Optimal Eavesdropping in Quantum Cryptography: Choice of Interaction Is Unique up to a Rotation of the Underlying Basis

Atanu Acharyya
Indian Statistical Institute, Kolkata–700108, India *

Abstract
A general framework of optimal eavesdropping on BB84 protocol was provided by Fuchs et al. in 1997. An upper bound on mutual information was derived in their work, which could be achieved by a specific type of interaction and the corresponding measurement. However, uniqueness of an optimal interaction was left as an open problem there. We resolve this problem here and establish the uniqueness. Nevertheless, the description of an optimal interaction changes as the basis of description gets rotated. The specific choice of optimal interaction by Fuchs et al. is shown to be a special case of the form provided in our work.

Keywords: BB84 Protocol, Key Distribution, Information Gain, Mutual Information, Optimal Eavesdropping, Optimal Interaction, Quantum Cryptography.

1 INTRODUCTION
Cryptography addresses the problem of sharing some secret information between two parties. The information (plain-text) is encrypted by the sender Alice using a secret key and send over an insecure channel to the receiver Bob. To decrypt the garbled message, Bob must have a key. In classical cryptography the key distribution is done by asymmetric key communication, where the encryption key is made public while the decryption key remains secret. However, the security of this system relies on high complexity of the inverse transformation (e.g., prime factorization of large integers) that could be broken by faster algorithms, for example, by Shor’s algorithm

*Email: pub.academy.15@gmail.com
in quantum computation. An alternative is provided by symmetric key cryptography, where the encryption and decryption keys are same. This system is secure if the key is not re-used, thus known as the “one-time pad”. But the main problem here is to distribute a secret key to the receiver. An eavesdropper Eve may get hold of the key, copy it and send it to the receiver. Quantum cryptography provides a safer way where key distribution makes use of quantum states while Eve cannot measure them without disturbing them and the disturbance is detectable by Bob. Moreover, no-cloning theorem ensures that Eve cannot make a copy of the state sent by Alice. Quantum key distribution (QKD) has thus gained importance over time.

The first and possibly the most celebrated QKD protocol is BB84 [1]. The protocol relies on the use of non-orthogonal states from one of the two conjugate bases, say, \(x-y\) and \(u-v\) to encode a bit string in qubits (e.g., polarized photons). Alice randomly selects one of the two orthogonal bases and encodes 0 and 1 respectively by a qubit prepared in one of the two states in each base - viz., Alice encodes 0 to \(|x\rangle\) or \(|u\rangle\), and 1 to \(|y\rangle\) or \(|v\rangle\), depending on the chosen basis. When Bob receives a state from Alice, he randomly selects a basis \(x-y\) or \(u-v\) and makes a measurement. Once the measurement is done for all the received qubits, Alice and Bob publicly announce the sequence of bases used by them and discard the bits where the bases do not match. The resulting bit string, followed by error correction and privacy amplification, becomes the common secret key. However, presence of an eavesdropper may disturb the state of a qubit sent by Alice for which Bob may get a wrong result even if the corresponding basis of measurement between Alice and Bob matches. To overcome this problem, Alice and Bob sacrifice some of the bits by comparing their values publicly.

Fuchs et al. in [4] provided a general framework of optimal eavesdropping on BB84 protocol. The authors there derived an upper bound on mutual information, described a specific type of interaction and the corresponding measurement that achieves the bound, and finally has explained an optimal strategy for Eve in interpreting her measurement. However, the optimal interaction described there was a specific choice and the uniqueness of an optimal interaction was left as an open problem. We address this problem here and establish the uniqueness.

The content of this paper is organized as follows: Section 2 explains basic terminologies used for optimal eavesdropping introduced in [4]. Section 3 contains certain results from [4] which are relevant to our work. Our results are explained in Section 4. The remaining portion discusses the connection of our results with [4] followed by a conclusion.

Our contribution mainly concerns the uniqueness of an optimal interac-
2 PRELIMINARIES

Alice and Bob want to share a secret key using BB84 protocol. Alice randomly chooses a basis from $\mathcal{B}_{xy} = \{ |x\rangle, |y\rangle \}$ and $\mathcal{B}_{uv} = \{ |u\rangle, |v\rangle \}$, such that

$$
|x\rangle = \frac{1}{\sqrt{2}} (|u\rangle + |v\rangle)
$$

$$
|y\rangle = \frac{1}{\sqrt{2}} (|u\rangle - |v\rangle)
$$

i.e., the bases are conjugate to each other. Then Alice encodes her key-bits, each as a polarized photon, and sends it to Bob. Suppose, an eavesdropper Eve interferes the communication while she lets a probe interact unitarily with the qubit sent by Alice.

