 + \sum_{\lambda} q_{\lambda} \sum_{i} Q_{i\lambda} \ln Q_{i\lambda} \\
&= \ln 2 + \sum_{\lambda} q_{\lambda} \left(Q_{x\lambda} \ln Q_{x\lambda} + Q_{y\lambda} \ln Q_{y\lambda} \right) \\
\texttt{Use:} \quad & \qquad Q_{x\lambda} = \tfrac{1}{2}\left(1 \pm G_{\lambda}\right), \qquad Q_{y\lambda} = \tfrac{1}{2}\left(1 \mp G_{\lambda}\right) \\
& \texttt{since,} \quad Q_{x\lambda} + Q_{y\lambda} = 1, \qquad Q_{x\lambda} - Q_{y\lambda} = \pm G_{\lambda}. \\
&= \frac{1}{2} \sum q_{\lambda} \phi\left(G_{\lambda}\right) \\
\texttt{Use:} \quad & \qquad \phi(z) \texttt{ is m.i.} \quad \Longrightarrow \quad \phi(x) \leq \phi(x+) \\
& \therefore \quad G_{\lambda} \leq 2\sqrt{d_{\lambda}\left(1-d_{\lambda}\right)} \quad \Longrightarrow \quad \phi(G_{\lambda}) \leq \phi\left[2\sqrt{d_{\lambda}\left(1-d_{\lambda}\right)}\right]. \\
&\leq \frac{1}{2} \sum_{\lambda} q_{\lambda} \underbrace{\phi\left[2\sqrt{d_{\lambda}\left(1-d_{\lambda}\right)}\right]} \\
&\leq \frac{1}{2} \phi\left[2\sqrt{\sum q_{\lambda} d_{\lambda}\left(1 - \sum q_{\lambda} d_{\lambda}\right)}\right] \\
&= \frac{1}{2} \phi\left[2\sqrt{D(1-D)}\right]
\end{aligned}$$

□

Relation between I^{opt} and G^{opt} is read as

$$I^{opt}_{\beta} = \frac{1}{2} \phi(G^{opt}_{\beta}), \quad \beta \in \{xy, uv\}.$$

3.4.5 Necessary and Sufficient Conditions to achieve $G_{opt}, I_{opt}, \Pr_{opt}(success)$

Here we discuss the conditions under which the maximum of IG, MI are achieved.

3.4.5.1 Necessary and Sufficient Conditions to Achieve $G^{\star}_{xy}$

Given a set of PIJSs in xy basis, are they optimal, i.e., do they achieve $G^{\star}_{xy}$ as in Ineq. (3.3.2.1)? The NSC (3.3.3.1, 3.3.3.2) is nothing but the compact form of the constraints under which the equality is attained in each of the five inequalities while deriving the bound on IG.

Proof. For five inequalities in the derivation of the upper bound on IG, the equality holds when the following happens.

1. Equality in [Ineq. 5] : all d_{λ}'s are equal.

2. Equality in [Ineq. 4] : $d_{\lambda u} = d_{\lambda v} = d_{\lambda} = D_{uv}$

3. Equality in [Ineq. 3] : $P_{\lambda u} = P_{\lambda v} = q_\lambda$

4. Equality in [Ineq. 2] : $\langle U_{\lambda u}|V_{\lambda u}\rangle, \langle U_{\lambda v}|V_{\lambda v}\rangle$ are real and same sign. Let their sign be σ_λ.

5. Equality in [Ineq. 1] : $|U_{\lambda u}\rangle \parallel |V_{\lambda u}\rangle$, $|U_{\lambda v}\rangle \parallel |V_{\lambda v}\rangle$, i.e.,

$$|V_{\lambda u}\rangle = \pm\, \mu\, |U_{\lambda u}\rangle = \varepsilon_\lambda\, \mu\, |U_{\lambda u}\rangle, \tag{3.4.5.1}$$

$$|V_{\lambda v}\rangle = \pm\, \nu\, |U_{\lambda v}\rangle = \varepsilon'_\lambda\, \nu\, |U_{\lambda v}\rangle. \tag{3.4.5.2}$$

These are the required NSCs while the following claims are also considered.

Claim 3.4.5.1. *For the following claim, proof is given in Sec. 3.4.10.3.*

$$\sigma_\lambda = \mathsf{sgn}\left(Q_{x\lambda} - Q_{y\lambda}\right), \tag{3.4.5.3}$$

$$\mu = \sqrt{\frac{D_{uv}}{1-D_{uv}}} = \nu, \tag{3.4.5.4}$$

$$\varepsilon_\lambda = \sigma_\lambda = \varepsilon'_\lambda. \tag{3.4.5.5}$$

Thus, using relations (3.4.5.4, 3.4.5.5, 3.4.5.3) in Eqs. (3.4.5.1, 3.4.5.2), we get:

$$\begin{aligned} |V_{\lambda u}\rangle &= \varepsilon_\lambda\, \mu\, |U_{\lambda u}\rangle, \\ |U_{\lambda v}\rangle &= \varepsilon_\lambda\, \mu\, |V_{\lambda v}\rangle. \end{aligned}$$

Hence proved. □

Note 3.1. *Since, for optimality, $d_\lambda = D_{uv}\ \forall\lambda$, we get,*

$$\forall\lambda,\ G_\lambda^{opt} = 2\sqrt{D_{uv}(1-D_{uv})} = G_{xy}^{opt} \tag{3.4.5.6}$$

Necessary and Sufficient Conditions to Achieve I_{opt}

When is the I_{opt} attained?

The proof for the upper bound of I essentially provides no extra condition than those to achieve G_{opt}. Therefore, the necessary and sufficient conditions for equality in Eq. (3.3.2.4) is same as in proposition 3.3, viz.,

Since the necessary and sufficient conditions to achieve I_{opt} is same as that of G_{opt}, therefore, **any condition (thus any measurement, any strategy) that maximize G, will also maximize I.**

Necessary and Sufficient Conditions to Achieve $G_{opt}, I_{opt}, \Pr_{opt}(success)$ are same

Following conclusions in sections 3.4.5.1, 3.4.5.1 and 3.4.8, we note that,

- Necessary and sufficient conditions to achieve $G_{opt}, I_{opt}, \Pr_{opt}(success)$ are same, viz., given by proposition 3.3.
- Hence, any measurement or strategy to maximize G, I or $\Pr(success)$ will remain same.

3.4.6 Optimal Interaction and Optimal POVM to maximize $G, I, \Pr(success)$

Here, both D_{xy} and D_{uv} are fixed independently.

3.4.6.1 Description of the Post-interaction States $|X\rangle, |Y\rangle$

Let's fix both D_{xy} and D_{uv} independently. The Schmidt decomposition of the post-interaction states must be of the form as in Eqs. (3.3.4.1, 3.3.4.3) for the respective bases. According the rules of Smith decomposition, the normalized states $|\xi_i\rangle, |\zeta_i\rangle$ follow the orthogonality relation as below.

$$\langle\xi_x|\zeta_x\rangle = \langle\xi_y|\zeta_y\rangle = 0 \quad \text{i.e.,} \quad |\xi_i\rangle \perp |\zeta_i\rangle \; \forall i \in \{x,y\},$$

and

$$\langle\xi_u|\zeta_u\rangle = \langle\xi_v|\zeta_v\rangle = 0 \quad \text{i.e.,} \quad |\xi_i\rangle \perp |\zeta_i\rangle \; \forall i \in \{u,v\}$$

for the respective bases.

The remaining relations between the $|\xi_i\rangle, |\zeta_i\rangle$ cannot be chosen arbitrarily. However, under proper assumption, the space of ξ and the space of ζ states can be made mutually orthogonal for each of the encoding bases. The relevant details are described below.

Remaining Relations between $|\xi_i\rangle, |\zeta_i\rangle$:

Proposition 3.4.6.1. *The orthogonality relation $\langle X|Y\rangle = 0$ induces the following restriction on the IVs.*

$$\langle\xi_x|\zeta_y\rangle + \langle\zeta_x|\xi_y\rangle = 0.$$

Proposition 3.4.6.2. *The conjugate relation between the two encoding bases induces a relation between the IVs across the bases as given by Eq.* (3.3.4.4). *We prove the following two relations in Sec. 3.4.10.4.*

$$\begin{aligned} 2\sqrt{1-D_{uv}}\,|\xi_u\rangle &= \sqrt{1-D_{xy}}\left(|\xi_x\rangle + |\xi_y\rangle\right) + \sqrt{D_{xy}}\left(|\zeta_x\rangle + |\zeta_y\rangle\right), \\ 2\sqrt{D_{uv}}\,|\zeta_u\rangle &= \sqrt{1-D_{xy}}\left(|\xi_x\rangle - |\xi_y\rangle\right) + \sqrt{D_{xy}}\left(|\zeta_y\rangle - |\zeta_x\rangle\right). \end{aligned}$$

Proposition 3.4.6.3. *Orthogonality between $|\xi_u\rangle$ and $|\zeta_u\rangle$ in the above two equations then induces further restrictions as following.*

$$\begin{aligned} \mathbf{Re}[\langle\xi_x|\zeta_y\rangle - \langle\zeta_x|\xi_y\rangle] &= 0, \\ \left(1-D_{xy}\right)\mathbf{Im}[\langle\xi_y|\xi_x\rangle] + D_{xy}\,\mathbf{Im}[\langle\zeta_x|\zeta_y\rangle] &= 0. \end{aligned}$$

[*Proof can be found in Sec 3.4.10.4.*]

Note 3.2. *Now, assume all inner products* $\langle \xi_i | \zeta_j \rangle$ *are real. Then*

$$\langle \xi_i | \zeta_j \rangle = 0, \ \forall i,j \qquad \textit{i.e.,}\ \{|\xi_x\rangle, |\xi_y\rangle\} \perp \{|\zeta_x\rangle, |\zeta_y\rangle\}.$$

From this orthogonality relation, one can say that Eve's probe lives in a Hilbert space having dimension at most four. Thus, her ancilla can taken to be 2 qubits(4 states). It is then convenient to introduce the same bases (xy and uv, used by Alice) for each of Eve's qubits.

3.4.6.2 Mixtures with Bob and Eve

The joint PIJSs are entangled. However, from the description of the PIJSs, it is possible to describe Bob and Eve's states as mixed state, viz., in terms of an density operator.

Bob's states are

$$\begin{aligned} \rho_x^{\mathcal{B}} &:= \mathsf{Tr}_{\mathrm{B}}\left(|X\rangle\langle X|\right) = (1-D_{xy})|x\rangle\langle x| + D_{xy}|y\rangle\langle y|, \\ \rho_y^{\mathcal{B}} &:= \mathsf{Tr}_{\mathrm{B}}\left(|Y\rangle\langle Y|\right) = (1-D_{xy})|y\rangle\langle y| + D_{xy}|x\rangle\langle x|. \end{aligned}$$

Eve's states are as follows.

$$\begin{aligned} \rho_x &:= \mathsf{Tr}_{\mathrm{A}}\left(|X\rangle\langle X|\right) = (1-D_{xy})|\xi_x\rangle\langle \xi_x| + D_{xy}|\zeta_x\rangle\langle \zeta_x| \\ &= (1-D_{xy})\widehat{\xi}_x + D_{xy}\widehat{\zeta}_x, && (3.4.6.1) \\ \rho_y &:= \mathsf{Tr}_{\mathrm{A}}\left(|Y\rangle\langle Y|\right) = (1-D_{xy})|\xi_y\rangle\langle \xi_y| + D_{xy}|\zeta_y\rangle\langle \zeta_y| \\ &= (1-D_{xy})\widehat{\xi}_y + D_{xy}\widehat{\zeta}_y. && (3.4.6.2) \end{aligned}$$

where

$$\begin{aligned} \widehat{\xi}_x &:= |\xi_x\rangle\langle \xi_x|, \qquad \widehat{\zeta}_x := |\zeta_x\rangle\langle \zeta_x|; \\ \widehat{\xi}_y &= |\xi_y\rangle\langle \xi_y|, \qquad \widehat{\zeta}_y := |\zeta_y\rangle\langle \zeta_y|. \end{aligned}$$

See [Calc in Sec 3.4.10.4] how to trace-out Alice or Eve.

3.4.6.3 The optimal POVM for the Chosen Interaction

Suppose Alice announces that a signal from the x-y basis was sent to Bob with **equal prior probabilities**. To achieve maximum G, Eve chooses an optimal POVM $\{E_\lambda\}$ to measure her probe (the post-interaction state $|X\rangle$).

Theorem 3.2. *For the optimal interactions* (3.3.7.3)*, the observable with Eve can be* *described as follows*

$$\Gamma_{xy} \;=\; 2\sqrt{D_{uv}(1-D_{uv})}\left[(1-D_{xy})\left(\mathbb{E}_{00}-\mathbb{E}_{33}\right)+D_{xy}\left(\mathbb{E}_{22}-\mathbb{E}_{11}\right)\right]. \tag{3.4.6.3}$$

where the projectors $\mathbb{E}_{ij} := |E_i\rangle\langle E_j|$ *correspond to the following states* $\{E_\lambda\}_{\lambda\in\{0,1,2,3\}}$.

$$\begin{aligned} |E_0\rangle &= |x\rangle|x\rangle, & |E_1\rangle &= |y\rangle|x\rangle \\ |E_2\rangle &= |x\rangle|y\rangle, & |E_3\rangle &= |y\rangle|y\rangle \end{aligned} \tag{3.4.6.4}$$

Clearly, the corresponding optimal POVM for these IVs are the eigenprojectors $E_\lambda = |E_\lambda\rangle\langle E_\lambda|$. *Considering*

$$\gamma_1 = 2(1-D_{xy})\sqrt{D_{uv}(1-D_{uv})}, \qquad \gamma_2 = D_{xy}k = 2D_{xy}\sqrt{D_{uv}(1-D_{uv})},$$

the eigenvalues and eigenvectors of Γ_{xy} *are listed as below.*

Eigenvalues:	γ_1	γ_2	-γ_2	-γ_1
Eigenvectors:	$\lvert E_0\rangle$	$\lvert E_2\rangle$	$\lvert E_1\rangle$	$\lvert E_3\rangle$

Proof. By theorem 3.3, the optimal POVM is an orthonormal eigenprojector of

$$\Gamma_{xy} := \rho_x - \rho_y$$

Using the density operators ρ_x, ρ_y of Eqs. (3.4.6.1, 3.4.6.2), we get

$$\Gamma_{xy} \;=\; (1-D_{xy})\left(\widehat{\xi}_x - \widehat{\xi}_y\right) + D_{xy}\left(\widehat{\zeta}_x - \widehat{\zeta}_y\right).$$

Now, for optimal IVs, i.e., for the choice of $|\xi_x\rangle, |\xi_y\rangle, |\zeta_x\rangle, |\zeta_y\rangle$ as in Eq. (3.3.7.3), we have [see Sec. 3.4.10.5 for calculations]

$$\begin{aligned} \widehat{\xi}_x &:= |\xi_x\rangle\langle\xi_x| = \mathscr{D}_{uv}^2\mathbb{E}_{00} + \overline{\mathscr{D}}_{uv}^2\mathbb{E}_{33} + \frac{1}{2}(1-2D_{uv})\left(\mathbb{E}_{03}+\mathbb{E}_{30}\right), \\ \widehat{\zeta}_x &:= |\zeta_x\rangle\langle\zeta_x| = \mathscr{D}_{uv}^2\mathbb{E}_{11} + \overline{\mathscr{D}}_{uv}^2\mathbb{E}_{22} + \frac{1}{2}(1-2D_{uv})\left(\mathbb{E}_{21}+\mathbb{E}_{12}\right). \end{aligned}$$

Here, $\mathbb{E}_{ij} := |E_i\rangle\langle E_j|$.Thus,

$$\Gamma_{xy} \;=\; 2\sqrt{D_{uv}(1-D_{uv})}\left[(1-D_{xy})\left(\mathbb{E}_{00}-\mathbb{E}_{33}\right)+D_{xy}\left(\mathbb{E}_{22}-\mathbb{E}_{11}\right)\right].$$

which turns out to be a diagonal matrix. The eigenbasis that diagonalizes it are the $\{|E_\lambda\rangle\}_{\lambda\in\{0,1,2,3\}}$ and are described in the theorem statement. □

3.4.6.4 Optimality of the Interaction

Here, we prove that the choice of interactions(i.e., $|\xi_i\rangle, |\zeta_j\rangle$ in Eq. (3.3.7.3)) is indeed optimal. We need to show that this specific interaction along with the optimal POVM $\{E_\lambda\}$ given by Eqs. (4.2.3, 3.4.6.4) for measurement achieves the bounds of G. For that, we need only to prove the proposition below:

Proposition 3.4.6.4. *The postinteraction states $|X\rangle, |Y\rangle$ in Eq.* (3.3.4.1) *along with the IVs $|\xi_i\rangle, |\zeta_j\rangle$ in Eq.* (3.3.7.3)*, and the measurement* (3.4.6.4)*, the necessary and sufficient conditions (3.3.2.1, 3.3.2.2) are satisfied. Thereby, the interaction is optimal.*

Proof. We prove the proposition for $\lambda = 1$ only. Other cases are similar.

Lemma 3.4.6.1. $\sqrt{E_\lambda} = E_\lambda$ *for the optimal E_λ chosen in Eq.* (3.4.6.4).

It is due to the facts that $\langle E_\lambda | E_\lambda \rangle = 1$, $E_\lambda^2 = E_\lambda$.

Lemma 3.4.6.2. *For the above choices, we can show that*

$$|V_{1u}\rangle \quad = \quad \varepsilon_1 \sqrt{\frac{D_{uv}}{1-D_{uv}}}\ |U_{1u}\rangle$$

where

$$\varepsilon_1 = +1$$

*[Proofsketch:]*We'll prove that

$$\begin{aligned} |U_{1u}\rangle &= \frac{1}{\sqrt{2}}\sqrt{1-D_{uv}}\sqrt{D_{xy}}\ |u\rangle|E_1\rangle \\ |V_{1u}\rangle &= \frac{1}{\sqrt{2}}\sqrt{D_{uv}}\sqrt{D_{xy}}\ |u\rangle|E_1\rangle \end{aligned}$$

and thus the above equation follows.

Proof.

$$\begin{aligned} |U_{1u}\rangle &= B_u \otimes \sqrt{E_1}\ |U\rangle \\ &= B_u \otimes E_1\ |U\rangle \\ &\Big\downarrow\ \begin{array}{l} B_u = |u\rangle\langle u|, E_1 = |E_1\rangle\langle E_1|, \text{eq } 3.3.4.3 \\ \qquad [\text{ Calc in Sec } 3.4.10.6\ ] \end{array} \\ &= \sqrt{1-D_{uv}}\ \langle E_1|\xi_u\rangle|u\rangle|E_1\rangle \\ &\Big\downarrow\ [\text{ Calc in Sec } 3.4.10.6\ ] \\ &= \frac{1}{\sqrt{2}}\sqrt{1-D_{uv}}\sqrt{D_{xy}}\ |u\rangle|E_1\rangle \end{aligned}$$

Similarly,

$$\begin{aligned} |V_{1u}\rangle &= B_u \otimes \sqrt{E_1}\ |V\rangle = B_u \otimes E_1\ |V\rangle \\ &= \sqrt{D_{uv}}\ \langle E_1|\zeta_v\rangle |u\rangle |E_1\rangle \\ &= \frac{1}{\sqrt{2}}\sqrt{D_{uv}}\sqrt{D_{xy}}\ |u\rangle |E_1\rangle \end{aligned}$$

Note 3.3. *It could be shown that*

$$\varepsilon_0 = +1, \qquad \varepsilon_1 = +1, \qquad \varepsilon_2 = -1, \qquad \varepsilon_3 = -1 \tag{3.4.6.5}$$

□

It completes the proof. □

Summary of the subsection

In this subsection, we have described an optimal POVM $\{E_\lambda\}$ given by Eq. 4.2.3 and (3.4.6.4) to measure the post-interaction states $|X\rangle, |Y\rangle$ given by Eq. (3.3.4.1) (where choice of $|\xi_x\rangle, |\xi_y\rangle, |\zeta_x\rangle, |\zeta_y\rangle$ is as described in Eq. (3.3.7.3)). This optimal POVM leads to maximization of G (and therefore $I, \Pr(success)$) once we fix an optimal strategy for Eve to assign a signal against a measurement outcome λ. In next section we fix the optimal strategy.

3.4.7 Eve's optimal strategy: Guess the state send by Alice

Eve's objective to maximize her information gained on the signal of Alice. It's enough to optimize IG, as it also optimizes MI and her probability of success. She must interpret her measurement outcome in such an way that assigns the best guess, that is her optimal strategy.

Mathematically, the task is to achieve G_{opt}, which in turn achieves I_{opt} and $\Pr_{opt}(success)$.

Let's do it for the xy basis. She performs an optimal measurement E_λ on her ancilla that is part of the optimal PIJS $|X\rangle, |Y\rangle$. Looking at the outcome λ, she must assign either x or y as a guess on Alice's signal s.t. G_{opt} is attained. The optimal strategy is given by

$$S_{opt}(\lambda) \quad = \quad \mathsf{arg\,max}\ \{Q_{x\lambda}, Q_{y\lambda}\}.$$

For the optimal states and the optimal measurement, we have to calculate these posterior probabilities for each measurement outcome, and choose the largest one.

Note that, for equal prior (i.e., $p_x = p_y = 1/2$),

$$Q_{x\lambda} = \frac{P_{\lambda x}}{P_{\lambda x} + P_{\lambda y}}, \qquad Q_{y\lambda} = \frac{P_{\lambda y}}{P_{\lambda x} + P_{\lambda y}}.$$

In that case,

$$\mathsf{arg\,max}\ \{Q_{x\lambda}, Q_{y\lambda}\} = \mathsf{arg\,max}\ \{P_{\lambda x}, P_{\lambda y}\}$$

We can apply the Born rule to calculate these probabilities in either of the following forms

$$\begin{aligned} P_{\lambda x} &= \langle X|\mathbb{1} \otimes E_\lambda|X\rangle &= \mathsf{Tr}(\rho_x E_\lambda), \\ P_{\lambda y} &= \langle Y|\mathbb{1} \otimes E_\lambda|Y\rangle &= \mathsf{Tr}(\rho_y E_\lambda). \end{aligned}$$

We use the optimal PIJSs $|X\rangle, |Y\rangle$ and the optimal measurements E_λ fromEq. (3.3.4.1), and Eq. (3.4.6.4), respectively. It's merely a matter of calculations that leads to the following results.

Table 3.1 | Eve's prior probabilities $P(\texttt{E}|\texttt{A})$ for the optimal states.

Values of $P_{\lambda x}, P_{\lambda y}$ for the optimal states

λ	$P_{\lambda x}$	$P_{\lambda y}$	$P_{\lambda x}+P_{\lambda y}$
0	$\left(\frac{1}{2}+\sqrt{D_{uv}(1-D_{uv})}\right)(1-D_{xy})$	$\left(\frac{1}{2}-\sqrt{D_{uv}(1-D_{uv})}\right)(1-D_{xy})$	$1-D_{xy}$
3	$\left(\frac{1}{2}-\sqrt{D_{uv}(1-D_{uv})}\right)(1-D_{xy})$	$\left(\frac{1}{2}+\sqrt{D_{uv}(1-D_{uv})}\right)(1-D_{xy})$	$1-D_{xy}$
2	$\left(\frac{1}{2}-\sqrt{D_{uv}(1-D_{uv})}\right)D_{xy}$	$\left(\frac{1}{2}+\sqrt{D_{uv}(1-D_{uv})}\right)D_{xy}$	D_{xy}
1	$\left(\frac{1}{2}+\sqrt{D_{uv}(1-D_{uv})}\right)D_{xy}$	$\left(\frac{1}{2}-\sqrt{D_{uv}(1-D_{uv})}\right)D_{xy}$	D_{xy}
$\Sigma\lambda$	1	1	2

The posterior probabilities are then calculated below using table 3.1.

Table 3.2 | Eve's posterior probabilities $P(\texttt{A}|\texttt{E})$ for the optimal states.

Values of $Q_{x\lambda}, Q_{y\lambda}$ for the optimal states

λ	$Q_{x\lambda}$	$Q_{y\lambda}$	$S_{opt}(\lambda)$
0	$\frac{1}{2}+\sqrt{D_{uv}(1-D_{uv})}$	$\frac{1}{2}-\sqrt{D_{uv}(1-D_{uv})}$	x
3	$\frac{1}{2}-\sqrt{D_{uv}(1-D_{uv})}$	$\frac{1}{2}+\sqrt{D_{uv}(1-D_{uv})}$	y
2	$\frac{1}{2}-\sqrt{D_{uv}(1-D_{uv})}$	$\frac{1}{2}+\sqrt{D_{uv}(1-D_{uv})}$	y
1	$\frac{1}{2}+\sqrt{D_{uv}(1-D_{uv})}$	$\frac{1}{2}-\sqrt{D_{uv}(1-D_{uv})}$	x

Eve's strategy: By the last column of the table 3.2, Eve's strategy becomes

assign signal x **against measurement outcome** E_0, E_1
assign signal y **against measurement outcome** E_2, E_3

Measurement Outcome	E_0, E_1	E_2, E_3
Eve assigns signal	x	y

Strategy of Eve remains same to attain optimality of $G, I, \Pr(\text{success})$. For this optimal strategy, optimal value of each of these parameters are tabulated below.

G_{xy}^{opt}	$\Pr_{xy}^{opt}(\text{success}) = \frac{1}{2} + \frac{1}{2}G_{xy}^{opt}$	$I_{xy}^{opt} = \frac{1}{2}\ \phi(G_{xy}^{opt})$
$2\sqrt{D_{uv}(1-D_{uv})}$	$\frac{1}{2} + \sqrt{D_{uv}(1-D_{uv})}$	$\frac{1}{2}\ \phi\left[2\sqrt{D_{uv}(1-D_{uv})}\right]$

An Interesting Observation: By Eq. (3.3.3.3),

$$\varepsilon_\lambda \;=\; \pm 1 = \mathsf{sgn}\left(Q_{x\lambda} - Q_{y\lambda}\right)$$

Thus,

$$\begin{aligned} \varepsilon_\lambda = +1 &\implies \mathsf{arg\,max}\left(Q_{x\lambda}, Q_{y\lambda}\right) = x \\ \varepsilon_\lambda = -1 &\implies \mathsf{arg\,max}\left(Q_{x\lambda}, Q_{y\lambda}\right) = y \end{aligned}$$

which indicates that value of ε_λ has 1-1 correspondence with $S_{opt}(\lambda) = \mathsf{arg\,max}\left(Q_{x\lambda}, Q_{y\lambda}\right)$. To be precise,

$\lambda =$	0	1	2	3
$\varepsilon_\lambda =$	$+1$	$+1$	-1	-1
$S_{opt}(\lambda) = \mathsf{arg\,max}\left(Q_{x\lambda}, Q_{y\lambda}\right) =$	x	x	y	y

3.4.8 Relation between G_{opt} and $\Pr_{opt}(success)$

Theorem 3.3. *Eve's probability to successfully guess Alice's signal in xy basis, for an optimal attack, is the following.*

$$\Pr_{xy}^{\star}(\text{success}) \;=\; \frac{1}{2} + \sqrt{D_{uv}(1-D_{uv})} = \frac{1}{2} + \frac{1}{2}G_{xy}^{\star}. \tag{3.4.8.1}$$

Therefore, the optimal states that optimize the IG (or MI), also optimize the success probability.

Proof.

$$\Pr_{xy}^{\star}(\text{success}) \;=\; F_{xy}\cdot\Pr(\text{success}|\text{undist}) + D_{xy}\cdot\Pr(\text{success}|\text{dist}).$$

The probability to distinguish two states can be given by Helstrom formula. Since the undisturbed states corresponds to $|\xi_x\rangle, |\xi_y\rangle$, and the disturbed states correspond to $|\zeta_x\rangle, |\zeta_y\rangle$, we can write

$$\Pr^{\star}_{xy}(\mathsf{success}) = F_{xy} \cdot \frac{1}{2}\left(1 + \sqrt{1 - |\langle \xi_x | \xi_y \rangle|^2}\right) + D_{xy} \cdot \frac{1}{2}\left(1 + \sqrt{1 - |\langle \zeta_x | \zeta_y \rangle|^2}\right).$$

Note that, for the optimal states, both the inner products become $1 - 2D_{uv}$. Since, $F_{xy} + D_{xy} = 1$, we get

$$\Pr^{\star}_{xy}(\mathsf{success}) = \frac{1}{2}\left(1 + \sqrt{1 - (1 - 2D_{uv})^2}\right).$$

A simplification leads to the desired result. □

3.4.9 Eavesdropping for equal error rates (i.e., $D_{xy} = D_{uv} = D$)

So far, we have got the maximum information, whether IG or MI, across the two bases. Maximum information in a basis is a function of the disturbance in the conjugate basis. A natural question is how much information did she have for both the bases? Since each of the bases are chosen 50% of times, we can consider the average IG and MI for an average of the disturbances, i.e., the following functions.

Average Information Gain:

$$G = \frac{1}{2}\left(G_{xy} + G_{uv}\right), \qquad (3.4.1)$$

Average Mutual Information:

$$I = \frac{1}{2}\left(I_{xy} + I_{uv}\right), \qquad (3.4.2)$$

Average Disturbance:

$$D = \frac{1}{2}\left(D_{xy} + D_{uv}\right). \qquad (3.4.3)$$

The objective is now to know Eve's best average information (IG, or, MI) for a fixed average disturbance D across the two bases. This is shown to be achieved with equal error rates, i.e., $D_{xy} = D_{uv}$.

Proposition 3.4.1. *Following upper bounds for average information are found.*

$$G \leq 2\sqrt{D(1-D)} \qquad (3.4.4)$$

$$I \leq \frac{1}{2}\,\phi\left[2\sqrt{D(1-D)}\right] \qquad (3.4.5)$$

Equality in each of the cases is achieved for $D_{xy} = D_{uv} = D$.

Proof. The idea is to use the concavity of the functions $\sqrt{x(1-x)}$, and $\phi\left[2\sqrt{x(1-x)}\right]$, respectively.

$$\begin{aligned} G=\frac{1}{2}\left(G_{xy}+G_{uv}\right) &\leq 2\left[\frac{1}{2}\sqrt{D_{xy}\left(1-D_{xy}\right)}+\frac{1}{2}\sqrt{D_{uv}(1-D_{uv})}\right] \\ &\leq 2\sqrt{\frac{1}{2}\left(D_{xy}+D_{uv}\right)\left(1-\frac{1}{2}\left(D_{xy}+D_{uv}\right)\right)}, \\ &\qquad \texttt{due to concavity of } \sqrt{x(1-x)}. \\ &\qquad \texttt{Equality holds for } D_{xy}=D_{uv}=D. \\ &= 2\sqrt{D(1-D)}. \end{aligned}$$

Similarly,

$$\begin{aligned} I=\frac{1}{2}\left(I_{xy}+I_{uv}\right) &\leq \frac{1}{2}\left[\frac{1}{2}\phi\left[2\sqrt{D_{xy}\left(1-D_{xy}\right)}\right]+\frac{1}{2}\phi\left[2\sqrt{D_{uv}(1-D_{uv})}\right]\right] \\ &\leq \frac{1}{2}\,\phi\left[2\sqrt{\frac{1}{2}\left(D_{xy}+D_{uv}\right)\left(1-\frac{1}{2}\left(D_{xy}+D_{uv}\right)\right)}\right], \\ &\qquad \texttt{due to concavity of } \phi\left[2\sqrt{x(1-x)}\right]. \\ &\qquad \texttt{Equality holds when } D_{xy}=D_{uv}=D. \\ &= \frac{1}{2}\phi\left[2\sqrt{D(1-D)}\right]. \end{aligned}$$

The equality of the error rates match the intuition of a symmetrical attack. □

Now, to see how the bounds could be achieved, one needs to find the suitable sates that can saturate the bound.

Consider the case when the legitimate parties agree with x-y basis. Eve is left with the following two density operators that she must distinguish in order to identify the signal sent by Alice.

$$\begin{aligned} \rho_x &= (1-D)|\xi_x\rangle\langle\xi_x| + D|\zeta_x\rangle\langle\zeta_x| \qquad (3.4.6) \\ \rho_y &= (1-D)|\xi_y\rangle\langle\xi_y| + D|\zeta_y\rangle\langle\zeta_y| \qquad (3.4.7) \end{aligned}$$

Note that, the ξ-states occur w.p. $(1-D)$, while the ζ-states occur w.p. D. Although, these two states are mutually orthogonal, at that point, there is no way to distinguish them.

So, consider the ξ,ζ-states in such a manner that the orthogonality constraint is met, as well, some (real) non-zero overlap is consumed by the states within a set. Following states may be considered (again, uniqueness is a question).

$$\begin{aligned} &|\xi_x\rangle = |x\rangle|x\rangle, \qquad |\xi_y\rangle = \big(\cos\alpha|x\rangle + \sin\alpha|y\rangle\big)\,|x\rangle; \\ &|\zeta_x\rangle = |x\rangle|y\rangle, \qquad |\zeta_y\rangle = \big(\cos\beta|x\rangle + \sin\beta|y\rangle\big)\,|y\rangle. \qquad (3.4.8) \end{aligned}$$

Apart from orthogonality, note that $\langle\xi_x|\xi_y\rangle = \cos\alpha, \langle\zeta_y|\zeta_x\rangle = \cos\beta$.

If Eve measures in computational basis the 2nd qubit for these states, she can distinguish ξ,ζ: if she gets $|x\rangle$ (or, $|y\rangle$), she considers the state to be from ξ (or, ζ) set. Once distinguished the sets, she now wants to distinguish the states within a set, having non-zero overlaps $\cos\alpha$ and $\cos\beta$, respectively.

The averaged success probability to distinguish the states then become

$$\begin{aligned} \mathrm{Pr}^{\star}_{xy}(\mathsf{success}) &= F\cdot \mathrm{Pr}(\mathsf{success}|\mathsf{undist}) + D\cdot \mathrm{Pr}(\mathsf{success}|\mathsf{dist}) \\ &= F\cdot\frac{1}{2}\left(1+\sqrt{1-|\langle\xi_x|\xi_y\rangle|^2}\right) + D\cdot\frac{1}{2}\left(1+\sqrt{1-|\langle\zeta_x|\zeta_y\rangle|^2}\right), \\ &\qquad\qquad \texttt{by Helstrom rule} \\ &= F\cdot\frac{1}{2}(1+\sin\alpha) + D\cdot\frac{1}{2}(1+\sin\beta) \\ &= \frac{1}{2}(1+F\cdot\sin\alpha + D\cdot\sin\beta). \qquad (3.4.9) \end{aligned}$$

Also, since the success probability in the individual states are $\frac{1}{2}(1+\sin\alpha)$, and $\frac{1}{2}(1+\sin\beta)$, her MI for these two sets become

$$\begin{aligned} I_\xi &= \frac{1}{2}(1+\sin\alpha)\ln(1+\sin\alpha) + \frac{1}{2}(1-\sin\alpha)\ln(1-\sin\alpha) = \frac{1}{2}\phi(\sin\alpha), \\ I_\zeta &= \frac{1}{2}\big(1+\sin\beta\big)\ln\big(1+\sin\beta\big) + \frac{1}{2}\big(1-\sin\beta\big)\ln\big(1-\sin\beta\big) = \frac{1}{2}\phi(\sin\beta). \end{aligned}$$

The average MI over the two sets becomes

$$I_{\xi} \;=\; F\cdot I_{\xi}+D\cdot I_{\zeta}=\frac{1}{2}\left(F\cdot\phi(\sin\alpha)+D\cdot\phi(\sin\beta)\right). \qquad (3.4.10)$$

The maximum for both Eq. (3.4.9) and Eq. (3.4.10) occur for $\alpha=\beta$, with $\sin\alpha = 2\sqrt{D(1-D)}$. We have shown the proof for the success probability, other one would be similar. [*Proof can be found in Sec 3.4.10.7.*]

To prove it, it's worth to note that the fidelity and disturbance can be written as a function of the angles α,β. One can consider any of the interrelation between the IVs across the two bases, and get the following relation.

Proposition 3.4.2. *For the choice of ξ_i,ζ_j in Eq.* (3.4.8)*, we get*

$$D \;=\; \frac{1-\cos\alpha}{2-\cos\alpha+\cos\beta} \qquad (3.4.11)$$

[*Proof can be found in Sec 3.4.10.7.*]

Remarks

Proposition 3.4.3. *For small D, $I_{AE}\leq 2D$. At the other extreme, $I_{AE}^{\max}=\ln 2$.*

[*Proof can be found in Sec 3.4.10.7.*].

Proposition 3.4.4. *One can verify easily the following*

$$I_{AB} \;=\; \ln 2+D\ln D+(1-D)\ln(1-D)=\frac{1}{2}\phi(1-2D). \qquad (3.4.12)$$

Thus, the threshold noise level for a potentially safe channel can be derived now.

Proposition 3.4.5. *The channel is considered inappropriate for key generation when the key-rate becomes zero. It happens for error-rate $D\geq 0.146447$.*

Proof. Key-rate becomes zero when $I_{\mathtt{AB}}=I_{\mathtt{AE}}$, i.e.,

$$\begin{aligned}
&\frac{1}{2}\phi(1-2D)=\frac{1}{2}\phi\left[2\sqrt{D(1-D)}\right]\\
\implies\;& |1-2D|=2\sqrt{D(1-D)}\\
\implies\;& 8D^2-8D+1 \;=\; 0\\
\implies\;& D=\frac{1}{2}-\frac{\sqrt{2}}{4}\approx 0.146447.
\end{aligned}$$

□

3.4.10 Illustration between the lines (IBtL)

3.4.10.1 IBtL Sec. 3.4.3: Proving bounds for IG

Calculation 3.4.10.1.1. *P.T.*

$$\left|\langle X|\mathbb{1}\otimes E_\lambda|X\rangle - \langle Y|\mathbb{1}\otimes E_\lambda|Y\rangle\right| = \left|\langle U|\mathbb{1}\otimes E_\lambda|V\rangle + \langle V|\mathbb{1}\otimes E_\lambda|U\rangle\right|.$$

Use: $|X\rangle = \frac{1}{\sqrt{2}}\Big(|U\rangle + |V\rangle\Big)$, $|Y\rangle = \frac{1}{\sqrt{2}}\Big(|U\rangle - |V\rangle\Big)$

Proof.

$$\begin{aligned} & \left|\langle X|\mathbb{1}\otimes E_\lambda|X\rangle - \langle Y|\mathbb{1}\otimes E_\lambda|Y\rangle\right| \\ = \; & \frac{1}{2}\left|\left(\langle U| + \langle V|\right)\left(\mathbb{1}\otimes E_\lambda\right)\Big(|U\rangle + |V\rangle\Big) - \left(\left(\langle U| - \langle V|\right)\left(\mathbb{1}\otimes E_\lambda\right)\left(|U\rangle - |V\rangle\right)\right)\right| \\ = \; & \frac{1}{2}\Big|\big(\langle U|\mathbb{1}\otimes E_\lambda|U\rangle + \langle V|\mathbb{1}\otimes E_\lambda|V\rangle + \langle U|\mathbb{1}\otimes E_\lambda|V\rangle + \langle V|\mathbb{1}\otimes E_\lambda|U\rangle\big) \\ & - \big(\langle U|\mathbb{1}\otimes E_\lambda|U\rangle + \langle V|\mathbb{1}\otimes E_\lambda|V\rangle - \langle U|\mathbb{1}\otimes E_\lambda|V\rangle - \langle V|\mathbb{1}\otimes E_\lambda|U\rangle\big)\Big| \\ = \; & \left|\langle U|\mathbb{1}\otimes E_\lambda|V\rangle + \langle V|\mathbb{1}\otimes E_\lambda|U\rangle\right| \end{aligned}$$

□

Calculation 3.4.10.1.2. *P.T.*

$$\left|\langle U|\mathbb{1}\otimes E_\lambda|V\rangle + \langle V|\mathbb{1}\otimes E_\lambda|U\rangle\right| = 2\;\left|\mathbf{Re}\langle U|B_u\otimes E_\lambda|V\rangle + \mathbf{Re}\langle V|B_v\otimes E_\lambda|U\rangle\right|$$

Use: $\mathbb{1} = B_u + B_v :\; B_u = |u\rangle\langle u|,\; B_v = |v\rangle\langle v|$ *Note:* $B_u^2 = B_u$

Proof.

$$\begin{aligned} & \left|\langle U|\mathbb{1}\otimes E_\lambda|V\rangle + \langle V|\mathbb{1}\otimes E_\lambda|U\rangle\right| \\ = \; & \left|\langle U|\left(B_u + B_v\right)\otimes E_\lambda|V\rangle + \langle V|\left(B_u + B_v\right)\otimes E_\lambda|U\rangle\right| \\ = \; & \Big|\left[\langle U|B_u\otimes E_\lambda|V\rangle + \langle V|B_u\otimes E_\lambda|U\rangle\right] + \left[\langle U|B_v\otimes E_\lambda|V\rangle + \langle V|B_v\otimes E_\lambda|U\rangle\right]\Big| \\ & \textbf{Use: } \langle U|B_u\otimes E_\lambda|V\rangle + \langle V|B_u\otimes E_\lambda|U\rangle = 2\,\mathbf{Re}\langle U|B_u\otimes E_\lambda|V\rangle \\ = \; & 2\;\left|\mathbf{Re}\langle U|B_u\otimes E_\lambda|V\rangle + \mathbf{Re}\langle V|B_v\otimes E_\lambda|U\rangle\right| \end{aligned}$$

□

Calculation 3.4.10.1.3. *P.T.*

$$\left|\mathbf{Re}\langle U|B_u\otimes E_\lambda|V\rangle + \mathbf{Re}\langle V|B_v\otimes E_\lambda|U\rangle\right| \le \left|\langle U_{\lambda u}|V_{\lambda u}\rangle\right| + \left|\langle U_{\lambda v}|V_{\lambda v}\rangle\right|.$$

Equality takes place when

1. $\mathbf{Re}\langle U|B_u \otimes E_\lambda|V\rangle = \langle U|B_u \otimes E_\lambda|V\rangle, \qquad \mathbf{Re}\langle V|B_v \otimes E_\lambda|U\rangle = \langle V|B_v \otimes E_\lambda|U\rangle$

2. $\mathbf{Re}\langle U|B_u \otimes E_\lambda|V\rangle, \mathbf{Re}\langle V|B_v \otimes E_\lambda|U\rangle$ *are of same sign.*

Proof.

$$\begin{aligned}
& \left|\mathbf{Re}\langle U|B_u \otimes E_\lambda|V\rangle + \mathbf{Re}\langle V|B_v \otimes E_\lambda|U\rangle\right| \\
\leq\ & \left|\langle U|B_u \otimes E_\lambda|V\rangle\right| + \left|\langle V|B_v \otimes E_\lambda|U\rangle\right| \\
& \quad \texttt{Equality: when } \mathbf{Re\langle U|B_u\otimes E_\lambda|V\rangle=\langle U|B_u\otimes E_\lambda|V\rangle}, \qquad \mathbf{Re\langle V|B_v\otimes E_\lambda|U\rangle=\langle V|B_v\otimes E_\lambda|U\rangle} \\
& \qquad\qquad \texttt{and } \mathbf{Re\langle U|B_u\otimes E_\lambda|V\rangle,Re\langle V|B_v\otimes E_\lambda|U\rangle} \texttt{ are of same sign} \\
\Big\downarrow\ & \quad \texttt{Use: } \mathbf{B_u^2=B_u,B_v^2=B_v} \\
=\ & \left|\langle U|B_u^2 \otimes E_\lambda|V\rangle\right| + \left|\langle V|B_v^2 \otimes E_\lambda|U\rangle\right| \\
=\ & \left|\langle U|\left(B_u \otimes \sqrt{E_\lambda}\right)\left(B_u \otimes \sqrt{E_\lambda}\right)|V\rangle\right| + \left|\langle V|\left(B_v \otimes \sqrt{E_\lambda}\right)\left(B_v \otimes \sqrt{E_\lambda}\right)|U\rangle\right| \\
\Big\downarrow\ & \quad \texttt{Use: } \begin{matrix} U_{\lambda u} = B_u \otimes \sqrt{E_\lambda}\ \ |U\rangle & V_{\lambda u} = B_u \otimes \sqrt{E_\lambda}\ \ |V\rangle \\ U_{\lambda v} = B_v \otimes \sqrt{E_\lambda}\ \ |U\rangle & V_{\lambda v} = B_v \otimes \sqrt{E_\lambda}\ \ |V\rangle \end{matrix} \\
=\ & \left|\langle U_{\lambda u}|V_{\lambda u}\rangle\right| + \left|\langle V_{\lambda v}|U_{\lambda v}\rangle\right| \ = \ \left|\langle U_{\lambda u}|V_{\lambda u}\rangle\right| + \left|\langle U_{\lambda v}|V_{\lambda v}\rangle\right|.
\end{aligned}$$

□

Proposition 3.4.10.1.1. *P.T.*

$$f(w) = \sqrt{\left(d_\lambda + w\right)\left(1 - d_\lambda + w\right)} + \sqrt{\left(d_\lambda - w\right)\left(1 - d_\lambda - w\right)}$$

is maximum at $w = 0$.