Let, Alice has chosen a signal, say, $|x\rangle$ (corresponding density operator $\rho_A^x = |x\rangle\langle x|$, in the basis $\mathcal{B}_{xy}$). Eve lets a probe, initially in state $|\psi_0\rangle$ (corresponding density operator $\rho_E^0 = |\psi_0\rangle\langle \psi_0|$), interact unitarily (realized by a unitary operator $U$) with the qubit sent by Alice. The post-interaction joint state $|X\rangle$ between Alice and Eve, which is an entangled state of the probe of Eve and the photon sent by Alice, is realized by

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

Bob receives a simple mixture of the two basis vectors (here $\mathcal{B}_{xy}$) chosen by Alice, i.e., Bob’s density matrix is always diagonal in the basis chosen by Alice. Thus, Schmidt decomposition of the post-interaction joint state $|X\rangle$ must be of the form:

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

such that

$$
|\xi_x\rangle \perp |\zeta_x\rangle
$$

where $|\xi_x\rangle, |\zeta_x\rangle$ are component of Eve’s part of the joint state after the interaction.
2 PRELIMINARIES

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

$$|\xi_y\rangle \perp |\zeta_y\rangle$$

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

$$\rho_{AE}^X = |X\rangle\langle X| = U\left(\rho_A^x \otimes \rho_E^0\right)U^{\dagger}$$

Eve’s description of the system will be:

$$\rho_x := \rho_E^x = \text{tr}_A\left(\rho_{AE}^x\right) = \text{tr}_A\left(|X\rangle\langle X|\right)$$

where $\text{tr}_A$ represents partial trace over Alice’s qubit.

Since the interaction is unitary, it follows, using Equations (1) and (7) that

$$|X\rangle = \frac{1}{\sqrt{2}}\left(|U\rangle + |V\rangle\right)$$

$$|Y\rangle = \frac{1}{\sqrt{2}}\left(|U\rangle - |V\rangle\right)$$

**Measurement of Eve:** Before performing any measurement Eve waits till Alice declares her choice of basis publicly. Eve’s measurement is considered to be a Positive Operator-Valued Measure (POVM) $\{E_\lambda\}$ or $\{F_\lambda\}$ depending on whether Alice’s choice is $x$-$y$ or $u$-$v$ basis. Note that the operators $\{E_\lambda\}$ satisfy two properties [4, 5]:

\[
\langle \gamma | E_\lambda | \gamma \rangle \geq 0, \quad \forall |\gamma\rangle
\]

and satisfy the completeness relation

$$\sum_\lambda E_\lambda = 1$$

Suppose, Alice sends a signal in $x$-$y$ (or, $u$-$v$) basis with prior probabilities $p_x, p_y$ (or, $p_u, p_v$) respectively. Once Alice reveals her basis to be $x$-$y$, Eve uses a POVM $\{E_\lambda\}$ to perform a measurement on her probe. Considering $A, B, E$ be the random variables corresponding to the signal sent
by Alice, signal received by Bob, and measurement outcome of Eve, the conditional probability of the various outcomes $\lambda$ of that measurement is

$$P_{\lambda x} := \Pr[\mathcal{E} = \lambda | \mathcal{A} = x] = \langle X | \mathbb{1} \otimes E_\lambda | X \rangle = \text{tr} (\rho_x E_\lambda) \tag{12}$$

$$P_{\lambda y} := \Pr[\mathcal{E} = \lambda | \mathcal{A} = y] = \langle Y | \mathbb{1} \otimes E_\lambda | Y \rangle = \text{tr} (\rho_y E_\lambda) \tag{13}$$

Henceforth, we use the notation $:= \text{to denote “defined as”}$.

The probability that Eve gets outcome $\lambda$, when Alice uses $x$-$y$ basis is thus

$$q_\lambda := \Pr[\mathcal{E} = \lambda | \mathcal{A} \in \{x, y\}] = P_{\lambda x} p_x + P_{\lambda y} p_y \tag{14}$$

Looking at outcome $\lambda$ Eve assigns a guess for the signal sent by Alice following some strategy. The posterior probability $Q_{x\lambda}$ (or $Q_{y\lambda}$) that Eve assigns $x$ (or $y$) is, by Bayes’ theorem,

$$Q_{x\lambda} := \Pr[\mathcal{A} = x | \mathcal{E} = \lambda] = \frac{P_{\lambda x} p_x}{q_\lambda} \tag{15}$$

$$Q_{y\lambda} := \Pr[\mathcal{A} = y | \mathcal{E} = \lambda] = \frac{P_{\lambda y} p_y}{q_\lambda} \tag{16}$$

A simple way that Eve can utilize these likelihoods to assign a signal is given by the following function:

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

A convenient measure of Eve’s information gain for an outcome $\lambda$, as proposed in [4], is

$$G_\lambda := |Q_{x\lambda} - Q_{y\lambda}| \tag{17}$$

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

$$G_{xy} := \sum_\lambda q_\lambda G_\lambda = \sum_\lambda \left| P_{\lambda x} p_x - P_{\lambda y} p_y \right| \tag{18}$$

For equiprobable signals,

$$G_{xy} = \frac{1}{2} \sum_\lambda \left| P_{\lambda x} - P_{\lambda y} \right| \tag{19}$$

A more sophisticated data processing by Eve is mutual information [2] - with equal prior probabilities, this is given by,

$$I_{xy} := \ln 2 + \sum_\lambda q_\lambda \left( Q_{x\lambda} \ln Q_{x\lambda} + Q_{y\lambda} \ln Q_{y\lambda} \right) \tag{20}$$
3 Revisit to Optimal Eavesdropping by Fuchs et al.

In this section we recollect the results given by [4]: an upper bound on information gain ($G$) and mutual information ($I$), followed by a necessary and sufficient condition to achieve the bounds and finally an optimal interaction (and the corresponding POVM) for unequal and equal error rates.

3.1 An Upper Bound on Information Gain ($G$) and Mutual Information ($I$)

Fuchs et al. [4] provided an upper bound on information gain ($G$). This bound was used to provide with an upper bound on mutual information ($I$). A necessary and sufficient condition to achieve the bounds was given there. We recollect these results here.

An Upper Bound on Information Gain ($G$):

**Proposition 1.** [4, Equation (23,24)] For a given POVM $\{E_\lambda\}$,

$$G_{xy} \leq 2\sqrt{D_{uv}(1-D_{uv})}$$ (22)

Indices $xy, uv$ are introduced to emphasize that Eve’s information gain refers to signals sent in the $x$-$y$ basis, and Bob’s error rate refers to signals sent in the $u$-$v$ basis. Similarly,

$$G_{uv} \leq 2\sqrt{D_{xy}(1-D_{xy})}$$ (23)

Notation: We consider $G^*_{xy}$ (or, $G^*_{uv}$) to denote optimal value of information gain.
An Upper Bound on Mutual Information (I):

**Proposition 2.** [4, Equation (31,32)] For a given POVM \( \{E_\lambda\} \),

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

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

where

\[
\phi(z) = (1 + z) \ln (1 + z) + (1 - z) \ln (1 - z)
\]

Subscripts are added to emphasize that the mutual information and error rate refer to signals sent in two different bases.

Notation: We consider \( I^*_{xy} \) (or, \( I^*_{uv} \)) to denote optimal value of mutual information.

Necessary and Sufficient Conditions to Maximize Information Gain:

**Proposition 3.** [4, Equation (38,39)] The necessary and sufficient conditions for equality in Equation (22) are

\[
|V_{\lambda u}\rangle = \varepsilon_\lambda \sqrt{\frac{D_{uv}}{1 - D_{uv}}} |U_{\lambda u}\rangle
\]

and

\[
|U_{\lambda v}\rangle = \varepsilon_\lambda \sqrt{\frac{D_{uv}}{1 - D_{uv}}} |V_{\lambda v}\rangle
\]

where

\[
\varepsilon_\lambda = \pm 1 = \text{sgn} \left( Q_{x\lambda} - Q_{y\lambda} \right)
\]

and

\[
U_{\lambda u} = B_u \otimes \sqrt{E_\lambda} |U\rangle, \quad V_{\lambda u} = B_u \otimes \sqrt{E_\lambda} |V\rangle; \quad B_u = |u\rangle\langle u| \]

\[
U_{\lambda v} = B_v \otimes \sqrt{E_\lambda} |U\rangle, \quad V_{\lambda v} = B_v \otimes \sqrt{E_\lambda} |V\rangle; \quad B_v = |v\rangle\langle v|
\]