Proof. Let

$$g(w) = \sqrt{(\pi_1 + w)(\pi_2 + w)} \quad \text{s.t. } \pi_1 + \pi_2 = 1$$

Then,

$$f(w) = g(w) + g(-w), \quad \text{where, } \pi_1 = d_\lambda, \pi_2 = 1 - d_\lambda$$

Let's say,

$$f_+(w) = g(w) + g(-w), \qquad f_-(w) = g(w) - g(-w)$$

Claim 3.4.10.1.1.

$$\begin{aligned}
1. &\quad g^2(w) = \pi_1\pi_2 + w + w^2 \\
2. &\quad g'(w) = \frac{1 + 2w}{2g(w)} \\
3. &\quad g''(w) = \frac{4\pi_1\pi_2 - 1}{4g^3(w)}
\end{aligned}$$

Clearly, for max $f(w)$,

$$\begin{aligned} 0 = f'(w) &= g'(w) - g'(-w) \\ \implies 2w &= \frac{f_-(w)}{f_+(w)} \end{aligned} \tag{3.4.1}$$

Also,

$$2w = g^2(w) - g^2(-w) = \big(g(w) - g(-w)\big)\big(g(w) + g(-w)\big) = f_-(w) f_+(w) \tag{3.4.2}$$

From Eq. (3.4.1) and (3.4.2) we get,

$$0 = f_-(w)\left[1 - f_+^2(w)\right] \tag{3.4.3}$$

Case 1:

$$0 = f_-(w) \implies w = 0$$

Case 2:

$$1 = f_+^2(w) \implies 1 = 4\pi_1\pi_2$$

Thus $f(w)$ is max at $w = 0$. □

3.4.10.2 IBtL Sec. 3.4.4: Proving bounds for MI

We need to use the following two results:

1. $z(x) = \sqrt{x(1-x)}$ is a concave function of x.

2. $\phi(z) = (1+z)\ln(1+z) + (1-z)\ln(1-z)$ is m.i.

Calculation 3.4.10.2.1. *P.T.* $\phi\left[2\sqrt{x(1-x)}\right]$ *is a concave function of x.*

Note 3.4. *Observations:*

1. $z(x) = 2\sqrt{x(1-x)}$ *is concave.*

2. $\phi(z) = (1+z)\ln(1+z) + (1-z)\ln(1-z)$ *is a non-concave function of z, since* $\frac{\mathrm{d}^2\phi}{\mathrm{d}z^2} > 0$.

Does it imply that $\phi(z(x))$ *is concave over x?*

Hint 1. *Let* $z(x) = 2\sqrt{x(1-x)}$. *Then T.P.T.*
$\phi(z) = (1+z)\ln(1+z) + (1-z)\ln(1-z)$ *is a concave function of x.*

Proof. Derivatives of $z(x)$:

$$z(x) = 2\sqrt{x(1-x)}, \qquad z'(x) = \frac{1-2x}{z(x)}, \qquad z''(x) = -\frac{4}{z^3(x)}$$

Derivatives of $\phi(z)$:

$$\begin{aligned} \phi(z) &= (1+z)\ln(1+z) + (1-z)\ln(1-z) \\ \frac{\mathrm{d}\phi}{\mathrm{d}z} &= \ln\left(\frac{1+z}{1-z}\right) \\ \frac{\mathrm{d}^2\phi}{\mathrm{d}z^2} &= \frac{2}{1-z^2} \end{aligned}$$

Thus, for $\phi = \phi(z(x))$,

$$\frac{\mathrm{d}^2\phi}{\mathrm{d}x^2} = \frac{4}{z^3}\left[2z - \ln\left(\frac{1+z}{1-z}\right)\right] = \frac{4}{z^3}\psi(z), \text{ say.}$$

But, $\psi(z)$ is m.d., since $\psi'(z) < 0$, while $\psi(0) = 0$.
Thus, $\psi(z) < 0$, ensuring $\frac{\mathrm{d}^2\phi}{\mathrm{d}z^2} < 0$ and thereby, $\phi(z(x))$ is concave. □

3.4.10.3 IBtL Sec. 3.4.5: Proving NSC for IG, MI

Calculation 3.4.10.1. *P.T.* $\sigma_\lambda = \mathsf{sgn}\left(Q_{x\lambda} - Q_{y\lambda}\right)$.

Proof.

$$\begin{aligned} \sigma_\lambda &= \mathsf{sgn}\left(\langle U_{\lambda u}|V_{\lambda u}\rangle + \langle U_{\lambda v}|V_{\lambda v}\rangle\right) \\ &= \mathsf{sgn}\left(q_\lambda G_\lambda\right) = \mathsf{sgn}\left(P_{\lambda x} - P_{\lambda y}\right) = \mathsf{sgn}\left(Q_{x\lambda} - Q_{y\lambda}\right) \end{aligned} \tag{3.4.4}$$

□

Calculation 3.4.10.2. *P.T.* $\varepsilon_\lambda = \sigma_\lambda = \varepsilon'_\lambda$.

Proof.

$$\begin{array}{rcl} & |V_{\lambda u}\rangle & = \varepsilon_\lambda \underbrace{\mu}_{\texttt{+ve}} |U_{\lambda u}\rangle \\ \Longrightarrow & \underbrace{\langle U_{\lambda u}|V_{\lambda u}\rangle}_{\substack{\texttt{Real} \\ \sigma_\lambda \texttt{times +ve quantity}}} & = \varepsilon_\lambda \underbrace{\mu}_{\texttt{+ve}} \underbrace{\langle U_{\lambda u}|U_{\lambda u}\rangle}_{\texttt{+ve}} \\ \Longrightarrow & \varepsilon_\lambda & = \sigma_\lambda \\ \text{Similarly} & \varepsilon'_\lambda & = \sigma_\lambda \end{array}$$

□

Calculation 3.4.10.3. *P.T.* $\mu = \sqrt{\frac{D_{uv}}{1-D_{uv}}} = \nu$

Proof.

$$\begin{aligned}
|V_{\lambda u}\rangle &= \pm\, \mu\ |U_{\lambda u}\rangle = \varepsilon_\lambda\ \mu\ |U_{\lambda u}\rangle \\
|V_{\lambda v}\rangle &= \pm\, \nu\ |U_{\lambda v}\rangle = \varepsilon'_\lambda\ \nu\ |U_{\lambda v}\rangle \\
\Longrightarrow\ \underbrace{\langle V_{\lambda u}|V_{\lambda u}\rangle}_{\substack{=\ \mathbf{P_{\lambda v} d_{\lambda v}} \\ =\ \mathbf{q_\lambda D_{uv}}}} &= \mu^2\ \ \underbrace{\langle U_{\lambda u}|U_{\lambda u}\rangle}_{\substack{=\ \mathbf{P_{\lambda u}(1-d_{\lambda u})} \\ =\ \mathbf{q_\lambda (1-D_{uv})}}} \\
\Longrightarrow\ \mu &= \sqrt{\frac{D_{uv}}{1-D_{uv}}} = \nu
\end{aligned}$$

□

3.4.10.4 IBtL Sec. 3.4.6.1: PIJS

Proposition 3.4.10.4.1. *P.T.*

$$\begin{aligned}
2\sqrt{1-D_{uv}}\ |\xi_u\rangle &= \sqrt{1-D_{xy}}\left(|\xi_x\rangle + |\xi_y\rangle\right) + \sqrt{D_{xy}}\left(|\zeta_x\rangle + |\zeta_y\rangle\right) \\
2\sqrt{D_{uv}}\ |\zeta_u\rangle &= \sqrt{1-D_{xy}}\left(|\xi_x\rangle - |\xi_y\rangle\right) + \sqrt{D_{xy}}\left(|\zeta_y\rangle - |\zeta_x\rangle\right)
\end{aligned}$$

Proof. In the PIJSs $|X\rangle, |Y\rangle$ of Eqs. (3.3.4.1) , (3.3.4.3), use the conjugate relations (2.1) ,(3.5).

Now, due to

$$\begin{aligned}
\frac{1}{\sqrt{2}}\left(|U\rangle + |V\rangle\right) &= |X\rangle = \sqrt{1-D_{xy}}\ |x\rangle|\xi_x\rangle + \sqrt{D_{xy}}\ |y\rangle|\zeta_x\rangle \\
&= \sqrt{1-D_{xy}}\ \frac{1}{\sqrt{2}}\left(|u\rangle + |v\rangle\right)|\xi_x\rangle + \sqrt{D_{xy}}\ \frac{1}{\sqrt{2}}\left(|u\rangle - |v\rangle\right)|\zeta_x\rangle,
\end{aligned}$$

we get,

$$|U\rangle + |V\rangle = |u\rangle\left(\sqrt{1-D_{xy}}\ |\xi_x\rangle + \sqrt{D_{xy}}\ |\zeta_x\rangle\right) + |v\rangle\left(\sqrt{1-D_{xy}}\ |\xi_x\rangle - \sqrt{D_{xy}}\ |\zeta_x\rangle\right). \tag{3.4.5}$$

Similarly, due to

$$\begin{aligned}
\frac{1}{\sqrt{2}}\left(|U\rangle - |V\rangle\right) &= |Y\rangle = \sqrt{1-D_{xy}}\ |y\rangle|\xi_y\rangle + \sqrt{D_{xy}}\ |x\rangle|\zeta_y\rangle \\
&= \sqrt{1-D_{xy}}\ \frac{1}{\sqrt{2}}\left(|u\rangle - |v\rangle\right)|\xi_y\rangle + \sqrt{D_{xy}}\ \frac{1}{\sqrt{2}}\left(|u\rangle + |v\rangle\right)|\zeta_y\rangle
\end{aligned}$$

we get,

$$|U\rangle - |V\rangle \;=\; |u\rangle\left(\sqrt{1-D_{xy}}\;|\xi_y\rangle + \sqrt{D_{xy}}\;|\zeta_y\rangle\right) - |v\rangle\left(\sqrt{1-D_{xy}}\;|\xi_y\rangle - \sqrt{D_{xy}}\;|\zeta_y\rangle\right). \tag{3.4.6}$$

Adding equations (3.4.5) and (3.4.6) we get

$$\begin{aligned} 2|U\rangle \;=\; & |u\rangle\left(\sqrt{1-D_{xy}}\left(|\xi_x\rangle + |\xi_y\rangle\right) + \sqrt{D_{xy}}\left(|\zeta_x\rangle + |\zeta_y\rangle\right)\right) + \\ & |v\rangle\left(\sqrt{1-D_{xy}}\left(|\xi_x\rangle - |\xi_y\rangle\right) + \sqrt{D_{xy}}\left(|\zeta_y\rangle - |\zeta_x\rangle\right)\right). \end{aligned}$$

Using the expression of $|U\rangle$ in Eq. 3.3.4.3, we get

$$\begin{aligned} & 2\left(\sqrt{1-D_{uv}}\;|u\rangle|\xi_u\rangle + \sqrt{D_{uv}}\;|v\rangle|\zeta_u\rangle\right) \\ & \qquad = |u\rangle\left(\sqrt{1-D_{xy}}\left(|\xi_x\rangle + |\xi_y\rangle\right) + \sqrt{D_{xy}}\left(|\zeta_x\rangle + |\zeta_y\rangle\right)\right) + \\ & \qquad\qquad |v\rangle\left(\sqrt{1-D_{xy}}\left(|\xi_x\rangle - |\xi_y\rangle\right) + \sqrt{D_{xy}}\left(|\zeta_y\rangle - |\zeta_x\rangle\right)\right). \end{aligned}$$

Comparing co-eff of $|u\rangle, |v\rangle$ (they are $l.i.$) we get the desired expressions. □

Proposition 3.4.10.4.2. *T.P.T.*

$$\begin{aligned}
\mathbf{Re}[\langle\xi_x|\zeta_y\rangle - \langle\zeta_x|\xi_y\rangle] &= 0 \\
\big(1-D_{xy}\big)\,\mathbf{Im}[\langle\xi_y|\xi_x\rangle] + D_{xy}\,\mathbf{Im}[\langle\zeta_x|\zeta_y\rangle] &= 0
\end{aligned} \tag{3.4.7}$$

Proof. Use the orthogonality $\langle\xi_u|\zeta_u\rangle = 0$ with the relevant states from Eq. (3.3.4.4)

$$\begin{aligned}
2\sqrt{1-D_{uv}}\,|\xi_u\rangle &= \sqrt{1-D_{xy}}\,\big(|\xi_x\rangle + |\xi_y\rangle\big) + \sqrt{D_{xy}}\,\big(|\zeta_x\rangle + |\zeta_y\rangle\big) \\
2\sqrt{D_{uv}}\,|\zeta_u\rangle &= \sqrt{1-D_{xy}}\,\big(|\xi_x\rangle - |\xi_y\rangle\big) + \sqrt{D_{xy}}\,\big(|\zeta_y\rangle - |\zeta_x\rangle\big)\,,
\end{aligned}$$

we get

$$\begin{aligned}
0 &= 2\sqrt{1-D_{uv}}\,2\sqrt{D_{uv}}\,\langle\xi_u||\zeta_u\rangle \\
&= \Big[\sqrt{1-D_{xy}}\,\big(\langle\xi_x| + \langle\xi_y|\big) + \sqrt{D_{xy}}\,\big(\langle\zeta_x| + \langle\zeta_y|\big)\Big] \\
&\quad \Big[\sqrt{1-D_{xy}}\,\big(|\xi_x\rangle - |\xi_y\rangle\big) + \sqrt{D_{xy}}\,\big(|\zeta_y\rangle - |\zeta_x\rangle\big)\Big] \\
&= \big(1-D_{xy}\big)\left[\underbrace{\langle\xi_x|\xi_x\rangle}_{=1} - \langle\xi_x|\xi_y\rangle + \langle\xi_y|\xi_x\rangle - \underbrace{\langle\xi_y|\xi_y\rangle}_{=1}\right] \\
&\quad + D_{xy}\left[\langle\zeta_x|\zeta_y\rangle - \underbrace{\langle\zeta_x|\zeta_x\rangle}_{=1} + \underbrace{\langle\zeta_y|\zeta_y\rangle}_{=1} - \langle\zeta_y|\zeta_x\rangle\right] \\
&\quad + \sqrt{D_{xy}\big(1-D_{xy}\big)}\left[\left(\underbrace{\langle\xi_x|\zeta_y\rangle}_{G1} - \underbrace{\langle\xi_x|\zeta_x\rangle}_{=0} + \underbrace{\langle\xi_y|\zeta_y\rangle}_{=0} - \overbrace{\langle\xi_y|\zeta_x\rangle}^{G2}\right)\right. \\
&\quad\quad \left. + \left(\underbrace{\langle\zeta_x|\xi_x\rangle}_{=0} - \overbrace{\langle\zeta_x|\xi_y\rangle}^{G2} + \underbrace{\langle\zeta_y|\xi_x\rangle}_{G1} - \underbrace{\langle\zeta_y|\xi_y\rangle}_{=0}\right)\right] \\
&\Big\downarrow \quad [\textbf{Use: } \langle z_1|z_2\rangle + \langle z_2|z_1\rangle = 2\,\mathbf{Re}\langle z_1|z_2\rangle, \qquad \langle z_1|z_2\rangle - \langle z_2|z_1\rangle = i\cdot 2\,\mathbf{Re}\langle z_1|z_2\rangle] \\
&= i\Big[\big(1-D_{xy}\big)\,2\,\mathbf{Im}\big(\langle\xi_y|\xi_x\rangle\big) + D_{xy}\,2\,\mathbf{Im}\big(\langle\zeta_x|\zeta_y\rangle\big)\Big] \\
&\quad + \sqrt{D_{xy}\big(1-D_{xy}\big)}\;2\,\big[\mathbf{Re}\langle\xi_x|\zeta_y\rangle - \mathbf{Re}\langle\zeta_x|\xi_y\rangle\big]
\end{aligned}$$

Equating real and imag. parts we get the desired result. □

Calculation 3.4.10.4.1. *P.T.* $\rho_x = \mathsf{Tr}_{Alice}\left(|X\rangle\langle X|\right) = (1-D_{xy})|\xi_x\rangle\langle\xi_x| + D_{xy}|\zeta_x\rangle\langle\zeta_x|$

Proof. With

$$|X\rangle \;=\; \sqrt{1-D_{xy}}\;|x\rangle|\xi_x\rangle + \sqrt{D_{xy}}\;|y\rangle|\zeta_x\rangle,$$

$$\begin{aligned}|X\rangle\langle X| \;&=\; (1-D_{xy})\;|x\rangle\langle x|\otimes|\xi_x\rangle\langle\xi_x| + D_{xy}\;|y\rangle\langle y|\otimes|\zeta_x\rangle\langle\zeta_x| \\ &\quad + \sqrt{D_{xy}(1-D_{xy})}\;\left(|x\rangle\langle y|\otimes|\xi_x\rangle\langle\zeta_x| + |y\rangle\langle x|\otimes|\zeta_x\rangle\langle\xi_x|\right)\end{aligned}$$

$$\begin{aligned}\mathsf{Tr}_{\text{Alice}}\left(|X\rangle\langle X|\right) \;&=\; (1-D_{xy})\;\mathsf{Tr}\left(|x\rangle\langle x|\right)|\xi_x\rangle\langle\xi_x| + D_{xy}\;\mathsf{Tr}\left(|y\rangle\langle y|\right)|\zeta_x\rangle\langle\zeta_x| \\ &\quad + \sqrt{D_{xy}(1-D_{xy})}\;\Big(\mathsf{Tr}\left(|x\rangle\langle y|\right)|\xi_x\rangle\langle\zeta_x| + \mathsf{Tr}\left(|y\rangle\langle x|\right)\otimes|\zeta_x\rangle\langle\xi_x|\Big)\end{aligned}$$

Using

$$\begin{aligned}\mathsf{Tr}\left(|x\rangle\langle x|\right) = \langle x|x\rangle = 1 = \langle y|y\rangle = \mathsf{Tr}\left(|y\rangle\langle y|\right), \\ \mathsf{Tr}\left(|x\rangle\langle y|\right) = \langle x|y\rangle = 0 = \langle y|x\rangle = \mathsf{Tr}\left(|y\rangle\langle x|\right),\end{aligned}$$

we get the desired result. □

3.4.10.5 IBtL Sec. 3.4.6.3: Optimal POVM

Proposition 3.4.10.5.1. *We'll show that, for the given IVs, Eve's observable becomes*

$$\Gamma_{xy} \;=\; = (1-D_{xy})\hat{\xi} + D_{xy}\hat{\zeta}$$

With $k := 2\sqrt{D_{uv}(1-D_{uv})}$*,*

$$\begin{aligned}\hat{\xi} \;&= |\xi_x\rangle\langle\xi_x| - |\xi_y\rangle\langle\xi_y| \;&&= k\left[|\phi^+_{xy}\rangle\langle\phi^-_{xy}| + |\phi^-_{xy}\rangle\langle\phi^+_{xy}|\right] \;&&= k\,(E_0 - E_3) \\ \hat{\zeta} \;&= |\zeta_x\rangle\langle\zeta_x| - |\zeta_y\rangle\langle\zeta_y| \;&&= k\left[|\psi^+_{xy}\rangle\langle\psi^-_{xy}| + |\psi^-_{xy}\rangle\langle\psi^+_{xy}|\right] \;&&= -k\,(E_1 - E_2)\end{aligned}$$

Proof. For the first equality, we proceed as follows. With the IVs

$$|\xi_x\rangle = \sqrt{1-D_{uv}}\;|\Phi^+_{xy}\rangle + \sqrt{D_{uv}}\;|\Phi^-_{xy}\rangle, \qquad |\xi_y\rangle = \sqrt{1-D_{uv}}\;|\Phi^+_{xy}\rangle - \sqrt{D_{uv}}\;|\Phi^-_{xy}\rangle,$$

$$\begin{aligned}|\xi_x\rangle\langle\xi_x| \;&=\; (1-D_{uv})\;|\phi^+_{xy}\rangle\langle\phi^+_{xy}| + D_{uv}\;|\phi^-_{xy}\rangle\langle\phi^-_{xy}| \\ &\quad + \sqrt{D_{uv}(1-D_{uv})}\;\left(|\phi^+_{xy}\rangle\langle\phi^-_{xy}| + |\phi^-_{xy}\rangle\langle\phi^+_{xy}|\right), \\ |\xi_y\rangle\langle\xi_y| \;&=\; (1-D_{uv})\;|\phi^+_{xy}\rangle\langle\phi^+_{xy}| + D_{uv}\;|\phi^-_{xy}\rangle\langle\phi^-_{xy}| \\ &\quad - \sqrt{D_{uv}(1-D_{uv})}\;\left(|\phi^+_{xy}\rangle\langle\phi^-_{xy}| + |\phi^-_{xy}\rangle\langle\phi^+_{xy}|\right).\end{aligned}$$

$$|\xi_x\rangle\langle\xi_x| - |\xi_y\rangle\langle\xi_y| \;=\; 2\sqrt{D_{uv}(1-D_{uv})}\,\left(|\phi^+_{xy}\rangle\langle\phi^-_{xy}| + |\phi^-_{xy}\rangle\langle\phi^+_{xy}|\right).$$

Similarly, with

$$|\zeta_x\rangle = \sqrt{1-D_{uv}}\,|\Psi^+_{xy}\rangle - \sqrt{D_{uv}}\,|\Psi^-_{xy}\rangle, \qquad |\zeta_y\rangle = \sqrt{1-D_{uv}}\,|\Psi^+_{xy}\rangle + \sqrt{D_{uv}}\,|\Psi^-_{xy}\rangle,$$

$$\begin{aligned}|\zeta_x\rangle\langle\zeta_x| &= (1-D_{uv})\,|\psi^+_{xy}\rangle\langle\psi^+_{xy}| + D_{uv}\,|\psi^-_{xy}\rangle\langle\psi^-_{xy}| \\ &\quad - \sqrt{D_{uv}(1-D_{uv})}\,\left(|\psi^+_{xy}\rangle\langle\psi^-_{xy}| + |\psi^-_{xy}\rangle\langle\psi^+_{xy}|\right),\end{aligned}$$

$$\begin{aligned}|\zeta_y\rangle\langle\zeta_y| &= (1-D_{uv})\,|\psi^+_{xy}\rangle\langle\psi^+_{xy}| + D_{uv}\,|\psi^-_{xy}\rangle\langle\psi^-_{xy}| \\ &\quad + \sqrt{D_{uv}(1-D_{uv})}\,\left(|\psi^+_{xy}\rangle\langle\psi^-_{xy}| + |\psi^-_{xy}\rangle\langle\psi^+_{xy}|\right).\end{aligned}$$

$$|\zeta_x\rangle\langle\zeta_x| - |\zeta_y\rangle\langle\zeta_y| \;=\; -2\sqrt{D_{uv}(1-D_{uv})}\,\left(|\psi^+_{xy}\rangle\langle\psi^-_{xy}| + |\psi^-_{xy}\rangle\langle\psi^+_{xy}|\right).$$

For the second equality, i.e. to prove

$$|\phi^+_{xy}\rangle\langle\phi^-_{xy}| + |\phi^-_{xy}\rangle\langle\phi^+_{xy}| = E_0 - E_3; \qquad |\psi^+_{xy}\rangle\langle\psi^-_{xy}| + |\psi^-_{xy}\rangle\langle\psi^+_{xy}| = E_1 - E_2,$$

we proceed as follows.

We use

$$\begin{aligned}|\Phi^\pm_{xy}\rangle &:= \frac{1}{\sqrt{2}}\left(|x\rangle|x\rangle \pm |y\rangle|y\rangle\right) = \frac{1}{\sqrt{2}}\left(|E_0\rangle \pm |E_3\rangle\right), \\ Psixy\pm &:= \frac{1}{\sqrt{2}}\left(|x\rangle|y\rangle \pm |y\rangle|x\rangle\right) = \frac{1}{\sqrt{2}}\left(|E_2\rangle \pm |E_1\rangle\right).\end{aligned}$$

Considering the notations $E_{ij} := |E_i\rangle\langle E_j|, E_i := |E_i\rangle\langle E_i|$,

$$\begin{aligned}|\phi^+_{xy}\rangle\langle\phi^-_{xy}| &= \frac{1}{2}\left(E_{00} - E_{33} - E_{03} + E_{30}\right) \\ |\phi^-_{xy}\rangle\langle\phi^+_{xy}| &= \frac{1}{2}\left(E_{00} - E_{33} + E_{03} - E_{30}\right) \\ |\psi^+_{xy}\rangle\langle\psi^-_{xy}| &= \frac{1}{2}\left(E_{11} - E_{22} + E_{12} - E_{21}\right) \\ |\psi^-_{xy}\rangle\langle\psi^+_{xy}| &= \frac{1}{2}\left(E_{11} - E_{22} - E_{12} + E_{21}\right).\end{aligned}$$

□

3.4.10.6 IBtL Sec. 3.4.6.4: Optimality of the interactions

Calculation 3.4.10.6.1. *P.T.* $B_u \otimes E_1\,|U\rangle = \sqrt{1-D_{uv}}\,\langle E_1|\xi_u\rangle|u\rangle|E_1\rangle$

Use: $B_u = |u\rangle\langle u|$, $E_1 = |E_1\rangle\langle E_1|$, $|U\rangle = \sqrt{1-D_{uv}}\ |u\rangle|\xi_u\rangle + \sqrt{D_{uv}}\ |v\rangle|\zeta_u\rangle$.

Proof.

$$\begin{aligned}
B_u \otimes E_1\ |U\rangle &= B_u \otimes E_1\ \left(\sqrt{1-D_{uv}}\ |u\rangle|\xi_u\rangle + \sqrt{D_{uv}}\ |v\rangle|\zeta_u\rangle\right) \\
&= \sqrt{1-D_{uv}}\ \left(B_u|u\rangle\right) \otimes \left(E_1|\xi_u\rangle\right) + \sqrt{D_{uv}}\ \left(B_u|v\rangle\right) \otimes \left(E_1|\zeta_v\rangle\right) \\
&\quad \left[\ \textbf{Use: } B_u|u\rangle = |u\rangle,\ E_1|\xi_u\rangle = |E_1\rangle\langle E_1||\xi_u\rangle; B_u|v\rangle = 0.\ \right] \\
&= \sqrt{1-D_{uv}}\ \langle E_1|\xi_u\rangle|u\rangle|E_1\rangle
\end{aligned}$$

□

Calculation 3.4.10.6.2. *P.T.* $\langle E_1|\xi_u\rangle = \frac{1}{\sqrt{2}}\sqrt{D_{xy}}$

Use: $|\xi_u\rangle = \sqrt{1-D_{xy}}\ \ |\Phi^+_{uv}\rangle + \sqrt{D_{xy}}\ \ |\Phi^-_{uv}\rangle$, $|\Phi^{\pm}_{uv}\rangle = \frac{1}{\sqrt{2}}\left(|u\rangle|u\rangle \pm |v\rangle|v\rangle\right) = \frac{1}{\sqrt{2}}\left(|F_0\rangle \pm |F_3\rangle\right)$.
Then,

$$\langle E_1|F_0\rangle = \frac{1}{2},\ \langle E_1|F_3\rangle = -\frac{1}{2} \quad \Longrightarrow \quad \langle E_1|\Phi^+_{uv}\rangle = 0,\ \langle E_1|\Phi^-_{uv}\rangle = \frac{1}{\sqrt{2}}.$$

3.4.10.7 IBtL: Equal-error

Proposition 3.4.10.1. *P.T.* $$D = \frac{1-\cos\alpha}{2-\cos\alpha+\cos\beta}.$$

Proof. We use the normalization constraint $\langle\zeta_u|\zeta_u\rangle = 1$, with the overlaps

$$\langle\xi_x|\xi_y\rangle = \cos\alpha, \langle\zeta_y|\zeta_x\rangle = \cos\beta; \qquad \langle\xi_x|\zeta_y\rangle = 0 = \langle\xi_y|\zeta_x\rangle; \qquad \langle\xi_i|\xi_i\rangle = 1 = \langle\zeta_j|\zeta_j\rangle.$$

For equal error rates $_{xy} = D_{uv} = D$, consider the following relation

$$2\sqrt{D_{uv}}\ |\zeta_u\rangle \quad = \quad \sqrt{1-D_{xy}}\left(|\xi_x\rangle - |\xi_y\rangle\right) + \sqrt{D_{xy}}\left(|\zeta_y\rangle - |\zeta_x\rangle\right).$$

Then, taking inner product of both sides, we get

$$4D \quad = \quad (1-D)\left[2-2\cos\alpha\right] + D\left[2-2\cos\beta\right] = 2-2\cos\alpha + 2D\left(\cos\alpha - \cos\beta\right).$$

Simplifying, we get the desired relation. □

Proposition 3.4.10.2 (Maximizing the success probability of Eve)**.** *The maximum for Eq. 3.4.9 occur for $\alpha = \beta$, with $\sin\alpha = 2\sqrt{D(1-D)}$.*

Proof. It's enough to maximize the following function.

$$f(\alpha,\beta) \quad := \quad F\cdot\sin\alpha + D\cdot\sin\beta.$$

Note that, $\sin\alpha$ is concave in $[0,\pi/2]$. So,

$$f(\alpha,\beta) \leq \sin\left(F\cdot\alpha + D\cdot\beta\right).$$

Since the equality occurs for $\alpha=\beta$,

$$f^{\star}(\alpha,\beta) := \sin\alpha.$$

In that case, the expression for disturbance becomes

$$D = \frac{1-\cos\alpha}{2},$$

giving rise to

$$\cos\alpha = 1-2D, \qquad \sin\alpha = 2\sqrt{D(1-D)}.$$

□

Proposition 3.4.10.3. *1. For small D,*

$$I_{AE} \leq 2D$$

2. At the other extreme, $I_{AE}^{\max} = \ln 2$

Proof. Consider D small.

$$\begin{aligned}
I_{\max} &= \frac{1}{2}\phi\left[2\sqrt{D(1-D)}\right]\\
\phi(z) &= (1+z)\ln(1+z)+(1-z)\ln(1-z)\\
\ln(1+z) &\approx z-\frac{z^2}{2}, \text{ for small } z\\
\ln(1-z) &\approx -z-\frac{z^2}{2}, \text{ for small } z\\
\phi(z) &\approx (1+z)\left(z-\frac{z^2}{2}\right)+(1-z)\left(-z-\frac{z^2}{2}\right) = 2z^2-z^2=z^2\\
\phi\left[2\sqrt{D(1-D)}\right] &\approx \phi(2\sqrt{D}) \approx 4D\\
I_{\max} &\approx 2D
\end{aligned}$$

At the other extreme, $I_{\mathrm{AE}}^{\max} = \ln 2$ and it occurs when $D=1/2$ at $\alpha = \frac{\pi}{2} = \beta$. In that case, choice of ξ_i, ζ_j becomes

$$|\xi_x\rangle = |x\rangle|x\rangle, \quad |\xi_y\rangle = |y\rangle|x\rangle, \quad |\zeta_x\rangle = |x\rangle|y\rangle, \quad |\zeta_y\rangle = |y\rangle|y\rangle.$$

which is an orthonormal basis, i.e., optimal information gathering measurement is defined by the 4-dim computational basis. □

3.5 Conclusion

We have discussed the nuts-and-bolts of the optimal attack [FGG$^+$97] on the BB84 protocol. It uses a two-qubit ancilla (four dimensional probe) to glean information from the senders signal. The unitary evolution entangles the two systems which is later measured in a suitable four-dimensional POVM. To quantify the amount of information gathered by Eve, two functions are considered: IG, and MI. Both are maximized simultaneously for the same optimal measurement. The attack can leave a further scope for secret key distillation for a QBER up to 14.64%. The secret key-rate is plotted in this window of disturbance.

The attack model is developed for the generalized asymmetric error rates across the two MUBs. A symmetric attack is then considered separately. For the asymmetric model, a necessary and sufficient condition is also given to testify optimality of any interaction. The maximum information is shown achievable by a judiciously chosen interaction, each in the asymmetric case and in the symmetric case.

Interestingly, such an optimal attack on the *p&m* scheme has clear connection with the Bell violation in the equivalent entanglement-based scheme. We have shown explicitly that the optimal states of the joint system can also be obtained by an optimal phase-covariant cloning mechanism, and vice versa.

It is left open an exercise in [FGG$^+$97] to find out the population of candidate interactions (IVs) that can achieve the maximum amount of information. For practical purposes, an eavesdropper requires the optimal unitary to evolve the joint system and the corresponding measurement that she must perform to glean the optimal information. We will discuss these issues in the subsequent chapters.

CHAPTER 4

CHARACTERIZING THE OPTIMAL INTERACTIONS

We have already discussed a general framework of optimal eavesdropping on the 4s protocol owing to Fuchs *et al.* [FGG$^+$97]. An upper bound for mutual information was derived, that was shown to be achievable by a specific kind of interaction-measurement combination.

However, it was left open-ended whether such an interaction is unique or not. We have shown [AP17] that there are infinitely many such interactions which are all derived from the first principle. They are however unique up to some kind of isomorphism that arise due to various rotations of the same measurement setup in the four-dimensional space. The specific choice for the optimal interaction by Fuchs *et al.* is shown to be a particular instantiation of the generalized expression derived in our work.

We also discuss the optimal POVMs. We first consider the second interaction from [FGG$^+$97], which is shown to arise from our general expression due to a specific choice of the measurement basis. Thereby we show by example how to find the optimal POVM from a given interaction. Then, for the sake of illustration, we considered some special instantiations from our generalized expression. We described the associated optimal POVMs as well.

The properties of the Hermitian (observable) providing the optimal measurements are discussed for more insight. We have described the expression of the Hermitian in the measurement basis, found their eigenvalues. We also have connected the sign parameter of the necessary and sufficient condition in [FGG$^+$97] with the sign of these eigenvalues.

Overall, we have developed a mathematical framework to tackle the derivation of the generalized optimal interactions from the scratch.

4.1 A brief overview

Note that, Fuchs *et al.* [FGG$^+$97] had an intelligent guess to come up with the expression for optimal interaction. On the other hand, in the paper [AP17], we explicitly derived a general form of the expression of any possible optimal interactions.

4.1.1 New notations

Following notations, viz., $\mathscr{D}_\beta, \overline{\mathscr{D}}_\beta$ for the encoding bases $\beta \in \{xy, uv\}$, are useful to deal with the interaction vectors.

Notations 4.1. *For any basis $\beta \in \{xy, uv\}$ we define the following notations.*

$$\begin{aligned} \mathscr{D}_\beta &:= \frac{\sqrt{1-D_\beta}+\sqrt{D_\beta}}{\sqrt{2}}, \\ \overline{\mathscr{D}}_\beta &:= \frac{\sqrt{1-D_\beta}-\sqrt{D_\beta}}{\sqrt{2}}. \end{aligned} \tag{4.1.1}$$

The square of the quantities form a probability distribution. The following relations appear to be useful.

$$\begin{aligned} \mathscr{D}_\beta \cdot \overline{\mathscr{D}}_\beta &= \frac{1}{2}\left(1-2D_\beta\right), \\ \mathscr{D}_\beta^2 + \overline{\mathscr{D}}_\beta^2 &= 1, \\ \mathscr{D}_\beta^2 - \overline{\mathscr{D}}_\beta^2 &= 2\sqrt{D_\beta\left(1-D_\beta\right)}. \end{aligned} \tag{4.1.2}$$

4.1.2 Chapter organization

The content of this chapter is organized as follows. Section 4.2 contains the basic results from [FGG$^+$97] required for the derivations that we have done in [AP17]. Our results are then explained in Sec. 4.3. Section 4.4 is devoted to discuss the connection of our results with [FGG$^+$97] and followed by a conclusion.

The main results are briefly described in Sec. 4.3. It includes the derivation of the optimal IVs, the associated optimal POVMs, the measurement observable and its eigenvalues etc. We tabulate the IV-POVM combinations for various instantiations.

4.2 Required ingredients

Let's recollect the tools from [FGG$^+$97] required to derive our results. We need the probabilities: prior, posterior etc. The maximum amount of information (IG and MI), with the IG-lets (IG for each measurement outcome). We require the NSC. We need to rewrite the optimal interactions and measurements using our own notations, one for unequal error rates and yet another for same error rate. Since all these results hold for equal prior, the rest of the sections follow the same assumption unless explicitly mentioned.

The following expressions will be useful in our derivation of the optimal interactions.

An upper bound on the information gain (G) from Eqs. (3.3.2.1, 3.3.2.2).

$$IG^{\star}_{\beta} \;=\; 2\sqrt{D_{\bar{\beta}}\left(1-D_{\bar{\beta}}\right)}.$$

The following bound on the IG-lets from Eq. (3.3.2.3) will be useful.

$$G^{\star}_{xy}(\lambda) \;=\; 2\sqrt{d_{uv}(\lambda)\left[1-d_{uv}(\lambda)\right]} \quad \forall\lambda.$$

The maximum IG upper bounds the mutual information (I), as in Eqs. (3.3.2.4, 3.3.2.5).

$$MI^{\star}_{\beta} \;=\; \tfrac{1}{2}\,\phi\left(IG^{\star}_{\beta}\right),$$

for the concave function

$$\phi(z) \;:=\; (1+z)\ln(1+z)+(1-z)\ln(1-z)\,.$$

Subscripts in the bounds emphasize that the mutual information and the error rates correspond to signals sent in two different bases.

A formal verification of the optimal interactions will require to testify the NSC as in Eqs. (3.3.3.1, 3.3.3.2, 3.3.3.3, 5.2.1).

One of the conditions that must hold to achieve the upper bounds in xy basis is the following [FGG^{+}97, Eq. (33)]:

$$d_{\lambda u}=d_{\lambda v}=d_{uv}(\lambda)=D_{uv},\;\; \forall\lambda.$$

An analogous condition holds good for signals sent in uv basis.

Then, the post-interaction state of the joint ancilla-signal system is as described in Eqs. (3.3.4.1, 3.3.4.3) for the two encoding bases.

Eve uses a 2-qubit (4 states) ancilla. Thus, she has four IVs for each of the encoding bases. For asymmetric error rates, Eq. (3.3.7.3) describes the four IVs in the xy basis. Similarly, for equal rates, Eq. (3.4.8) serves the purpose. We rewrite these IVs in terms of our notations $\mathscr{D}_{\beta},\overline{\mathscr{D}}_{\beta}$ for the encoding bases $\beta \in \{xy,uv\}$.

For unequal error rates (i.e., $D_{xy}\neq D_{uv}$), the following vectors work. Consider a canonical basis for Eve's probe as $\{|\mathcal{E}_0\rangle,|\mathcal{E}_1\rangle,|\mathcal{E}_2\rangle,|\mathcal{E}_3\rangle\}$. Without loss of generality (w.l.o.g.),

$$|\mathcal{E}_0\rangle=|x\rangle|x\rangle,|\mathcal{E}_1\rangle=|y\rangle|x\rangle,|\mathcal{E}_2\rangle=|x\rangle|y\rangle,|\mathcal{E}_3\rangle=|y\rangle|y\rangle. \qquad (4.2.1)$$

To describe the interaction vectors $|\xi_i\rangle, |\zeta_j\rangle$, we introduce the new notations $\mathscr{D}_{uv}, \overline{\mathscr{D}}_{uv}$ to rewrite Eq. (3.3.7.3) as below.

$$\begin{aligned}
|\xi_x\rangle &= \mathscr{D}_{uv}\,|\mathcal{E}_0\rangle + \overline{\mathscr{D}}_{uv}\,|\mathcal{E}_3\rangle, \\
|\xi_y\rangle &= \overline{\mathscr{D}}_{uv}\,|\mathcal{E}_0\rangle + \mathscr{D}_{uv}\,|\mathcal{E}_3\rangle, \\
|\zeta_x\rangle &= \overline{\mathscr{D}}_{uv}\,|\mathcal{E}_2\rangle + \mathscr{D}_{uv}\,|\mathcal{E}_1\rangle, \\
|\zeta_y\rangle &= \mathscr{D}_{uv}\,|\mathcal{E}_2\rangle + \overline{\mathscr{D}}_{uv}\,|\mathcal{E}_1\rangle, \qquad (4.2.2)
\end{aligned}$$

The corresponding optimal POVM, as shown in [FGG$^+$97, Eqs. (55,56)], is described below.

$$E_\lambda = |E_\lambda\rangle\langle E_\lambda|,$$

where

$$|E_0\rangle = |\mathcal{E}_0\rangle,\ |E_1\rangle = |\mathcal{E}_1\rangle,\ |E_2\rangle = |\mathcal{E}_2\rangle,\ |E_3\rangle = |\mathcal{E}_3\rangle. \qquad (4.2.3)$$

Similar expression for the optimal IVs hold for uv basis as well.