It is intriguing to note that the set of conditions that optimize \( G \) also optimize \( I \). Therefore, the necessary and sufficient conditions for equality in Equation (24) is same as in Proposition 3. Hence, any measurement or strategy to maximize \( G,I \) will remain same.
3.2 Description of the Post-interaction States $|X\rangle, |Y\rangle$

For optimal $G_{xy}$, the post-interaction states are:

$$|X\rangle = \sqrt{1 - D_{xy}} |x\rangle |\xi_x\rangle + \sqrt{D_{xy}} |y\rangle |\zeta_x\rangle$$

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

(30)

Assuming that all inner products $\langle \xi_i | \zeta_j \rangle$ are real, the restriction on $|\xi_i\rangle, |\zeta_j\rangle$ in Equations (4) and (6) becomes more restricted as

$$\langle \{\xi_x\rangle, |\xi_y\rangle \rangle \perp \langle \{\zeta_x\rangle, |\zeta_y\rangle \rangle$$

(31)

Similarly, for optimal $G_{uv}$, the post-interaction states are:

$$|U\rangle = \sqrt{1 - D_{uv}} |u\rangle |\xi_u\rangle + \sqrt{D_{uv}} |v\rangle |\zeta_u\rangle$$

$$|V\rangle = \sqrt{1 - D_{uv}} |v\rangle |\xi_v\rangle + \sqrt{D_{uv}} |u\rangle |\zeta_v\rangle$$

(32)

Since the bases $B_{xy}$ and $B_{uv}$ are conjugate to each other, we expect to get a relationship between $|\xi_i\rangle, |\zeta_j\rangle$ in $u-v$ basis and those in $x-y$ basis which is described below:

$$2\sqrt{1 - D_{uv}} |\xi_u\rangle = \sqrt{1 - D_{xy}} (|\xi_x\rangle + |\xi_y\rangle) + \sqrt{D_{xy}} (|\zeta_x\rangle + |\zeta_y\rangle)$$

$$2\sqrt{D_{uv}} |\zeta_u\rangle = \sqrt{1 - D_{xy}} (|\xi_x\rangle - |\xi_y\rangle) + \sqrt{D_{xy}} (|\zeta_y\rangle - |\zeta_x\rangle)$$

(33)

Similarly,

$$2\sqrt{1 - D_{uv}} |\xi_v\rangle = \sqrt{1 - D_{xy}} (|\xi_x\rangle + |\xi_y\rangle) - \sqrt{D_{xy}} (|\zeta_x\rangle + |\zeta_y\rangle)$$

$$2\sqrt{D_{uv}} |\zeta_v\rangle = \sqrt{1 - D_{xy}} (|\xi_x\rangle - |\xi_y\rangle) - \sqrt{D_{xy}} (|\zeta_y\rangle - |\zeta_x\rangle)$$

(34)

From the orthogonality relation (31), we can say that Eve’s probe lives in a Hilbert space of at most 4-dimensions, thus, taken to be 2 qubits (4 states). It is therefore convenient to introduce same bases ($x-y$ and $u-v$, used by Alice) for each of Eve’s qubits.

3.3 Optimal Interaction to maximize $G, I$ - A Specific Choice

A particular choice of Eve’s interaction ($|\xi_i\rangle, |\zeta_j\rangle$) needs to be fixed now. Any choice that leads to optimality could then be chosen. In [4, Section III: Equations (50,51)], one such specific choice was made for unequal error rates, which was shown to be a correct choice (correct in the sense that the choice leads to optimality). Similarly, for equal error rates, another specific choice was made in [4, Section IV, Equation (69)]. However, uniqueness of the choice was left as an open problem - as [4, Section III, first paragraph] states that “It is easy to check that the solution here is correct, but the extent to which it is unique (aside from trivial changes of basis and of phase) remains unknown”. In this paper we conclude the problem.
3.3.1 For Unequal Error Rates, i.e., \(D_{xy} \neq D_{uv}\)