For equal error rates ($D_{xy} = D_{uv} = D$), another set of optimal IVs was described in [FGG$^+$97] as described below.

$$\begin{aligned}
|\xi_x\rangle &= |\mathcal{E}_0\rangle, \\
|\xi_y\rangle &= 2\mathscr{D}\overline{\mathscr{D}}\,|\mathcal{E}_0\rangle + \left(\mathscr{D}^2 - \overline{\mathscr{D}}^2\right)|\mathcal{E}_1\rangle, \\
|\zeta_x\rangle &= |\mathcal{E}_2\rangle, \\
|\zeta_y\rangle &= 2\mathscr{D}\overline{\mathscr{D}}\,|\mathcal{E}_2\rangle + \left(\mathscr{D}^2 - \overline{\mathscr{D}}^2\right)|\mathcal{E}_3\rangle. \qquad (4.2.4)
\end{aligned}$$

However, the corresponding optimal POVM was not shown explicitly in [FGG$^+$97], which we establish in Sec. 4.3.1.4.

Given an interaction, to identify an optimal POVM is already known, as discussed earlier. For an optimal interaction, the density operators ρ_x, ρ_y with Eve are such that a set of the eigenprojectors $\{E_\lambda\}$ onto the orthonormal eigenbasis $\{|E_\lambda\rangle\}$ of the Hermitian operator $\Gamma_{xy} := \frac{1}{2}\left(\rho_x - \rho_y\right)$ becomes the optimal measurement. It optimizes both IG and MI across the two MUBs.

Although, both interactions (4.2.2, 4.2.4) lead to optimality, the way they were proposed in [FGG$^+$97] seems to be a judicious guesswork. This leaves open a few interesting questions:

1. Rather than guessing an interaction and verifying its optimality, can we derive it from the first principle?

2. Are there alternate optimal interactions than the two specific ones?

3. If so, is it possible to characterize all the possible interactions?

We answer these questions in the following section.

4.3 Our Results [AP17]

Here we derive a general expression for an interaction by Eve that leads to optimal information gain. Eventually, we show that the expression is unique in the measurement basis. Associated optimal POVMs follow automatically.

4.3.1 Optimal interaction to maximize information gain (G): A generic form of optimal $|\xi_i\rangle, |\zeta_j\rangle$

4.3.1.1 The basic ingredients

We use the following result with equal priors to find an expression of $|\xi_i\rangle, |\zeta_j\rangle$ for optimal interaction.

Lemma 4.1. *Optimality conditions for G_{xy} ensure that each $G^{\star}_{xy}(\lambda)$ is equal to $G^{\star}_{xy}$ and the corresponding optimal value is given by*

$$G^{\star}_{xy} = 2\sqrt{D_{uv}(1-D_{uv})} = G^{\star}_{xy}(\lambda), \qquad \forall\lambda. \tag{4.3.1}$$

Proof. For signal sent in xy basis, the optimal information gain, by Eq. (3.3.2.1), is

$$G^{\star}_{xy} \;=\; 2\sqrt{D_{uv}(1-D_{uv})}.$$

By Eq. (3.3.2.3), for measurement outcome λ of Eve,

$$G^{\star}_{xy}(\lambda) \;=\; 2\sqrt{d_{uv}(\lambda)\left[1-d_{uv}(\lambda)\right]}$$

In order to satisfy optimality, the necessary and sufficient conditions in Proposition 3.3 must be satisfied. By [FGG$^+$97, Eq. (33)], this requires

$$d_{uv}(\lambda) = D_{uv},\ \forall\lambda$$

which ensures that the lemma is proved. □

Note 4.1. *As we considered equal prior probabilities, we use the following working formula of $G_{xy}(\lambda)$ while we derive the general form of an optimal interaction,*

$$G_{xy}(\lambda) \;=\; \left|Q_{x\lambda}-Q_{y\lambda}\right| = \frac{\left|P_{\lambda x}-P_{\lambda y}\right|}{P_{\lambda x}+P_{\lambda y}}. \tag{4.3.2}$$

Here we describe an expression of $P_{\lambda x}, P_{\lambda y}$ in terms of $|\xi_i\rangle, |\zeta_j\rangle$ and a POVM $\{E_\lambda\}$.

Theorem 4.1. *Given the postinteraction joint sates* (3.3.4.1)*, and a POVM* $\{E_\lambda\}_{\lambda\in\{0,1,2,3\}}$,

$$\begin{aligned} P_{\lambda x} &= (1-D_{xy})\langle\xi_x|E_\lambda\rangle^2 + D_{xy}\langle\zeta_x|E_\lambda\rangle^2, \\ P_{\lambda y} &= (1-D_{xy})\langle\xi_y|E_\lambda\rangle^2 + D_{xy}\langle\zeta_y|E_\lambda\rangle^2. \end{aligned} \tag{4.3.3}$$

Proof. Using Eq. (3.3.4.1) in Eq. (3.4.6.1), we get,

$$\rho_x = \mathsf{Tr}_{\mathsf{A}}\left(|X\rangle\langle X|\right) = (1-D_{xy})\widehat{\xi}_x + D_{xy}\widehat{\zeta}_x, \tag{4.3.4}$$

where

$$\widehat{\xi}_x := |\xi_x\rangle\langle\xi_x|, \qquad \widehat{\zeta}_x := |\zeta_x\rangle\langle\zeta_x|.$$

By Eq. (3.6),

$$\begin{aligned} P_{\lambda x} &= \mathsf{Tr}\left(\rho_x E_\lambda\right) \\ &= (1-D_{xy})\mathsf{Tr}\left(\widehat{\xi}_x E_\lambda\right) + D_{xy}\mathsf{Tr}\left(\widehat{\zeta}_x E_\lambda\right) \\ &= (1-D_{xy})\langle\xi_x|E_\lambda\rangle^2 + D_{xy}\langle\zeta_x|E_\lambda\rangle^2. \end{aligned}$$

Similarly, we can derive an expression for $P_{\lambda y}$. □

We now have all the required ingredients in place to derive the optimal interactions.

4.3.1.2 The main result

First we understand the difficulty to performing the derivation if we consider the interaction vectors in terms of the canonical basis $\{|\mathcal{E}_\lambda\rangle\}$ states. We notice that the expressions (4.3.3) of $P_{\lambda x}, P_{\lambda y}$ dependent on the measurement directions $|E_\lambda\rangle$. Thus, if we can express the IVs $|\xi_i\rangle, |\zeta_j\rangle$ in terms of the measurement directions $|E_\lambda\rangle$, then we can compute the probabilities $P_{\lambda x}, P_{\lambda y}$, and thereby compute the IG. This is the main difficulty that we resolve here.

With this understanding about the way of describing the interaction vectors, we start with a general form (4.3.6) of the IVs $|\xi_i\rangle, |\zeta_j\rangle$ expressed in the associated orthonormal measurement basis $\{|E_\lambda\rangle\}$, while abiding by the orthogonality restriction (3.3.4.2). Subsequently, we plug-in the expression (4.3.6) of the interaction vectors into Eq. (4.3.3) to get the probabilities $P_{\lambda x}, P_{\lambda y}$. Then we substitute these probabilities into Eq. (4.3.2) to get the values of $G_{xy}(\lambda)$. Finally, we compare these values with their optimal counterparts in Eq. (4.3.1), and derive the general form of an optimal interaction $|\xi_i\rangle, |\zeta_j\rangle$ expressed in the eigenbasis $\{|E_\lambda\rangle\}$.

This way of expressing interaction vectors (in measurement basis than in the computational basis) not only helps us deriving the optimal interactions, but, as we will realize shortly, all the optimal interactions eventually lead to a unique expression.

Theorem 4.2. *Let* $\{|E_\lambda\rangle\}$ *be an orthonormal eigenbasis of the observable* Γ_{xy} *pertaining to the arbitrary choice of the interaction vectors* $|\xi_i\rangle, |\zeta_j\rangle$ *in Eq.* (3.3.4.1) *of the postinteraction states while abiding by the orthogonality restriction* (3.3.4.2). *Then, for an interaction done optimally, the general form of the IVs* $|\xi_i\rangle, |\zeta_j\rangle$ *described in that measurement basis becomes*

$$\begin{aligned}
|\xi_x\rangle &= \mathscr{D}_{uv}\,|E_0\rangle + \overline{\mathscr{D}}_{uv}\,|E_1\rangle, \\
|\xi_y\rangle &= \overline{\mathscr{D}}_{uv}\,|E_0\rangle + \mathscr{D}_{uv}\,|E_1\rangle, \\
|\zeta_x\rangle &= \mathscr{D}_{uv}\,|E_2\rangle + \overline{\mathscr{D}}_{uv}\,|E_3\rangle, \\
|\zeta_y\rangle &= \overline{\mathscr{D}}_{uv}\,|E_2\rangle + \mathscr{D}_{uv}\,|E_3\rangle, \qquad (4.3.5)
\end{aligned}$$

where $\mathscr{D}_{uv}, \overline{\mathscr{D}}_{uv}$ *are as defined in Eq.* (5.1.1).

Proof. First we need to fix an orthonormal basis to describe $|\xi_i\rangle, |\zeta_j\rangle$ following restriction (3.3.4.2). For that purpose, there is no harm to choose the above eigenbasis to describe $|\xi_i\rangle, |\zeta_j\rangle$. Orthogonality restriction (3.3.4.2) is automatically satisfied if we choose $|\xi_i\rangle \in \mathsf{span}\{|E_0\rangle, |E_1\rangle\}$ and $|\zeta_j\rangle \in \mathsf{span}\{|E_2\rangle, |E_3\rangle\}$ [1]. So the general form of $|\xi_i\rangle, |\zeta_j\rangle$ becomes

$$\begin{aligned}
|\xi_x\rangle &= \sqrt{\alpha}\,|E_0\rangle + \sqrt{1-\alpha}\,|E_1\rangle, \\
|\xi_y\rangle &= \sqrt{\beta}\,|E_0\rangle + \sqrt{1-\beta}\,|E_1\rangle, \\
|\zeta_x\rangle &= \sqrt{\mu}\,|E_2\rangle + \sqrt{1-\mu}\,|E_3\rangle, \\
|\zeta_y\rangle &= \sqrt{\nu}\,|E_2\rangle + \sqrt{1-\nu}\,|E_3\rangle. \qquad (4.3.6)
\end{aligned}$$

Using this form of $|\xi_i\rangle, |\zeta_j\rangle$ in Eq. (4.3.3), we find values of $G_{xy}(\lambda)$ as shown in Table 4.1.

By Lemma 4.1 , for optimal G_{xy}, the values of $G_{xy}(\lambda)$ are all equal. Equating $G_{xy}(0), G_{xy}(1)$ in Table 4.1, we get,

$$\alpha + \beta = 1, \qquad G_{xy}(0) = G_{xy}(1) = |2\alpha - 1|.$$

Similarly, equating $G_{xy}(2), G_{xy}(3)$ in Table 4.1, we get,

$$\mu + \nu = 1, \qquad G_{xy}(2) = G_{xy}(3) = \left|2\mu - 1\right|.$$

[1] This is a choice. There are other orthogonality choices indeed. However, it doesn't matter, as we establish in the next chapter.

Table 4.1 | Ingredients to derive optimal interactions.

For the general form of the IVs $|\xi_i\rangle, |\zeta_j\rangle$ as in Eq. (4.3.6), tabulated the values of the probabilities $P_{\lambda x}, P_{\lambda y}$, and the IG-lets $G_{xy}(\lambda)$ for various measurement outcomes λ. Here, $F := 1 - D$.

λ	$P_{\lambda x}$	$P_{\lambda y}$	$G_{xy}(\lambda) = \frac{\lvert P_{\lambda x} - P_{\lambda y} \rvert}{P_{\lambda x} + P_{\lambda y}}$
0	$F_{xy}\langle\xi_x\|E_0\rangle^2 = F_{xy}\alpha$	$F_{xy}\langle\xi_y\|E_0\rangle^2 = F_{xy}\beta$	$\lvert\alpha-\beta\rvert/(\alpha+\beta)$
1	$F_{xy}\langle\xi_x\|E_1\rangle^2 = F_{xy}(1-\alpha)$	$F_{xy}\langle\xi_y\|E_1\rangle^2 = F_{xy}(1-\beta)$	$\lvert\beta-\alpha\rvert/(1-\alpha+1-\beta)$
2	$D_{xy}\langle\zeta_x\|E_2\rangle^2 = D_{xy}\mu$	$D_{xy}\langle\zeta_y\|E_2\rangle^2 = D_{xy}\nu$	$\lvert\mu-\nu\rvert/(\mu+\nu)$
3	$D_{xy}\langle\zeta_x\|E_3\rangle^2 = D_{xy}(1-\mu)$	$D_{xy}\langle\zeta_y\|E_3\rangle^2 = D_{xy}(1-\nu)$	$\lvert\nu-\mu\rvert/(1-\mu+1-\nu)$

Together, equating $G_{xy}(0), G_{xy}(2)$, we get,

$$\mu = \alpha, \qquad \nu = \beta = 1 - \alpha. \tag{4.3.7}$$

Thus,

$$G^{\star}_{xy}(0) = \mathscr{D}^2_{uv} - \overline{\mathscr{D}}^2_{uv} = 2\mathscr{D}^2_{uv} - 1 = |2\alpha - 1|$$

gives rise to

$$\sqrt{\alpha} = \mathscr{D}_{uv}, \qquad \sqrt{1-\alpha} = \overline{\mathscr{D}}_{uv}. \tag{4.3.8}$$

Using Eqs. (4.3.8, 4.3.7) in Eq. (4.3.6), we get a generic form for optimal $|\xi_i\rangle, |\zeta_j\rangle$ as in Eq. (4.3.5). □

Analogous to Eq. (4.3.5), a set of optimal interaction vectors exist in the uv basis as following.

$$\begin{aligned} |\xi_u\rangle &= \mathscr{D}_{xy}\,|F_0\rangle + \overline{\mathscr{D}}_{xy}\,|F_1\rangle, \\ |\xi_v\rangle &= \overline{\mathscr{D}}_{xy}\,|F_0\rangle + \mathscr{D}_{xy}\,|F_1\rangle, \\ |\zeta_u\rangle &= \mathscr{D}_{xy}\,|F_2\rangle + \overline{\mathscr{D}}_{xy}\,|F_3\rangle, \\ |\zeta_v\rangle &= \overline{\mathscr{D}}_{xy}\,|F_2\rangle + \mathscr{D}_{xy}\,|F_3\rangle. \end{aligned} \tag{4.3.9}$$

The most interesting aspect with the expression (4.3.5) of the optimal IVs is that it has a unique form capturing all the optimal interactions while realized in the measurement basis (orthonormal eigenbasis of the associated Γ_{xy}).

Since any rotation (in general, unitary transformation) **U** of the canonical basis $\{|\mathcal{E}_\lambda\rangle\}$ produces an orthonormal basis $\{|E_\lambda\rangle\}$, it can be considered as a measurement basis.

Remark 4.1. *An optimal interaction for equal error rates could be described by an expression analogous to Eq.* (4.3.5) *while* $\mathscr{D}_{uv}, \overline{\mathscr{D}}_{uv}$ *are replaced by* $\mathscr{D}, \overline{\mathscr{D}}$ *respectively.*

4.3.1.3 The optimal observable

The general expression of the observable Γ_{xy} corresponding to the interactions (4.3.5) along with the PIJSs (3.3.4.1) can easily be found.

Theorem 4.3. *For an optimal interaction (4.3.5, 3.3.4.1), and its optimal POVM* $\{E_\lambda\}$*,*

$$\Gamma_{xy} \;=\; \frac{1}{2}(\mathscr{D}_{uv}^2 - \overline{\mathscr{D}}_{uv}^2)\,[(1-D_{xy})(E_0-E_1)+D_{xy}(E_2-E_3)] \tag{4.3.10}$$

Proof. By Eq. (4.3.4) and its analogue for signal y,

$$2\Gamma_{xy} \;=\; \rho_x - \rho_y = (1-D_{xy})\left(\widehat{\xi}_x - \widehat{\xi}_y\right) + D_{xy}\left(\widehat{\zeta}_x - \widehat{\zeta}_y\right).$$

Using expressions of $|\xi_i\rangle, |\zeta_j\rangle$ in Eq. (4.3.5), and denoting $\mathbb{E}_{ij} := |E_i\rangle\langle E_j|$, we get,

$$\begin{aligned}
\widehat{\xi}_x &= \mathscr{D}_{uv}^2\,\mathbb{E}_{00} + \overline{\mathscr{D}}_{uv}^2\,\mathbb{E}_{11} + 2\mathscr{D}_{uv}\overline{\mathscr{D}}_{uv}\left(\mathbb{E}_{01}+\mathbb{E}_{10}\right),\\
\widehat{\xi}_y &= \overline{\mathscr{D}}_{uv}^2\,\mathbb{E}_{00} + \mathscr{D}_{uv}^2\,\mathbb{E}_{11} + 2\mathscr{D}_{uv}\overline{\mathscr{D}}_{uv}\left(\mathbb{E}_{01}+\mathbb{E}_{10}\right),\\
\widehat{\zeta}_x &= \mathscr{D}_{uv}^2\,\mathbb{E}_{22} + \overline{\mathscr{D}}_{uv}^2\,\mathbb{E}_{33} + 2\mathscr{D}_{uv}\overline{\mathscr{D}}_{uv}\left(\mathbb{E}_{23}+\mathbb{E}_{32}\right),\\
\widehat{\zeta}_y &= \overline{\mathscr{D}}_{uv}^2\,\mathbb{E}_{22} + \mathscr{D}_{uv}^2\,\mathbb{E}_{33} + 2\mathscr{D}_{uv}\overline{\mathscr{D}}_{uv}\left(\mathbb{E}_{23}+\mathbb{E}_{32}\right).
\end{aligned}$$

which leads to the desired form of Γ_{xy}. □

Remark 4.2. *The eigenvalues of the observable* Γ_{xy} *in* (4.3.10) *are*

$$\begin{aligned}
\gamma_0 = \frac{1}{2}\left(\mathscr{D}_{uv}^2 - \overline{\mathscr{D}}_{uv}^2\right)\left(1-D_{xy}\right), &\qquad \gamma_1 = -\gamma_0,\\
\gamma_2 = \frac{1}{2}\left(\mathscr{D}_{uv}^2 - \overline{\mathscr{D}}_{uv}^2\right)D_{xy}, &\qquad \gamma_3 = -\gamma_2.
\end{aligned} \tag{4.3.11}$$

Note 4.2. *It is interesting to note here that, for the IVs in Eq.* (4.3.5)*, the optimal value* $\mathscr{D}_{uv}^2 - \overline{\mathscr{D}}_{uv}^2$ *of* G_{xy} *in Eq.* (3.3.2.1) *agrees with the upper bound* $\sum_\lambda |\gamma_\lambda|$ *of* G_{xy} *in Lemma 3.3.*

4.3.1.4 Optimal POVM for the optimal interaction of [FGG+97, equal error rates]

In [FGG+97], for equal error rates (i.e., $D_{xy} = D_{uv} = D$), only the optimal states [as described in Eq. (4.2.4)] were mentioned but not the associated optimal POVM. We show here how to find it by simply comparing these states with our general expression (4.3.5).

Theorem 4.4. *For the optimal interaction* (4.2.4) *expressed in the computational basis* (4.2.1)*, the optimal POVM* $\{E_\lambda\}$ *corresponds to the eigenprojectors* $E_\lambda = |E_\lambda\rangle\langle E_\lambda|$ *defined in the following eigenbasis.*

$$\begin{aligned} |E_0\rangle &= \mathscr{D}|\mathcal{E}_0\rangle - \overline{\mathscr{D}}|\mathcal{E}_1\rangle, \qquad & |E_1\rangle &= \overline{\mathscr{D}}|\mathcal{E}_0\rangle + \mathscr{D}|\mathcal{E}_1\rangle, \\ |E_2\rangle &= \mathscr{D}|\mathcal{E}_2\rangle - \overline{\mathscr{D}}|\mathcal{E}_3\rangle, \qquad & |E_3\rangle &= \overline{\mathscr{D}}|\mathcal{E}_2\rangle + \mathscr{D}|\mathcal{E}_3\rangle. \end{aligned} \tag{4.3.12}$$

Proof. Comparing a special form of $|\xi_x\rangle, |\xi_y\rangle$ given by Eq. (4.2.4) and the general form of $|\xi_x\rangle, |\xi_y\rangle$ described in Eq. (4.3.5) but for equal error rates, we get

$$\begin{aligned} \mathscr{D}\,|E_0\rangle + \overline{\mathscr{D}}\,|E_1\rangle &= |\mathcal{E}_0\rangle, \\ \overline{\mathscr{D}}\,|E_0\rangle + \mathscr{D}\,|E_1\rangle &= 2\mathscr{D}\overline{\mathscr{D}}\,|\mathcal{E}_0\rangle + \left(\mathscr{D}^2 - \overline{\mathscr{D}}^2\right)|\mathcal{E}_1\rangle. \end{aligned}$$

Solving for $|E_0\rangle$ and $|E_1\rangle$, one can arrive at the first two expressions of Eq. (4.3.12). The remaining two expressions of Eq. (4.3.12) can be derived by comparing the expressions of the IVs $|\zeta_x\rangle, |\zeta_y\rangle$ in Eqs. (4.2.4, 4.3.5). □

4.3.1.5 Interrelation between optimal POVMs

We have deduced infinitely many optimal interactions in the *xy* basis. Similarly, there are infinitely many optimal interactions in the *uv* basis as well. There is an one-to-one correspondence between them as the associated POVMs are interrelated. We provide an interrelation between the optimal POVMs across the two MUBs in the next chapter. Here, we provide a differently posed interrelation that serves the purpose of verifying the NSCs.

4.3.2 Verifying optimality of general interactions

The optimal interactions (4.3.5) in *xy* basis along with their counterpart in *uv* basis, should lead to optimal information gain by virtue of our construction of optimal interactions. Thereby, they saturate the upper bounds of IVs in Eqs. (3.3.2.1, 3.3.2.2), as well MI-bounds in Eqs. (3.3.2.4, 3.3.2.5).

Thus, if the derivation is correct, the derived IVs should satisfy the necessary and sufficient conditions for optimality given by Proposition 3.3. However, the verification process demands the knowledge on the overlaps between the IVs in a basis with the POVMs in the conjugate basis. To find the overlaps, say, $\langle E_\lambda|\xi_u\rangle$, we require to rewrite $|\xi_u\rangle$ in terms of the eigenbasis $|E_\lambda\rangle$ than in terms of $|F_\lambda\rangle$. We do it fist before checking the NSCs.[2]

[2] One can also find the interrelation between the measurement directions across the two MUBs to accomplish the job, that we do later.

4.3.2.1 Expressing the IVs of one basis w.r.t. the measurement direction in the conjugate basis

Remark 4.3. *We can rewrite the optimal IVs in Eq.* (4.3.5) *w.r.t. another orthonormal basis* $\{|\tilde{E_\lambda}\rangle\}$ *as follows.*

$$\begin{aligned}
|\xi_x\rangle &= \sqrt{1-D_{uv}}\,|\tilde{E_0}\rangle + \sqrt{D_{uv}}\,|\tilde{E_1}\rangle,\\
|\xi_y\rangle &= \sqrt{1-D_{uv}}\,|\tilde{E_0}\rangle - \sqrt{D_{uv}}\,|\tilde{E_1}\rangle,\\
|\zeta_x\rangle &= \sqrt{1-D_{uv}}\,|\tilde{E_2}\rangle + \sqrt{D_{uv}}\,|\tilde{E_3}\rangle,\\
|\zeta_y\rangle &= \sqrt{1-D_{uv}}\,|\tilde{E_2}\rangle - \sqrt{D_{uv}}\,|\tilde{E_3}\rangle,
\end{aligned} \tag{4.3.1}$$

Here, $\{|\tilde{E_\lambda}\rangle\}$ *is nothing but the Bell basis over the eigenspace spanned by* $\{E_\lambda\}$.

$$\begin{aligned}
|\tilde{E_0}\rangle &= \frac{1}{\sqrt{2}}\left(|E_0\rangle + |E_1\rangle\right), \qquad |\tilde{E_1}\rangle = \frac{1}{\sqrt{2}}\left(|E_0\rangle - |E_1\rangle\right),\\
|\tilde{E_2}\rangle &= \frac{1}{\sqrt{2}}\left(|E_2\rangle + |E_3\rangle\right), \qquad |\tilde{E_3}\rangle = \frac{1}{\sqrt{2}}\left(|E_2\rangle - |E_3\rangle\right),
\end{aligned} \tag{4.3.2}$$

Clearly, these IVs are analogous to [FGG$^+$97, Eqs. (51) and (50)].

We can plug-in the IVs of Eqs. (4.3.1, 4.3.2) into the interrelation between the IVs across the MUBs as in Eqs. (3.3.4.4, 3.3.4.5), and derive the IVs of *uv* basis expressed in $\{|\tilde{E_\lambda}\rangle\}$ basis as follows.

Lemma 4.2. *For achieving the maximum information gain, we must have*

$$\begin{aligned}
|\xi_u\rangle &= \sqrt{1-D_{xy}}\,|\tilde{E_0}\rangle + \sqrt{D_{xy}}\,|\tilde{E_2}\rangle\\
|\xi_v\rangle &= \sqrt{1-D_{xy}}\,|\tilde{E_0}\rangle - \sqrt{D_{xy}}\,|\tilde{E_2}\rangle\\
|\zeta_u\rangle &= \sqrt{1-D_{xy}}\,|\tilde{E_1}\rangle - \sqrt{D_{xy}}\,|\tilde{E_3}\rangle\\
|\zeta_v\rangle &= \sqrt{1-D_{xy}}\,|\tilde{E_1}\rangle + \sqrt{D_{xy}}\,|\tilde{E_3}\rangle
\end{aligned} \tag{4.3.3}$$

where the basis $\{|\tilde{E_\lambda}\rangle\}$ *is as described in Eq.* (4.3.2).

Remark 4.4. *To get expressions of the IVs* $|\xi_i\rangle, |\zeta_j\rangle$ *in uv basis symmetric to those in xy basis, e.g., like [FGG$^+$97, Eq. (52)], one must consider the canonical basis states in the order* $|\mathcal{E}_0\rangle = |x\rangle|x\rangle, |\mathcal{E}_1\rangle = |y\rangle|y\rangle, |\mathcal{E}_2\rangle = |x\rangle|y\rangle, |\mathcal{E}_3\rangle = |y\rangle|x\rangle$, *compatible with [FGG$^+$97].*

4.3.2.2 Verifying the NSCs

Theorem 4.5. *The interaction given by Eqs. (4.3.5, 3.3.4.1) and a POVM corresponding to the eigenbasis given by Eq.* (3.4.6.4) *satisfy the necessary and sufficient conditions stated in Proposition 3.3 and therefore attain both the optimal information gain and optimal mutual information.*

Proof. From Eq. (5.2.1), we have

$$\begin{aligned}|U_{\lambda u}\rangle &= B_u \otimes \sqrt{E_\lambda}\ |U\rangle = B_u \otimes E_\lambda\ |U\rangle \\ &= \sqrt{1-D_{uv}}\ \big(B_u|u\rangle\big) \otimes \big(E_\lambda|\xi_u\rangle\big) \\ &\quad +\sqrt{D_{uv}}\ \big(B_u|v\rangle\big) \otimes \big(E_\lambda|\zeta_v\rangle\big)\text{ , by Eq. (3.3.4.3).}\end{aligned}$$

Since $B_u|u\rangle = |u\rangle, B_u|v\rangle = \mathbf{0}$, and $E_\lambda|\xi_u\rangle = \langle E_\lambda|\xi_u\rangle|E_\lambda\rangle$, we get,

$$|U_{\lambda u}\rangle = \sqrt{1-D_{uv}}\ \langle E_\lambda|\xi_u\rangle|u\rangle|E_\lambda\rangle.$$

Similarly,

$$|V_{\lambda u}\rangle = \sqrt{D_{uv}}\ \langle E_\lambda|\zeta_v\rangle|u\rangle|E_\lambda\rangle.$$

Here, we want equality in magnitude between $\langle E_\lambda|\xi_u\rangle$ and $\langle E_\lambda|\zeta_v\rangle$. Now, by Eq. (4.3.3), $\langle E_\lambda|\xi_u\rangle$ takes values

$$\frac{1}{\sqrt{2}}\sqrt{1-D_{xy}}, \frac{1}{\sqrt{2}}\sqrt{1-D_{xy}}, \frac{1}{\sqrt{2}}\sqrt{D_{xy}}, \frac{1}{\sqrt{2}}\sqrt{D_{xy}}\ ;$$

whereas, $\langle E_\lambda|\zeta_v\rangle$ takes values

$$\frac{1}{\sqrt{2}}\sqrt{1-D_{xy}}, -\frac{1}{\sqrt{2}}\sqrt{1-D_{xy}}, \frac{1}{\sqrt{2}}\sqrt{D_{xy}}, -\frac{1}{\sqrt{2}}\sqrt{D_{xy}}\ ,$$

respectively for $\lambda = 0,1,2,3$. Therefore,

$$|V_{\lambda u}\rangle = \varepsilon_\lambda \sqrt{\frac{D_{uv}}{1-D_{uv}}}\ |U_{\lambda u}\rangle,$$

where,

$$\varepsilon_0 = +1, \quad \varepsilon_1 = -1, \quad \varepsilon_2 = +1, \quad \varepsilon_3 = -1. \tag{4.3.4}$$

Similarly, one may calculate to verify that

$$|U_{\lambda v}\rangle = \varepsilon_\lambda \sqrt{\frac{D_{uv}}{1-D_{uv}}}\ |V_{\lambda v}\rangle,$$

for the same combination of ε_λ as in Eq. (4.3.4). This completes the proof of the theorem. □

4.3.2.3 Yet another signature of optimality

Further, we take the opportunity to establish a direct relation between the sign parameter ε_λ as in the NSCs and the signs of eigenvalues γ_λ. It's yet another indicator for optimality.

Lemma 4.3. *For optimal G,*

$$\varepsilon_\lambda = \mathsf{sgn}\ \gamma_\lambda. \tag{4.3.5}$$

Proof. For optimal G, Γ is a diagonal matrix with diagonal entries γ_λ. Thus, for signals sent in xy basis,

$$\gamma_\lambda = \mathsf{tr}(\Gamma_{xy}E_\lambda) = \frac{1}{2}[\mathsf{tr}(\rho_x E_\lambda) - \mathsf{tr}(\rho_y E_\lambda)] = \frac{1}{2}(P_{\lambda x} - P_{\lambda y}).$$

By Eq. (3.3.3.3),

$$\varepsilon_\lambda \ = \ \mathsf{sgn}\left(Q_{x\lambda} - Q_{y\lambda}\right) = \mathsf{sgn}\left(P_{\lambda x} - P_{\lambda y}\right) = \mathsf{sgn}\ \gamma_\lambda,$$

which establishes the relation. □

Remark 4.5. *By Lemma 4.3, another cross-checking for optimality is that Eq.* (4.3.4) *should match with the signs of the eigenvalues γ_λ of Γ_{xy} as in Eq.* (4.3.11).

An interesting observation

- Given an interaction by Eve, if it is optimal, her IVs should carry the quality of optimality themselves, irrespective of whether the optimal POVM is known or not.
- The optimal states in [AP17] exhibit an interesting property: the overlap between the two ξ states (undisturbed counterpart) are same as the overlap between the two ζ states (disturbed counterpart) and are equal to $2\mathscr{D}\overline{\mathscr{D}}$. We show in the next chapter that this is precisely the criteria for an arbitrary set of IVs to be optimal.

4.4 Generating some specific optimal IVs: Connecting [FGG$^+$97]

Here we show that the instances of an optimal interaction presented in [FGG$^+$97] is a particular instance of the generalized unique expression of the optimal interactions that we have derived. Moreover, for better understanding, we generate a new instance (different from the two instances of [FGG$^+$97]) of the optimal interaction.

Some orthogonal rotation of the computational basis $\{|\mathcal{E}_\lambda\rangle\}$ [as in Eq. (4.2.1)] can be considered as the measurement basis. An one-parametric family is considered here.

$$\begin{aligned}
|E_0\rangle &= \sqrt{a}|\mathcal{E}_0\rangle - \sqrt{1-a}|\mathcal{E}_1\rangle, \\
|E_1\rangle &= \sqrt{1-a}|\mathcal{E}_0\rangle + \sqrt{a}|\mathcal{E}_1\rangle, \\
|E_2\rangle &= \sqrt{a}|\mathcal{E}_2\rangle - \sqrt{1-a}|\mathcal{E}_3\rangle, \\
|E_3\rangle &= \sqrt{1-a}|\mathcal{E}_2\rangle + \sqrt{a}|\mathcal{E}_3\rangle.
\end{aligned} \tag{4.4.1}$$

For this eigenbasis, the optimal IVs of Eq. (4.3.5) becomes

$$\begin{aligned}
|\xi_x\rangle &= \left(\mathscr{D}_{uv}\sqrt{a}+\overline{\mathscr{D}}_{uv}\sqrt{1-a}\right)|\mathcal{E}_0\rangle+\left(\overline{\mathscr{D}}_{uv}\sqrt{a}-\mathscr{D}_{uv}\sqrt{1-a}\right)|\mathcal{E}_1\rangle,\\
|\xi_y\rangle &= \left(\overline{\mathscr{D}}_{uv}\sqrt{a}+\mathscr{D}_{uv}\sqrt{1-a}\right)|\mathcal{E}_0\rangle+\left(\mathscr{D}_{uv}\sqrt{a}-\overline{\mathscr{D}}_{uv}\sqrt{1-a}\right)|\mathcal{E}_1\rangle,\\
|\zeta_x\rangle &= \left(\mathscr{D}_{uv}\sqrt{a}+\overline{\mathscr{D}}_{uv}\sqrt{1-a}\right)|\mathcal{E}_2\rangle+\left(\overline{\mathscr{D}}_{uv}\sqrt{a}-\mathscr{D}_{uv}\sqrt{1-a}\right)|\mathcal{E}_3\rangle,\\
|\zeta_y\rangle &= \left(\overline{\mathscr{D}}_{uv}\sqrt{a}+\mathscr{D}_{uv}\sqrt{1-a}\right)|\mathcal{E}_2\rangle+\left(\mathscr{D}_{uv}\sqrt{a}-\overline{\mathscr{D}}_{uv}\sqrt{1-a}\right)|\mathcal{E}_3\rangle.
\end{aligned}\tag{4.4.2}$$

For unequal error rates, Eq. (4.2.2) is a special case (other than a permutation of the measurement basis states) with $a=1$ in Eq. (4.4.2). Similarly, for equal error rates ($D_{xy}=D_{uv}=D$), Eq. (4.2.4) is a specific instance with $a=\mathscr{D}^2$ in Eq. (4.4.2). One may generate various optimal interactions along with the associated optimal POVM by tuning the rotation parameter a in the range $[0,1]$. One such example is given here for unequal error rates.

Example 4.1. *Let $a=\frac{1}{2}$. Thus the optimal interaction in Eq.* (4.4.2) *becomes*

$$\begin{aligned}
|\xi_x\rangle &= \sqrt{1-D_{uv}}\,|\mathcal{E}_0\rangle-\sqrt{D_{uv}}\,|\mathcal{E}_1\rangle,\\
|\xi_y\rangle &= \sqrt{1-D_{uv}}\,|\mathcal{E}_0\rangle+\sqrt{D_{uv}}\,|\mathcal{E}_1\rangle,\\
|\zeta_x\rangle &= \sqrt{1-D_{uv}}\,|\mathcal{E}_2\rangle-\sqrt{D_{uv}}\,|\mathcal{E}_3\rangle,\\
|\zeta_y\rangle &= \sqrt{1-D_{uv}}\,|\mathcal{E}_2\rangle+\sqrt{D_{uv}}\,|\mathcal{E}_3\rangle,
\end{aligned}\tag{4.4.3}$$

and the corresponding optimal POVM is captured by

$$\begin{aligned}
|E_0\rangle &= \frac{1}{\sqrt{2}}\left(|\mathcal{E}_0\rangle-|\mathcal{E}_1\rangle\right), \qquad |E_1\rangle = \frac{1}{\sqrt{2}}\left(|\mathcal{E}_0\rangle+|\mathcal{E}_1\rangle\right),\\
|E_2\rangle &= \frac{1}{\sqrt{2}}\left(|\mathcal{E}_2\rangle-|\mathcal{E}_3\rangle\right), \qquad |E_3\rangle = \frac{1}{\sqrt{2}}\left(|\mathcal{E}_2\rangle+|\mathcal{E}_3\rangle\right).
\end{aligned}\tag{4.4.4}$$

4.4.1 Sidewaytable: optimal interactions and optimal POVMs listed

Here, we enlist various optimal interactions and the corresponding optimal POVMs discussed throughout the chapter.

Table 4.2 describes the general form of the optimal interaction and also shows its four specific instantiations, of which the first two coincide with those of [FGG+97]. The associated optimal POVMs are also mentioned.

The first one represent our [AP17] general form of optimal interactions as in Eq. (4.3.5).

The second and third ones are due to [FGG+97], for asymmetric and symmetric error rate, respectively. These are indeed special cases for our generalized expression. The (interaction, POVM) tuples are given by Eqs. (4.2.2, 4.2.3), and Eqs. (4.2.4, 4.3.12).

Finally, we consider two special cases of our [AP17] generalized expression. The first of them represents a set of optimal interactions as in Eq. (4.4.2) that depends on one parameter, while the corresponding optimal POVMs are in Eq. (4.4.1). A specific instance is then listed: the optimal interaction as in Eq. (4.4.3) along with its optimal POVM from Eq. (4.4.4).

We could establish that there exists infinitely many possible instances of an optimal interaction when represented in a canonical basis. However, they all have a unique representation while expressed in the measurement basis. Feeding an optimal measurement to the unique form of the optimal IVs produces a specific instance of an optimal interaction.

Clearly, the general expression of the optimal interaction vectors derived here yields different choices of those in [FGG$^+$97]. Moreover, the implementation is independent of equal or unequal error rates.

Cases	General form [AP17]	Asymm-err [FGG$^+$97]	Symm-err [FGG$^+$97]	One Parametric [AP17]	Specific [AP17]
Optimal Interaction	IVs: Eq. (4.3.5)	IVs Eq. (4.2.2)	IVs: Eq. (4.2.4)	IVs: Eq. (4.4.2)	IVs: Eq. (4.4.3)
$\|\xi_x\rangle$	$\mathscr{D}_{uv}\,\|E_0\rangle+\overline{\mathscr{D}}_{uv}\,\|E_1\rangle$	$\mathscr{D}_{uv}\,\|\mathcal{E}_0\rangle+\overline{\mathscr{D}}_{uv}\,\|\mathcal{E}_3\rangle$	$\|\mathcal{E}_0\rangle$	$\left(\mathscr{D}_{uv}\sqrt{a}+\overline{\mathscr{D}}_{uv}\sqrt{1-a}\right)\|\mathcal{E}_0\rangle+\left(\overline{\mathscr{D}}_{uv}\sqrt{a}-\mathscr{D}_{uv}\sqrt{1-a}\right)\|\mathcal{E}_1\rangle$	$\sqrt{1-D_{uv}}\,\|\mathcal{E}_0\rangle-\sqrt{D_{uv}}\,\|\mathcal{E}_1\rangle$
$\|\xi_y\rangle$	$\overline{\mathscr{D}}_{uv}\,\|E_0\rangle+\mathscr{D}_{uv}\,\|E_1\rangle$	$\overline{\mathscr{D}}_{uv}\,\|\mathcal{E}_0\rangle+\mathscr{D}_{uv}\,\|\mathcal{E}_3\rangle$	$2\mathscr{D}\overline{\mathscr{D}}\,\|\mathcal{E}_0\rangle+\left(\mathscr{D}^2-\overline{\mathscr{D}}^2\right)\|\mathcal{E}_1\rangle$	$\left(\overline{\mathscr{D}}_{uv}\sqrt{a}+\mathscr{D}_{uv}\sqrt{1-a}\right)\|\mathcal{E}_0\rangle+\left(\mathscr{D}_{uv}\sqrt{a}-\overline{\mathscr{D}}_{uv}\sqrt{1-a}\right)\|\mathcal{E}_1\rangle$	$\sqrt{1-D_{uv}}\,\|\mathcal{E}_0\rangle+\sqrt{D_{uv}}\,\|\mathcal{E}_1\rangle$
$\|\zeta_x\rangle$	$\mathscr{D}_{uv}\,\|E_2\rangle+\overline{\mathscr{D}}_{uv}\,\|E_3\rangle$	$\overline{\mathscr{D}}_{uv}\,\|\mathcal{E}_2\rangle+\mathscr{D}_{uv}\,\|\mathcal{E}_1\rangle$	$\|\mathcal{E}_2\rangle$	$\left(\mathscr{D}_{uv}\sqrt{a}+\overline{\mathscr{D}}_{uv}\sqrt{1-a}\right)\|\mathcal{E}_2\rangle+\left(\overline{\mathscr{D}}_{uv}\sqrt{a}-\mathscr{D}_{uv}\sqrt{1-a}\right)\|\mathcal{E}_3\rangle$	$\sqrt{1-D_{uv}}\,\|\mathcal{E}_2\rangle-\sqrt{D_{uv}}\,\|\mathcal{E}_3\rangle$
$\|\zeta_y\rangle$	$\overline{\mathscr{D}}_{uv}\,\|E_2\rangle+\mathscr{D}_{uv}\,\|E_3\rangle$	$\mathscr{D}_{uv}\,\|\mathcal{E}_2\rangle+\overline{\mathscr{D}}_{uv}\,\|\mathcal{E}_1\rangle$	$2\mathscr{D}\overline{\mathscr{D}}\,\|\mathcal{E}_2\rangle+\left(\mathscr{D}^2-\overline{\mathscr{D}}^2\right)\|\mathcal{E}_3\rangle$	$\left(\overline{\mathscr{D}}_{uv}\sqrt{a}+\mathscr{D}_{uv}\sqrt{1-a}\right)\|\mathcal{E}_2\rangle+\left(\mathscr{D}_{uv}\sqrt{a}-\overline{\mathscr{D}}_{uv}\sqrt{1-a}\right)\|\mathcal{E}_3\rangle$	$\sqrt{1-D_{uv}}\,\|\mathcal{E}_2\rangle+\sqrt{D_{uv}}\,\|\mathcal{E}_3\rangle$
Optimal POVM	General	Eq. (4.2.3)	Eq. (4.3.12)	Eq. (4.4.1)	Eq. (4.4.4)
$\|E_0\rangle$	$\|E_0\rangle$	$\|\mathcal{E}_0\rangle$	$\mathscr{D}\|\mathcal{E}_0\rangle-\overline{\mathscr{D}}\|\mathcal{E}_1\rangle$	$\sqrt{a}\|\mathcal{E}_0\rangle-\sqrt{1-a}\|\mathcal{E}_1\rangle$	$\frac{1}{\sqrt{2}}\left(\|\mathcal{E}_0\rangle-\|\mathcal{E}_1\rangle\right)$
$\|E_1\rangle$	$\|E_1\rangle$	$\|\mathcal{E}_1\rangle$	$\overline{\mathscr{D}}\|\mathcal{E}_0\rangle+\mathscr{D}\|\mathcal{E}_1\rangle$	$\sqrt{1-a}\|\mathcal{E}_0\rangle+\sqrt{a}\|\mathcal{E}_1\rangle$	$\frac{1}{\sqrt{2}}\left(\|\mathcal{E}_0\rangle+\|\mathcal{E}_1\rangle\right)$
$\|E_2\rangle$	$\|E_2\rangle$	$\|\mathcal{E}_2\rangle$	$\mathscr{D}\|\mathcal{E}_2\rangle-\overline{\mathscr{D}}\|\mathcal{E}_3\rangle$	$\sqrt{a}\|\mathcal{E}_2\rangle-\sqrt{1-a}\|\mathcal{E}_3\rangle$	$\frac{1}{\sqrt{2}}\left(\|\mathcal{E}_2\rangle-\|\mathcal{E}_3\rangle\right)$
$\|E_3\rangle$	$\|E_3\rangle$	$\|\mathcal{E}_3\rangle$	$\overline{\mathscr{D}}\|\mathcal{E}_2\rangle+\mathscr{D}\|\mathcal{E}_3\rangle$	$\sqrt{1-a}\|\mathcal{E}_2\rangle+\sqrt{a}\|\mathcal{E}_3\rangle$	$\frac{1}{\sqrt{2}}\left(\|\mathcal{E}_2\rangle+\|\mathcal{E}_3\rangle\right)$

Table 4.2 Optimal interaction and corresponding optimal POVM - unique general form and its specific instances.

4.5 Conclusion

For the BB84 quantum protocol, we have established a unique form describing the optimal interaction vectors of Eve. The corresponding optimal measurements follow automatically. To attain the optimal information gain for a given average disturbance, she can perform such an interaction followed by the associated measurements signal by signal. We have shown that the choice of optimal interaction in [FGG$^+$97], for equal as well as unequal error rates, is a special case of the optimal expression provided by us.

Although we have derived infinitely many candidate interactions for optimality, it is arguably not clear whether they are the all possible optimal interactions. A more convincing approach would be to derive them as part of a necessary and sufficient condition, which we'll discuss in the following chapter.

CHAPTER 5

A NEW NECESSARY AND SUFFICIENT CONDITION FOR OPTIMALITY AND DERIVING OPTIMAL INTERACTION VECTORS

Given an arbitrary interaction, to check whether It is optimal or not require a certificate. For example, a necessary and sufficient condition (NSC) by [FGG$^+$97], where the verification involves the PIJSs in the joint Hilbert space. Here we suggest [AP21] a refined NSC involving the states of Eve only that makes the verification easier.

It reveals that the optimal (non-zero) overlaps between the attackers post-interactions states must be equal and numerically same as the difference between the fidelity and the disturbance at the receiving end. That amount turns out to be same as the reduction (factor) in Bell violation when estimated for the equivalent entanglement-based protocol.

We move further with these NSCs to derive the optimal IVs. Thus, those are unambiguously the only and all possible interactions. Surprisingly, they are unitarily equivalent to the optimal interactions derived in [AP17]. We show that these optimal states are same as the outputs of an optimal phase-covariant cloner.

Moreover, we also have established an interrelation between the optimal POVMs for the two MUBs. These relations are useful for practical purposes, in the sense that whenever Eve chooses her measurement setups for the two MUBs corresponding to her suitable choice of the unitary evolution.

5.1 A brief overview

Fuchs *et al.* [FGG$^+$97] provided with a *necessary and sufficient condition*(NSC) involving the joint Hilbert space of the sender and the attacker. On the other hand, we suggest here a necessary and sufficient condition for optimality involving the Hilbert space of the eavesdropper only. This newly proposed verification is easier to perform than the earlier one [FGG$^+$97]. When interpreted the physical significance, this new criteria depicts explicitly the geometry of the optimal states with the attacker. We could figure out a direct connection of the optimal overlap with the equivalent entanglement-based protocol (reduction in Bell violation) and with optimal phase-covariant (pc) cloner [BCMDM00].

To be specific, an optimal incoherent attack is characterized by the non-zero overlaps between various non-orthogonal post-interaction states of Eve's ancilla. The amount of optimal overlap must be equal to the difference between the fidelity and the disturbance incurred at Bob's end. We have shown that this amount equates the reduction (factor) in the CHSH sum [CHSH69, Cir80] for the equivalent entanglement-based scheme. From geometrical perspective, it amounts to the contraction in the Bloch vectors that Bob finds in his received states due to eavesdropping affects.

We moved the process further through a chain of NSCs and derived infinitely many optimal interactions as a NSC only. Therefore, without any ambiguity, the resulted IVs are the only and all possible optimal interactions. These newly derived optimal IVs are unitarily the same as the optimal states derived earlier by us [AP17]. As our optimal states are described in terms of the measurement directions, an optimal PIJS clearly exhibits an one-to-one correspondence with the optimal measurement of Eve. Thereby, from mathematical perspective, specifying Eves measurement setup determines her IVs and vice versa. A set of optimal IVs corresponding to one encoding basis is interrelated with a set of optimal IVs in the conjugate basis: the associated relation between Eve's optimal measurements across the two MUBs are established. The optimal PIJSs (when measured in the computational basis) are in sync with the outputs obtained by an optimal pc-cloner [BCMDM00]. However, the states we derived are much more general in the sense that it doesn't depend on the specification of the measurement basis a priori.

Notations in use

To recall the optimal IVs from [AP17], we redefine the notations $\mathscr{D}^{+}_{uv}$ and $\mathscr{D}^{-}_{uv}$ as follows.

$$\mathscr{D}^{\pm}_{\beta} \quad := \quad \frac{\sqrt{1-D_\beta} \pm \sqrt{D_\beta}}{\sqrt{2}}. \tag{5.1.1}$$

Chapter organization

The section-wise work-flow is as follows. Firstly, we recollect the required results from the earlier chapters as described in Sec. 5.2. Then, we discuss the main results briefly in Sec. 5.3. The derivations are later detailed in Sec. 5.5. We conclude the chapter by summarizing the new findings and also discuss some further scopes to explore.

The new necessary and sufficient conditions along with the optimal states are discussed in Sec. 5.3. It starts from the existed NSC in [FGG+97], and moves through a series of NSCs that finally ended by deriving the optimal IVs as part of the process. An unitary equivalence with the earlier IVs [AP17] is established then. We also explain the NSCs and their physical significance.

The optimal IVs for an encoding basis are related to their optimal counterpart in the conjugate basis. We establish the one-to-one correspondence between the optimal states across the two MUBs. The during in turn depicts an interrelation between the optimal POVMs. We discuss these results in Sec. 5.4.

5.2 Required ingredients

In earlier chapter, we have derived infinitely many optimal interactions. We wish to know whether they cover the whole population or not. Thus, we develop a series of NSCs to derive the whole population of optimal interactions. Then we compare the newer population with the earlier ones.

We recollect here the optimal interactions in terms of the new notations. We also need to remember the NSC in [FGG+97].

5.2.1 The optimal states of Eve's ancilla after an interaction

Let's rewrite the optimal interactions from [AP17] in terms of the new notations $\mathscr{D}_\beta^\pm$.

When Alice's encoding basis is xy basis, the optimal IVs of Eve can be represented in her orthonormal measurement basis $\{|E_\lambda\rangle\}$ as follows.

$$\begin{aligned} |\xi_x^\star\rangle &= \mathscr{D}_{uv}^+|E_0\rangle + \mathscr{D}_{uv}^-|E_1\rangle, & |\xi_y^\star\rangle &= \mathscr{D}_{uv}^-|E_0\rangle + \mathscr{D}_{uv}^+|E_1\rangle, \\ |\zeta_x^\star\rangle &= \mathscr{D}_{uv}^+|E_2\rangle + \mathscr{D}_{uv}^-|E_3\rangle, & |\zeta_y^\star\rangle &= \mathscr{D}_{uv}^-|E_2\rangle + \mathscr{D}_{uv}^+|E_3\rangle. \end{aligned} \tag{5.2.1}$$

Note that, any of the optimal IVs is a superposition of two (out of four) measurement directions having amplitudes $\mathscr{D}_{uv}^+$ and $\mathscr{D}_{uv}^-$.

Similarly, the general expression of the optimal IVs in uv basis are as follows.

$$\begin{aligned} |\xi_u^\star\rangle &= \mathscr{D}_{xy}^+|F_0\rangle + \mathscr{D}_{xy}^-|F_1\rangle, & |\xi_v^\star\rangle &= \mathscr{D}_{xy}^-|F_0\rangle + \mathscr{D}_{xy}^+|F_1\rangle, \\ |\zeta_u^\star\rangle &= \mathscr{D}_{xy}^+|F_2\rangle + \mathscr{D}_{xy}^-|F_3\rangle, & |\zeta_v^\star\rangle &= \mathscr{D}_{xy}^-|F_2\rangle + \mathscr{D}_{xy}^+|F_3\rangle. \end{aligned} \tag{5.2.2}$$

Any specification of the measurement basis $\{|E_\lambda\rangle\}$ (or $\{|F_\lambda\rangle\}$) provides a specific instance of the optimal IVs in terms of the computational basis: *e.g.*, the optimal IVs due to Fuchs *et al.* [FGG$^+$97]. Due to various choices of the eigenbasis, there are infinitely many configurations of the optimal IVs when expressed in the computational basis.

As expected, there is an one-to-correspondence between these optimal IVs across the two MUBs. We discuss it to Sec. 5.4.

5.2.2 An existing Necessary and Sufficient Condition

An constraints of optimality of an interaction induces a restriction on the PIJSs and thereby restricting the nature of the IVs. Optimal IVs must satisfy certain necessary and sufficient conditions [FGG$^+$97]. Given a set of IVs, this verification is a routine task. However, deriving optimal IVs from these NSCs remained a harder task, and we have tackled this issue herein.

Consider the optimality of the post-interaction states (3.3.4.1, 3.3.4.3). Denote Alices state-symbol by $a^\beta \in \{x,y,u,v\}$, and denote the PIJS symbol S_a^β as X,Y,U,V, respectively. The existing NSC [FGG$^+$97, Eqs. (38,39)] for optimality in the xy basis involves the following four states defined over the joint Hilbert space of Bob and Eve.

$$|W_{\lambda a}\rangle := B_a \otimes \sqrt{E_\lambda}\,|W\rangle, \tag{5.2.1}$$

with $W \in \{U,V\}$ and $a \in \{u,v\}$; Bob uses the projective measurements $B_a := |a\rangle\langle a|$.

For optimal knowledge gain in xy basis (KG_{xy}), the inner products $\langle U_{\lambda u}|V_{\lambda u}\rangle$ and $\langle U_{\lambda v}|V_{\lambda v}\rangle$ must be real and have the same sign [1] $\varepsilon_\lambda^0 \in \pm 1$. Checking optimality is essentially to verify the following parallelism:

$$|U_{\lambda u}\rangle \parallel |V_{\lambda u}\rangle \text{ and } |U_{\lambda v}\rangle \parallel |V_{\lambda v}\rangle.$$

To be more specific, the PIJSs $|X\rangle, |Y\rangle$ are optimal for Eve having a POVM $\{E_\lambda\}$ *iff* the following equations are satisfied:

$$\sqrt{D_{uv}}\,|U_{\lambda u}\rangle = \varepsilon_\lambda \sqrt{1-D_{uv}}\,|V_{\lambda u}\rangle, \tag{5.2.2.u}$$

$$\sqrt{D_{uv}}\,|V_{\lambda v}\rangle = \varepsilon_\lambda \sqrt{1-D_{uv}}\,|U_{\lambda v}\rangle. \tag{5.2.2.v}$$

Similarly, analogous conditions hold for the optimality of the PIJSs in uv basis.

[1] Henceforth, we use the notations λ^β and $\varepsilon_\lambda^\beta$ to denote the eigenvalues and their signs [AP17] in a basis β.

5.3 A new necessary and sufficient condition towards completely characterizing Eve's optimal states

Now, we move from this existing NSC to derive refined ones involving Eve's system only. In this pursuit, we move through a series of *iff* conditions which eventually derives the optimal IVs w.r.t. the optimal measurement basis. One of these NSCs appeared to be of utmost interest: the following observation will help finding that refined certificate for optimality.

Lemma 5.1. *The post-interaction states of Eve exhibit an interrelation involving the overlap between the two undisturbed states and that between the two disturbed states.*

$$\left(1-D_{xy}\right)\langle\xi_x|\xi_y\rangle+D_{xy}\langle\zeta_x|\zeta_y\rangle=2\mathscr{D}^{+}_{uv}\mathscr{D}^{-}_{uv}.$$

The result follows by considering the inter-relations (5.5.1) between the IVs across the two MUBs, while imposing the normalization constraint on the IV $|\xi_u\rangle$.

In the following, we derive a series of *iff* conditions for an interaction to become optimal. The following criteria are equivalent.

Theorem 5.1. *The set of interaction vectors IV_{xy} is optimal along with the projectors $E_\lambda := |E_\lambda\rangle\langle E_\lambda|$ for measurement iff any of the following conditions hold:*

1. *The overlap between the measurement direction $|E_\lambda\rangle$ in xy basis and the IVs in uv basis are related in the following way:*

$$\begin{aligned}\langle E_\lambda|\xi_u\rangle &= \varepsilon^0_\lambda\,\langle E_\lambda|\zeta_v\rangle,\\ \langle E_\lambda|\xi_v\rangle &= \varepsilon^0_\lambda\,\langle E_\lambda|\zeta_u\rangle.\end{aligned}\tag{5.3.1}$$

   Corollary 1 *The overlap between the IVs in xy basis satisfy the following condition:*

$$\langle\xi_x|\xi_y\rangle=\langle\zeta_x|\zeta_y\rangle=1-2D_{uv}.\tag{5.3.2}$$

2. *The overlaps between the measurement direction $|E_\lambda\rangle$ in xy basis and the IVs in the same basis must maintain the following ratio:*

$$\frac{\langle E_\lambda|\xi_x\rangle}{\langle E_\lambda|\xi_y\rangle}=\frac{\langle E_\lambda|\zeta_x\rangle}{\langle E_\lambda|\zeta_y\rangle}=\frac{\mathscr{D}^{(+,\varepsilon^0_\lambda)}_{uv}}{\mathscr{D}^{(-,\varepsilon^0_\lambda)}_{uv}}=\left(\frac{\mathscr{D}^{+}_{uv}}{\mathscr{D}^{-}_{uv}}\right)^{\varepsilon^0_\lambda}.\tag{5.3.3}$$

   Here, we improvise to the following notation

$$\mathscr{D}^{(\sigma,\varepsilon^0_\lambda)}_{uv} = \frac{1}{\sqrt{2}}\left(\sqrt{1-D_{uv}}+\sigma\varepsilon^0_\lambda\sqrt{D_{uv}}\right),\tag{5.3.4}$$

with the sign parameter $\sigma = \pm 1$. It becomes $\mathscr{D}^{+}_{uv}$ or $\mathscr{D}^{-}_{uv}$, depending on whether the product $\sigma\varepsilon^{0}_{\lambda}$ becomes plus or minus, respectively.

3. *The interaction vectors in the xy basis can be expressed in an orthonormal basis $\{|E^{+}_{\lambda\xi}\rangle, |E^{-}_{\lambda\xi}\rangle, |E^{+}_{\lambda\zeta}\rangle, |E^{-}_{\lambda\zeta}\rangle\}$ as follows:*

$$\begin{aligned}
|\xi_x\rangle &= \mathscr{D}^{+}_{uv}|E^{+}_{\lambda\xi}\rangle + \mathscr{D}^{-}_{uv}|E^{-}_{\lambda\xi}\rangle, \\
|\xi_y\rangle &= \mathscr{D}^{-}_{uv}|E^{+}_{\lambda\xi}\rangle + \mathscr{D}^{+}_{uv}|E^{-}_{\lambda\xi}\rangle, \\
|\zeta_x\rangle &= \mathscr{D}^{+}_{uv}|E^{+}_{\lambda\zeta}\rangle + \mathscr{D}^{-}_{uv}|E^{-}_{\lambda\zeta}\rangle, \\
|\zeta_y\rangle &= \mathscr{D}^{-}_{uv}|E^{+}_{\lambda\zeta}\rangle + \mathscr{D}^{+}_{uv}|E^{-}_{\lambda\zeta}\rangle. \qquad (5.3.5)
\end{aligned}$$

The basis vectors $|E^{\pm}_{\lambda\xi}\rangle, |E^{\pm}_{\lambda\zeta}\rangle$ correspond to some unitary transform $\mathbf{R}^{\pm}$ of those two measurement directions $|E_{\lambda}\rangle$ that provide ±ve outcomes.

In the above-said list of NSCs, It is worthy to notice the change of basis while describing the overlap between Eve's measurement directions and the IVs. While Eve's measurements are considered in the *xy* basis, the IVs are considered in i) the *uv* basis for Eq. (5.3.1), and ii) the *xy* basis for Eq. (5.3.3).

5.3.1 Workflow exhibiting the equivalence between the IFF conditions

The four *iff* conditions as mentioned in Thm. 5.1 are equivalent in the sense that any of them can be derived [see Sec. 5.5] from the other one: directly, or via some of the remaining conditions. The inter-connections between them are sketched below in the Workflow 1.

5.3.2 Explaining the *iff* conditions

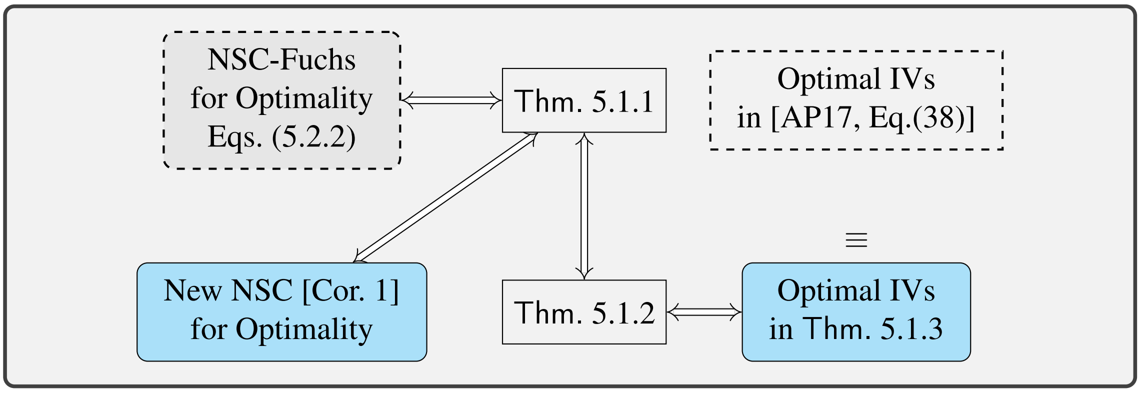
Here we explain the gross essence of the four *iff* conditions described in Thm. 5.1 indicating the optimality of the four IVs in the *xy* basis.

The 1^{st} *iff* condition says that the overlap between a measurement direction $|E_{\lambda}\rangle$ and a fidelity state (Eve's states corresponding to undisturbed counterpart of Bob's state) corresponding to Alice's signal u (or v) is same in magnitude as the overlap between that measurement direction and the disturbed state (Eve's states corresponding to disturbed counterpart of Bob's state) corresponding to Alice's signal v (or u), except that they differ in sign $\varepsilon^{0}_{\lambda}$.

The 2^{nd} *iff* condition says that the ratio of the overlaps between a measurement direction and Eve's undisturbed component are same as the ratio of the overlaps between that measurement direction and the disturbed component. The ratio becomes $\mathscr{D}^{+}_{uv}/\mathscr{D}^{-}_{uv}$ or its inverse depending on whether the measurement outcome is positive or negative in sign.

Workflow 1 | Workflow for Necc-suff-conditions and optimal IVs.

We started with the NSC due to [FGG$^+$97] and derived two important things: i) a new NSC involving the attackers Hilbert space only, and, ii) the optimal states with Eve. Both the results are found through a series of NSCs. These IVs are shown to be equivalent to those found earlier in Chap. 4.



The 3^{rd} *iff* condition provides the optimal interaction vectors, and therefore are the only and all possible optimal IVs without ambiguity. They are proved to be unitarily equivalent to those in [AP17, Eq.(38)]: see Sec. 5.5.2 for details.

The *iff* condition in Corollary 1, which is a byproduct of the 1^{st} *iff* condition of Thm. 5.1, restricts Eve's optimal states to have a specific orientation in the four-dimensional Hilbert space. To be more precise, when Alice encodes is the xy basis, the overlap between the two fidelity states must be same as the overlap between the two disturbed states and is equal to $(1-2D_{uv})$. This overlap-value appears in a lot many other crucial aspects, that we'll discuss shortly.

5.3.3 Physical significance of the new NSC

The necessary and sufficient condition in Corollary 1 can be used as a working formula to verify whether a given set of IVs is optimal or not. It's efficient due to easy verification, it's simple as it involves Eve's states only than the joint Hilbert space as in [FGG$^+$97], it's intuitive as it demands a specific configuration of the states in Eve's Hilbert space.

An optimal attack is essentially characterized by the optimal overlap, called here as *optimal syndrome*, that amounts to $1-2D$ for a symmetric (error-rate) attack. It exhibits interesting links between various other aspects for eavesdropping. Although, the connection between Bell violation and optimal state discrimination is known [FGG$^+$97], we find the connection more explicit here with respect to the optimal syndrome. For a specific error-rate D, the fraction of reduction in the optimal CHSH-sum in an *eb* scheme is precisely the optimal syndrome in the *p&m* scheme. On the other hand, the Bloch vectors at

Bob's end shrinks by the same factor $(1-2D)$.

5.3.4 Equivalence between the optimal interactions found here and found earlier

In this chapter, we wanted to find the optimal interactions as part of of a necessary and sufficient condition, which is done so far. Therefore, these are the only and all possible optimal interactions in the four dimensional Hilbert space.

The immediate question that comes to mind is whether these collection of IVs is more than what we have found earlier in the previous chapter. Or, are they same up to some isomorphism?

As it turns out, and surprising enough, that the OLD and the NEW collection of IVs are same up to some unitary equivalence. We establish the equivalence between the optimal IV_{xy} in Eq. (5.2.1) and those in Eq. (5.3.5). The proof is given in Sec. 5.5.

Since they are found to be equivalent, we can now safely use the later whenever they appear more handy to deal with. For instance, we use them in the very next section to establish one-to-one correspondence between the optimal IVs across the two MUBs.

5.4 Interrelation between the optimal POVMs across the two MUBs

Any specification of the orthonormal basis $\{|E_\lambda\rangle\}$ (or $\{|F_\lambda\rangle\}$) provides a specific instance of optimal IVs in computational basis, *e.g.*, the optimal IVs due to Fuchs *et al.* [FGG$^+$97]. Due to varied choices of the eigenbasis, there are infinitely many setups of the optimal IVs when expressed in computational basis. A one-to-one correspondence between the optimal IVs in each basis can be established (Sec. 5.5) since the optimal measurement directions $\{|E_\lambda\rangle\}$ in xy basis are interrelated to the optimal measurement directions $\{|F_\lambda\rangle\}$ in uv basis as follows:

$$\begin{aligned}
2|F_0\rangle &= |E_0\rangle+|E_1\rangle+|E_2\rangle+|E_3\rangle,\\
2|F_1\rangle &= |E_0\rangle+|E_1\rangle-|E_2\rangle-|E_3\rangle,\\
2|F_2\rangle &= |E_0\rangle-|E_1\rangle-|E_2\rangle+|E_3\rangle,\\
2|F_3\rangle &= |E_0\rangle-|E_1\rangle+|E_2\rangle-|E_3\rangle. \qquad (5.4.1)
\end{aligned}$$

For instance, the measurement basis $\{|E_\lambda\rangle\}=\{|00\rangle,|11\rangle,|10\rangle,|01\rangle\}$ fixes the measurement basis $\{|F_\lambda\rangle\}=\{|\bar{0}\bar{0}\rangle,|\bar{1}\bar{1}\rangle,|\bar{1}\bar{0}\rangle,|\bar{0}\bar{1}\rangle\}$ for Eve. These choice of the permuted computational basis retains the symmetry in Eve's measurement basis $(\{|E_\lambda\rangle\},\{|F_\lambda\rangle\})$ across the two encoding bases (xy,uv). The corresponding IVs in Eqs. (5.2.1, 5.2.2) represent the optimal states chosen by Fuchs *et al.*

The proof is deferred to Sec. 5.5. The basic idea is to recall the interrelation (Eq. (5.5.1)) between the optimal IVs across the two MUBs as provided by [FGG$^+$97]. Then to feed the generic form of their candidates (Eqs. (5.2.1, 5.2.2)) across the two MUBs into these relations. A little manipulation leads to the desired result.

The transformation matrix: It's interesting to observe that the two measurement setups in the conjugate bases are connected via the following unitary transformation

$$\widetilde{\mathbf{H}}_2 \;:=\; \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \tag{5.4.2}$$

Note that, It is a permutation (1,3,4,2) of the columns of the Hadamard matrix $\mathbf{H}_2 := \mathbb{H}^{\otimes 2}$ of order four. The effect of the unitary evolution that entangles the joint system seems to have some inner-connection to leverage that transformation.

5.4.1 Optimal PIJSs in the p&m scheme versus the output of an optimal phase-covariant cloner

For symmetric attack leveraging QBER=*D*, the PIJSs becomes

$$\begin{aligned}
\mathcal{U}|0\rangle|e\rangle &= \sqrt{1-D}|0\rangle\left(\mathscr{D}^{+}|E_0\rangle+\mathscr{D}^{-}|E_1\rangle\right)+\sqrt{D}|1\rangle\left(\mathscr{D}^{+}|E_2\rangle+\mathscr{D}^{-}|E_3\rangle\right),\\
\mathcal{U}|1\rangle|e\rangle &= \sqrt{1-D}|1\rangle\left(\mathscr{D}^{-}|E_0\rangle+\mathscr{D}^{+}|E_1\rangle\right)+\sqrt{D}|0\rangle\left(\mathscr{D}^{-}|E_2\rangle+\mathscr{D}^{+}|E_3\rangle\right).
\end{aligned}$$

Now, the amplitude of each eigenstate at the max. tolerable disturbance $D^{\star} := \frac{1}{2}\left(1-\frac{1}{\sqrt{2}}\right)$ are calculated as follows.

$$\begin{aligned}
\sqrt{1-D}\;\;\mathscr{D}^{+} &= \tfrac{1}{\sqrt{2}}\left(1-D+\sqrt{D(1-D)}\right) &&= \tfrac{1}{2}\left(1+\tfrac{1}{\sqrt{2}}\right),\\
\sqrt{1-D}\;\;\mathscr{D}^{-} &= \tfrac{1}{\sqrt{2}}\left(1-D-\sqrt{D(1-D)}\right) &&= \tfrac{1}{2\sqrt{2}},\\
\sqrt{D}\;\;\mathscr{D}^{+} &= \tfrac{1}{\sqrt{2}}\left(D+\sqrt{D(1-D)}\right) &&= \tfrac{1}{2\sqrt{2}},\\
\sqrt{D}\;\;\mathscr{D}^{-} &= \tfrac{1}{\sqrt{2}}\left(D-\sqrt{D(1-D)}\right) &&= \tfrac{1}{2}\left(1-\tfrac{1}{\sqrt{2}}\right).
\end{aligned}$$

Thus, at threshold $D^{\star}$, the optimal PIJSs w.r.t. a measurement basis are as follows.

$$\begin{aligned}
\mathcal{U}|0\rangle|e\rangle &= \frac{1}{2}\left(1+\frac{1}{\sqrt{2}}\right)|0\rangle|E_0\rangle + \frac{1}{2\sqrt{2}}\left(|0\rangle|E_1\rangle + |1\rangle|E_2\rangle\right) + \frac{1}{2}\left(1-\frac{1}{\sqrt{2}}\right)|1\rangle|E_3\rangle,\\
\mathcal{U}|1\rangle|e\rangle &= \frac{1}{2}\left(1+\frac{1}{\sqrt{2}}\right)|1\rangle|E_1\rangle + \frac{1}{2\sqrt{2}}\left(|1\rangle|E_0\rangle + |0\rangle|E_3\rangle\right) + \frac{1}{2}\left(1-\frac{1}{\sqrt{2}}\right)|0\rangle|E_2\rangle.
\end{aligned}$$

For the measurement setup $\{|E_0\rangle, |E_1\rangle, |E_2\rangle, |E_3\rangle\} = \{|00\rangle, |11\rangle, |01\rangle, |10\rangle\}$, which in disguise is the Fuchs basis (except the scuffle in last two) the optimal PIJSs can then be written in the computational basis as follows.

$$\begin{aligned}
\mathcal{U}|0\rangle|e\rangle &= \frac{1}{2}\left(1+\frac{1}{\sqrt{2}}\right)|0\rangle|00\rangle + \frac{1}{2\sqrt{2}}\left(|0\rangle|11\rangle + |1\rangle|01\rangle\right) + \frac{1}{2}\left(1-\frac{1}{\sqrt{2}}\right)|1\rangle|10\rangle,\\
\mathcal{U}|1\rangle|e\rangle &= \frac{1}{2}\left(1+\frac{1}{\sqrt{2}}\right)|1\rangle|11\rangle + \frac{1}{2\sqrt{2}}\left(|1\rangle|00\rangle + |0\rangle|10\rangle\right) + \frac{1}{2}\left(1-\frac{1}{\sqrt{2}}\right)|0\rangle|01\rangle.
\end{aligned}$$

They are same as the outputs of optimal phase-covariant cloner as in [BCMDM00, Eq. (36)].

5.5 Technical Details

Here we sketch a broad outline to prove the claims in the earlier sections.

5.5.1 Proving the necessary and sufficient conditions

Here we prove Thm. 5.1. The following relations involving the amplitudes $\mathscr{D}^{+}_{uv}$ and $\mathscr{D}^{-}_{uv}$ defined in Eq. (5.1.1) are heavily used in the derivations here.

$$(\mathscr{D}^{+}_{uv})^2 - (\mathscr{D}^{-}_{uv})^2 = 2\sqrt{D_{uv}(1-D_{uv})}, \quad (\mathscr{D}^{+}_{uv})^2 + (\mathscr{D}^{-}_{uv})^2 = 1,$$
$$2\mathscr{D}^{+}_{uv}\mathscr{D}^{-}_{uv} = 1 - 2D_{uv}.$$

Proof of the iff condition 1 of Thm. 5.1. The catch here is to unfold the states in Eq. (5.2.1) for the projectors E_λ while using the Schmidt form of the PIJSs, and use them in Eq. (5.2.2). In Eq. (5.2.1), for $a = u$,

$$\begin{aligned}
|U_{\lambda 0_u}\rangle &= B_u \otimes E_\lambda\ |U\rangle\\
&= \sqrt{1-D_{uv}}\ \langle E_\lambda|\xi_u\rangle \left(|u\rangle|E_\lambda\rangle\right).\\
|V_{\lambda 0_u}\rangle &= B_u \otimes E_\lambda\ |V\rangle\\
&= \sqrt{D_{uv}}\ \langle E_\lambda|\zeta_v\rangle \left(|u\rangle|E_\lambda\rangle\right).
\end{aligned}$$

Feeding them back into Eq. (5.2.2.u) leads to the first of the equations (5.3.1). The other relation can similarly be derived from the *iff* condition (5.2.2.v) while using the Schmidt form of the PIJSs and unfolding the states in Eq. (5.2.1) for $a = v$. □

Proof of the iff condition 2 of Thm. 5.1. The *iff* conditions in Eq. (5.3.1) can be grouped as follows:

$$\langle E_\lambda|(|\xi_u\rangle \pm |\xi_v\rangle) \quad = \quad \varepsilon^0_\lambda \, \langle E_\lambda|(|\zeta_v\rangle \pm |\zeta_u\rangle).$$

Now, we look back to the interrelations between the IVs in xy and uv basis, viz., use Eqs. (5.5.2.F+, 5.5.2.D+). Taking the inner product of the IVs in each of these equations with the measurement direction $|E_\lambda\rangle$, and then taking the ratio of the like sides, we get,

$$\frac{\langle E_\lambda|\xi_x\rangle + \langle E_\lambda|\xi_y\rangle}{\langle E_\lambda|\xi_x\rangle - \langle E_\lambda|\xi_y\rangle} \quad = \quad \frac{\sqrt{F_{uv}}}{\sqrt{D_{uv}}} \frac{\langle E_\lambda|\xi_u\rangle + \langle E_\lambda|\xi_v\rangle}{\langle E_\lambda|\zeta_u\rangle + \langle E_\lambda|\zeta_v\rangle} = \frac{\sqrt{F_{uv}}}{\sqrt{D_{uv}}} \varepsilon^0_\lambda.$$

By componendo and dividendo, we get,

$$\frac{\langle E_\lambda|\xi_x\rangle}{\langle E_\lambda|\xi_y\rangle} = \frac{\mathscr{D}_{uv}^{(+\varepsilon^0_\lambda)}}{\mathscr{D}_{uv}^{(-\varepsilon^0_\lambda)}} = \left(\frac{\mathscr{D}_{uv}}{\overline{\mathscr{D}}_{uv}}\right)^{\varepsilon^0_\lambda}.$$

We used here the improvised notation of Eq. (5.3.4). The ratio $\mathscr{D}_{uv}^{(+\varepsilon^0_\lambda)}/\mathscr{D}_{uv}^{(-\varepsilon^0_\lambda)}$ becomes $\mathscr{D}_{uv}/\overline{\mathscr{D}}_{uv}$ or its inverse depending on whether the sign ε^0_λ of the eigenvalue assumes $+1$ or -1, respectively.

Similarly, to establish the other ratio $\langle E_\lambda|\zeta_x\rangle/\langle E_\lambda|\zeta_y\rangle$ of Eq. (5.3.3), we consider Eqs. (5.5.2.F–, 5.5.2.D–) and follow the same procedure as above. □

Proof of the iff condition 3 of Thm. 5.1. The proof follows from the *iff* condition 2, viz., Eq. (5.3.3). The overlaps in the ratio $\langle E_\lambda|\xi_x\rangle/\langle E_\lambda|\xi_y\rangle$ can be unfolded using some (complex) constant of proportion $r_{\lambda,\xi}$ as follows.

$$\langle E_\lambda|\xi_x\rangle = r_{\lambda,\xi} \, \mathscr{D}_{uv}^{(+\varepsilon^0_\lambda)}, \qquad \langle E_\lambda|\xi_y\rangle = r_{\lambda,\xi} \, \mathscr{D}_{uv}^{(-\varepsilon^0_\lambda)}.$$

Note that, these overlaps constitute the components of the fidelity states when expressed in the eigenbasis $\{|E_\lambda\rangle\}$.

Similarly, in the ratio $\langle E_\lambda|\zeta_x\rangle/\langle E_\lambda|\zeta_y\rangle$, the overlaps can be written, for some complex number $r_{\lambda,\zeta}$, in the following way.

$$\langle E_\lambda|\zeta_x\rangle = r_{\lambda,\zeta} \, \mathscr{D}_{uv}^{(+\varepsilon^0_\lambda)}, \qquad \langle E_\lambda|\zeta_y\rangle = r_{\lambda,\zeta} \, \mathscr{D}_{uv}^{(-\varepsilon^0_\lambda)}.$$

These are the components of the disturbed states when expressed in the eigenbasis $\{|E_\lambda\rangle\}$.

Then we can write down the IVs with respect to the eigenbasis $\{|E_\lambda\rangle\}$ as follows.

$$\begin{aligned}
|\xi_x\rangle &= \sum_\lambda r_{\lambda,\xi}\ \mathscr{D}_{uv}^{(+\varepsilon_\lambda^0)}\ |E_\lambda\rangle, \\
|\xi_y\rangle &= \sum_\lambda r_{\lambda,\xi}\ \mathscr{D}_{uv}^{(-\varepsilon_\lambda^0)}\ |E_\lambda\rangle, \\
|\zeta_x\rangle &= \sum_\lambda r_{\lambda,\zeta}\ \mathscr{D}_{uv}^{(+\varepsilon_\lambda^0)}\ |E_\lambda\rangle, \\
|\zeta_y\rangle &= \sum_\lambda r_{\lambda,\zeta}\ \mathscr{D}_{uv}^{(-\varepsilon_\lambda^0)}\ |E_\lambda\rangle.
\end{aligned}$$

But, we observe that, $\mathscr{D}_{uv}^{(+\varepsilon_\lambda^0)} = \mathscr{D}_{uv}^+, \mathscr{D}_{uv}^-$ for $\varepsilon_\lambda^0 = +1, -1$ respectively. Similarly, $\mathscr{D}_{uv}^{(-\varepsilon_\lambda^0)} = \mathscr{D}_{uv}^-, \mathscr{D}_{uv}^+$ for $\varepsilon_\lambda^0 = +1, -1$ respectively. Thereby, in the expression of the IVs, we can group the basis vectors $|E_\lambda\rangle$ according to the sign of the measurement outcome. For instance, each of the fidelity states get two groups: $|E_{\lambda\xi}^\pm\rangle$ groups the measurement directions for ±ve outcomes. Similarly, the two groups for the disturbed states correspond to $|E_{\lambda\zeta}^\pm\rangle$. The following equation captures the grouping:

$$\begin{aligned}
|E_{\lambda\xi}^\pm\rangle &:= \sum_{\lambda:\ \pm\text{ve outcomes}} r_{\lambda,\xi}|E_\lambda\rangle, \\
|E_{\lambda\zeta}^\pm\rangle &:= \sum_{\lambda:\ \pm\text{ve outcomes}} r_{\lambda,\zeta}|E_\lambda\rangle.
\end{aligned}$$

With these grouping, the IVs can be described as in Eq. (5.3.5). That the vectors $\{|E_{\lambda\xi}^+\rangle, |E_{\lambda\xi}^-\rangle, |E_{\lambda\zeta}^+\rangle, |E_{\lambda\zeta}^-\rangle\}$ form an orthonormal basis, can be argued as follows. As defined, the states in $E_\lambda^+ := \{|E_{\lambda\xi}^+\rangle, |E_{\lambda\zeta}^+\rangle\}$ are mutually orthogonal to the states in $E_\lambda^- := \{|E_{\lambda\xi}^-\rangle, |E_{\lambda\zeta}^-\rangle\}$. Then, the normalization constraint on the fidelity (or disturbed) states together induces the normalization constraint on the states in E_λ^+ (or E_λ^-). Moreover, the orthogonality between the fidelity states and the disturbed states inherits the orthogonality within the states in E_λ^+ as well the orthogonality within the states in E_λ^-.

The final piece of the proof is the fact that each of the states $\{|E_{\lambda\xi}^+\rangle, |E_{\lambda\xi}^-\rangle, |E_{\lambda\zeta}^+\rangle, |E_{\lambda\zeta}^-\rangle\}$ can be expressed in terms of exactly two of the measurement directions $\{|E_\lambda\rangle\}$. It is so because, the sign of the measurement outcomes are evenly distributed for an optimal interaction: two +ve outcomes, and two -ve outcomes. Had it not been this way, then, w.l.o.g, let's assume the possibility for only one +ve outcome. Then, each of the states $|E_{\lambda\xi}^+\rangle, |E_{\lambda\zeta}^+\rangle$ should have only one of the measurement directions $|E_\lambda\rangle$ in their description. While the normalization constraint on these states indicate the coefficients $r_{\lambda,\xi}, r_{\lambda,\zeta}$ to be unimodular, their mutual orthogonality enforces one of these coefficients to be zero, leading to a contradiction. □

Proof of Corollary 1 of Thm. 5.1. The proof follows from condition 1 of the same theorem and Lem. 5.1.

Clearly, an equality of the overlaps in Lem. 5.1 lead to the desired result (5.3.2). To establish this equality, we consider the *iff* conditions (5.3.1), but for optimality in uv basis, viz.

$$\begin{aligned}\langle F_\lambda|\xi_x\rangle &= \varepsilon^1_\lambda\,\langle F_\lambda|\zeta_y\rangle,\\ \langle F_\lambda|\xi_y\rangle &= \varepsilon^1_\lambda\,\langle F_\lambda|\zeta_x\rangle.\end{aligned}$$

Multiplying the like sides of these two equations and adding over the measurement outcomes λ in uv basis, we get,

$$\Sigma_\lambda\langle\xi_x|F_\lambda\rangle\langle F_\lambda|\xi_y\rangle=\Sigma_\lambda\langle\zeta_x|F_\lambda\rangle\langle F_\lambda|\zeta_y\rangle.$$

Since the projectors F_λ consist a POVM, their completeness relation leads to the equality between the two overlaps $\langle\xi_x|\xi_y\rangle$ and $\langle\zeta_x|\zeta_y\rangle$, and consequently the desired result follows from Lem. 5.1. □

5.5.2 The two representations of the optimal IVs are unitarily equivalent

To establish the equivalence of the optimal IV$_{xy}$ in Eq. (5.2.1) and those in Eq. (5.3.5) we make a matrix-vector representation of the IVs. We introduce a few notations for that in Table 5.1. Here, OLD denotes IVs in Eq. (5.2.1) and NEW denotes IVs in Eq. (5.3.5).

Table 5.1 Notations for matrix-vector form of optimal IV

$$\begin{aligned}
&& \mathbf{a}_{xy} &:= (|\xi_x\rangle,|\xi_y\rangle,|\zeta_x\rangle,|\zeta_y\rangle),\\
\mathbf{M}^{\mathsf{NEW}}_{xy} &\equiv & \mathbf{M}^{+-+-}_{\xi,\xi,\zeta,\zeta} &:= (|E^+_{\lambda\xi}\rangle,|E^-_{\lambda\xi}\rangle,|E^+_{\lambda\zeta}\rangle,|E^-_{\lambda\zeta}\rangle),\\
&& \mathbf{M}^{++--}_{\xi,\zeta,\xi,\zeta} &:= (|E^+_{\lambda\xi}\rangle,|E^+_{\lambda\zeta}\rangle,|E^-_{\lambda\xi}\rangle,|E^-_{\lambda\zeta}\rangle),\\
&& \mathbf{M}^{++--}_{0,2,1,3} &:= (|E^+_0\rangle,|E^+_2\rangle,|E^-_1\rangle,|E^-_3\rangle),\\
\mathbf{M}^{\mathsf{OLD}}_{xy} &\equiv & \mathbf{M}^{+-+-}_{0,1,2,3} &:= (|E^+_0\rangle,|E^-_1\rangle,|E^+_2\rangle,|E^-_3\rangle),\\
&& \mathbb{D}_{uv} &:= \mathbb{1}_2\otimes\left(\mathscr{D}^+_{uv}\mathbb{1}_2+\mathscr{D}^-_{uv}\sigma_x\right).
\end{aligned}$$

Those optimal IVs can be expressed in matrix-vector form as follows:

$$\begin{aligned}(\mathbf{a}^\star_{xy})^{\mathsf{OLD}} &= \mathbf{M}^{\mathsf{OLD}}_{xy}\,\mathbb{D}_{uv},\\ (\mathbf{a}^\star_{xy})^{\mathsf{NEW}} &= \mathbf{M}^{\mathsf{NEW}}_{xy}\,\mathbb{D}_{uv}.\end{aligned}$$

To establish the equivalence, It is enough to show that $\mathbf{M}_{xy}^{\mathsf{NEW}}$ is unitarily equivalent to $\mathbf{M}_{xy}^{\mathsf{OLD}}$. The intermediate transformations are as follows:

$$\begin{tikzcd}
\mathbf{M}_{xy}^{\mathsf{OLD}} \arrow[d, phantom, "\rotatebox{90}{$\equiv$}"] & & & \mathbf{M}_{xy}^{\mathsf{NEW}} \arrow[d, phantom, "\rotatebox{90}{$\equiv$}"] \\
\mathbf{M}_{0,1,2,3}^{+-+-} \arrow[r, "S_w"] & \mathbf{M}_{0,2,1,3}^{++--} \arrow[r, "\mathbf{R}"] & \mathbf{M}_{\xi,\zeta,\xi,\zeta}^{++--} \arrow[r, "S_w"] & \mathbf{M}_{\xi,\xi,\zeta,\zeta}^{+-+-}
\end{tikzcd}$$

All the three maps are post-multiplication to transform the column-space, *e.g.*, $\mathbf{M}_{0,2,1,3}^{++--} = \mathbf{M}_{0,1,2,3}^{+-+-}\ S_w$ etc. The swap operation S_w exchanges two qubit states.