Equations (50), (51) of [1] Section III are restated here.
Consider a canonical basis for Eve’s probe as \(\{|E_0\rangle, |E_1\rangle, |E_2\rangle, |E_3\rangle\}\), w.l.g.,

\[
|E_0\rangle = |x\rangle|x\rangle, \quad |E_1\rangle = |y\rangle|x\rangle, \quad |E_2\rangle = |x\rangle|y\rangle, \quad |E_3\rangle = |y\rangle|y\rangle
\] (35)

To describe \(|\xi_i\rangle, |\zeta_j\rangle\), [4] considered an orthonormal set - two standard (maximally) entangled bases, a Bell Basis w.r.t. \(x,y\):

\[
|\Phi_{xy}^\pm \rangle := \frac{1}{\sqrt{2}} (|x\rangle|y\rangle \pm |y\rangle|x\rangle) = \frac{1}{\sqrt{2}} (|E_0\rangle \pm |E_3\rangle)
\]

\[
|\Psi_{xy}^\pm \rangle := \frac{1}{\sqrt{2}} (|x\rangle|y\rangle \pm |y\rangle|x\rangle) = \frac{1}{\sqrt{2}} (|E_2\rangle \pm |E_1\rangle)
\] (36)

In terms of Bell basis vectors for Eve’s probe, the interaction was chosen such that

\[
|\xi_x\rangle = \sqrt{1 - D_{uv}} |\Phi_{xy}^+ \rangle + \sqrt{D_{uv}} |\Phi_{xy}^- \rangle
\]

\[
|\xi_y\rangle = \sqrt{1 - D_{uv}} |\Psi_{xy}^+ \rangle - \sqrt{D_{uv}} |\Psi_{xy}^- \rangle
\]

\[
|\zeta_x\rangle = \sqrt{1 - D_{uv}} |\Psi_{xy}^+ \rangle - \sqrt{D_{uv}} |\Phi_{xy}^- \rangle
\]

\[
|\zeta_y\rangle = \sqrt{1 - D_{uv}} |\Phi_{xy}^+ \rangle + \sqrt{D_{uv}} |\Psi_{xy}^- \rangle
\] (37)

The corresponding optimal POVM, as in [1] Equations (55,56), are described below:

\[
E_\lambda = |E_\lambda\rangle\langle E_\lambda|
\] (38)

where

\[
|E_0\rangle = |E_0\rangle, \quad |E_1\rangle = |E_1\rangle, \quad |E_2\rangle = |E_2\rangle, \quad |E_3\rangle = |E_3\rangle
\] (39)

A Closed form of \(|\xi_i\rangle, |\zeta_j\rangle\): Using our notations \(\varrho_{uv}, \varphi_{uv}\). We can rewrite the expressions of \(|\xi_x\rangle, |\xi_y\rangle, |\zeta_x\rangle, |\zeta_y\rangle\) as below:

\[
|\xi_x\rangle = \varrho_{uv} |E_0\rangle + \varphi_{uv} |E_3\rangle
\]

\[
|\xi_y\rangle = \varphi_{uv} |E_0\rangle + \varrho_{uv} |E_3\rangle
\]

\[
|\zeta_x\rangle = \varphi_{uv} |E_2\rangle + \varrho_{uv} |E_1\rangle
\]

\[
|\zeta_y\rangle = \varrho_{uv} |E_2\rangle + \varphi_{uv} |E_1\rangle
\] (40)

where

\[
\varrho_{uv} := \frac{\sqrt{1 - D_{uv}} + \sqrt{D_{uv}}}{\sqrt{2}}, \quad \varphi_{uv} := \frac{\sqrt{1 - D_{uv}} - \sqrt{D_{uv}}}{\sqrt{2}}
\] (41)

Note 1. Following relations appear to be useful:

\[
\varrho_{uv}^2 + \varphi_{uv}^2 = 1, \quad \varrho_{uv} \cdot \varphi_{uv} = \frac{1}{2} (1 - 2D_{uv})
\]

\[
\varrho_{uv}^2 - \varphi_{uv}^2 = 2\sqrt{D_{uv}} (1 - D_{uv}) = G_\lambda^*
\] (42)
3.3.2 For Equal Error Rates, i.e., $D_{xy} = D_{uv} = D$