The unitary $\mathbf{R} := \mathsf{diag}(\mathbf{R}^+, \mathbf{R}^-)$ works on the measurement directions in order to affect unitarily the two subspaces, one for positive outcomes and the other for negative outcomes. To be specific, the measurement directions $\{|E_0^{\pm}\rangle, |E_2^{\pm}\rangle\}$ go through an unitary transformation $\mathbf{R}^{\pm}$ in that subspace.

Therefore, we get the following interrelation between the POVMs associated with the NEW and OLD optimal IVs.

$$\mathbf{M}_{xy}^{\mathsf{NEW}} = \mathbf{M}_{xy}^{\mathsf{OLD}}\ S_w\, \mathbf{R}\, S_w.$$

Hence the equivalence follows.

5.5.3 Proving the interrelation between optimal POVMs

Here we sketch an outline to prove Eq. (5.4.1) that inter-relates two optimal POVMs associated with the two MUBs. Note that the conjugate relation between the two encoding bases gets inherited to a similar conjugate relation between the PIJSs across two MUBs. The later conjugate relation in turn produces an inter-relation between the two sets of IVs across the two MUBs [FGG+97]. In this relation, one can directly plug in the optimal IVs as expressed in Eqs. (5.2.1, 5.2.2) and a simple algebra would eventually lead to Eq. (5.4.1). The technical details are as follows.

Since the conjugate relation for the encoding bases inherits to the PIJSs, the IVs in each of the encoding bases gets interrelated as follows.

$$\begin{aligned}
2\sqrt{F_{uv}}|\xi_u\rangle &= \sqrt{F_{xy}}(|\xi_x\rangle + |\xi_y\rangle) + \sqrt{D_{xy}}(|\zeta_x\rangle + |\zeta_y\rangle), \\
2\sqrt{F_{uv}}|\xi_v\rangle &= \sqrt{F_{xy}}(|\xi_x\rangle + |\xi_y\rangle) - \sqrt{D_{xy}}(|\zeta_x\rangle + |\zeta_y\rangle), \\
2\sqrt{D_{uv}}|\zeta_u\rangle &= \sqrt{F_{xy}}(|\xi_x\rangle - |\xi_y\rangle) + \sqrt{D_{xy}}(|\zeta_y\rangle - |\zeta_x\rangle), \\
2\sqrt{D_{uv}}|\zeta_v\rangle &= \sqrt{F_{xy}}(|\xi_x\rangle - |\xi_y\rangle) - \sqrt{D_{xy}}(|\zeta_y\rangle - |\zeta_x\rangle).
\end{aligned} \tag{5.5.1}$$

The sum and difference between the fidelity states (and similarly for the disturbed states) in *uv* basis are written in terms of the Eve's states in *xy* basis.

$$\sqrt{F_{uv}}\left(|\xi_u\rangle+|\xi_v\rangle\right)=\sqrt{F_{xy}}\left(|\xi_x\rangle+|\xi_y\rangle\right), \tag{5.5.2.F+}$$
$$\sqrt{F_{uv}}\left(|\xi_u\rangle-|\xi_v\rangle\right)=\sqrt{D_{xy}}\left(|\zeta_x\rangle+|\zeta_y\rangle\right), \tag{5.5.2.F–}$$
$$\sqrt{D_{uv}}\left(|\zeta_u\rangle+|\zeta_v\rangle\right)=\sqrt{F_{xy}}\left(|\xi_x\rangle-|\xi_y\rangle\right), \tag{5.5.2.D+}$$
$$\sqrt{D_{uv}}\left(|\zeta_u\rangle-|\zeta_v\rangle\right)=\sqrt{D_{xy}}\left(|\zeta_y\rangle-|\zeta_x\rangle\right). \tag{5.5.2.D–}$$

Now, we use the optimal IVs for *xy* and *uv* basis as in Eqs. (5.2.1, 5.2.2) to find the sum and difference of the parity IVs (disturbed or undisturbed).

When Alice encodes in *xy* basis, Eve's optimal IVs are grouped as follows: the sum (difference) of the two fidelity states are proportional to the sum (difference) of the measurement directions for outcomes 0,1. Similarly, the disturbed states exhibit the similar relations involving the measurement directions for outcomes 2,3.

$$\begin{aligned}
|\xi_x^\star\rangle+|\xi_y^\star\rangle &= 2\sqrt{F_{uv}}\left(|E_0\rangle+|E_1\rangle\right),\\
|\xi_x^\star\rangle-|\xi_y^\star\rangle &= 2\sqrt{D_{uv}}\left(|E_0\rangle-|E_1\rangle\right),\\
|\zeta_x^\star\rangle+|\zeta_y^\star\rangle &= 2\sqrt{F_{uv}}\left(|E_2\rangle+|E_3\rangle\right),\\
|\zeta_x^\star\rangle-|\zeta_y^\star\rangle &= 2\sqrt{D_{uv}}\left(|E_2\rangle-|E_3\rangle\right).
\end{aligned} \tag{5.5.3}$$

Similarly, for optimal IVs in *uv*-basis, we get

$$\begin{aligned}
|\xi_u^\star\rangle+|\xi_v^\star\rangle &= \sqrt{2}\sqrt{F_{xy}}\left(|F_0\rangle+|F_1\rangle\right),\\
|\xi_u^\star\rangle-|\xi_v^\star\rangle &= \sqrt{2}\sqrt{D_{xy}}\left(|F_0\rangle-|F_1\rangle\right),\\
|\zeta_u^\star\rangle+|\zeta_v^\star\rangle &= \sqrt{2}\sqrt{F_{xy}}\left(|F_2\rangle+|F_3\rangle\right),\\
|\zeta_u^\star\rangle-|\zeta_v^\star\rangle &= \sqrt{2}\sqrt{D_{xy}}\left(|F_2\rangle-|F_3\rangle\right).
\end{aligned} \tag{5.5.4}$$

Finally, feeding back the Eqs. (5.5.3, 5.5.4) for optimal IVs into Eq. (5.5.2), we get the following relations: and feed them back into Eq. (5.5.2) to get the following relations:

$$\begin{aligned}
|F_0\rangle+|F_1\rangle &= |E_0\rangle+|E_1\rangle, & |F_2\rangle+|F_3\rangle &= |E_0\rangle-|E_1\rangle,\\
|F_0\rangle-|F_1\rangle &= |E_2\rangle+|E_3\rangle, & |F_2\rangle+|F_3\rangle &= |E_3\rangle-|E_2\rangle.
\end{aligned}$$

Finally, solving these equations, one can easily get the desired relation between the optimal measurement directions as in Eq. (5.4.1).

5.6 Conclusion

In this chapter, we have characterized the optimal individual attacks (i.e., characterizing the optimal IVs) on the BB84 protocol exhaustively. For the analysis, we have considered the generalized asymmetric error rates across the two MUBs in order to uncover all possible alternatives for an attacker, while a symmetric (error) attack automatically becomes a special case. A series of necessary and sufficient conditions is derived here to testify the optimality of an interaction performed by an eavesdropper. The NSCs involves the attackers Hilbert space only. As it unveils, an optimal attack corresponds to a specific configuration of the attacker's post-interaction states: that the overlap between the two disturbed states are same as the overlap between the two undisturbed states and is equal to the difference between the fidelity and the disturbance at the receiving end. Interestingly enough, the optimal overlap is same as the reduction (factor) in Bell violation in the equivalent entanglement-based scheme. We have shown explicitly that the optimal states of the joint system (when the measurement basis is the computational basis) can also be obtained by an optimal phase-covariant cloning mechanism.

From practical implementation perspectives, It is important to know the optimal unitary evolution that is required to evolve the joint system. We address this issue in the next chapter.

CHAPTER 6

CHARACTERIZING THE OPTIMAL UNITARY EVOLUTIONS

To mount an optimal attack on the BB84 protocol following the eavesdropping mechanism in [FGG$^+$97], an eavesdropper first need to know the optimal unitary evolution (apart from the optimal measurement) that she requires to evolve the joint system.

Given a specific set of interaction vectors, finding an optimal unitary is not that difficult a task to be addressed. One may consider a numerical approach. But, for parameter-based IVs, we have discussed here a rudimentary basis-completion method [AP21] to get an optimal unitary described in terms of the same parameter. However, this approach sometimes face difficulty to find the unknown basis vectors in an eight-dimensional Hilbert space out of a couple of known ones. Mainly, the parametric expressions need some elegant algebraic manipulation. We remove such difficulty by developing new mathematical tools to deal with generalized parametric expressions of the IVs and can derive parametric expressions of optimal unitary evolution in a general setup.

Some more generalizations are further addressed: like, finding all possible unitaries from a given interaction (specified by IVs), finding an unitary for a different initial state from a known unitary for a given IS, and tracking the changes in an unitary when measurement setup changes.

These techniques can also be improvised for any other protocol with any attack model where the optimal IVs are known. An interested reader may delve further towards that direction.

6.1 A brief overview

We consider the task of characterizing the optimal unitary attacks, i.e., to derive the optimal unitary evolutions that will lead to the general expression of (infinitely many) optimal interactions that we derived earlier. For that, it is enough to consider the optimal states in one of the MUBs.

First, we describe the rudimentary approach to find an optimal unitary for a given pair of optimal PIJSs, which faces some technical difficulties like basis completion issue in an arbitrary measurement basis (not a specific instance, but variable) for Eve. We bypass these hurdles in an elegant analytical approach to obtain an optimal unitary fit for an *initial state* (IS) when Eve measures in the computational basis.

However, given a set of specific IVs, the corresponding joint unitary is not unique. We have developed the methods to find any of its infinitely many siblings. Further, we explain the ways to get optimal unitaries for arbitrary IS and then for arbitrary measurement basis used by Eve. We exemplify these methods to understand the intricacies. Essentially we have characterized the whole space of optimal unitary attacks.

Chapter organization

The section wise work-flow is as follows. The main results are briefly described in Sec. 6.2, while their derivations are deferred until in Sec. 6.3. Sec. 6.2 deals with characterizing optimal unitary evolutions. We conclude by summarizing the new findings and also discuss further scopes to explore.

6.2 Characterizing optimal unitary evolutions

Given the optimal PIJSs $|X^\star\rangle, |Y^\star\rangle$, the objective is to find an optimal unitary for a suitable initial state $|\psi_0\rangle$ of Eve's ancilla. Mathematically speaking, the task is to solve the following two equations.

$$\mathcal{U}^{AE}_{\psi_0}|0\rangle_A|\psi_0\rangle_E = |X^\star\rangle, \qquad \mathcal{U}^{AE}_{\psi_0}|1\rangle_A|\psi_0\rangle_E = |Y^\star\rangle. \tag{6.2.1}$$

Although the same unitary serves the purpose in the conjugate basis, the measurement setup generally differs.

Getting a specific optimal unitary $\mathcal{U}_{\psi_0}$ from a given pair of PIJSs, i.e., solving the Eq. (6.2.1), can be done by the following basis completion technique. By introducing some *auxiliary states*, an unitary evolution $\mathcal{U}_{\psi_0}$ can be viewed as a linear transformation that maps an orthonormal basis $\{|0\rangle_A|\psi_i\rangle_E, |1\rangle_A|\psi_i\rangle_E\}_{i\in\{0,1,2,3\}}$ involving the IS to the orthonormal basis $\{|X_i\rangle, |Y_i\rangle\}_{i\in\{0,1,2,3\}}$ involving the PIJSs, where $|X_0\rangle = |X^\star\rangle, |Y_0\rangle = |Y^\star\rangle$.

Thus, we get the following eight equations to solve.

$$\mathcal{U}_{\psi_0}|0\rangle_A|\psi_i\rangle_E = |X_i\rangle, \qquad \mathcal{U}_{\psi_0}|1\rangle_A|\psi_i\rangle_E = |Y_i\rangle,$$
$$\forall i \in \{0,1,2,3\}.$$

Then, a solution for the optimal unitary can be expressed as follows

$$\mathcal{U}_{\psi_0} = \sum_{i=0}^{3} \big(|X_i\rangle\langle 0_A| + |Y_i\rangle\langle 1_A|\big)\,\langle\psi_i|_E. \tag{6.2.2}$$

It can further be factored [see Sec. 6.3.2] in two unitaries as

$$\mathcal{U}_{\psi_0} = \mathcal{W}^{AE}_{X,Y}\,(\mathbb{1}^A_2 \otimes \mathcal{W}^{E\dagger}_{\psi_0}). \tag{6.2.3}$$

The first unitary $\mathcal{W}_{X,Y}$, that depends on the PIJSs $|X^\star\rangle, |Y^\star\rangle$, is defined as

$$\mathcal{W}_{X,Y} := \sum_{i=0}^{3} |X_i\rangle\langle 0_A|\langle i_E| + |Y_i\rangle\langle 1_A|\langle i_E|, \tag{6.2.4}$$

which has the following matrix representation

$$\big[\ |X_0\rangle, |X_1\rangle, |X_2\rangle, |X_3\rangle,\ |Y_0\rangle, |Y_1\rangle, |Y_2\rangle, |Y_3\rangle\ \big]. \tag{6.2.4.Mat}$$

In the second unitary, the local unitary $\mathcal{W}$ depends on the initial state $|\psi_0\rangle$, and is defined as follows

$$\mathcal{W}^E_{\psi_0} = \sum_{i=0}^{3} |\psi_i\rangle\langle i_E|, \tag{6.2.5}$$

which has the following matrix representation

$$\big[\ |\psi_0\rangle\ \ |\psi_1\rangle\ \ |\psi_2\rangle\ \ |\psi_3\rangle\ \big]. \tag{6.2.5.Mat}$$

We observe from Eq. (6.2.3) that an optimal unitary is a product of two unitaries. The individual effect of these unitaries are explained as follows. In order to evolve the joint system from the initial state $|a\rangle|e\rangle$, the part of it (the 2nd part) first transforms Eve's initial state to $|00\rangle$ leaving Alice's part invariant, and then the other part (1st component) creates the required entanglement between Alice and Eve's states.

Eqs. (3.3.4.1, 5.2.1) together depicts that Eve's measurement setup $\mathbf{M}$ is in one-to-one correspondence with the PIJSs $|X\rangle^{\mathbf{M}}, |Y\rangle^{\mathbf{M}}$. Moreover, the factorization in Eq. (6.2.3) indicates that the joint unitary $\mathcal{U}$ depends on the initial state **IS** of Eve's ancilla, and Eve's measurement setup **M**. While the earlier one (IS) controls the unitary $\mathcal{W}_{\psi_0}$, the later one (**M**) determines $\mathcal{W}_{X,Y}$. Nevertheless, $\mathcal{U} \equiv \mathcal{U}^{\mathbf{M}}_{\mathbf{IS}}$ represent an infinite collection of unitaries.

When Eve measures in the four-dimensional computational basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, the corresponding optimal PIJSs are denoted by $|X\rangle^{\mathcal{C}}, |Y\rangle^{\mathcal{C}}$ and are described in Table 6.1. Let's consider the problem of getting an optimal unitary that evolves an IS into these PIJSs. As discussed earlier, it corresponds to the problem of basis completion: once in the eight dimensional space of optimal PIJSs, and once in the four dimensional space of the initial state. To avoid any technical difficulty (*e.g.*, numerical, or, trial-and-error approach) with basis completion, we address here briefly a unique analytical approach.

The crux is to view the PIJSs $|X\rangle^{\mathcal{C}}$ and $|Y\rangle^{\mathcal{C}}$ mathematically as an outcome of an action of two sub-matrices $\mathcal{U}_x := \big(|00\rangle \;\; |11\rangle\big) \otimes \mathbb{1}_2$ and $\mathcal{U}_y := (|10\rangle \;\; |01\rangle) \otimes \sigma_x$, respectively, on some specific initial state

$$|\Delta^{\mathbb{H}}\rangle_E \;\; := \;\; |\Delta_{xy}\rangle_{E_1} |\Delta^{\mathbb{H}}_{uv}\rangle_{E_2}, \tag{6.2.6}$$

where

$$\begin{aligned} |\Delta_\beta\rangle \;\; &:= \;\; \sqrt{F_\beta}|0\rangle + \sqrt{D_\beta}|1\rangle, \\ |\Delta^{\mathbb{H}}_\beta\rangle \;\; &:= \;\; \mathbb{H}|\Delta_\beta\rangle = \mathscr{D}^+_\beta|0\rangle + \mathscr{D}^-_\beta|1\rangle. \end{aligned} \tag{6.2.7}$$

The optimal unitary $\mathcal{U}$ eventually becomes the partitioned matrix $\big[\, \mathcal{U}_x \;\; \mathcal{U}_y \,\big]$, which, in its block-matrix form, looks as follows

$$\mathcal{U}^c_{\Delta^{\mathbb{H}}} = \begin{bmatrix} \mathbb{1}_2 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \sigma_x \\ \cdot & \cdot & \sigma_x & \cdot \\ \cdot & \mathbb{1}_2 & \cdot & \cdot \end{bmatrix}. \tag{6.2.8}$$

Each of the constituent two-dimensional unitaries (e.g., σ_x) will operate on the second qubit of Eve's probe.

However, the same PIJSs $|X\rangle^{\mathcal{C}}$, and $|Y\rangle^{\mathcal{C}}$ can also be produced by some other unitary (due to the choice of free variables) acting on the same joint initial states $|0\rangle_A|\Delta^{\mathbb{H}}\rangle_E$ and $|1\rangle_A|\Delta^{\mathbb{H}}\rangle_E$, respectively. Eventually, there are infinitely many such unitary evolutions for the same PIJSs and the same IS, as evident from Eq. (6.2.3). The arbitration is two-fold: on $\mathcal{W}_{X,Y}$ or on $\mathcal{W}_{\psi_0}$, which in turn corresponds to various choices of the auxiliary states $\{|X_i\rangle, |Y_i\rangle\}_{i=1,2,3}$ or $|\psi_i\rangle_{i=1,2,3}$, respectively. The later arbitration, for instance, is technically achieved by post-multiplying the unitary in Eq. (6.2.8) by some $\mathbb{1}_2 \otimes \Gamma_{\psi_0^\perp}$, where the local unitary $\Gamma_{\psi_0^\perp}$ leaves $|\psi_0\rangle$ unchanged while considers new choices for the auxiliary states $|\psi_i\rangle_{i=1,2,3}$. For instance, the following choice

$$\Gamma_{\psi_0^\perp} = \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & \frac{1}{\sqrt{2}} & \cdot & \frac{1}{\sqrt{2}} \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \frac{1}{\sqrt{2}} & \cdot & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

will affect the (2nd, 4th) and (6th, 8th) columns of the unitary $\mathcal{U}^{\mathcal{C}}_{\Delta^{\mathbb{H}}}$. Naturally, the optimal unitary in Eq. (6.2.8) is the simplest among its infinitely many siblings.

Moreover, the same PIJSs $|X\rangle^{\mathcal{C}}$, and $|Y\rangle^{\mathcal{C}}$ can be reproduced by a newer family of unitary evolutions if we consider a different IS of Eve's ancilla. Getting an optimal unitary for a different IS corresponds to pin-point a local unitary that transforms the earlier IS to the newer one, as formulated in Eq. (6.3.1). For instance, consider the task of finding an optimal unitary for the IS $|00\rangle_E$ which in turn is a unitary tweak $A_{xy} \otimes A_{uv}\mathbb{H}$ of the IS in Eq. (6.2.6) for the 2-dimensional unitary $A_\beta = \sqrt{1-D_\beta}\sigma_z + \sqrt{D_\beta}\sigma_x$. The desired optimal unitary $\mathcal{U}^{\mathcal{C}}_{00}$ is given in Table 6.4.1.

We can now interrelate our unique approach with the rudimentary basis completion method. One can now directly read the auxiliary basis states $\{|X_i\rangle, |Y_i\rangle\}$ from the columns of the unitary $\mathcal{U}^{\mathcal{C}}_{00}$ by fixing the other auxiliary basis states $\{|\psi_i\rangle\}$ to enforce $\mathcal{W}_{00}$ to be the identity matrix. It is so because the unitary $\mathcal{U}^{\mathcal{C}}_{00}$ represents the matrix in Eq. (6.2.4.Mat).

Finally, consider finding optimal unitaries in a different measurement basis other than the computational basis. Eve's new measurement basis $\{|E_\lambda\rangle\}$ is a unitary transformation $|E_\lambda\rangle = \mathbf{M}_{xy}|\lambda\rangle$ of the computational basis. For instance, consider the unitary transformation $M_{xy} = (|E_0\rangle \;\; |E_1\rangle \;\; |E_2\rangle \;\; |E_3\rangle) = \mathcal{R}$, as mentioned in Eq. (6.4.1). Corresponding optimal PIJSs $(|X\rangle^{\mathcal{M}}, |Y\rangle^{\mathcal{M}})$ and an optimal unitary $(\mathcal{U}^{\mathcal{M}})$ can be obtained by operating $(\mathbb{1}_2 \otimes \mathbf{M}_{xy})$ on those in the computational basis, and are described in Table 6.2 and Eq. (6.4.1), respectively.

6.3 Technical Details

Here we sketch a broad outline to prove the claims in the earlier sections.

6.3.1 Getting an optimal unitary along with an initial state when Eve measures in the computational basis

First, we find the optimal IVs and the optimal PIJSs as Eve measures in computational basis.

Her optimal IVs can be expressed as follows:

$$|\xi_x^\star\rangle^{\mathcal{C}} = |0\rangle_{E_1}|\Delta_{uv}^{\mathbb{H}}\rangle_{E_2}, \qquad |\xi_y^\star\rangle^{\mathcal{C}} = (\mathbb{1}_2^{E_1}\otimes\sigma_x^{E_2})|\xi_x^\star\rangle^{\mathcal{C}},$$
$$|\zeta_x^\star\rangle^{\mathcal{C}} = |1\rangle_{E_1}|\Delta_{uv}^{\mathbb{H}}\rangle_{E_2}, \qquad |\zeta_y^\star\rangle^{\mathcal{C}} = (\mathbb{1}_2^{E_1}\otimes\sigma_x^{E_2})|\zeta_x^\star\rangle^{\mathcal{C}}.$$

Here the state $|\Delta_{uv}^{\mathbb{H}}\rangle$ is as defined in Eq. (6.2.7).

Therefore, the corresponding optimal PIJS$_{xy}^{AE}$ can be expressed as follows

$$\begin{aligned} |X^\star\rangle^{\mathcal{C}} &= |\Phi_{D_{xy}}^+\rangle_{AE_1}|\Delta_{uv}^{\mathbb{H}}\rangle_{E_2}, \\ |Y^\star\rangle^{\mathcal{C}} &= |\Psi_{D_{xy}}^+\rangle_{AE_1}\otimes\sigma_x^{E_2}|\Delta_{uv}^{\mathbb{H}}\rangle_{E_2}. \end{aligned}$$

where

$$\begin{aligned} |\Phi_{D_{xy}}^+\rangle_{AE_1} &= \sqrt{1-D_{xy}}|00\rangle_{AE_1}+\sqrt{D_{xy}}|11\rangle_{AE_1}, \\ |\Psi_{D_{xy}}^+\rangle_{AE_1} &= \sqrt{1-D_{xy}}|10\rangle_{AE_1}+\sqrt{D_{xy}}|01\rangle_{AE_1}. \end{aligned}$$

To get an optimal unitary out of these equations, we need to rewrite the PIJSs in matrix-vector form. First, note that the entangled states from the subsystem AE$_1$ can be expressed in matrix-vector form as follows

$$\begin{aligned} |\Phi_{D_{xy}}^+\rangle_{AE_1} &= W_x^{AE_1}|\Delta_{xy}\rangle_{E_1}, \\ |\Psi_{D_{xy}}^+\rangle_{AE_1} &= W_y^{AE_1}|\Delta_{xy}\rangle_{E_1}, \end{aligned}$$

with the 4×2 matrices

$$\begin{aligned} \mathbf{W}_x^{AE_1} &= |00\rangle_{AE_1}\langle 0_{E_1}|+|11\rangle_{AE_1}\langle 1_{E_1}|, \\ \mathbf{W}_y^{AE_1} &= |10\rangle_{AE_1}\langle 0_{E_1}|+|01\rangle_{AE_1}\langle 1_{E_1}|. \end{aligned}$$

Thereby, the optimal PIJS$_{xy}^{AE}$ can be expressed in matrix-vector form as follows:

$$\begin{aligned} |X^\star\rangle^{\mathcal{C}} &= \mathcal{U}_x^{AE}\,|\Delta_{xy}\rangle_{E_1}|\Delta_{uv}^{\mathbb{H}}\rangle_{E_2}, \\ |Y^\star\rangle^{\mathcal{C}} &= \mathcal{U}_y^{AE}\,|\Delta_{xy}\rangle_{E_1}|\Delta_{uv}^{\mathbb{H}}\rangle_{E_2}, \end{aligned}$$

with the 8×4 matrices

$$\mathcal{U}_x^{AE} = \mathbf{W}_x^{AE_1}\otimes\mathbb{1}_2^{E_2}, \qquad \mathcal{U}_y^{AE} = \mathbf{W}_y^{AE_1}\otimes\sigma_x^{E_2}.$$

Then, for an initial state

$$|\Delta^{\mathbb{H}}\rangle_E \;:=\; |\Delta_{xy}\rangle_{E_1}|\Delta_{uv}^{\mathbb{H}}\rangle_{E_2},$$

an optimal unitary can be given as

$$\begin{aligned}
\mathcal{U}^{AE}_{\Delta^{\mathbb{H}}} &= \mathcal{U}^{AE}_{x}\langle 0|_A + \mathcal{U}^{AE}_{y}\langle 1|_A \\
&= \mathbf{W}^{AE_1}_{x}\langle 0|_A \otimes \mathbb{1}^{E_2}_{2} + \mathbf{W}^{AE_1}_{y}\langle 1|_A \otimes \sigma^{E_2}_{x} \\
&= (|00\rangle_{AE_1}\langle 00| + |11\rangle_{AE_1}\langle 01|) \otimes \mathbb{1}^{E_2}_{2} \\
&\quad +(|10\rangle_{AE_1}\langle 10| + |01\rangle_{AE_1}\langle 11|) \otimes \sigma^{E_2}_{x}.
\end{aligned}$$

6.3.2 Factorization of an optimal unitary

The optimal unitary in Eq. (6.2.2) can be factored into two in the following way

$$\begin{aligned}
\mathcal{U}_{\psi_0} &= \sum_{a=0}^{1}\sum_{i=0}^{3}|S_a\rangle\langle a_A|\langle\psi_i|_E \\
&= \sum_{a=0}^{1}\sum_{i=0}^{3}|S_a\rangle\langle a_A|\langle i|_E|i\rangle\langle\psi_i|_E \\
&= \sum_{a=0}^{1}\sum_{i=0}^{3}|S_a\rangle\langle a_A|\langle i_E| \;\times\; \sum_{i=0}^{3}\mathbb{1}_2\otimes|i\rangle_E\langle\psi_i|.
\end{aligned}$$

6.3.3 Alternate solutions for optimal unitaries for a fixed IS

Here we explain how to find alternate optimal unitaries for a given IS by already knowing an optimal unitary for that IS. We completely characterize this bi-level arbitration.

Theorem 6.1. *For a given initial state $|\psi_0\rangle$, let an optimal unitary is known as $\mathcal{U}_{\psi_0}$. For the same initial state, a new optimal unitary $\mathcal{U}'_{\psi_0}$ can be found in one of the following ways.*

1. *A change in the auxiliary states spanning the orthogonal complement of the IS $|\psi_0\rangle$, such that*

$$\mathcal{U}'_{\psi_0} = \mathcal{U}_{\psi_0}(\mathbb{1}_2\otimes\Gamma_{\psi_0^{\perp}}).$$

The local unitary $\Gamma_{\psi_0^{\perp}} = \begin{bmatrix} 1 & \cdot \\ \cdot & T^{\dagger}_{\psi_0^{\perp}} \end{bmatrix}$ makes an alternate choice $\mathcal{W}'_{\psi_0}$ for $\mathcal{W}_{\psi_0}$:

$$\mathcal{W}'_{\psi_0} = \mathcal{W}_{\psi_0}\Gamma_{\psi_0^{\perp}} = \Big[\; |\psi_0\rangle \;\; |\psi_1\rangle' \;\; |\psi_2\rangle' \;\; |\psi_3\rangle' \;\Big].$$

The three dimensional unitary $T_{\psi_0^{\perp}}$ transforms the auxiliary states $|\psi_i\rangle_{i=1,2,3}$ to a newer one, while $\Gamma_{\psi_0^{\perp}}$ leaves $|\psi_0\rangle$ intact.

2. *Changing the auxiliary states spanning the orthogonal complement of the PIJSs*

$|X^\star\rangle, |Y^\star\rangle$, *such that*

$$\mathcal{U}'_{\psi_0} \;=\; \mathcal{W}'_{XY}\,\mathcal{W}_{\psi_0} \;=\; \mathcal{W}_{XY}\,\Gamma_{X^\perp Y^\perp}\,\mathcal{W}_{\psi_0}.$$

The global unitary

$$\Gamma_{X^\perp Y^\perp} \;=\; diag(\Gamma_{X^\perp}, \Gamma_{Y^\perp})$$

transforms $\mathcal{U}_{XY}$ *to a new one* $\mathcal{U}'_{XY} = \mathcal{U}_{XY}\,\Gamma_{X^\perp Y^\perp}$ *having the following matrix representation*

$$\begin{bmatrix} |X^\star\rangle & |X'_1\rangle & |X'_2\rangle & |X'_3\rangle & |Y^\star\rangle & |Y'_1\rangle & |Y'_2\rangle & |Y'_3\rangle \end{bmatrix},$$

by changing the auxiliary states $\{|X_i\rangle, |Y_i\rangle\} \mapsto \{|X'_i\rangle, |Y'_i\rangle\}$ *for* $i = 1,2,3$ *while leaving the optimal PIJSs* $|X^\star\rangle, |Y^\star\rangle$ *intact.*

3. *due to a change in both of the above auxiliary states.*

Note that, the first rule doesn't require the knowledge of the factorization. In that case, given an optimal unitary, an alternate solution can be found by simply post-multiplying the former (known one) by $\mathbb{1}_2 \otimes \Gamma_{\psi_0^\perp}$.

6.3.4 Finding an optimal unitary when Eve's initial state changes

The global unitary evolves the joint system as follows:

$$\mathcal{U}^{\mathbf{M}}_{\mathbf{IS}=e}\,|a\rangle_A|e\rangle_E \;=\; |S_a\rangle^{\mathbf{M}}_{AE}.$$

For $a \in \{x,y\}$, the PIJSs $S_a \in \{X,Y\}$ gets fixed by fixing the measurement setup $\mathbf{M}$. However, the same PIJS $|S\rangle^{\mathbf{M}}_{AE}$ can be produced for a different **IS** and by a different unitary:

$$\mathcal{U}^{\mathbf{M}}_{\mathbf{IS}=f}\,|a\rangle_A|f\rangle_E \;=\; |S_a\rangle^{\mathbf{M}}_{AE}.$$

Given an unitary $\mathcal{U}^{\mathbf{M}}_{\mathbf{IS}=e}$, one can find an unitary $\mathcal{U}^{\mathbf{M}}_{\mathbf{IS}=f}$ by knowing the local unitary that transforms $|e\rangle \to |f\rangle$.

Theorem 6.2. *If an unitary* $\mathcal{U}_e$ *is known for some initial state* $|e\rangle$*, one can find an unitary* $\mathcal{U}_f$ *for some other IS* $|f\rangle$*, just by knowing the local unitary* T_{ef} *that transforms* $|e\rangle \mapsto |f\rangle$*.*

$$\mathcal{U}_f = \mathcal{U}_e\left(\mathbb{1}^A_2 \otimes T^{E\dagger}_{ef}\right). \tag{6.3.1}$$

Proof. Since $|f\rangle = T_{ef}|e\rangle$, we get

$$\begin{aligned}\mathcal{U}_f\ |a\rangle_A|f\rangle_E = |S_a\rangle_{AE} &= \mathcal{U}_e\ |a\rangle_A|e\rangle_E \\ &= \mathcal{U}_e\ |a\rangle_A \otimes T_{ef}^{\dagger}|f\rangle_E \\ &= \mathcal{U}_e\left(\mathbb{1}_2^A \otimes T_{ef}^{E\dagger}\right)\ |a\rangle_A|f\rangle_E.\end{aligned}$$

And the result follows. □

Let's understand the theorem through some examples.

For instance, consider the task to find an optimal unitary for the IS $|\Delta\rangle_E := |\Delta_{xy}\rangle_{E_1}|\Delta_{uv}\rangle_{E_2}$, which is a small tweak $T_{ef} = \mathbb{1}_2 \otimes \mathbb{H}\ :\ |\Delta^{\mathbb{H}}\rangle_E \mapsto |\Delta\rangle_E$ of the earlier IS $|\Delta^{\mathbb{H}}\rangle_E$ (6.2.6). Then, the global unitary is transformed as follows

$$\mathcal{U}_{\Delta}^{\mathcal{C}} \ = \ \mathcal{U}_{\Delta^{\mathbb{H}}}^{\mathcal{C}}(\mathbb{1}_2^A \otimes \mathbb{1}_2^{E_1} \otimes \mathbb{H}^{E_2}).$$

The corresponding matrix is a tweak of the one in Eq. (6.2.8) while each inner sub-matrix $(\mathbb{1}_2, \sigma_x)$ gets post-multiplied by the Hadamard transformation $\mathbb{H}$.

A more involved example would be finding an optimal unitary for the IS $|\phi_{xy}^{+}\rangle := \frac{|00\rangle+|11\rangle}{\sqrt{2}}$ which is maximally entangled Bell-state. Since this new state can be obtained by applying an unitary $T = c\text{-}\sigma_x \cdot (\mathbb{H} \otimes \mathbb{1}_2)$ (a Hadamard on the first qubit followed by a CNOT operation) on the state $|00\rangle$, an optimal unitary for the former state ($\mathcal{U}_{\phi^+}^{\mathcal{C}}$) can be obtained from an optimal unitary for the later state ($\mathcal{U}_{00}^{\mathcal{C}}$) by applying the following unitary transform

$$\frac{1}{\sqrt{2}}\begin{bmatrix}\mathbb{1}_2 & \cdot \\ \cdot & \sigma_x\end{bmatrix}\begin{bmatrix}\mathbb{1}_2 & \mathbb{1}_2 \\ \mathbb{1}_2 & -\mathbb{1}_2\end{bmatrix} \ = \ \frac{1}{\sqrt{2}}\begin{bmatrix}\mathbb{1}_2 & \mathbb{1}_2 \\ \sigma_x & -\sigma_x\end{bmatrix}.$$

6.3.5 Finding an optimal unitary when Eve's measurement setup changes

Here we wish to understand the change in the optimal unitary as Eve's measurement setup changes from the computational basis to a different one.

Theorem 6.3. *Consider a different measurement basis $\{|E_\lambda\rangle\}$ which is a unitary transformation $|E_\lambda\rangle = \mathbf{M}_{xy}|\lambda\rangle$ of the computational basis chosen earlier. Then, the following retrospective effects could be observed on the optimal IVs, the optimal PIJSs, and the optimal global unitary.*

1. *The optimal IVs of Eve are changed as follows:*

$$|\boldsymbol{IV}_{xy}^{\star}\rangle^{\mathbf{M}} \ = \ \mathbf{M}_{xy}|\boldsymbol{IV}_{xy}^{\star}\rangle^{\mathcal{C}}. \tag{6.3.1}$$

2. *The optimal PIJSs are transformed as follows:*

$$|S_a\rangle^{\mathbf{M}} \;=\; (\mathbb{1}_2 \otimes \mathbf{M}_{xy})|S_a\rangle^{\mathcal{C}}, \;\; a = x, y. \tag{6.3.2}$$

3. *The global unitary gets tweaked as follows:*

$$\mathcal{U}^{\mathbf{M}} \;=\; (\mathbb{1}_2 \otimes \mathbf{M}_{xy})\,\mathcal{U}^{\mathcal{C}}. \tag{6.3.3}$$

Proof. The first two claims are straight-forward, while the last claim is proved below.

$$\begin{aligned}\mathcal{U}^{\mathbf{M}}\,|0\rangle_A|\psi_0\rangle_E = |X\rangle^{\mathbf{M}} &= (\mathbb{1}_2 \otimes \mathbf{M}_{xy})\,|X\rangle^{\mathcal{C}} \\ &= (\mathbb{1}_2 \otimes \mathbf{M}_{xy})\,\mathcal{U}^{\mathcal{C}}\,|0\rangle_A|\psi_0\rangle_E.\end{aligned}$$

We have used $\mathbf{M}$ and $\mathbf{M}_{xy}$ interchangeably for Eve's measurement. □

Note that, a permutation of the measurement basis (*e.g.*, $\mathbf{M}_{xy} = [|00\rangle, |11\rangle, |10\rangle, |01\rangle]$) or a phase shift (*e.g.*, $|E_\lambda\rangle \mapsto -|E_\lambda\rangle$) doesn't change the measurement statistics, despite such a small tweak ($\mathbb{1}_2 \otimes \mathbf{M}_{xy}$) on the optimal unitary. Thus, we consider the overall effect due to such changes as equivalent. An effective change in the measurement basis corresponds to those unitary transformations on the 4-dimensional computational basis, which consists of at least a row having more than one non-zero entries.

6.4 Tables: PIJSs and Optimal unitaries

Consider the eavesdropping setup when the initial state of Eve's ancilla is $|00\rangle$. For this IS, consider two different (four-dimensional) measurement setups for Eve:

1. she measures in the computational basis, denoted by $\mathcal{C}$
2. she measures in some $\mathcal{R}$-rotated computational basis, denoted by $\mathcal{M}$

The corresponding optimal PIJSs and the optimal unitaries are listed here for these two different scenario.

The rotation $\mathcal{R}$ is chosen here as

$$\mathcal{R} \;=\; \begin{pmatrix} \frac{1}{3} & -\frac{\sqrt{2}}{3} & -\frac{\sqrt{2}}{3} & \frac{2}{3} \\ \frac{\sqrt{2}}{3} & \frac{1}{3} & -\frac{2}{3} & -\frac{\sqrt{2}}{3} \\ \frac{\sqrt{2}}{3} & -\frac{2}{3} & \frac{1}{3} & -\frac{\sqrt{2}}{3} \\ \frac{2}{3} & \frac{\sqrt{2}}{3} & \frac{\sqrt{2}}{3} & \frac{1}{3} \end{pmatrix} \tag{6.4.1}$$

One can experiment with their own choices though.

6.4.1 When Eve measures in the 4-d computational basis

The PIJSs and the optimal unitary is listed here.

6.4.1.1 The PIJSs when Eve measures in the computational basis

Consider the vectorial representation of the eight-dimensional PIJSs in *xy* basis in terms of the QBER across the two bases.

Table 6.1 | The PIJSs when Eve measures in the computational basis.

The IS is $|00\rangle$. The PIJSs $|X\rangle^{\mathcal{C}}, |Y\rangle^{\mathcal{C}}$ are described in the xy basis.

$\|X\rangle^{\mathcal{C}}$	$\|Y\rangle^{\mathcal{C}}$
$\begin{pmatrix} \sqrt{F_{xy}}\ \ \mathscr{D}_{uv}^{+} \\ \sqrt{F_{xy}}\ \ \mathscr{D}_{uv}^{-} \\ 0 \\ 0 \\ 0 \\ 0 \\ \sqrt{D_{xy}}\ \ \mathscr{D}_{uv}^{+} \\ \sqrt{D_{xy}}\ \ \mathscr{D}_{uv}^{-} \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ \sqrt{D_{xy}}\ \ \mathscr{D}_{uv}^{-} \\ \sqrt{D_{xy}}\ \ \mathscr{D}_{uv}^{+} \\ \sqrt{F_{xy}}\ \ \mathscr{D}_{uv}^{-} \\ \sqrt{F_{xy}}\ \ \mathscr{D}_{uv}^{+} \\ 0 \\ 0 \end{pmatrix}$

6.4.1.2 The optimal unitary when Eve measures in the computational basis

An optimal unitary $\mathcal{U}_{00}^{\mathcal{C}}$ for the IS as $|00\rangle$ and Eve measures in the four-dimensional computational basis. Note that, it is parameterized in terms of the QBER across the two bases.

These are the same parameters to describe the PIJSs.

$$\mathcal{U}^{\mathcal{C}}_{00} = \begin{pmatrix} \sqrt{F_{xy}} \begin{bmatrix} \mathscr{D}^{+}_{uv} & -\mathscr{D}^{-}_{uv} \\ \mathscr{D}^{-}_{uv} & \mathscr{D}^{+}_{uv} \end{bmatrix} & \sqrt{D_{xy}} \begin{bmatrix} \mathscr{D}^{+}_{uv} & -\mathscr{D}^{-}_{uv} \\ \mathscr{D}^{-}_{uv} & \mathscr{D}^{+}_{uv} \end{bmatrix} & \mathbb{O}_2 & \mathbb{O}_2 \\ \mathbb{O}_2 & \mathbb{O}_2 & \sqrt{D_{xy}} \begin{bmatrix} \mathscr{D}^{-}_{uv} & \mathscr{D}^{+}_{uv} \\ \mathscr{D}^{+}_{uv} & -\mathscr{D}^{-}_{uv} \end{bmatrix} & -\sqrt{F_{xy}} \begin{bmatrix} \mathscr{D}^{-}_{uv} & \mathscr{D}^{+}_{uv} \\ \mathscr{D}^{+}_{uv} & -\mathscr{D}^{-}_{uv} \end{bmatrix} \\ \mathbb{O}_2 & \mathbb{O}_2 & \sqrt{F_{xy}} \begin{bmatrix} \mathscr{D}^{-}_{uv} & \mathscr{D}^{+}_{uv} \\ \mathscr{D}^{+}_{uv} & -\mathscr{D}^{-}_{uv} \end{bmatrix} & \sqrt{D_{xy}} \begin{bmatrix} \mathscr{D}^{-}_{uv} & \mathscr{D}^{+}_{uv} \\ \mathscr{D}^{+}_{uv} & -\mathscr{D}^{-}_{uv} \end{bmatrix} \\ \sqrt{D_{xy}} \begin{bmatrix} \mathscr{D}^{+}_{uv} & -\mathscr{D}^{-}_{uv} \\ \mathscr{D}^{-}_{uv} & \mathscr{D}^{+}_{uv} \end{bmatrix} & -\sqrt{F_{xy}} \begin{bmatrix} \mathscr{D}^{+}_{uv} & -\mathscr{D}^{-}_{uv} \\ \mathscr{D}^{-}_{uv} & \mathscr{D}^{+}_{uv} \end{bmatrix} & \mathbb{O}_2 & \mathbb{O}_2 \end{pmatrix}$$

Here, $\mathbb{O}_2$ represents null submatrix of order 2.

One can now easily read the eight-dimensional PIJSs along with their auxiliary orthonormal counterparts in the same Hilbert space from the columns of the unitary. Orthonormality is an easy verification here.

6.4.2 When Eve measures in the $\mathcal{R}$-rotated 4-d computational basis

The PIJSs and the optimal unitary is listed.

6.4.2.1 PIJSs when Eve measures in the $\mathcal{R}$-rotated computational basis

Notice how the nature of the rotation on the measurement setup ensures that the PIJSs share non-zero overlap with every direction in the eight-dimensional Hilbert space (i.e., the 8d-vector has more non-zero components due to complicated rotations). For such PIJSs, it is quite difficult to merely guess their orthonormal auxiliary counterparts. We have discussed earlier how to tackle such situations.

Table 6.2 | The PIJSs when Eve measures in the $\mathcal{R}$-rotated computational basis.

The IS is $|00\rangle$. The PIJSs $|X\rangle^{\mathcal{M}}, |Y\rangle^{\mathcal{M}}$ are described in the xy basis.

$\vert X\rangle^{\mathcal{M}}$	$\vert Y\rangle^{\mathcal{M}}$
$\begin{pmatrix} \sqrt{F_{xy}}\ (\frac{1}{3}\mathscr{D}^{+}_{uv} - \frac{\sqrt{2}}{3}\mathscr{D}^{-}_{uv}) \\ \sqrt{F_{xy}}\ (\frac{\sqrt{2}}{3}\mathscr{D}^{+}_{uv} + \frac{1}{3}\mathscr{D}^{-}_{uv}) \\ \sqrt{F_{xy}}\ (\frac{\sqrt{2}}{3}\mathscr{D}^{+}_{uv} - \frac{2}{3}\mathscr{D}^{-}_{uv}) \\ \sqrt{F_{xy}}\ (\frac{2}{3}\mathscr{D}^{+}_{uv} + \frac{\sqrt{2}}{3}\mathscr{D}^{-}_{uv}) \\ \sqrt{D_{xy}}\ (-\frac{\sqrt{2}}{3}\mathscr{D}^{+}_{uv} + \frac{2}{3}\mathscr{D}^{-}_{uv}) \\ \sqrt{D_{xy}}\ (-\frac{2}{3}\mathscr{D}^{+}_{uv} - \frac{\sqrt{2}}{3}\mathscr{D}^{-}_{uv}) \\ \sqrt{D_{xy}}\ (\frac{1}{3}\mathscr{D}^{+}_{uv} - \frac{\sqrt{2}}{3}\mathscr{D}^{-}_{uv}) \\ \sqrt{D_{xy}}\ (\frac{\sqrt{2}}{3}\mathscr{D}^{+}_{uv} + \frac{1}{3}\mathscr{D}^{-}_{uv}) \end{pmatrix}$	$\begin{pmatrix} \sqrt{D_{xy}}\ (-\frac{\sqrt{2}}{3}\mathscr{D}^{-}_{uv} + \frac{2}{3}\mathscr{D}^{+}_{uv}) \\ \sqrt{D_{xy}}\ (-\frac{2}{3}\mathscr{D}^{-}_{uv} - \frac{\sqrt{2}}{3}\mathscr{D}^{+}_{uv}) \\ \sqrt{D_{xy}}\ (\frac{1}{3}\mathscr{D}^{-}_{uv} - \frac{\sqrt{2}}{3}\mathscr{D}^{+}_{uv}) \\ \sqrt{D_{xy}}\ (\frac{\sqrt{2}}{3}\mathscr{D}^{-}_{uv} + \frac{1}{3}\mathscr{D}^{+}_{uv}) \\ \sqrt{F_{xy}}\ (\frac{1}{3}\mathscr{D}^{-}_{uv} - \frac{\sqrt{2}}{3}\mathscr{D}^{+}_{uv}) \\ \sqrt{F_{xy}}\ (\frac{\sqrt{2}}{3}\mathscr{D}^{-}_{uv} + \frac{1}{3}\mathscr{D}^{+}_{uv}) \\ \sqrt{F_{xy}}\ (\frac{\sqrt{2}}{3}\mathscr{D}^{-}_{uv} - \frac{2}{3}\mathscr{D}^{+}_{uv}) \\ \sqrt{F_{xy}}\ (\frac{2}{3}\mathscr{D}^{-}_{uv} + \frac{\sqrt{2}}{3}\mathscr{D}^{+}_{uv}) \end{pmatrix}$

6.4.2.2 The optimal unitary when Eve measures in the $\mathcal{R}$-rotated computational basis

An optimal unitary $\mathcal{U}_{00}^{\mathcal{M}}$ for IS $|00\rangle$ and Eve measures in the $\mathcal{R}$-rotated computational basis.

Let us rewrite the optimal unitary $\mathcal{U}_{00}^{\mathcal{C}}$ as a block-matrix

$$\mathcal{U}_{00}^{\mathcal{C}} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

Then, we can express the optimal unitary $\mathcal{U}_{00}^{\mathcal{M}}$ as the following block-matrix

$$\mathcal{U}_{00}^{\mathcal{M}} = \begin{pmatrix} \mathcal{R}A & \mathcal{R}B \\ \mathcal{R}C & \mathcal{R}D \end{pmatrix} \tag{6.4.1}$$

For instance, the submatrix $\mathcal{R}A$ can be written as the following.

$$\mathcal{R}A = \begin{pmatrix} \sqrt{F_{xy}}\,(\frac{1}{3}\mathscr{D}_{uv}^{+} - \frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{-}) & \sqrt{F_{xy}}\,(-\frac{1}{3}\mathscr{D}_{uv}^{-} - \frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{+}) & \sqrt{D_{xy}}\,(\frac{1}{3}\mathscr{D}_{uv}^{+} - \frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{-}) & \sqrt{D_{xy}}\,(-\frac{1}{3}\mathscr{D}_{uv}^{-} - \frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{+}) \\ \sqrt{F_{xy}}\,(\frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{+} + \frac{1}{3}\mathscr{D}_{uv}^{-}) & \sqrt{F_{xy}}\,(-\frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{-} + \frac{1}{3}\mathscr{D}_{uv}^{+}) & \sqrt{D_{xy}}\,(\frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{+} + \frac{1}{3}\mathscr{D}_{uv}^{-}) & \sqrt{D_{xy}}\,(-\frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{-} + \frac{1}{3}\mathscr{D}_{uv}^{+}) \\ \sqrt{F_{xy}}\,(\frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{+} - \frac{2}{3}\mathscr{D}_{uv}^{-}) & \sqrt{F_{xy}}\,(-\frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{-} - \frac{2}{3}\mathscr{D}_{uv}^{+}) & \sqrt{D_{xy}}\,(\frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{+} - \frac{2}{3}\mathscr{D}_{uv}^{-}) & \sqrt{D_{xy}}\,(-\frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{-} - \frac{2}{3}\mathscr{D}_{uv}^{+}) \\ \sqrt{F_{xy}}\,(\frac{2}{3}\mathscr{D}_{uv}^{+} + \frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{-}) & \sqrt{F_{xy}}\,(-\frac{2}{3}\mathscr{D}_{uv}^{-} + \frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{+}) & \sqrt{D_{xy}}\,(\frac{2}{3}\mathscr{D}_{uv}^{+} + \frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{-}) & \sqrt{D_{xy}}\,(-\frac{2}{3}\mathscr{D}_{uv}^{-} + \frac{\sqrt{2}}{3}\mathscr{D}_{uv}^{+}) \end{pmatrix}$$

Notice how the optimal unitary is now filled up with all non-zero entries, parameterized by the QBERs. A symmetric attack is dealt trivially.

6.5 Conclusion

For practical purposes, all an eavesdropper requires in hand is i) the optimal unitary to evolve the joint system and ii) the corresponding measurement that she must perform to glean the maximum possible information within the error-threshold. We have developed the methods to fully characterize the optimal unitaries and demonstrated the techniques via examples. Our approach could figure out the simplest one out of the infinite family of optimal unitaries in a most natural fashion.

As an optimal unitary is parameterized by the error-rate, an attacker may first fix the QBER she wishes to introduce and accordingly design an optimal unitary (not unique) for a specific choice of her measurement setup and the initial state (note that the measurement for the other (conjugate basis) gets fixed automatically). An attacker would like to choose such unitaries which, if feasible, is easier to design than its siblings, considering the IS-unitary designing trade-off. As a further work, an interested reader may explore the design of the optimal unitaries in terms of universal quantum gates.

So far, including the earlier chapters, we essentially have characterized the optimal attacks on the BB84 protocol exhaustively, where an attacker entangles a four dimensional probe per transmitted qubit. An optimal attack consists of two main components: the optimal unitary evolution, and the optimal measurement. The intermediate results for the analysis purposes involve the study of optimal interaction vectors. We have characterized all these components of an optimal attack exhaustively. We have considered the generalized asymmetric error rates across the two MUBs, while a symmetric attack automatically becomes a special case.

CHAPTER 7

POST-PROCESSING AND COMPARATIVE STUDY

If the disturbance D is within the threshold, the one-way classical post-processing can be considered to filtrate into a secret key. Here, we calculate and plot some post-processing related parameters for the sake of completeness. Although we have already calculated the key-rate for the sifted key, the final key-rate is also relevant to compute. For our optimal states [AP17, AP21, FGG$^+$97] for incoherent eavesdropping on the 4s [BB84, BB14] protocol, we have computed the related parameters. Then, we graphically compared them with the 6s protocol [Bru98] as well.

7.0.1 Classical post-processing

It's a two-step procedure as stated below.

Making Alice and Bob's bit-streams in sync

The sifted key with Bob may disagree with that of Alice in the presence of noise (eavesdropping). A simple method is discussed here to locate such erroneous positions.[1] They may either discard or correct those bits.

1. They divide their strings into blocks of some agreed length. The length L is so chosen that each block is unlikely to have more than one error ($DL \ll 1$, for QBER=D). For each block, they locate an error as follows.

2. They compute the XOR (*parity*) of the bits in a block. If the parity disagree, they know that an error occurred, and proceed as follows to locate it.

[1] More involved approach can be found in [BS94].

3. They perform a binary search in the same manner as above until they locate an error if any. They halve the block in two sub-blocks, compute the parity and tally it publicly. In not matched, they continue with that sub-block for a binary search of the erroneous bit. .

4. They locate any error in each of the blocks.

5. Several such rounds of varying block length may require to locate all the errors with high probability.

If they can afford an encrypted authenticated communication, they can correct the errors without leaking any information. Otherwise, they discard the errors. In case they correct errors, Alice can simply send Bob a correct bit against each error-bit. For err-discard, as their parity checking is public, they should discard a bit of the block under consideration throughout the recursive process.

For QBER=D, Bob knows $I_{AB} = 1 - h(D)$ fraction of Alice's bit-string correctly. The remaining $H(D)$ fraction is erroneous due to Eve, and is either discarded or corrected. In case of err-discard, they remove at least $H(D)$ fraction of the sifted key. In practice, the procedure reduces the string even more as they need to sacrifice some bits due to public discussion.

Making Eve's knowledge arbitrarily small

This is the last step of classical post-processing to virtually eliminate any knowledge of Eve on the key. The estimated value of the QBER tells the legitimate parties about the maximum amount of knowledge (in number of bits, say k) that Eve has on the key (from I_{AE}^R). For an n-bit key, the legitimate parties can then reduce the key to an $(n-k-s)$-bit string on which Eve is left with merely $\mathcal{O}(2^{-s})$ bits of information. To be precise [BBCM95],

$$I_{\mathrm{E}}^{\mathrm{fin}} \;\leq\; \log_2(2^{-s}+1) \approx \frac{2^{-s}}{\ln 2}.$$

Consequently, Eve's information can be made arbitrarily small by tuning the security parameter s. For $s = 0$, Eve still knows 1 bit of information. Thus, $s > 0$ is a safe choice. Following is a simple way to compress the key. Alice and Bob can choose at random $n-k-s$ subsets of the key, and replace each subset by their XOR value to arrive at the final key. Choosing the subsets is a challenge. In practice, choosing a hash function randomly from a family of 2-universal hash function does the job well.

The key-rate for the final key, for individual attacks, depends on the two things: $MI_{\mathtt{AB}}$, and the discarded fraction $\tau(D)$. The later one is defined by the amount of fraction discarded during privacy amplification. It could be calculated from the collision probability

P_c of the reconciled key as follows

$$\tau(D) \quad := \quad 1 + \log_2 \langle P_c \rangle.$$

The average collision probability per bit of the reconciled key is defined on the posterior distribution of the reconciled key as follows

$$\langle P_c \rangle \quad := \quad \sum \Pr(\text{outcome}) \sum P_{post}^2.$$

Using the bounds on P_c, and expression for $\tau(D)$ one can calculate the key-rate for the final key as follows.

$$\begin{aligned} R_{corr} &:= I_{\mathtt{AB}}(D) - \tau(D), \\ R_{disc} &:= I_{\mathtt{AB}}(D) - (1-D)\tau(D) - D, \\ I_{\mathtt{AB}} &= 1 - H(D). \end{aligned}$$

To get the key-rate on the transmitted signal-stream, one can simply multiply the above rates by $\frac{1}{2}$ for B92,BB84, and by $\frac{1}{3}$ for 6s protocol.

For individual attacks with err-discard, the tolerable error-rates become $4\%, 10.5\%, 12\%$ for the three protocols respectively.

Note: Even when Bob's information becomes less than that of Eve's, there is some scope to filter a secret key, albeit inefficient and leading to less throughput. The legitimate parties can perform a two-way process called *advantage distillation* to allow Bob accumulating more information than Eve. Then, they can perform the classical post-processing as usual.

For instance, Alice can mask the bits in a block by a random bit and sends to Bob who then XOR his block with that of Alice. The secret bit is then retrieved only when the blocks are identical. In such cases the masking bits for various blocks define a secret key.

7.1 Rényi information, Discarded Fraction, final Key-rate [AP21]

The literature on the classical post-processing is found to lack clarity, and often encountered with conflicting ideas. We found that Rényi information is sometimes computed apart from Shannon information without any further inference (e.g., [BPG99] etc). It appears in the bound during privacy amplification. During classical post-processing, *discarded fraction* (DF) is a parameter found in use, to be calculated in order to compute the final key-rate.

Shannon information (SI) may not necessarily provide the whole picture of mutual information for quantum measurements [BZ01], and Rényi information [Rén61] is another relevant measure. Some extra care is necessary to deal with it [BBCM95], as it looses some usual notions like drop in entropy, symmetry etc. More alternatives can be found in [Ras19].

7.1.1 Rényi information for the 4s and 6s protocols

Rényi information (RI) between Eve and Alice (or Bob) is defined as follows.

$$RI = R_0 - \sum_{\lambda} P_\lambda R_\lambda .$$

Here, R_0 and R_λ denotes the Rényi entropy before and after Eve measures, respectively. P_λ denotes the frequency of each outcome λ with Eve.

We consider here the Rényi information of order 2. When the signal was prepared in the xy basis, R_0 and R_λ are defined in terms of prior (p_s) and posterior ($P_{s|\lambda}$) probabilities as follows.

$$\begin{aligned} R_0 &= -\log_2\left(p_x^2 + p_y^2\right), \\ R_\lambda &= -\log_2\left(P_{x|\lambda}^2 + P_{y|\lambda}^2\right). \end{aligned}$$

For 4s BB84 and 6s protocol, due to equal prior ($p_x = p_y = \frac{1}{2}$), $R_0 = 1$. Then,

$$RI_{xy} = 1 - \sum_{\lambda} P_\lambda R_\lambda .$$

For the 4s protocol with [AP21, FGG$^+$97] optimal states in xy basis, the posterior probabilities become

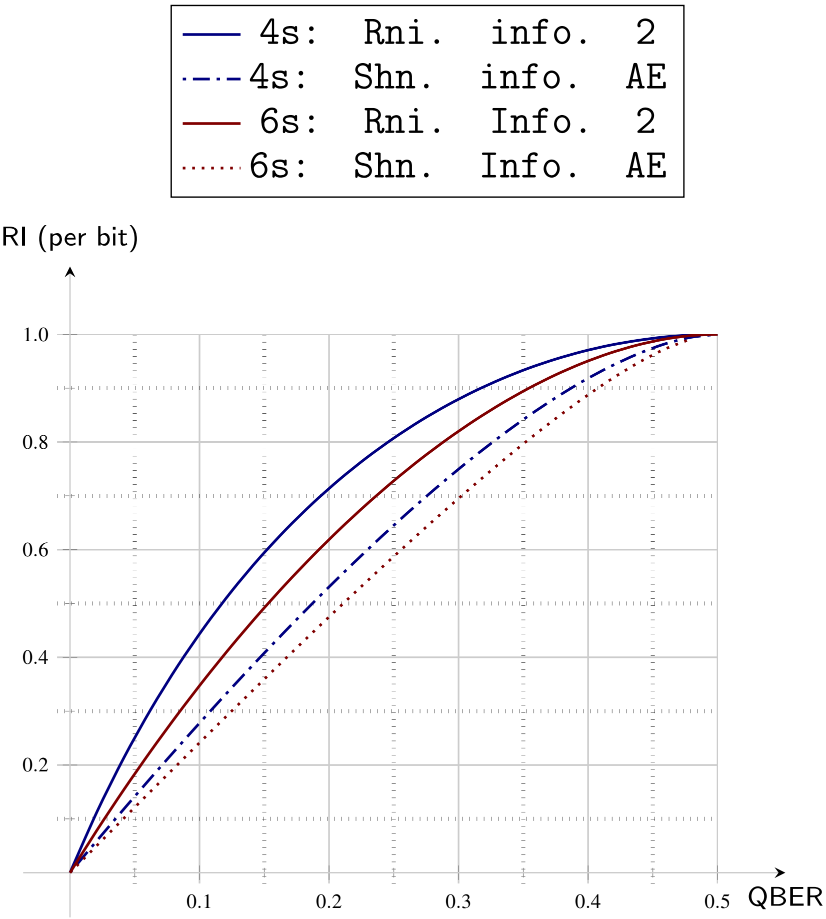
$$(P_{x|\lambda}, P_{y|\lambda}) = \left(\frac{1}{2} \pm \sqrt{D_{uv}(1-D_{uv})},\ \frac{1}{2} \mp \sqrt{D_{uv}(1-D_{uv})}\right), \quad \forall\lambda.$$

Thus, the Rényi information between Eve and Bob becomes

$$
\begin{aligned}
RI_{xy}^{(2)} &= 1+\log_2\left(\frac{1}{2}+2D_{uv}(1-D_{uv})\right)\\
&= \log_2(1+4D_{uv}-4D_{uv}^2).
\end{aligned}
$$

Figure 7.1 | Rényi Information (RI) for incoherent attack on 4s, 6s protocol.

Considered the optimal states of [AP21, FGG$^+$97] for the 4s protocol and those for the 6s [BPG99] protocol. For 4s: the Rényi Information of order 2 meets the Lütkenhaus bound for discarded fraction.



For equal errors across the two MUBs, Rényi information of 2nd order between Eve and Bob for 4s protocol with [AP21, FGG$^+$97] optimal states becomes

$$
RI_{4s}^{(2)} = \log_2(1+4D-4D^2).
$$

For 6s [BPG99, Incoeh.] optimal states,

$$\begin{aligned} RI_{6s}^{(2)} &= 1+(1-D)\log_2(p_f^2+p_s^2), \\ p_f &= \frac{1}{2}\left[1-\frac{\sqrt{D(2-3D)}}{1-D}\right], \\ p_s &= 1-p_f. \end{aligned}$$

Here, p_f, p_s are failure and success probabilities of Eve.

The SI *versus* RI for optimal attack on 4s and 6s protocol are plotted in Fig. 7.1.

7.1.2 Discarded fraction for our optimal states

Denoted by τ, it is defined as the fraction by which the key is shortened during privacy amplification. Based on a criteria for strong security [BBCM95], a condition on the DF in order to reduce Eve's knowledge (Shannon information) on the final key is as follows [HBHP08, Lüt96]. For individual attack, for an error-rate D, Eve knows less than $\frac{1}{\ln 2}$ bits of the final key provided

$$\tau_D \geq 1+\log_2\langle P_2^{(c)}\rangle.$$

Here, $\langle P_2^{(c)}\rangle$ is the maximum *average collision probability* of Eve's knowledge per bit of the reconciled key. When Alice prepares her signals in the xy basis, the collision probability of order 2 averaged over Eve's measurement knowledge λ, is defined as follows.

$$\langle P_2^{(c)}\rangle = \sum_{\lambda} P_\lambda \left(P_{x|\lambda}^2+P_{y|\lambda}^2\right).$$

However, in the general scenario of individual attacks, the DF is bounded above by Lütkenhaus bound [HBHP08, Lüt99]

$$\tau_D \leq \log_2(1+4D-4D^2). \tag{7.1.1}$$

Note that the lower bound on the *discarded fraction* coincides with the RI of order 2. Therefore, for the [AP21, FGG$^+$97] optimal states, the maximum discarded fraction saturates the Lütkenhaus bound.

Success probability on the final key Consider a simple post-processing strategy [BPG99, 6s-coeh]. Let the legitimate parties consider blocks of size two. In the reconciled key, they replace the two bits by their XOR sum. Eve's probability to correctly guess the bit is

$$\begin{aligned}
\mathsf{Pr(success)}^{fin} &= P_s^2 + P_f^2 \\
&= (\frac{1}{2} + \sqrt{D(1-D)})^2 + (\frac{1}{2} - \sqrt{D(1-D)})^2 \\
&= \frac{1}{2}(1+4D-4D^2).
\end{aligned}$$

7.1.3 Final Key-rate for our optimal states

From the expression of $\tau(D)$ as calculated above, one can calculate the key-rate (scaled w.r.t. raw-key-len) for the final key as follows.

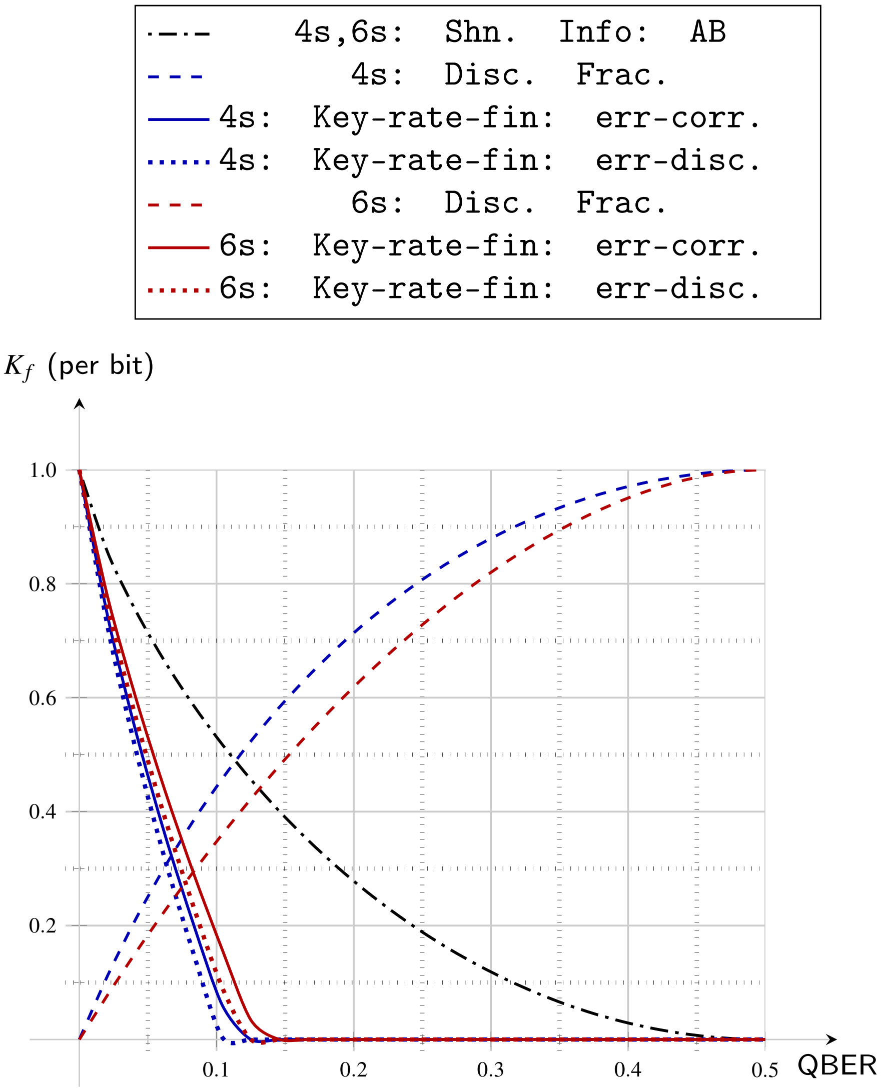
$$\begin{aligned}
R_{corr} &:= I_{\mathtt{AB}}(D) - \tau(D) \\
&= 1 - h(D) - \log_2(1+4D-4D^2), \quad \texttt{for 4s protocol} \\
R_{disc} &:= I_{\mathtt{AB}}(D) - (1-D)\tau(D) - D \\
&= 1 - h(D) - (1-D)\log_2(1+4D-4D^2) - D, \quad \texttt{for 4s protocol.}
\end{aligned}$$

With the bounds on $\tau(D)$ as calculated for the 4s and the 6s protocols, we have plotted here the key-rates. As evident, the key-rates become the following: i) when error discarded: 10.5% for 4s, 12% for 6s, and ii) when error corrected: 11.5% for 4s, 13% for 6s.

These are plotted in Fig. 7.2.

Figure 7.2 | **Final key-rate for 4s and 6s protocols.**

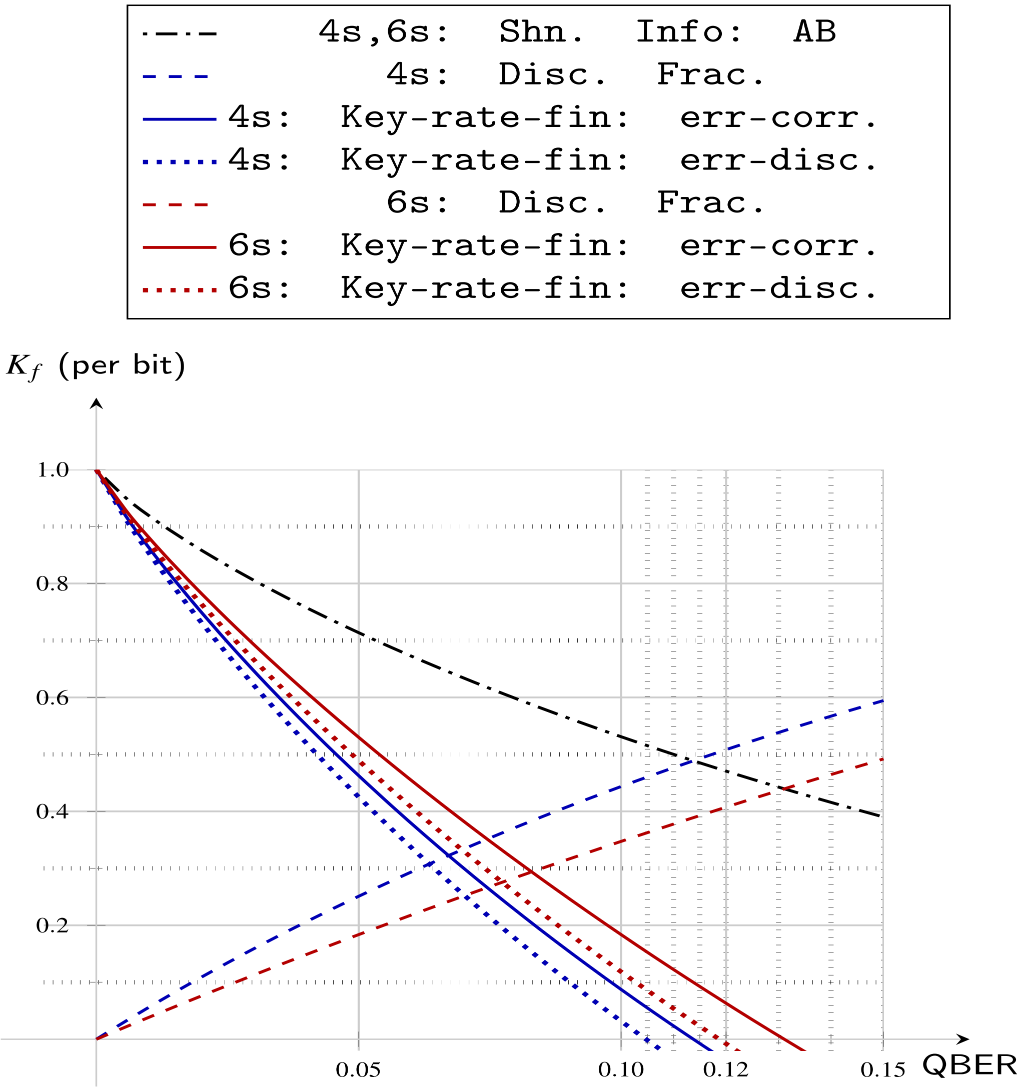
When error discarded and corrected



A more closer view is plotted in Fig. 7.3.

Figure 7.3 | Final key-rate for 4s and 6s: close look.

When error discarded: 10.5% for 4s, 12% for 6s When error corrected: 11.5% for 4s, 13% for 6s



7.2 Comparative study: across protocols

Here we graphically compare the bipartite mutual informations (between Alice-Eve and Alice-Bob) for the 4s and 6s protocols. Then, Rényi information of order 2 is also compared for these two protocols.

7.2.1 Mutual information (AE,AB) for 4s, 6s, and 4MUBs with high-dimensional states

We consider the following protocols to study the mutual information comparatively for a given QBER D.

For BB84, maximum bipartite mutual information are as follows [FGG$^+$97, 4s-Indv.]:

$$\begin{aligned} I_{\mathtt{AB}} &= 1+D\log D+(1-D)\log(1-D), \\ I_{\mathtt{AE}} &= \frac{1}{2}\phi(2\sqrt{D(1-D)}), \\ &\quad \text{where,} \quad \phi(x) = (1-x)\log(1-x)+(1+x)\log(1+x). \end{aligned}$$

For 6s protocol, maximum bipartite mutual information are as follows [Bru98, 6s-Indv.]:

$$\begin{aligned} I_{\mathtt{AB}} &= 1+D\log D+(1-D)\log(1-D), \\ I_{\mathtt{AE}} &= 1+(1-D)\left[f(D)\log f(D)+(1-f(D))\log(1-f(D))\right], \\ &\quad \text{where,} \quad f(D) = \frac{1}{2}\left(1-\frac{1}{1-D}\sqrt{D(2-3D)}\right). \end{aligned}$$

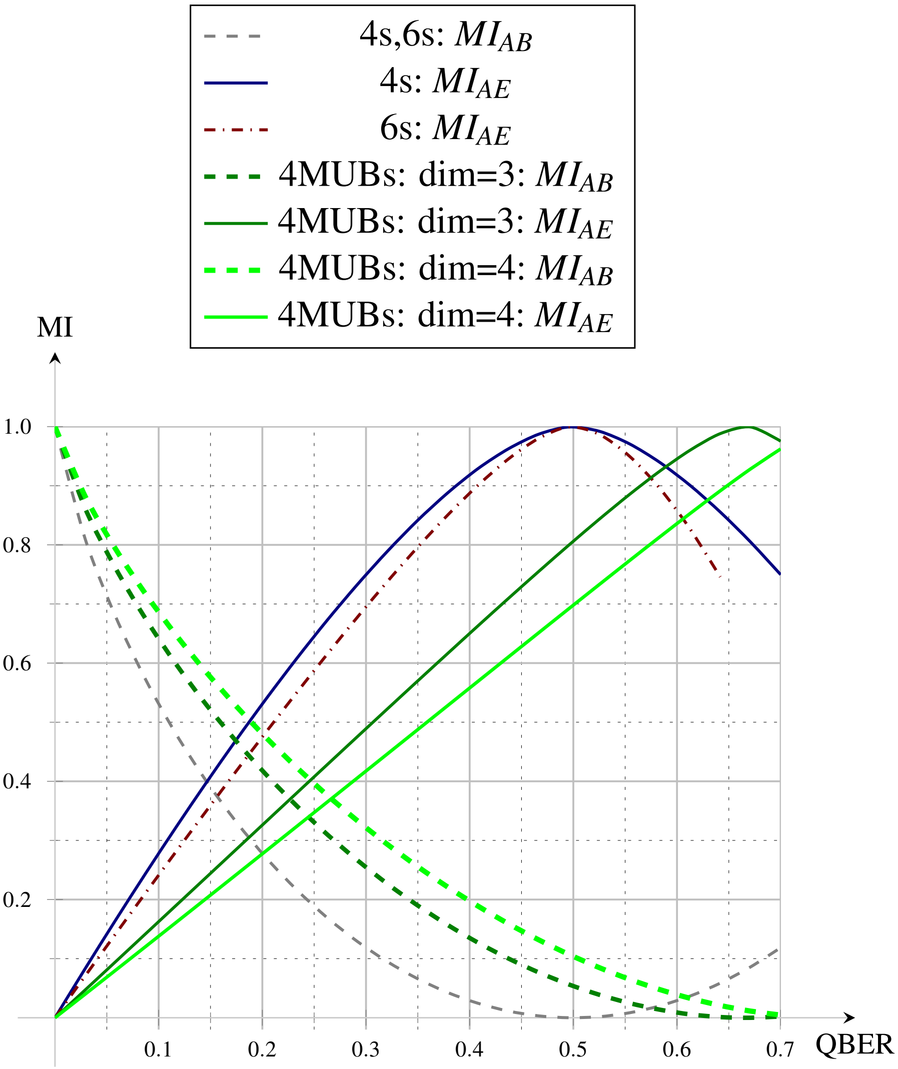
For a protocol with 4 MUBs having d-dimensional states (qudits) in each basis, maximum bipartite mutual information are as follows [BM02].

$$\begin{aligned} I_{\mathtt{AB}} &= 1+D\log_d(\frac{D}{d-1})+(1-D)\log_d(1-D), \\ I_{\mathtt{AE}} &= 1+(1-D)\left[f_d(D)\log_d f_d(D)+(1-f_d(D))\log_d\left(\frac{1-f_d(D)}{d-1}\right)\right], \\ &\quad \text{where,} \quad f_d(D) = \frac{d-2D+\sqrt{(d-2D)^2-d^2(1-2D)^2}}{d^2(1-D)}. \end{aligned}$$

So, more degrees of freedom by the legitimate parties decreases Eve's mutual information.

Figure 7.4 | Comparative mutual informations for BB84, 6s protocol, and a protocol with 4MUBs having d-dim bases.

Intersection of MI_{AB} and MI_{AE} provides the maximum tolerable disturbance for one-way post-processing. Although the tolerance level increases as dimension d increases, the key-rate also drops.



CHAPTER 8

COHERENT EAVESDROPPING

8.1 Eavesdropping on 4s protocol

Let's consider the four-state BB84 protocol. Now, consider Eve attacking two qubits at a time, i.e., coherently with an ancilla. A symmetric attack (i.e., same QBER across all bases) was first briefed in [CG97]. We sketch an outline of the attack, and then explain elaborately some of the intermediate results, particularly characterizing the parameters that defines the unitary. These are derivation of the results up to our understanding that didn't appear so straightforward.

Consider Eve attacking two qubits with an ancilla. In the incoherent case, while Eve has four possible states in a specific encoding basis, in the two-qubit coherent case, Eve has sixteen different states and therefore 256 inner products. However, we'll see shortly that the unitary is ultimately characterized by only five parameters. We'll discuss in depth-and-breath on that side.

The importance of that characterization lies on determining the optimal measurement for Eve, her probability to successfully find the message, the strategy that gives her the maximum amount of information etc.

8.2 Overview: the unitary is characterized by five real parameters

An unitary is defined by its action on the states in some basis. For a 2-qubit state, four such states are enough to be considered. Any such state, when evolved with some ancillary state, it produces a superposition of four different states. Thus, the attack on the basis-states produce sixteen possible states with Eve. The inner products between them define the unitary: 256 IPs are there.

Mere restriction of unitarity reduces the count to 240, not much useful. Like incoherent attack, an orthogonal grouping is possible that reduces the count dramatically. Four such groups, each with four states, produce 24 real IPs.

Then, the rules of the symmetry of the attack reduces the count to ten, and then further to five.

$$256 \quad \rightarrow \quad 24 \quad \rightarrow 10 \rightarrow 5$$

We'll first sketch an outline of those results and the detailed derivations will follow in a separate section.

8.2.1 Notations and conventions

8.2.1.1 Alice's bases and states

Alice uses one of the two bases z, and x. The states in these bases are denoted as $\{|z\rangle, |-z\rangle\}$, and $\{|x\rangle, |-x\rangle\}$, respectively.

For a 2-qubit chunk, the states in the zz basis are then $\{|zz\rangle, |z-z\rangle, |-zz\rangle, |-z-z\rangle\}$.

One may consider the binary notations as well. In any case, a basis may be denoted as $\beta \equiv \beta_1\beta_2$.

8.2.1.2 Bob's and Eve's states after an unitary interaction

For Alice's pair of qubits in state $|a_1a_2\rangle$, due to the unitary entanglement, Bob receives one of the basis states, say, $|b_1b_2\rangle$, and Eve's ancilla is in one of the states $|E^{a_1a_2}_{b_1b_2}\rangle$.

In decimal, the equivalent symbols are $|a\rangle$, $|b\rangle$, and $|E^a_b\rangle$, respectively.

Following [BPG99], it is sometimes useful to denote Eve's post-interaction states in the form of $|\psi^{\text{Send}}_{\text{\# Err, Err location}}\rangle$ and symbolize as $|\psi^{\text{a}}_{d,q}\rangle$. Thus,

$$|E^{\text{Send}}_{\text{Received}}\rangle \quad \equiv \quad |\psi^{\text{Send}}_{\text{\# Err, Err location}}\rangle$$

We'll follow a more useful convention as we find suitable for calculations while attaching the orthogonal group with its name.

8.2.1.3 Naming convention

Consider the four states in any basis, on which the unitary will act when combined with the ancillary system. Consider, for simplicity, the four-dimensional computational basis. The four states are as follows.

$$|00\rangle,\ |10\rangle,\ |01\rangle,\ |11\rangle.$$

Following the convention of [CG97], the ordering of the qubits are as follows.

$$|\mathrm{Q}_1\rangle|\mathrm{Q}_2\rangle.$$

Then, the decimal representation of the states will respectively be

$$|0\rangle,\ |1\rangle,\ \ |2\rangle,\ |3\rangle.$$

Now, consider an unitary interaction on any of these basis states and the resulting states of Eve's evolved ancilla.

8.2.2 Post-interaction states: attack 2 qubits coherently

An unitary attack $\mathcal{U}$, coherently on 2 qubits, can completely be specified by its action on the elements of any 2-dim orthonormal basis $\beta = \beta_1\beta_2$ that consists of 4 elements. (e.g., the basis $\beta_1\beta_2 = 00$):

$$\mathcal{B}_\beta \quad := \quad \{|00\rangle^{\beta_1\beta_2}, |01\rangle^{\beta_1\beta_2}, |10\rangle^{\beta_1\beta_2}, |11\rangle^{\beta_1\beta_2}\}.$$

To attack one such pair of qubits in state $|a\rangle^\beta = |a_1a_2\rangle^{\beta_1\beta_2}$, Eve attaches an ancilla having initial state $|E\rangle$, and evolves the joint system unitarily with $\mathcal{U}$. The resulting state of the joint system is in superposition of 4 states.

$$\begin{aligned}\mathcal{U}|a\rangle^\beta|E\rangle \quad &\equiv \quad \mathcal{U}|a_1a_2\rangle^{\beta_1\beta_2}|E\rangle \qquad (\text{ for each } \ a_1a_2 \in \{00,01,10,11\}\ ) \\ &= \quad \sum_{d=0,1,2} |b_1b_2\rangle^{\beta_1\beta_2}|\psi^a_{d,q}\rangle, \quad \forall\ b_1b_2 \in \{00,01,10,11\} \\ &= \quad \sum_{b_1b_2\in\{00,01,10,11\}} |b_1b_2\rangle^{\beta_1\beta_2}|E^{a_1a_2}_{b_1b_2}\rangle.\end{aligned}$$

Here, d indicates # disturbances; q indicates (only for single disturbance) which qubit is disturbed.

Moreover, $b_1 = a_1 + d_1, b_2 = a_2 + d_2$ with $d_1, d_2 \in \{0,1\}$. Then, $d = d_1 + d_2$, and $q = 1^{d_1}2^{d_2}$ for $d = 1$ only.

Eve's states $|\psi^a_{d,q}\rangle \equiv |E^{a_1a_2}_{b_1b_2}\rangle$ are not normalized.

To elaborate,

$$\begin{aligned}&\mathcal{U}|a\rangle^\beta|E\rangle \qquad \equiv \mathcal{U}\ |a_1a_2\rangle^{\beta_1\beta_2}|E\rangle \\ &= |a_1a_2\rangle\ |\psi^a_0\rangle \quad +|a_1{+}1,a_2\rangle\ |\psi^a_{1,1}\rangle \quad +|a_1,a_2{+}1\rangle\ |\psi^a_{1,2}\rangle \quad +|a_1{+}1,a_2{+}1\rangle\ |\psi^a_2\rangle, \\ &= |a_1a_2\rangle\ |e^{a_1a_2}_{a_1a_2}\rangle\ +|a_1{+}1,a_2\rangle\ |E^{a_1a_2}_{a_1+1,a_2}\rangle\ +|a_1,a_2{+}1\rangle\ |E^{a_1a_2}_{a_1,a_2+1}\rangle\ +|a_1{+}1,a_2{+}1\rangle\ |E^{a_1a_2}_{a_1+1,a_2+1}\rangle.\end{aligned}$$

The basis stamp is dropped in the R.H.S.

Action of the unitary on 4 basis-states produces 16 post-interaction states (PIS) of Eve.

8.2.2.1 Action of the unitary on the zz basis, i.e., the basis $\beta_1\beta_2 = 00$.

Now, consider the action of the unitary on the four states $\{|zz\rangle, |z-z\rangle, |-zz\rangle, |-z-z\rangle\}$ of the basis zz. The action of the unitary on the basis states are described below, while Eve's states are represented like $|\psi^{\text{Send}}_{\text{\# Err, location}}\rangle$.

$$\begin{aligned}
\mathcal{U}\,|\mathbf{zz}\rangle|\psi\rangle &= |\textbf{zz}\rangle\,|\psi_0^{\mathbf{zz}}\rangle + |\textbf{z-z}\rangle\,|\psi_{1,2}^{\mathbf{zz}}\rangle + |\textbf{-zz}\rangle\,|\psi_{1,1}^{\mathbf{zz}}\rangle + |\textbf{-z-z}\rangle\,|\psi_2^{\mathbf{zz}}\rangle,\\
\mathcal{U}\,|\mathbf{z-z}\rangle|\psi\rangle &= |\textbf{z-z}\rangle\,|\psi_0^{\mathbf{z-z}}\rangle + |\textbf{zz}\rangle\,|\psi_{1,2}^{\mathbf{z-z}}\rangle + |\textbf{-z-z}\rangle\,|\psi_{1,1}^{\mathbf{z-z}}\rangle + |\textbf{-zz}\rangle\,|\psi_2^{\mathbf{z-z}}\rangle,\\
\mathcal{U}\,|\mathbf{-zz}\rangle|\psi\rangle &= |\textbf{-zz}\rangle\,|\psi_0^{\mathbf{-zz}}\rangle + |\textbf{-z-z}\rangle\,|\psi_{1,2}^{\mathbf{-zz}}\rangle + |\textbf{zz}\rangle\,|\psi_{1,1}^{\mathbf{-zz}}\rangle + |\textbf{z-z}\rangle\,|\psi_2^{\mathbf{-zz}}\rangle,\\
\mathcal{U}\,|\mathbf{-z-z}\rangle|\psi\rangle &= |\textbf{-z-z}\rangle\,|\psi_0^{\mathbf{-z-z}}\rangle + |\textbf{-zz}\rangle\,|\psi_{1,2}^{\mathbf{-z-z}}\rangle + |\textbf{z-z}\rangle\,|\psi_{1,1}^{\mathbf{-z-z}}\rangle + |\textbf{zz}\rangle\,|\psi_2^{\mathbf{-z-z}}\rangle.
\end{aligned}$$

We re-write the equations while considering the binary representation $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ of Alice's basis states prepared in the basis $\beta = 00$, while Eve's states are written as $|E^{\text{Send}}_{\text{Received}}\rangle$ in binary. The basis stamp is dropped customarily.

$$\begin{aligned}
\mathcal{U}\,|\mathbf{00}\rangle^{\beta=00}|E\rangle &= |\mathbf{00}\rangle\,|E^{\mathbf{00}}_{\mathbf{00}}\rangle + |\mathbf{01}\rangle\,|E^{\mathbf{00}}_{\mathbf{01}}\rangle + |\mathbf{10}\rangle\,|E^{\mathbf{00}}_{\mathbf{10}}\rangle + |\mathbf{11}\rangle\,|E^{\mathbf{00}}_{\mathbf{11}}\rangle,\\
\mathcal{U}\,|\mathbf{01}\rangle^{\beta=00}|E\rangle &= |\mathbf{01}\rangle\,|E^{\mathbf{01}}_{\mathbf{01}}\rangle + |\mathbf{00}\rangle\,|E^{\mathbf{01}}_{\mathbf{00}}\rangle + |\mathbf{11}\rangle\,|E^{\mathbf{01}}_{\mathbf{11}}\rangle + |\mathbf{10}\rangle\,|E^{\mathbf{01}}_{\mathbf{10}}\rangle,\\
\mathcal{U}\,|\mathbf{10}\rangle^{\beta=00}|E\rangle &= |\mathbf{10}\rangle\,|E^{\mathbf{10}}_{\mathbf{10}}\rangle + |\mathbf{11}\rangle\,|E^{\mathbf{10}}_{\mathbf{11}}\rangle + |\mathbf{00}\rangle\,|E^{\mathbf{10}}_{\mathbf{00}}\rangle + |\mathbf{01}\rangle\,|E^{\mathbf{10}}_{\mathbf{01}}\rangle,\\
\mathcal{U}\,|\mathbf{11}\rangle^{\beta=00}|E\rangle &= |\mathbf{11}\rangle\,|E^{\mathbf{11}}_{\mathbf{11}}\rangle + |\mathbf{10}\rangle\,|E^{\mathbf{11}}_{\mathbf{10}}\rangle + |\mathbf{01}\rangle\,|E^{\mathbf{11}}_{\mathbf{01}}\rangle + |\mathbf{00}\rangle\,|E^{\mathbf{11}}_{\mathbf{00}}\rangle.
\end{aligned}$$

One may re-order the R.H.S. according to the states received by Bob. Alice's states are prepared in the basis $\beta = 00$, while Eve's states are written as $|E^{\text{Send}}_{\text{Received}}\rangle$ in binary.

$$\begin{aligned}
\mathcal{U}\,|\mathbf{00}\rangle^{\beta=00}|E\rangle &= |\mathbf{00}\rangle\,|E^{\mathbf{00}}_{\mathbf{00}}\rangle + |\mathbf{01}\rangle\,|E^{\mathbf{00}}_{\mathbf{01}}\rangle + |\mathbf{10}\rangle\,|E^{\mathbf{00}}_{\mathbf{10}}\rangle + |\mathbf{11}\rangle\,|E^{\mathbf{00}}_{\mathbf{11}}\rangle,\\
\mathcal{U}\,|\mathbf{01}\rangle^{\beta=00}|E\rangle &= |\mathbf{00}\rangle\,|E^{\mathbf{01}}_{\mathbf{00}}\rangle + |\mathbf{01}\rangle\,|E^{\mathbf{01}}_{\mathbf{01}}\rangle + |\mathbf{10}\rangle\,|E^{\mathbf{01}}_{\mathbf{10}}\rangle + |\mathbf{11}\rangle\,|E^{\mathbf{01}}_{\mathbf{11}}\rangle,\\
\mathcal{U}\,|\mathbf{10}\rangle^{\beta=00}|E\rangle &= |\mathbf{00}\rangle\,|E^{\mathbf{10}}_{\mathbf{00}}\rangle + |\mathbf{01}\rangle\,|E^{\mathbf{10}}_{\mathbf{01}}\rangle + |\mathbf{10}\rangle\,|E^{\mathbf{10}}_{\mathbf{10}}\rangle + |\mathbf{11}\rangle\,|E^{\mathbf{10}}_{\mathbf{11}}\rangle,\\
\mathcal{U}\,|\mathbf{11}\rangle^{\beta=00}|E\rangle &= |\mathbf{00}\rangle\,|E^{\mathbf{11}}_{\mathbf{00}}\rangle + |\mathbf{01}\rangle\,|E^{\mathbf{11}}_{\mathbf{01}}\rangle + |\mathbf{10}\rangle\,|E^{\mathbf{11}}_{\mathbf{10}}\rangle + |\mathbf{11}\rangle\,|E^{\mathbf{11}}_{\mathbf{11}}\rangle.
\end{aligned}$$

Sometimes, the decimal representation is useful for calculations, re-order according to the states received by Bob.

$$\begin{aligned}
\mathcal{U}\,|\mathbf{0}\rangle|\psi\rangle &= |\mathbf{0}\rangle\,|\psi_0^{\mathbf{0}}\rangle + |\mathbf{2}\rangle\,|\psi_{1,2}^{\mathbf{0}}\rangle + |\mathbf{1}\rangle\,|\psi_{1,1}^{\mathbf{0}}\rangle + |\mathbf{3}\rangle\,|\psi_2^{\mathbf{0}}\rangle,\\
\mathcal{U}\,|\mathbf{2}\rangle|\psi\rangle &= |\mathbf{0}\rangle\,|\psi_{1,2}^{\mathbf{2}}\rangle + |\mathbf{2}\rangle\,|\psi_0^{\mathbf{2}}\rangle + |\mathbf{1}\rangle\,|\psi_2^{\mathbf{2}}\rangle + |\mathbf{3}\rangle\,|\psi_{1,1}^{\mathbf{2}}\rangle,\\
\mathcal{U}\,|\mathbf{1}\rangle|\psi\rangle &= |\mathbf{0}\rangle\,|\psi_{1,1}^{\mathbf{1}}\rangle + |\mathbf{2}\rangle\,|\psi_2^{\mathbf{1}}\rangle + |\mathbf{1}\rangle\,|\psi_0^{\mathbf{1}}\rangle + |\mathbf{3}\rangle\,|\psi_{1,2}^{\mathbf{1}}\rangle,\\
\mathcal{U}\,|\mathbf{3}\rangle|\psi\rangle &= |\mathbf{0}\rangle\,|\psi_2^{\mathbf{3}}\rangle + |\mathbf{2}\rangle\,|\psi_{1,1}^{\mathbf{3}}\rangle + |\mathbf{1}\rangle\,|\psi_{1,2}^{\mathbf{3}}\rangle + |\mathbf{3}\rangle\,|\psi_0^{\mathbf{3}}\rangle.
\end{aligned}$$

8.2.3 Reduction in number of IPs: 256 → 5

Characterize the Unitary $\equiv$ characterize the IPs.

8.2.3.1 24 IPs from 4 orthogonal groups

The 16 post-interaction states with Eve can be classified into 4 mutually orthogonal sets $S_0, S_{1,1}, S_{1,2}, S_2$ as follows. The subscripts indicate disturbance, while the superscript in the states indicate the binary labeling for the state sent by Alice.

$$\begin{aligned}
S_0 &:= \{\, |\psi_0^{00}\rangle, |\psi_0^{01}\rangle, |\psi_0^{10}\rangle, |\psi_0^{11}\rangle \,\},\\
S_{1,1} &:= \{\, |\psi_{1,1}^{00}\rangle, |\psi_{1,1}^{01}\rangle, |\psi_{1,1}^{10}\rangle, |\psi_{1,1}^{11}\rangle \,\},\\
S_{1,2} &:= \{\, |\psi_{1,2}^{00}\rangle, |\psi_{1,2}^{01}\rangle, |\psi_{1,2}^{10}\rangle, |\psi_{1,2}^{11}\rangle \,\},\\
S_2 &:= \{\, |\psi_2^{00}\rangle, |\psi_2^{01}\rangle, |\psi_2^{10}\rangle, |\psi_2^{11}\rangle \,\}.
\end{aligned}$$

With this representation, it is clear that the sets are classified for orthogonality according to the number of disturbances. For instance, the first set contains those states which correspond to no disturbance, etc. A hint towards arguing the orthogonality is given in Sec. 8.4.1. An alternate representation for the groupings could be found therein.

It is clear that the 4 states in each grouping lead to 6 IPs when they are considered real. Thus, there are total 24 IPs from 4 groupings.

We rename here the states of Eve by attaching the orthogonality label to the states and calling them with the decimal representation.

Table 8.1 Relabeled states in 4 $\perp$ groupings: $S_0, S_{1,1}, S_{1,2}, S_2$.

Alice send ↓		Eve's states			
$\lvert 0\rangle$	→	$\lvert 0\rangle^{S_0}$	$\lvert 0\rangle^{S_{1,1}}$	$\lvert 0\rangle^{S_{1,2}}$	$\lvert 0\rangle^{S_2}$
$\lvert 2\rangle$	→	$\lvert 2\rangle^{S_0}$	$\lvert 2\rangle^{S_{1,1}}$	$\lvert 2\rangle^{S_{1,2}}$	$\lvert 2\rangle^{S_2}$
$\lvert 1\rangle$	→	$\lvert 1\rangle^{S_0}$	$\lvert 1\rangle^{S_{1,1}}$	$\lvert 1\rangle^{S_{1,2}}$	$\lvert 1\rangle^{S_2}$
$\lvert 3\rangle$	→	$\lvert 3\rangle^{S_0}$	$\lvert 3\rangle^{S_{1,1}}$	$\lvert 3\rangle^{S_{1,2}}$	$\lvert 3\rangle^{S_2}$

With this new representation, we rewrite the action of the unitary considering Alice's states (2-qubits) in the zz basis.

$$\begin{aligned}
\mathcal{U}\,|0\rangle|E\rangle &= |0\rangle\,|0\rangle^{S_0} && +|2\rangle\,|0\rangle^{S_{1,2}} && +|1\rangle\,|0\rangle^{S_{1,1}} && +|3\rangle\,|0\rangle^{S_2},\\
\mathcal{U}\,|2\rangle|E\rangle &= |0\rangle\,|2\rangle^{S_{1,2}} && +|2\rangle\,|2\rangle^{S_0} && +|1\rangle\,|2\rangle^{S_2} && +|3\rangle\,|2\rangle^{S_{1,1}},\\
\mathcal{U}\,|1\rangle|E\rangle &= |0\rangle\,|1\rangle^{S_{1,1}} && +|2\rangle\,|1\rangle^{S_2} && +|1\rangle\,|1\rangle^{S_0} && +|3\rangle\,|1\rangle^{S_{1,2}},\\
\mathcal{U}\,|3\rangle|E\rangle &= |0\rangle\,|3\rangle^{S_2} && +|2\rangle\,|3\rangle^{S_{1,1}} && +|1\rangle\,|3\rangle^{S_{1,2}} && +|3\rangle\,|3\rangle^{S_0}.
\end{aligned} \tag{8.2.1}$$

R.H.S. is ordered according to the states received by Bob.

8.2.3.2 Further reduction: Rules of symmetry

IPs remain unchanged even after the following changes made:

Rule 1.1 [$\bar{\mathbf{1}}, \bar{\mathbf{2}}$]

Bit flip (i.e., $|0\rangle \leftrightarrow |1\rangle$) in one position

$\bar{\mathbf{1}}$. Bit flip in 1st position: $|a_1 a_2\rangle \rightarrow |a_1+1, a_2\rangle$.
e.g., $\langle\psi_0^{00}|\psi_0^{01}\rangle = \langle\psi_0^{10}|\psi_0^{11}\rangle$.

$\bar{\mathbf{2}}$. Bit flip in 2nd position: $|a_1 a_2\rangle \rightarrow |a_1, a_2+1\rangle$).
e.g., $\langle\psi_0^{00}|\psi_0^{10}\rangle = \langle\psi_0^{01}|\psi_0^{11}\rangle$.

Rule 1.2 [$\overset{\leftrightarrow}{\mathbf{12}}$]

Exchange of bits: $|a_1 a_2\rangle \leftrightarrow |a_2 a_1\rangle$. e.g., $\langle\psi_0^{00}|\psi_0^{01}\rangle = \langle\psi_0^{00}|\psi_0^{10}\rangle$.

Rule 2 [$\mathbf{z} \leftrightarrow \mathbf{x}$]

Basis change: A basis looks like $\alpha\beta$ s.t., $\alpha, \beta \in \{z, x\}$. So, there are 4 possible bases as follows.

$$zz \leftrightarrow zx \leftrightarrow xz \leftrightarrow xx.$$

Any mutual exchange among them can be considered.

$$\begin{array}{ll} \text{e.g.} & \langle\psi_0^{zz}|\psi_0^{z-z}\rangle = \langle\psi_0^{xx}|\psi_0^{x-x}\rangle \\ \text{i.e.,} & \langle\psi_0^{00}|\psi_0^{01}\rangle^{00} = \langle\psi_0^{00}|\psi_0^{01}\rangle^{11} \end{array}$$

The outside superscript stands for the binary representation of a basis:

$$zz = 00, \;\; xx = 11, \;\; xz = 10, \;\; zx = 01.$$

8.2.3.3 10 IPs: 7 equivalence classes.

Rules of symmetry: bit-flip, bit-swap.

Three of them are due to normalization of the states in each orthogonal sets S_0, S_2, $S_{1,1}$, $S_{1,2}$. Rest of the seven are due to applying rules of symmetry 1, i.e., bit flip $(\bar{\mathbf{1}}, \bar{\mathbf{2}})$ and exchange $(\overset{\leftrightarrow}{\mathbf{12}})$, in between two different states in each group. Two classes (A_1, A_2) from S_0, two (C_1, C_2) from S_2, and another three (B_1, B_2, B_3) from $S_{1,1}$, $S_{1,2}$. Thus, 24 IPs from the four orthogonal sets form 7 equivalence classes: A1, A2, B1, B2, B3, C1, C2.

Table 8.2 Equivalence classes of the IPs

Sets	# error	# Eqv Cls	Equivalence classes	# IPs
S_0	0	2	A_1, A_2	4,2
S_2	2	2	C_1, C_2	4,2
$S_{1,1}$, $S_{1,2}$	1	3	B_1, B_2, B_3	4,4,4

Applying the symmetric rules 1, i.e., bit flip $(\bar{\mathbf{1}}, \bar{\mathbf{2}})$ and exchange $(\overset{\leftrightarrow}{\mathbf{12}})$, we get

$$\begin{aligned}
A_1 &= \langle\psi_0^{zz}|\psi_0^{z-z}\rangle = \langle\psi_0^{zz}|\psi_0^{-zz}\rangle = \langle\psi_0^{-z-z}|\psi_0^{z-z}\rangle = \langle\psi_0^{-z-z}|\psi_0^{-zz}\rangle,\\
A_2 &= \langle\psi_0^{zz}|\psi_0^{-z-z}\rangle = \langle\psi_0^{z-z}|\psi_0^{-zz}\rangle,\\
C_1 &= \langle\psi_2^{zz}|\psi_2^{z-z}\rangle = \langle\psi_2^{zz}|\psi_2^{-zz}\rangle = \langle\psi_2^{z-z}|\psi_2^{-z-z}\rangle = \langle\psi_2^{-zz}|\psi_2^{-z-z}\rangle,\\
C_2 &= \langle\psi_2^{zz}|\psi_2^{-z-z}\rangle = \langle\psi_2^{z-z}|\psi_2^{-zz}\rangle,\\
B_1 &= \langle\psi_{12}^{zz}|\psi_{12}^{-zz}\rangle = \langle\psi_{12}^{z-z}|\psi_{12}^{-z-z}\rangle = \langle\psi_{11}^{zz}|\psi_{11}^{z-z}\rangle = \langle\psi_{11}^{-zz}|\psi_{11}^{-z-z}\rangle,\\
B_2 &= \langle\psi_{12}^{zz}|\psi_{12}^{z-z}\rangle = \langle\psi_{12}^{-zz}|\psi_{12}^{-z-z}\rangle = \langle\psi_{11}^{zz}|\psi_{11}^{-zz}\rangle = \langle\psi_{11}^{z-z}|\psi_{11}^{-z-z}\rangle,\\
B_3 &= \langle\psi_{12}^{zz}|\psi_{12}^{-z-z}\rangle = \langle\psi_{12}^{-zz}|\psi_{12}^{z-z}\rangle = \langle\psi_{11}^{zz}|\psi_{11}^{-z-z}\rangle = \langle\psi_{11}^{z-z}|\psi_{11}^{-zz}\rangle.
\end{aligned}$$

$$\begin{aligned}
A &= \langle\psi_0^{zz}|\psi_0^{zz}\rangle = \langle\psi_0^{z-z}|\psi_0^{z-z}\rangle = \langle\psi_0^{-zz}|\psi_0^{-zz}\rangle = \langle\psi_0^{-z-z}|\psi_0^{-z-z}\rangle,\\
B &= \langle\psi_{1,1}^{zz}|\psi_{1,1}^{zz}\rangle = \langle\psi_{1,1}^{z-z}|\psi_{1,1}^{z-z}\rangle = \langle\psi_{1,1}^{-zz}|\psi_{1,1}^{-zz}\rangle = \langle\psi_{1,1}^{-z-z}|\psi_{1,1}^{-z-z}\rangle\\
&= \langle\psi_{1,2}^{zz}|\psi_{1,2}^{zz}\rangle = \langle\psi_{1,2}^{z-z}|\psi_{1,2}^{z-z}\rangle = \langle\psi_{1,2}^{-zz}|\psi_{1,2}^{-zz}\rangle = \langle\psi_{1,2}^{-z-z}|\psi_{1,2}^{-z-z}\rangle,\\
C &= \langle\psi_2^{zz}|\psi_2^{zz}\rangle = \langle\psi_2^{z-z}|\psi_2^{z-z}\rangle = \langle\psi_2^{-zz}|\psi_2^{-zz}\rangle = \langle\psi_2^{-z-z}|\psi_2^{-z-z}\rangle.
\end{aligned}$$

24 IPs w.r.t 10 parameters

Here, we write the 24 IPs w.r.t. 10 parameters.

Recall that Eve's states are written as $|a\rangle^{\mathcal{S}}$, where a denotes the Alice's pair of bits in decimal, and $\mathcal{S}$ represents the orthogonal set.

Table 8.3 Eve's 16 states as 4 ⊥ groupings.

Alice send ↓		Eve's states			
$\|0\rangle$	→	$\|0\rangle^{S_0}$	$\|0\rangle^{S_{1,1}}$	$\|0\rangle^{S_{1,2}}$	$\|0\rangle^{S_2}$
$\|2\rangle$	→	$\|2\rangle^{S_0}$	$\|2\rangle^{S_{1,1}}$	$\|2\rangle^{S_{1,2}}$	$\|2\rangle^{S_2}$
$\|1\rangle$	→	$\|1\rangle^{S_0}$	$\|1\rangle^{S_{1,1}}$	$\|1\rangle^{S_{1,2}}$	$\|1\rangle^{S_2}$
$\|3\rangle$	→	$\|3\rangle^{S_0}$	$\|3\rangle^{S_{1,1}}$	$\|3\rangle^{S_{1,2}}$	$\|3\rangle^{S_2}$

24 IPs can then be described in terms of the 7 parameters (from the last section).

Table 8.4 24 IPs → 7 groupings

Set $S_{?,?}$	IP$_1$ $\langle 0\|2\rangle$	IP$_2$ $\langle 0\|1\rangle$	IP$_3$ $\langle 0\|3\rangle$	IP$_4$ $\langle 2\|1\rangle$	IP$_5$ $\langle 2\|3\rangle$	IP$_6$ $\langle 1\|3\rangle$
S_0	A_1	A_1	$\mathbf{A_2}$	$\mathbf{A_2}$	A_1	A_1
$S_{1,1}$	$\mathbf{B_3}$	B_1	$\mathbf{B_2}$	$\mathbf{B_2}$	B_1	$\mathbf{B_3}$
$S_{1,2}$	B_1	$\mathbf{B_3}$	$\mathbf{B_2}$	$\mathbf{B_2}$	$\mathbf{B_3}$	B_1
S_2	$\mathbf{C_1}$	$\mathbf{C_1}$	C_2	C_2	$\mathbf{C_1}$	$\mathbf{C_1}$

Details with other notations are tabulated in Sec. 8.4.3.

All the 16 IPs in each group are then expressed in terms of the 10 parameters as follows.

Table 8.5 All 16 IPs in each ⊥ group

IPs within a grp				IPs for S_0				IPs for S_2				IPs for $S_{1,1}$				IPs for $S_{1,2}$			
0·0	0·2	0·1	0·3	A	A_1	A_1	A_2	C	C_1	C_1	C_2	B	B_3	B_1	B_2	B	B_1	B_3	B_2
2·0	2·2	2·1	2·3	A_1	A	A_2	A_1	C_1	C	C_2	C_1	B_3	B	B_2	B_1	B_1	B	B_2	B_3
1·0	1·2	1·1	1·3	A_1	A_2	A	A_1	C_1	C_2	C	C_1	B_1	B_2	B	B_3	B_3	B_2	B	B_1
3·0	3·2	3·1	3·3	A_2	A_1	A_1	A	C_2	C_1	C_1	C	B_2	B_1	B_3	B	B_2	B_3	B_1	B

8.2.3.4 5 IPs due to 5 equations. Rules of symmetry = Basis change.

So far, we could reduce 256 IPs to 10 following the symmetry rules. It is possible to further reduce them to five, as five inter-relations between them follow due to the rules of basis-change.

Proposition 8.1. *The following five relations between the 10 IPs hold.*

$$
\begin{aligned}
A+2B+C &= 1, && (8.2.2)\\
A_1+B_1 &= A-B, && (8.2.3)\\
B_3+C_1 &= B-C, && (8.2.4)\\
B_2+B_3 &= A_1-A_2, && (8.2.5)\\
C_1+C_2 &= B_1-B_2. && (8.2.6)
\end{aligned}
$$

However, it is not so easy to find out these interrelations, as it is not known which basis-change rule will help. For instance, if we consider the basis-change $zz \mapsto xx$, then we do not get the desired relations. In fact easier relations exists due to other basis-changes. Moreover, for a given basis change, which state of Alice should be considered in another difficulty.

Here, we describe a possible approach to get these relations. The first relation follows due to normalization restriction of any of the L.H.S.-states in Eq. (8.2.2). The remaining relations emanate when we consider the basis change $zz \mapsto xz$.

We consider Alice's states (2 qubits) in the zz basis. Then, in this basis, we analyze the evolution for Alice's two states $|0\rangle$ and $|2\rangle$.

8.3 Eavesdropping on 6-st Protocol

The 6-st Protocol works with 3 bases in the 2-dim Hilbert space.

8.3.1 2-qubit attack

For a 2-qubit attack, learning the unitary is to study the action of the unitary on a basis (e.g., zx basis), which contains 4 states. Again, 256 inner products are possible. It can ultimately be reduced to only 2 real parameters.

$$256 \to 10 \to 2.$$

A 2-qubit attack slightly increases Eve's chance to successfully guess the 2-qubits together than that for an incoherent attack. However, it doesn't increase the Shannon mutual information. However, it slightly increases the Rényi information.

8.3.2 3-qubit attack

A 3-qubit attack doesn't improve Eve's Shannon information, nor does it improve her chance to guess correctly the bits that Bob receives correctly. However, for the bits those are received undisturbed by Bob, Eve gains full information, since all the eight states of Eve in the ψ_3 grouping are mutually orthogonal. Thus, the 3-qubit attack improves (not negligible) the chance that Eve guesses all the 3 qubits correctly.

For attacking more qubits, the parametrization task goes arduous.

8.4 Details

Some detailed calculations of the earlier sections are placed here. The interesting one is the orthogonal grouping where we find an alternative solution. It may not possibly change the cryptographic bounds except possibly the way how things take place.

8.4.1 Orthogonal groupings (unpublished) of Eve's states

16 states of Eve → 4 mutually orthogonal groups, each having 4 states.

The binary-decimal representation of the send-received states are described in Table 8.6.

Table 8.6 Send vs received bits in binary. Subscripts indicate decimal.

Alice Send	Bob Received			
00_0	00_0	01_1	10_2	11_3
01_1	01_1	00_0	11_3	10_2
10_2	10_2	11_3	00_0	01_1
11_3	11_3	10_2	01_1	00_0

Now, consider the decimal representation: a, b represents Alice and Bob's bits. Then,

$$\begin{aligned} \mathcal{U}|a\rangle|E\rangle &= \sum_{b=0}^{3}|b\rangle|E_{ab}\rangle, \\ \mathcal{U}|a'\rangle|E\rangle &= \sum_{b=0}^{3}|b\rangle|E_{a'b}\rangle. \end{aligned}$$

Then, due to orthogonality of the above two states, we get

$$0 = \sum_{b=0}^{3}\langle E_{ab}|E_{a'b}\rangle.$$

One way to make the expression zero is to consider each term to be zero, i.e., $\langle E_{ab}|E_{a'b}\rangle = 0$ for each b. If a bin corresponds to an orthogonal grouping, then $|E_{ab}\rangle$ and $|E_{a'b}\rangle$ cannot occupy the same bin.

Following this logic, Table 8.7 explains the bin-allocation logic. The last row indicate the choice due to [CG97]. Table 8.8 shows an alternative allocation following a trial-and-error approach. contains the two groupings.

Table 8.7 | Orthogonal groupings: natural.

Four bins are allocated for Eve's four states as Alice sends 0. For the remaining 12 states, the corresponding bin is indicated by the restriction where it cannot go: e.g., all the ~ 3 need be replaced by different indices from 0,1,2 and so on. A natural allocation then follows as in the last row, as chosen by [CG97].

Send	Bob	Bin#	Bob	Bin#	Bob	Bin#	Bob	Bin#
0	0	0	1	1	2	2	3	3
1	1	~ 1	0	~ 0	3	~ 3	2	~ 2
2	2	~ 2	3	~ 3	0	~ 0	1	~ 1
3	3	~ 3	2	~ 2	1	~ 1	0	~ 0
Allocation:		0		1		2		3

Table 8.8 Orthogonal groupings: alternative!

Send	Bob	Bin#	Bob	Bin#	Bob	Bin#	Bob	Bin#
0	0	0	1	1	2	2	3	3
1	1	3	0	2	3	1	2	0
2	2	1	3	2	0	3	1	0
3	3	0	2	3	1	2	0	1

The two different groupings are then listed. A grouping allocates Eve's 16 states in 4 mutually orthogonal sets.

Natural grouping

Bin 0	Bin 1	Bin 2	Bin 3
$\|E_{0,0}\rangle$	$\|E_{0,1}\rangle$	$\|E_{0,2}\rangle$	$\|E_{0,3}\rangle$
$\|E_{1,1}\rangle$	$\|E_{1,0}\rangle$	$\|E_{1,3}\rangle$	$\|E_{1,2}\rangle$
$\|E_{2,2}\rangle$	$\|E_{2,3}\rangle$	$\|E_{2,0}\rangle$	$\|E_{2,1}\rangle$
$\|E_{3,3}\rangle$	$\|E_{3,2}\rangle$	$\|E_{3,1}\rangle$	$\|E_{3,0}\rangle$

Possible alternative!

Bin 0	Bin 1	Bin 2	Bin 3
$\|E_{0,0}\rangle$	$\|E_{0,1}\rangle$	$\|E_{0,2}\rangle$	$\|E_{0,3}\rangle$
$\|E_{1,2}\rangle$	$\|E_{1,3}\rangle$	$\|E_{1,0}\rangle$	$\|E_{1,1}\rangle$
$\|E_{2,1}\rangle$	$\|E_{2,2}\rangle$	$\|E_{2,3}\rangle$	$\|E_{2,0}\rangle$
$\|E_{3,3}\rangle$	$\|E_{3,0}\rangle$	$\|E_{3,1}\rangle$	$\|E_{3,2}\rangle$

8.4.2 16 post-interaction states of Eve: 4 mutually orthogonal groupings

Action of the unitary on the basis elements in a basis $\beta_1\beta_2$ (e.g., zz basis) can be seen as follows. Eve's states after interaction are given in 3 different conventions: $|\psi^{\text{Send}}_{\text{\# Err, Err location}}\rangle$, $|E^{\text{Send}}_{\text{Received}}\rangle$ in binary, $|E_{\text{Send,Received}}\rangle$ in decimal, respectively in the first, second and third rows for each of Alice's states.

Alice's 2 qubits	Eve's post-interaction states				
$\|a_1a_2\rangle^{\beta_1\beta_2}$		0 err	1 err: Q1	1 err: Q2	2 err
$\|00\rangle^{\beta_1\beta_2}$		$\|\psi_0^{00}\rangle$	$\|\psi_{1,1}^{00}\rangle$	$\|\psi_{1,2}^{00}\rangle$	$\|\psi_2^{00}\rangle$
	$=$	$\|E_{00}^{00}\rangle$	$\|E_{10}^{00}\rangle$	$\|E_{01}^{00}\rangle$	$\|E_{11}^{00}\rangle$
$\|\mathbf{0}\rangle$	$=$	$\|E_{0,0}\rangle$	$\|E_{0,1}\rangle$	$\|E_{0,2}\rangle$	$\|E_{0,3}\rangle$
$\|01\rangle^{\beta_1\beta_2}$		$\|\psi_0^{01}\rangle$	$\|\psi_{1,1}^{01}\rangle$	$\|\psi_{1,2}^{01}\rangle$	$\|\psi_2^{01}\rangle$
	$=$	$\|E_{01}^{01}\rangle$	$\|E_{11}^{01}\rangle$	$\|E_{00}^{01}\rangle$	$\|E_{10}^{01}\rangle$
$\|\mathbf{2}\rangle$	$=$	$\|E_{2,2}\rangle$	$\|E_{2,3}\rangle$	$\|E_{2,0}\rangle$	$\|E_{2,1}\rangle$
$\|10\rangle^{\beta_1\beta_2}$		$\|\psi_0^{10}\rangle$	$\|\psi_{1,1}^{10}\rangle$	$\|\psi_{1,2}^{10}\rangle$	$\|\psi_2^{10}\rangle$
	$=$	$\|E_{10}^{10}\rangle$	$\|E_{00}^{10}\rangle$	$\|E_{11}^{10}\rangle$	$\|E_{01}^{10}\rangle$
$\|\mathbf{1}\rangle$	$=$	$\|E_{1,1}\rangle$	$\|E_{1,0}\rangle$	$\|E_{1,3}\rangle$	$\|E_{1,2}\rangle$
$\|11\rangle^{\beta_1\beta_2}$		$\|\psi_0^{11}\rangle$	$\|\psi_{1,1}^{11}\rangle$	$\|\psi_{1,2}^{11}\rangle$	$\|\psi_2^{11}\rangle$
	$=$	$\|E_{11}^{11}\rangle$	$\|E_{01}^{11}\rangle$	$\|E_{10}^{11}\rangle$	$\|E_{00}^{11}\rangle$
$\|\mathbf{3}\rangle$	$=$	$\|E_{3,3}\rangle$	$\|E_{3,2}\rangle$	$\|E_{3,1}\rangle$	$\|E_{3,0}\rangle$
		S_0	$S_{1,1}$	$S_{1,2}$	S_2

One can classify these 16 states into 4 mutually orthogonal groupings: $S_0, S_{1,1}, S_{1,2}, S_2$, as indicated in the last row.

8.4.3 24 IPs: 6 from each set

Set		IP$_1$	IP$_2$	IP$_3$	IP$_4$	IP$_5$	IP$_6$
S_0		A_1	A_1	$\mathbf{A_2}$	$\mathbf{A_2}$	A_1	A_1
		$\langle\psi_0^{00}\vert\psi_0^{01}\rangle$	$\langle\psi_0^{00}\vert\psi_0^{10}\rangle$	$\langle\psi_0^{00}\vert\psi_0^{11}\rangle$	$\langle\psi_0^{01}\vert\psi_0^{10}\rangle$	$\langle\psi_0^{01}\vert\psi_0^{11}\rangle$	$\langle\psi_0^{10}\vert\psi_0^{11}\rangle$
	$=$	$\langle E_{00}^{00}\vert E_{01}^{01}\rangle$	$\langle E_{00}^{00}\vert E_{10}^{10}\rangle$	$\langle E_{00}^{00}\vert E_{11}^{11}\rangle$	$\langle E_{01}^{01}\vert E_{10}^{10}\rangle$	$\langle E_{01}^{01}\vert E_{11}^{11}\rangle$	$\langle E_{10}^{10}\vert E_{11}^{11}\rangle$
	$=$	$\langle E_{0,0}\vert E_{2,2}\rangle$	$\langle E_{0,0}\vert E_{1,1}\rangle$	$\langle E_{0,0}\vert E_{3,3}\rangle$	$\langle E_{2,2}\vert E_{1,1}\rangle$	$\langle E_{2,2}\vert E_{3,3}\rangle$	$\langle E_{1,1}\vert E_{3,3}\rangle$
$S_{1,1}$		$\mathbf{B_3}$	B_1	$\mathbf{B_2}$	$\mathbf{B_2}$	B_1	$\mathbf{B_3}$
		$\langle\psi_{1,1}^{00}\vert\psi_{1,1}^{01}\rangle$	$\langle\psi_{1,1}^{00}\vert\psi_{1,1}^{10}\rangle$	$\langle\psi_{1,1}^{00}\vert\psi_{1,1}^{11}\rangle$	$\langle\psi_{1,1}^{01}\vert\psi_{1,1}^{10}\rangle$	$\langle\psi_{1,1}^{01}\vert\psi_{1,1}^{11}\rangle$	$\langle\psi_{1,1}^{10}\vert\psi_{1,1}^{11}\rangle$
	$=$	$\langle E_{10}^{00}\vert E_{11}^{01}\rangle$	$\langle E_{10}^{00}\vert E_{00}^{10}\rangle$	$\langle E_{10}^{00}\vert E_{01}^{11}\rangle$	$\langle E_{11}^{01}\vert E_{00}^{10}\rangle$	$\langle E_{11}^{01}\vert E_{01}^{11}\rangle$	$\langle E_{00}^{10}\vert E_{01}^{11}\rangle$
	$=$	$\langle E_{0,1}\vert E_{2,3}\rangle$	$\langle E_{0,1}\vert E_{1,0}\rangle$	$\langle E_{0,1}\vert E_{3,2}\rangle$	$\langle E_{2,3}\vert E_{1,0}\rangle$	$\langle E_{2,3}\vert E_{3,2}\rangle$	$\langle E_{1,0}\vert E_{3,2}\rangle$
$S_{1,2}$		B_1	$\mathbf{B_3}$	$\mathbf{B_2}$	$\mathbf{B_2}$	$\mathbf{B_3}$	B_1
		$\langle\psi_{1,2}^{00}\vert\psi_{1,2}^{01}\rangle$	$\langle\psi_{1,2}^{00}\vert\psi_{1,2}^{10}\rangle$	$\langle\psi_{1,2}^{00}\vert\psi_{1,2}^{11}\rangle$	$\langle\psi_{1,2}^{01}\vert\psi_{1,2}^{10}\rangle$	$\langle\psi_{1,2}^{01}\vert\psi_{1,2}^{11}\rangle$	$\langle\psi_{1,2}^{10}\vert\psi_{1,2}^{11}\rangle$
	$=$	$\langle E_{01}^{00}\vert E_{00}^{01}\rangle$	$\langle E_{01}^{00}\vert E_{11}^{10}\rangle$	$\langle E_{01}^{00}\vert E_{10}^{11}\rangle$	$\langle E_{00}^{01}\vert E_{11}^{10}\rangle$	$\langle E_{00}^{01}\vert E_{10}^{11}\rangle$	$\langle E_{11}^{10}\vert E_{10}^{11}\rangle$
	$=$	$\langle E_{0,2}\vert E_{2,0}\rangle$	$\langle E_{0,2}\vert E_{1,3}\rangle$	$\langle E_{0,2}\vert E_{3,1}\rangle$	$\langle E_{2,0}\vert E_{1,3}\rangle$	$\langle E_{2,0}\vert E_{3,1}\rangle$	$\langle E_{1,3}\vert E_{3,1}\rangle$
S_2		$\mathbf{C_1}$	$\mathbf{C_1}$	C_2	C_2	$\mathbf{C_1}$	$\mathbf{C_1}$
		$\langle\psi_2^{00}\vert\psi_2^{01}\rangle$	$\langle\psi_2^{00}\vert\psi_2^{10}\rangle$	$\langle\psi_2^{00}\vert\psi_2^{11}\rangle$	$\langle\psi_2^{01}\vert\psi_2^{10}\rangle$	$\langle\psi_2^{01}\vert\psi_2^{11}\rangle$	$\langle\psi_2^{10}\vert\psi_2^{11}\rangle$
	$=$	$\langle E_{11}^{00}\vert E_{10}^{01}\rangle$	$\langle E_{11}^{00}\vert E_{01}^{10}\rangle$	$\langle E_{11}^{00}\vert E_{00}^{11}\rangle$	$\langle E_{10}^{01}\vert E_{01}^{10}\rangle$	$\langle E_{10}^{01}\vert E_{00}^{11}\rangle$	$\langle E_{01}^{10}\vert E_{00}^{11}\rangle$
	$=$	$\langle E_{0,3}\vert E_{2,1}\rangle$	$\langle E_{0,3}\vert E_{1,2}\rangle$	$\langle E_{0,3}\vert E_{3,0}\rangle$	$\langle E_{2,1}\vert E_{1,2}\rangle$	$\langle E_{2,1}\vert E_{3,0}\rangle$	$\langle E_{1,2}\vert E_{3,0}\rangle$

There are total 24 IPs. But, there are 7 different IPs up to equivalence due to the rules of symmetry: i) bit-flip, ii) bit-swap.

8.4.4 7+3 Equivalence Classes of 24 IPs

24 IPs from the four orthogonal sets S$_0$, S$_2$, S$_{1,1}$, S$_{1,2}$ form 7 equivalence classes. Two classes (A$_1$, A$_2$) from S$_0$, two (C$_1$, C$_2$) from S$_2$, and another three (B$_1$, B$_2$, B$_3$) from S$_{1,1}$, S$_{1,2}$. Apply rules of symmetry: 1.1) bit flip $(\bar{\mathbf{1}},\bar{\mathbf{2}})$ and 1.2) exchange $(\overset{\leftrightarrow}{\mathbf{12}})$.

Table 8.9 | 24 IPs → 7+3 equivalence classes.

$\mathsf{S}_0 \rightarrow$ Two eqv. cls.: **A1, A2** $\qquad$ $\mathsf{S}_2 \rightarrow$ Two eqv. cls.: **C1, C2**

Class **A1**

$$\begin{tikzcd}
\begin{array}{c}\langle E_{0,0}|E_{2,2}\rangle\\ \langle E^{\mathbf{00}}_{00}|E^{\mathbf{01}}_{01}\rangle\\ \langle\psi_0^{zz}|\psi_0^{z-z}\rangle\end{array} \arrow[r, leftrightarrow, "\overset{\leftrightarrow}{12}"] \arrow[d, leftrightarrow, "\bar{1}"'] & \begin{array}{c}\langle \mathbf{E_{0,0}}|\mathbf{E_{1,1}}\rangle\\ \langle E^{\mathbf{00}}_{00}|E^{\mathbf{10}}_{10}\rangle\\ \langle\psi_0^{zz}|\psi_0^{-zz}\rangle\end{array} \arrow[d, leftrightarrow, "\bar{2}"] \\
\begin{array}{c}\langle E_{1,1}|E_{3,3}\rangle\\ \langle E^{\mathbf{10}}_{10}|E^{\mathbf{11}}_{11}\rangle\\ \langle\psi_0^{-zz}|\psi_0^{-z-z}\rangle\end{array} \arrow[r, leftrightarrow, "\overset{\leftrightarrow}{12}"'] & \begin{array}{c}\langle E_{2,2}|E_{3,3}\rangle\\ \langle E^{\mathbf{01}}_{01}|E^{\mathbf{11}}_{11}\rangle\\ \langle\psi_0^{z-z}|\psi_0^{-z-z}\rangle\end{array}
\end{tikzcd}$$

Class **C1**

$$\begin{tikzcd}
\begin{array}{c}\langle E_{0,3}|E_{2,1}\rangle\\ \langle E^{\mathbf{00}}_{11}|E^{\mathbf{01}}_{10}\rangle\\ \langle\psi_2^{zz}|\psi_2^{z-z}\rangle\end{array} \arrow[r, leftrightarrow, "\overset{\leftrightarrow}{12}"] \arrow[d, leftrightarrow, "\bar{1}"'] & \begin{array}{c}\langle \mathbf{E_{0,3}}|\mathbf{E_{1,2}}\rangle\\ \langle E^{\mathbf{00}}_{11}|E^{\mathbf{10}}_{01}\rangle\\ \langle\psi_2^{zz}|\psi_2^{-zz}\rangle\end{array} \arrow[d, leftrightarrow, "\bar{2}"] \\
\begin{array}{c}\langle E_{1,2}|E_{3,0}\rangle\\ \langle E^{\mathbf{10}}_{01}|E^{\mathbf{11}}_{00}\rangle\\ \langle\psi_2^{-zz}|\psi_2^{-z-z}\rangle\end{array} \arrow[r, leftrightarrow, "\overset{\leftrightarrow}{12}"'] & \begin{array}{c}\langle E_{2,1}|E_{3,0}\rangle\\ \langle E^{\mathbf{01}}_{10}|E^{\mathbf{11}}_{00}\rangle\\ \langle\psi_2^{z-z}|\psi_2^{-z-z}\rangle\end{array}
\end{tikzcd}$$

Class **A2**

$$\begin{tikzcd}
\begin{array}{c}\langle \mathbf{E_{0,0}}|\mathbf{E_{3,3}}\rangle\\ \langle E^{\mathbf{00}}_{00}|E^{\mathbf{11}}_{11}\rangle\\ \langle\psi_0^{zz}|\psi_0^{-z-z}\rangle\end{array} \arrow[r, leftrightarrow, "\bar{2}"] & \begin{array}{c}\langle E_{2,2}|E_{1,1}\rangle\\ \langle E^{\mathbf{01}}_{01}|E^{\mathbf{10}}_{10}\rangle\\ \langle\psi_0^{z-z}|\psi_0^{-zz}\rangle\end{array}
\end{tikzcd}$$

Class **C2**

$$\begin{tikzcd}
\begin{array}{c}\langle \mathbf{E_{0,3}}|\mathbf{E_{3,0}}\rangle\\ \langle E^{\mathbf{00}}_{11}|E^{\mathbf{11}}_{00}\rangle\\ \langle\psi_2^{zz}|\psi_2^{-z-z}\rangle\end{array} \arrow[r, leftrightarrow, "\bar{2}"] & \begin{array}{c}\langle E_{2,1}|E_{1,2}\rangle\\ \langle E^{\mathbf{01}}_{10}|E^{\mathbf{10}}_{01}\rangle\\ \langle\psi_2^{z-z}|\psi_2^{-zz}\rangle\end{array}
\end{tikzcd}$$

$\mathsf{S}_{1,1}, \mathsf{S}_{1,2} \rightarrow$ 3 Eqv. Cls.: **B1, B2, B3**

Class **B3**

$$\begin{tikzcd}
\begin{array}{c}\langle \mathbf{E_{0,1}}|\mathbf{E_{2,3}}\rangle\\ \langle E^{\mathbf{00}}_{10}|E^{\mathbf{01}}_{11}\rangle\\ \langle\psi_{11}^{zz}|\psi_{11}^{z-z}\rangle\end{array} \arrow[r, leftrightarrow, "\overset{\leftrightarrow}{12}"] \arrow[d, leftrightarrow, "\bar{1}"'] & \begin{array}{c}\langle E_{0,2}|E_{1,3}\rangle\\ \langle E^{\mathbf{00}}_{01}|E^{\mathbf{10}}_{11}\rangle\\ \langle\psi_{12}^{zz}|\psi_{12}^{-zz}\rangle\end{array} \arrow[d, leftrightarrow, "\bar{2}"] \\
\begin{array}{c}\langle E_{1,0}|E_{3,2}\rangle\\ \langle E^{\mathbf{10}}_{00}|E^{\mathbf{11}}_{01}\rangle\\ \langle\psi_{11}^{-zz}|\psi_{11}^{-z-z}\rangle\end{array} \arrow[r, leftrightarrow, "\overset{\leftrightarrow}{12}"'] & \begin{array}{c}\langle E_{2,0}|E_{3,1}\rangle\\ \langle E^{\mathbf{01}}_{00}|E^{\mathbf{11}}_{10}\rangle\\ \langle\psi_{12}^{z-z}|\psi_{12}^{-z-z}\rangle\end{array}
\end{tikzcd}$$

Class **B1**

$$\begin{tikzcd}
\begin{array}{c}\langle E_{0,2}|E_{2,0}\rangle\\ \langle E^{\mathbf{00}}_{01}|E^{\mathbf{01}}_{00}\rangle\\ \langle\psi_{12}^{zz}|\psi_{12}^{z-z}\rangle\end{array} \arrow[r, leftrightarrow, "\overset{\leftrightarrow}{12}"] \arrow[d, leftrightarrow, "\bar{1}"'] & \begin{array}{c}\langle \mathbf{E_{0,1}}|\mathbf{E_{1,0}}\rangle\\ \langle E^{\mathbf{00}}_{10}|E^{\mathbf{10}}_{00}\rangle\\ \langle\psi_{11}^{zz}|\psi_{11}^{-zz}\rangle\end{array} \arrow[d, leftrightarrow, "\bar{2}"] \\
\begin{array}{c}\langle E_{1,3}|E_{3,1}\rangle\\ \langle E^{\mathbf{10}}_{11}|E^{\mathbf{11}}_{10}\rangle\\ \langle\psi_{12}^{-zz}|\psi_{12}^{-z-z}\rangle\end{array} \arrow[r, leftrightarrow, "\overset{\leftrightarrow}{12}"'] & \begin{array}{c}\langle E_{2,3}|E_{3,2}\rangle\\ \langle E^{\mathbf{01}}_{11}|E^{\mathbf{11}}_{01}\rangle\\ \langle\psi_{11}^{z-z}|\psi_{11}^{-z-z}\rangle\end{array}
\end{tikzcd}$$

Class **B2**

$$\begin{tikzcd}
\begin{array}{c}\langle E_{0,2}|E_{3,1}\rangle\\ \langle E^{\mathbf{00}}_{01}|E^{\mathbf{11}}_{10}\rangle\\ \langle\psi_{12}^{zz}|\psi_{12}^{-z-z}\rangle\end{array} \arrow[r, leftrightarrow, "\overset{\leftrightarrow}{12}"] \arrow[d, leftrightarrow, "\bar{1}"'] & \begin{array}{c}\langle \mathbf{E_{0,1}}|\mathbf{E_{3,2}}\rangle\\ \langle E^{\mathbf{00}}_{10}|E^{\mathbf{11}}_{01}\rangle\\ \langle\psi_{11}^{zz}|\psi_{11}^{-z-z}\rangle\end{array} \arrow[d, leftrightarrow, "\bar{2}"] \\
\begin{array}{c}\langle E_{1,3}|E_{2,0}\rangle\\ \langle E^{\mathbf{10}}_{11}|E^{\mathbf{01}}_{00}\rangle\\ \langle\psi_{12}^{-zz}|\psi_{12}^{z-z}\rangle\end{array} \arrow[r, leftrightarrow, "\overset{\leftrightarrow}{12}"'] & \begin{array}{c}\langle E_{2,3}|E_{1,0}\rangle\\ \langle E^{\mathbf{01}}_{11}|E^{\mathbf{10}}_{00}\rangle\\ \langle\psi_{11}^{z-z}|\psi_{11}^{-zz}\rangle\end{array}
\end{tikzcd}$$

7+3 Equivalence Classes

$A_1 = \{\langle \mathbf{E_{0,0}}|\mathbf{E_{1,1}}\rangle, \langle E_{0,0}|E_{2,2}\rangle, \langle E_{2,2}|E_{3,3}\rangle, \langle E_{1,1}|E_{3,3}\rangle\},$ $\qquad A_2 = \{\langle \mathbf{E_{0,0}}|\mathbf{E_{3,3}}\rangle, \langle E_{2,2}|E_{1,1}\rangle\}.$

$C_1 = \{\langle \mathbf{E_{0,3}}|\mathbf{E_{1,2}}\rangle, \langle E_{0,3}|E_{2,1}\rangle, \langle E_{2,1}|E_{3,0}\rangle, \langle E_{1,2}|E_{3,0}\rangle\},$ $\qquad C_2 = \{\langle \mathbf{E_{0,3}}|\mathbf{E_{3,0}}\rangle, \langle E_{2,1}|E_{1,2}\rangle\}.$

$$B_1 = \{\langle E_{0,2}|E_{2,0}\rangle, \langle \mathbf{E_{0,1}}|\mathbf{E_{1,0}}\rangle, \langle E_{1,3}|E_{3,1}\rangle, \langle E_{2,3}|E_{3,2}\rangle\},$$
$$B_2 = \{\langle E_{0,2}|E_{3,1}\rangle, \langle \mathbf{E_{0,1}}|\mathbf{E_{3,2}}\rangle, \langle E_{1,3}|E_{2,0}\rangle, \langle E_{2,3}|E_{1,0}\rangle\},$$
$$B_3 = \{\langle \mathbf{E_{0,1}}|\mathbf{E_{2,3}}\rangle, \langle E_{0,2}|E_{1,3}\rangle, \langle E_{1,0}|E_{3,2}\rangle, \langle E_{2,0}|E_{3,1}\rangle\}.$$

$$A = \langle E_{a,a}|E_{a,a}\rangle^{S_0},\quad B = \langle E_{a,a}|E_{a,a}\rangle^{S_{1,1}}_{S_{1,2}},\quad C = \langle E_{a,a}|E_{a,a}\rangle^{S_2},\qquad \forall a \in \{0,2,1,3\}.$$

8.4.5 5 parameters: 4 inter-relations between the IPs due to basis change

Five inter-relations between the 10 IPs follow due to the rules of basis-change. The following four are to be proved.

$$\begin{aligned} A_1+B_1 &= A-B, \\ B_3+C_1 &= B-C, \\ B_2+B_3 &= A_1-A_2, \\ C_1+C_2 &= B_1-B_2. \end{aligned}$$

8.4.5.1 Action on zz basis, i.e., the basis $\beta_1\beta_2 = 00 = zz$.

Consider the Alice's states (2 qubits) in the zz basis. Eve's unitary evolves the joint system as the following. The subscript (zz basis) for each state is same and thus dropped.

$$\begin{array}{lllll} \mathcal{U}\,|0\rangle|\psi\rangle & =|0\rangle\,|0\rangle^{S_0} & +|2\rangle\,|0\rangle^{S_{1,2}} & +|1\rangle\,|0\rangle^{S_{1,1}} & +|3\rangle\,|0\rangle^{S_2} \\ \mathcal{U}\,|2\rangle|\psi\rangle & =|0\rangle\,|2\rangle^{S_{1,2}} & +|2\rangle\,|2\rangle^{S_0} & +|1\rangle\,|2\rangle^{S_2} & +|3\rangle\,|2\rangle^{S_{1,1}} \\ \mathcal{U}\,|1\rangle|\psi\rangle & =|0\rangle\,|1\rangle^{S_{1,1}} & +|2\rangle\,|1\rangle^{S_2} & +|1\rangle\,|1\rangle^{S_0} & +|3\rangle\,|1\rangle^{S_{1,2}} \\ \mathcal{U}\,|3\rangle|\psi\rangle & =|0\rangle\,|3\rangle^{S_2} & +|2\rangle\,|3\rangle^{S_{1,1}} & +|1\rangle\,|3\rangle^{S_{1,2}} & +|3\rangle\,|3\rangle^{S_0} \end{array} \qquad (8.4.1)$$

R.H.S. is ordered according to the states received by Bob.

Consider the evolution for Alice's two states $|0\rangle_{zz}$ and $|2\rangle_{zz}$ with Eve's initial state $|E\rangle$. Then, apply the basis change $zz \mapsto xz$.

8.4.5.2 $zz \mapsto xz$ basis change

It corresponds to the transformation $\mathbb{H} \otimes \mathbb{1}_2$, and changes the basis states as follows.

$$\begin{aligned} \sqrt{2}|0\rangle_{zz} &= |0\rangle_{xz}+|1\rangle_{xz}, \\ \sqrt{2}|2\rangle_{zz} &= |2\rangle_{xz}+|3\rangle_{xz}, \\ \sqrt{2}|1\rangle_{zz} &= |0\rangle_{xz}-|1\rangle_{xz}, \\ \sqrt{2}|3\rangle_{zz} &= |2\rangle_{xz}-|3\rangle_{xz}. \end{aligned} \qquad (8.4.2)$$

8.4.5.3 $|0\rangle_{zz}$ evolved

Consider the evolution of $|0\rangle_{zz}$ with $|E\rangle$:

$$\mathcal{U}|0\rangle_{zz}|E\rangle = |0\rangle_{zz}|0\rangle_{zz}^{S_0} + |1\rangle_{zz}|0\rangle_{zz}^{S_{1,1}} + |2\rangle_{zz}|0\rangle_{zz}^{S_{1,2}} + |3\rangle_{zz}|0\rangle_{zz}^{S_2}.$$

Use Eq. (8.4.2) to replace the basis elements $|i\rangle_{zz}$ in terms of the superposition of basis elements $|r\rangle_{xz}$.

$$\begin{aligned} &\mathcal{U}\left(|0\rangle_{xz} + |1\rangle_{xz}\right)|E\rangle \\ &= \left(|0\rangle_{xz} + |1\rangle_{xz}\right)|0\rangle_{zz}^{S_0} + \left(|2\rangle_{xz} + |3\rangle_{xz}\right)|0\rangle_{zz}^{S_{1,2}} + \left(|0\rangle_{xz} - |1\rangle_{xz}\right)|0\rangle_{zz}^{S_{1,1}} + \left(|2\rangle_{xz} - |3\rangle_{xz}\right)|0\rangle_{zz}^{S_2}. \end{aligned}$$

Then, using the expansion of various $\mathcal{U}|i\rangle_{xz}|E\rangle$ in the L.H.S., we get,

$$\begin{aligned} \text{L.H.S.} &= |0\rangle_{xz}\,|0\rangle_{xz}^{S_0} + |2\rangle_{xz}\,|0\rangle_{xz}^{S_{1,2}} + |1\rangle_{xz}\,|0\rangle_{xz}^{S_{1,1}} + |3\rangle_{xz}\,|0\rangle_{xz}^{S_2} \\ &+ |0\rangle_{xz}\,|1\rangle_{xz}^{S_{1,1}} + |2\rangle_{xz}\,|1\rangle_{xz}^{S_2} + |1\rangle_{xz}\,|1\rangle_{xz}^{S_0} + |3\rangle_{xz}\,|1\rangle_{xz}^{S_{1,2}} \end{aligned}$$

Comparing the coefficients of various $|i\rangle_{xx}$ from both sides, we get

$$\begin{aligned} |0\rangle_{zz}^{S_0} + |0\rangle_{zz}^{S_{1,1}} = |0\rangle_{xz}^{S_0} + |1\rangle_{xz}^{S_{1,1}}, &\qquad |0\rangle_{zz}^{S_{1,2}} + |0\rangle_{zz}^{S_2} = |0\rangle_{xz}^{S_{1,2}} + |1\rangle_{xz}^{S_2}, \\ |0\rangle_{zz}^{S_0} - |0\rangle_{zz}^{S_{1,1}} = |0\rangle_{xz}^{S_{1,1}} + |1\rangle_{xz}^{S_0}, &\qquad |0\rangle_{zz}^{S_{1,2}} - |0\rangle_{zz}^{S_2} = |0\rangle_{xz}^{S_2} + |1\rangle_{xz}^{S_{1,2}}. \end{aligned}$$

Solving, we get,

$$\begin{aligned} 2|0\rangle_{zz}^{S_0} &= |0\rangle_{xz}^{S_0} + |1\rangle_{xz}^{S_0} + |0\rangle_{xz}^{S_{1,1}} + |1\rangle_{xz}^{S_{1,1}}, \\ 2|0\rangle_{zz}^{S_{1,1}} &= |0\rangle_{xz}^{S_0} - |1\rangle_{xz}^{S_0} - |0\rangle_{xz}^{S_{1,1}} + |1\rangle_{xz}^{S_{1,1}}, \\ 2|0\rangle_{zz}^{S_{1,2}} &= |0\rangle_{xz}^{S_{1,2}} + |1\rangle_{xz}^{S_{1,2}} + |0\rangle_{xz}^{S_2} + |1\rangle_{xz}^{S_2}, \\ 2|0\rangle_{zz}^{S_2} &= |0\rangle_{xz}^{S_{1,2}} - |1\rangle_{xz}^{S_{1,2}} - |0\rangle_{xz}^{S_2} + |1\rangle_{xz}^{S_2}. \end{aligned}$$

Taking the inner products from both sides, we get,

$$\begin{aligned} 4A &= 2(A + A_1) + 2(B + B_1), \\ 4B &= 2(A - A_1) + 2(B - B_1), \\ 4B &= 2(B + B_3) + 2(C + C_1), \\ 4C &= 2(B - B_3) + 2(C - C_1). \end{aligned}$$

Solving the first and third equations, we get two relations,

$$A_1 + B_1 = A - B, \qquad B_3 + C_1 = B - C.$$

8.4.5.4 $|2\rangle_{zz}$ evolved

Consider the evolution of $|2\rangle_{zz}$ with $|E\rangle$:

$$\mathcal{U}|2\rangle_{zz}|E\rangle \;=\; |0\rangle_{zz}\,|2\rangle_{zz}^{S_{1,2}}+|2\rangle_{zz}\,|2\rangle_{zz}^{S_0}+|1\rangle_{zz}\,|2\rangle_{zz}^{S_2}+|3\rangle_{zz}\,|2\rangle_{zz}^{S_{1,1}}.$$

Use Eq. (8.4.2) to replace the basis elements $|i\rangle_{zz}$ in terms of the superposition of basis elements $|r\rangle_{xz}$.

$$\begin{aligned}&\mathcal{U}\left(|2\rangle_{xz}+|3\rangle_{xz}\right)|E\rangle\\ &= \left(|0\rangle_{xz}+|1\rangle_{xz}\right)\,|2\rangle_{zz}^{S_{1,2}}+\left(|2\rangle_{xz}+|3\rangle_{xz}\right)\,|2\rangle_{zz}^{S_0}+\left(|0\rangle_{xz}-|1\rangle_{xz}\right)\,|2\rangle_{zz}^{S_2}+\left(|2\rangle_{xz}-|3\rangle_{xz}\right)\,|2\rangle_{zz}^{S_{1,1}}.\end{aligned}$$

Then, using the expansion of various $\mathcal{U}|i\rangle_{xz}|E\rangle$ in the L.H.S., we get,

$$\begin{aligned}\text{L.H.S.} &= |0\rangle_{xz}\,|2\rangle_{xz}^{S_{1,2}}+\;|2\rangle_{xz}\,|2\rangle_{xz}^{S_0}+\;|1\rangle_{xz}\,|2\rangle_{xz}^{S_2}+\;|3\rangle_{xz}\,|2\rangle_{xz}^{S_{1,1}}\\ &+\;|0\rangle_{xz}\,|3\rangle_{xz}^{S_2}+\;|2\rangle_{xz}\,|3\rangle_{xz}^{S_{1,1}}+\;|1\rangle_{xz}\,|3\rangle_{xz}^{S_{1,2}}+\;|3\rangle_{xz}\,|3\rangle_{xz}^{S_0}\end{aligned}$$

Comparing the coefficients of various $|i\rangle_{xx}$ from both sides, we get

$$\begin{aligned}&|2\rangle_{zz}^{S_{1,2}}+|2\rangle_{zz}^{S_2}=|2\rangle_{xz}^{S_{1,2}}+|3\rangle_{xz}^{S_2}, &\quad& |2\rangle_{zz}^{S_0}+|2\rangle_{zz}^{S_{1,1}}=|2\rangle_{xz}^{S_0}+|3\rangle_{xz}^{S_{1,1}},\\ &|2\rangle_{zz}^{S_{1,2}}-|2\rangle_{zz}^{S_2}=|2\rangle_{xz}^{S_2}+|3\rangle_{xz}^{S_{1,2}}. && |2\rangle_{zz}^{S_0}-|2\rangle_{zz}^{S_{1,1}}=|2\rangle_{xz}^{S_{1,1}}+|3\rangle_{xz}^{S_0}.\end{aligned}$$

Solving, we get,

$$\begin{aligned}2|2\rangle_{zz}^{S_0} &= |2\rangle_{xz}^{S_0}+|3\rangle_{xz}^{S_0}+|2\rangle_{xz}^{S_{1,1}}+|3\rangle_{xz}^{S_{1,1}},\\ 2|2\rangle_{zz}^{S_{1,1}} &= |2\rangle_{xz}^{S_0}-|3\rangle_{xz}^{S_0}-|2\rangle_{xz}^{S_{1,1}}+|3\rangle_{xz}^{S_{1,1}},\\ 2|2\rangle_{zz}^{S_{1,2}} &= |2\rangle_{xz}^{S_{1,2}}+|3\rangle_{xz}^{S_{1,2}}+|2\rangle_{xz}^{S_2}+|3\rangle_{xz}^{S_2},\\ 2|2\rangle_{zz}^{S_2} &= |2\rangle_{xz}^{S_{1,2}}-|3\rangle_{xz}^{S_{1,2}}-|2\rangle_{xz}^{S_2}+|3\rangle_{xz}^{S_2}.\end{aligned}$$

But, here, we do not find any new relation merely by taking the inner products from the above equations. However, once we consider the states $|0\rangle^{?}$ from the earlier subsection and compute the inner products with these $|2\rangle^{?}$ states, we get new relations. For instance, taking the inner products between $|0\rangle_{zz}^{S_0}, |2\rangle_{zz}^{S_0}$, and between $|0\rangle_{zz}^{S_{1,1}}, |2\rangle_{zz}^{S_{1,1}}$, we get,

$$\begin{aligned}4A_1 &= 2(A_1+A_2)+2(B_2+B_3),\\ 4B_1 &= 2(B_1+B_2)+2(C_1+C_2).\end{aligned}$$

Thereby, we get two more relations:

$$B_2+B_3=A_1-A_2, \qquad C_1+C_2=B_1-B_2.$$

CHAPTER 9

CONCLUSION AND FINAL REMARKS

Here we conclude by summarizing the work done, and mentioning some possible future works.

9.1 Summary of the Work Done

In this thesis, we have addressed some open-ended issues and have extended completeness of some non-structured literature. For our working purpose, we have mainly considered the attack model where each single information carrier is learned separately by the attacker. From the attackers perspective, to mount and analyze the quality of the attack, it is important to know certain parameters: like,

- the unitary evolution that entangles her information gleaning system (ancilla) with Alice's information career,
- the nature of the resultant entanglement, particularly, whether it carries optimal information for Eve.

In our work, we started with the second problem first to characterize the optimality of Eve's states [AP17]. Using this information, we then focused on the first problem [AP21] [1]. However, the second work actually is an amalgamation of two separate works, namely,

- finding a new necessary and sufficient condition, and
- characterizing the space of optimal unitary evolutions.

Those two works evolved separately, and ultimately been combined to satisfy the urge of publishing timely. An expert can easily find some scars of that joint.

[1] For bravery, an enthusiast researcher may explore the other way round as well.

Now, let's try to understand the motive behind choosing the three different problems as described in our two papers [AP17, AP21]. We also highlight the motive behind the way we tackled the problems. We address them one-by-one.

9.1.1 Characterizing the optimal post-interaction joint states [AP17]

The motive to proceed with the generalization was actually from the paper [FGG$^+$97] itself. Firstly, they themselves posed the open-ended question whether their candidate optimal interaction is a lone witness to saturate the upper bound on the maximum achievable information. The answer-hint actually was hidden there in that paper itself, and was obscured due to analyzing the attack separately for equal and unequal QBER across the two MUBs. They basically did come up with two different optimal witnesses: one for the equal-error analysis, another for the unequal-error analysis. Naturally, their doubt was valid[2].

We started attacking the issue by trying to construct new witnesses which in turn exposed the individual bricks of the problem. With that insight, we tried to find them analytically by considering a generalized possible expression for the optimal states. The first inspection was that Eve's state-space consists of two mutually orthogonal subspaces, and thereby, if a basis-state is included to describe a state living in one of the subspaces, the same basis-state can be excluded while describing a state living in its orthogonal subspace [3]. Then, if one describe Eve's states in the measurement basis, then one can calculate the probabilities to find a specific measurement outcome for a given signal from Alice. These probabilies can then be used to calculate the information gain which can then be compared with the optimal value. This approach, that started with some generalized expression for the optimal states, could eventually expose an infinite collection of optimal states. Nevertheless, for different rotations of the measurement apparatus, one can get different witness for the optimal information. Two of them are in [FGG$^+$97], for which one needs to calculate the optimal measurement separately In our case, the optimal states are described in terms of the measurement basis, and therefore doesn't require calculating the measurement directions, rather one just chooses a measurement direction to define the optimal states. With these mathematical tools in hand, mounting an attack becomes simpler: choose a measurement setup (which in turn defines the optimal states), and find the associated optimal unitary evolution along with an initial state. The later problem is described below.

[2] To gain that insight, the first requirement was to understand the delicate fabrics of the work [FGG$^+$97] completely by own hand-held calculations and create more examples owing to that insight. We have dedicated Chapter 3 entirely to consolidate that extensive journey. Some of our own results done for our first work are placed in that chapter in order to maintain the flow of thoughts. So, it is a semi-contributory chapter

[3] Now, one may argue whether it exhaust the state-space or not, which remained a part of the motivations behind finding a new NSC [AP21] that in turn figures out the optimal states.

9.1.2 Finding a new and efficient NSC [AP21]

The urge of finding a new NSC is two-fold.

Firstly, whether we can characterize the optimal states as the outcome of an NSC. In that case, we can certainly conclude that those states exhaust the entire space of the optimal states. We have done that part successfully which in turn establish our first published work firmly [4].

Secondly, some deep thoughts following long-time works precipitated a few important observations which demands a new NSC, and are described as follows. While certifying optimality of a given set of PIJSs, the earlier extant approach [FGG^{+}97] required the knowledge (specification) of the optimal measurement that in turn is used to testify a set of necessary and sufficient conditions. Moreover, in that approach, to testify optimality for the states in a given basis, one needs to involve the PIJSs in the conjugate basis. We felt strongly that both these specifications could be avoided while the reason is as follows.

(i) Given a set of PIJSs, if they are optimal, the optimality exists irrespective of whether one knows an optimal measurement or not. Therefore, specifying the measurement can possibly be avoided.

(ii) The signature of optimality of the PIJSs for a given basis should be reflected in the constituent IVs with Eve, and therefore, can be testified with these states only. So, the involvement of the PIJSs or the IVs in the conjugate basis can be avoided in order to verify the NSCs for the given pair of PIJSs.

Having these shortcomings in the back of the mind, we started the venture to come up with a NSC that doesn't involve any specification of the measurement apparatus and doesn't unnecessarily involve Eves states (or PIJSs) in the conjugate basis. Moreover, when the signature of optimality in a pair of PIJSs are likely to be inherited in the constituent IVs, why don't we find a NSC involving these IVs only?

We have ultimately deduced such a NSC through a series of NSCs. Now, if one follow those intermediate NSCs closely, one can notice that some of those initial NSCs involve Eve's states in the conjugate basis, which is a part of the transition that took place from the then existing NSC to our desired NSC. The simplicity of that master NSC is that it needs only to calculate the overlaps between the IVs. The ingenuity of the NSC is that it depicts the geometry of the optimality, i.e., the orientation of the IVs in space. A more interesting fact is that the optimal states automatically emanate out of these NSCs, thereby characterizing the only and all of the optimal states.

[4] Although the approach in the first work [AP17] remained quite rudimentary, some pieces of the puzzle still remained a bit illusive and that incompleteness was eminent in the end part of the paper. The clarity was revealed a long later and is put in the prelude of the arxiv version of the second paper.

9.1.3 Characterizing the optimal unitary evolutions [AP21]

To mount an optimal attack, all an attacker requires for practical purposes is a specification of the following triplet:

(optimal unitary, ancilla-state, optimal measurement).

An intermediate byproduct of the attack are the optimal states with Eve, that we already have found exhaustively. We can then apply reverse-engineering to find the optimal unitaries exhaustively for those optimal states.

Unfortunately, the literature is silent to tackle such problems. For a given quad of optimal states, it is a mathematical problem from linear algebra to find the optimal unitary. For another set of such states, one needs to approach afresh to get the associated unitary. But, we have infinitely many such *quad of IVs* which in their general form is parametric in choosing the describing (measurement) basis. Thus, the above-said approach fails to tackle our infinite population of optimal states at a go. We had to face a lot of troubles to tackle that problem. A prolonged effort culminated in an unusual approach to get the first breakthrough (as in arxiv: the first version). However, it was not enough to tackle some more degrees of generalizations. Then we had a more rudimentary approach to tackle the problem again in a different way. Finally, we could link these two seemingly different approaches which gradually answered all our questions in hand.

Note: It is also quite encouraging to mention that we could maintain the generalized approach with error rates not necessarily same across the two bases. Although it didn't inculcate any immediate significant advantage (future will say whether there is any), it surely provides much more sublime intricacies than the general trend (found in the existing literature) of considering the analysis with equal errors across the bases. To switch from our generalized approach to the same-error mode, one can simply replace the two error-rates by a single variable (QBER).

9.2 Future Work

Although we have studied extensively the issues in hand, we have noticed some more complex problems that an interested researcher may explore.

Connections: Our work has established some connections involving various aspects of quantum cryptography more explicitly, e.g., some deep connections between a prepare-and-measure scheme and its entanglement-based counterpart, p&m scheme and pc-cloner etc. But, it is strongly felt that much more such connections could (and should) be established if one tries to approach differently the same problems, e.g., from the scratch. One can then literally represent the existing knowledge of the QKD like a collage, by establishing more hidden connections between various fields as a schematic view. This, however, seems to involve at least a researcher who has understood the bottomline of the existing quantum cryptography.

Unitary-optimality: Although we have characterized the space of optimal unitary evolutions, a broader question may be the following. Given an arbitrary unitary evolution, how to certify whether it is optimal or not.

If we follow our approach, we can consider an ancilla having some specific state, then calculate the resulted post-interaction joint states, and finally test for optimality using a necessary and sufficient condition. It is a bit calculation intensive work for an attacker indeed.

It is interesting to explore whether the intermediate knowledge of the PIJSs is essential or not in order to tackle this specific problem. As such the signature of the optimality is content within the unitary itself. So, one may also try to come up with some direct approach to certify the optimality of an unitary.

Given an arbitrary unitary, certifying optimality may address the following questions.

- Is there an ancillary state that allows it to produce optimal PIJSs?
- If yes, are there other ancillary states that serves the same purpose?
- If yes, does an arbitrary ancillary state serve the purpose?

Circuit: Irrespective of the feasibility within the current experimental setup, at least for the sake of theoretical quest, some domain specific researcher may consider the problem of designing plausible circuits for an unitary attack. The main difficulty here is due to the involvement of the eight dimensional space owing to three qubits. Given an arbitrary optimal unitary, and an ancillary state, whether one can design a circuit using universal quantum gates. Then one can consider the efficiency of the design. Can one come up with some generalized approach to tackle an arbitrary optimal unitary?

Generalize: In research, the scope to work on new issues is always there by trying to generalize an work in various ways. For our work, following are some such scopes.

- We have considered Eve's ancilla consisting of two qubits spanning four dimensions. One may wonder whether and how a single four-dimensional quantum bit serves the purpose.
- What if Alice considers her quantum encoder as a higher dimensional object, say, qudits. Can one come up with a necessary and sufficient condition for

optimality of the attack? Can one analyze the attack for unequal error-rates across the encoding bases?

- What happens when the number of encoding bases are increased? For instance, one can consider the six-state protocol involving three MUBs and try to incorporate the ideas from our work in order to analyze it from general perspective.

The above-mentioned problems seem to be quite challenging and remained out of our scope and maturity to tackle. We wish all the best for the future generation to fiddle with and pass quality time with those problems.

REFERENCES

[ABB+14] R. Alléaume, C. Branciard, J. Bouda, T. Debuisschert, M. Dianati, N. Gisin, M. Godfrey, P. Grangier, T. Länger, N. Lütkenhaus, C. Monyk, P. Painchault, M. Peev, A. Poppe, T. Pornin, J. Rarity, R. Renner, G. Ribordy, M. Riguidel, L. Salvail, A. Shields, H. Weinfurter, and A. Zeilinger. Using quantum key distribution for cryptographic purposes: A survey. *Theoretical Computer Science*, 560:62 – 81, 2014. Theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84. URL: http://www.sciencedirect.com/science/article/pii/S0304397514006963, doi:https://doi.org/10.1016/j.tcs.2014.09.018. [Page 1]

[AP17] Atanu Acharyya and Goutam Paul. Revisiting optimal eavesdropping in quantum cryptography: Optimal interaction is unique up to rotation of the underlying basis. *Phys. Rev. A*, 95:022326, Feb 2017. URL: https://link.aps.org/doi/10.1103/PhysRevA.95.022326, doi:10.1103/PhysRevA.95.022326. [Page iii, v, xvi, xvii, xviii, 2, 3, 47, 48, 51, 99, 100, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 117, 118, 119, 120, 123, 149, 179, 180, 181]

[AP21] Atanu Acharyya and Goutam Paul. A complete characterization of the optimal unitary attacks in quantum cryptography with a refined optimality criteria involving the attacker's Hilbert space only. *The European Physical Journal D*, 75(8):215, Jul 2021. URL: https://doi.org/10.1140/epjd/s10053-021-00203-7, doi:10.1140/epjd/s10053-021-00203-7. [Page iii, v, xvi, xvii, xviii, 2, 3, 117, 118, 120, 122, 124, 126, 128, 130, 132, 133, 134, 136, 138, 140, 142, 144, 146, 149, 152, 153, 154, 155, 157, 179, 180, 181, 182]

[BB84] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing*, volume 1 of *3*, pages 175–179, Bangalore, India, 1984. URL: http://researcher.watson.ibm.com/researcher/files/us-bennetc/BB84highest.pdf. [Page xiv, 33, 47, 149]

[BB14] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical Computer Science*, 560:7 – 11, 2014. Theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84. URL: http://www.sciencedirect.com/science/article/pii/S0304397514004241, doi:https://doi.org/10.1016/j.tcs.2014.05.025. [Page 33, 149]

[BBBW82] Charles H. Bennett, Gilles Brassard, Seth Breidbard, and Stephen Wiesner. Quantum cryptography, or unforgeable subway tokens. In *Advances in Cryptology: Proceedings of CRYPTO '82*, pages 267–275. Plenum, 1982. [Page xiv, xix, 42, 44]

[BBCM95] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer. Generalized privacy amplification. *IEEE Transactions on Information Theory*, 41(6):1915–1923, Nov 1995. [Page 150, 152, 154]

[BBG$^+$02] Alexios Beveratos, Rosa Brouri, Thierry Gacoin, André Villing, Jean-Philippe Poizat, and Philippe Grangier. Single photon quantum cryptography. *Phys. Rev. Lett.*, 89:187901, Oct 2002. URL: https://link.aps.org/doi/10.1103/PhysRevLett.89.187901, doi:10.1103/PhysRevLett.89.187901. [Page 1]

[BBM92] Charles H. Bennett, Gilles Brassard, and N. David Mermin. Quantum cryptography without Bell's theorem. *Phys. Rev. Lett.*, 68:557–559, Feb 1992. URL: https://link.aps.org/doi/10.1103/PhysRevLett.68.557, doi:10.1103/PhysRevLett.68.557. [Page xiv, 24, 33, 40]

[BCMDM00] Dagmar Bruß, Mirko Cinchetti, G. Mauro D'Ariano, and Chiara Macchiavello. Phase-covariant quantum cloning. *Phys. Rev. A*, 62:012302, Jun 2000. URL: https://link.aps.org/doi/10.1103/PhysRevA.62.012302, doi:10.1103/PhysRevA.62.012302. [Page 65, 118, 126]

[Ben92] Charles H. Bennett. Quantum cryptography using any two nonorthogonal states. *Phys. Rev. Lett.*, 68:3121–3124, May 1992. URL: https://link.aps.org/doi/10.1103/PhysRevLett.68.3121, doi:10.1103/PhysRevLett.68.3121. [Page xiv, 33, 41]

[BM02] D. Bruß and C. Macchiavello. Optimal eavesdropping in cryptography with three-dimensional quantum states. *Phys. Rev. Lett.*, 88:127901, Mar 2002. URL: https://link.aps.org/doi/10.1103/PhysRevLett.88.127901, doi:10.1103/PhysRevLett.88.127901. [Page 158]

[BPG99] H. Bechmann-Pasquinucci and N. Gisin. Incoherent and coherent eavesdropping in the six-state protocol of quantum cryptography. *Phys. Rev. A*, 59:4238–4248, Jun 1999. URL: https://link.aps.org/doi/10.1103/PhysRevA.59.4238, doi:10.1103/PhysRevA.59.4238. [Page 152, 153, 154, 162]

[Bru98] Dagmar Bruß. Optimal eavesdropping in quantum cryptography with six states. *Phys. Rev. Lett.*, 81:3018–3021, Oct 1998. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.3018, doi:10.1103/PhysRevLett.81.3018. [Page xiv, 33, 37, 149, 158]

[BS94] G. Brassard and L. Salvail. Secret-key reconciliation by public discussion. In *Helleseth T. (eds) Advances in Cryptology — EUROCRYPT '93. EUROCRYPT 1993. Lecture Notes in Computer Science*, volume 765, pages 175–179, Berlin, Heidelberg, 1994. Springer. URL: https://doi.org/10.1007/3-540-48285-7_35. [Page 149]

[BZ01] φCaslav Brukner and Anton Zeilinger. Conceptual inadequacy of the Shannon information in quantum measurements. *Phys. Rev. A*, 63:022113, Jan 2001. URL: https://link.aps.org/doi/10.1103/PhysRevA.63.022113, doi:10.1103/PhysRevA.63.022113. [Page 152]

[CG97] J.I Cirac and N Gisin. Coherent eavesdropping strategies for the four state quantum cryptography protocol. *Physics Letters A*, 229(1):1–7, 1997. URL: http://www.sciencedirect.com/science/article/pii/S037596019700176X, doi:https://doi.org/10.1016/S0375-9601(97)00176-X. [Page 161, 163, 171]

[CHSH69] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, Oct 1969. URL: https://link.aps.org/doi/10.1103/PhysRevLett.23.880, doi:10.1103/PhysRevLett.23.880. [Page xiv, 26, 118]

[Cir80] B. S. Cirel'son. Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics*, 4(2):93–100, Mar 1980. URL: https://doi.org/10.1007/BF00417500, doi:10.1007/BF00417500. [Page xiv, 28, 118]

[CK78] I. Csiszár and J. Körner. Broadcast channels with confidential messages. *IEEE Transactions on Information Theory*, 24(3):339–348, May 1978. doi:10.1109/TIT.1978.1055892. [Page 66]

[CT06] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, USA, 2006. [Page 54]

[EHPP94] Artur K. Ekert, Bruno Huttner, G. Massimo Palma, and Asher Peres. Eavesdropping on quantum-cryptographical systems. *Phys. Rev. A*, 50:1047–1056, Aug 1994. URL: https://link.aps.org/doi/10.1103/PhysRevA.50.1047, doi:10.1103/PhysRevA.50.1047. [Page 66]

[Eke91] Artur K. Ekert. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.*, 67:661–663, Aug 1991. URL: https://link.aps.org/doi/10.1103/

PhysRevLett.67.661, doi:10.1103/PhysRevLett.67.661. [Page xiv, 33, 38]

[Eng13] Berthold-Georg Englert. On quantum theory. *Eur. Phys. J. D*, 67:238, 2013. doi:10.1140/epjd/e2013-40486-5. [Page 1]

[FGG$^+$97] C. A. Fuchs, N. Gisin, R. B. Griffiths, C. S. Niu, and A. Peres. Optimal eavesdropping in quantum cryptography. I. information bound and optimal strategy. *Phys. Rev. A*, 56:1163–1172, Aug 1997. URL: https://link.aps.org/doi/10.1103/PhysRevA.56.1163, doi:10.1103/PhysRevA.56.1163. [Page xiv, xvi, 1, 2, 3, 42, 47, 48, 51, 52, 53, 55, 56, 57, 59, 60, 63, 67, 68, 70, 98, 99, 100, 101, 102, 103, 107, 109, 111, 112, 113, 114, 115, 117, 118, 119, 120, 123, 124, 125, 130, 133, 149, 152, 153, 154, 158, 180, 181]

[Fuc96] C. A. Fuchs. Information Gain vs. State Disturbance in Quantum Theory. 1996. arXiv:9611010. [Page 52, 59, 68, 69, 70]

[HBHP08] I. M. Herbauts, S. Bettelli, H. Hübel, and M. Peev. On the optimality of individual entangling-probe attacks against BB84 quantum key distribution. *The European Physical Journal D*, 46(2):395–406, 2008. URL: https://doi.org/10.1140/epjd/e2008-00002-x, doi:10.1140/epjd/e2008-00002-x. [Page 154]

[Hel69] Carl W. Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 1:231–252, Jun 1969. URL: https://doi.org/10.1007/BF01007479, doi:10.1007/BF01007479. [Page 65]

[Lüt96] Norbert Lütkenhaus. Security against eavesdropping in quantum cryptography. *Phys. Rev. A*, 54:97–111, Jul 1996. URL: https://link.aps.org/doi/10.1103/PhysRevA.54.97, doi:10.1103/PhysRevA.54.97. [Page 154]

[Lüt99] Norbert Lütkenhaus. Estimates for practical quantum cryptography. *Phys. Rev. A*, 59:3301–3319, May 1999. URL: https://link.aps.org/doi/10.1103/PhysRevA.59.3301, doi:10.1103/PhysRevA.59.3301. [Page 154]

[NC11] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, USA, 10th edition, 2011. [Page 7, 11, 22, 52]

[Ras19] Alexey E. Rastegin. Individual attacks with generalized discrimination and inadequacy of some information measures. *Quantum Information Processing*, 18(9):276, 2019. URL: https://doi.org/10.1007/s11128-019-2385-4, doi:10.1007/s11128-019-2385-4. [Page 152]

[Rén61] Alfréd Rényi. On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pages 547–561, Berkeley, Calif., 1961. University of California Press. URL: https://projecteuclid.org/euclid.bsmsp/1200512181. [Page 152]

[Ste04] J. Michael Steele. *The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities*. Cambridge University Press, 2004. doi:10.1017/CBO9780511817106. [Page 10]

[WW] Hui-Hua Wu and Shanhe Wu. Various proofs of the cauchy-schwarz inequality. URL: https://rgmia.org/papers/v12e/Cauchy-Schwarzinequality.pdf. [Page 10]

[WZ82] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned. *Nature*, 299(5886):802–803, Oct 1982. URL: https://doi.org/10.1038/299802a0, doi:10.1038/299802a0. [Page xiv, 9, 31]

[ZBB$^+$05] Zoller, P., Beth, Th., Binosi, D., Blatt, R., Briegel, H., Bruss, D., Calarco, T., Cirac, J. I., Deutsch, D., Eisert, J., Ekert, A., Fabre, C., Gisin, N., Grangiere, P., Grassl, M., Haroche, S., Imamoglu, A., Karlson, A., Kempe, J., Kouwenhoven, L., Kröll, S., Leuchs, G., Lewenstein, M., Loss, D., Lütkenhaus, N., Massar, S., Mooij, J. E., Plenio, M. B., Polzik, E., Popescu, S., Rempe, G., Sergienko, A., Suter, D., Twamley, J., Wendin, G., Werner, R., Winter, A., Wrachtrup, J., and Zeilinger, A. Quantum information processing and communication - strategic report on current status, visions and goals for research in europe. *Eur. Phys. J. D*, 36(2):203–228, 2005. URL: https://doi.org/10.1140/epjd/e2005-00251-1, doi:10.1140/epjd/e2005-00251-1. [Page 1]