For equal error rates, \cite{4} Section IV, Equation (69)] comes up with another choice of $|\xi_i\rangle, |\zeta_j\rangle$. We describe it as below:

\[
\begin{align*}
|\xi_x\rangle &= |x\rangle|x\rangle \\
|\xi_y\rangle &= (\cos \alpha |x\rangle + \sin \alpha |y\rangle) |x\rangle \\
|\zeta_x\rangle &= |x\rangle|y\rangle \\
|\zeta_y\rangle &= (\cos \beta |x\rangle + \sin \beta |y\rangle) |y\rangle
\end{align*}
\]

\( \text{(43)} \)

Optimality of $G$ (or $I$) is reached when

$\alpha = \beta$ and $\sin \alpha = 2\sqrt{D(1-D)} = \mathcal{D}^2 - \mathcal{D}'^2$

where notations $\mathcal{D}, \mathcal{D}'$ are analogous to those in Equation (42).

Thus the optimal interaction can be written as

\[
\begin{align*}
|\xi_x\rangle &= |\xi_0\rangle \\
|\xi_y\rangle &= 2\mathcal{D} \cdot \mathcal{D}' |\xi_0\rangle + (\mathcal{D}^2 - \mathcal{D}'^2) |\xi_1\rangle \\
|\zeta_x\rangle &= |\xi_2\rangle \\
|\zeta_y\rangle &= 2\mathcal{D} \cdot \mathcal{D}' |\xi_2\rangle + (\mathcal{D}^2 - \mathcal{D}'^2) |\xi_3\rangle
\end{align*}
\]

\( \text{(44)} \)

However, the corresponding optimal POVM was not shown explicitly in \cite{4}, which we establish in Section 4.3.

4 Our Results

In this section we show that the interaction by Eve (and the corresponding POVM) that leads to optimal information gain has an unique form in a fixed basis.

4.1 Optimal Measurement (POVM) to maximize Information Gain ($G$) - a Generic Condition

Let’s consider the problem below:

\[
\text{maximize } G := \sum_{\lambda} \left| P_{\lambda x}p_x - P_{\lambda y}p_y \right| \quad \text{over all POVM } \{E_{\lambda}\}
\]

In \cite{3}, an optimal measurement for this maximization was derived afresh. There, the maximization was done on Kolomogorov Variational Distance defined in Equation (130), the calculation was performed in Appendix (Section
7) which shown that the optimal measurement corresponds to a Hermitian operator given by Equation (21) and the optimal POVM is an orthonormal eigenbasis of that operator. We describe the result here with a proof in terms of maximizing $G$.

**Theorem 4.** An optimal POVM to attain $G^*_{xy}$ is an orthonormal projector $\{E_\lambda\}$ that diagonalize the Hermitian operator

$$\tilde{\Gamma}_{xy} := p_x \rho_x - p_y \rho_y$$

where $\rho_x$, as defined in Equation (8), is the partial trace of the post-interaction state $|X\rangle$.

**Proof.**

$$G_{xy} := \sum_\lambda |P_{\lambda x} p_x - P_{\lambda y} p_y|$$

$$= \sum_\lambda p_x \text{tr} (p_x E_\lambda) - p_y \text{tr} (p_y E_\lambda), \text{ using Eq (12)}$$

$$= \sum_\lambda \left| \text{tr} \left( \tilde{\Gamma} E_\lambda \right) \right|, \text{ using Eq (45)}$$

$$= \sum_\lambda \left| \sum_i \gamma_i \langle \gamma_i | E_\lambda | \gamma_i \rangle \right|,$$

where $\{|\gamma_i\rangle\}_i$ is a normalized eigenbasis for $\tilde{\Gamma}$ and $\{|\gamma_i\rangle\}_i$ are the associated eigenvalues

$$\leq \sum_\lambda \sum_i |\gamma_i| \langle \gamma_i | E_\lambda | \gamma_i \rangle,$$

equality occurs for orthogonal $\{|\gamma_i\rangle\}_i$

$$= \sum_i |\gamma_i| \langle \gamma_i | \sum_\lambda E_\lambda | \gamma_i \rangle$$

$$= \sum_i |\gamma_i| = \text{tr} \left| \tilde{\Gamma}_{xy} \right|$$

Thus, $G^*_{xy}$ is achieved by some POVM $\{E_\lambda\}$, the projectors onto an orthonormal eigenbasis of $\tilde{\Gamma}_{xy}$.

**Remark 1.** Since we consider equal prior probabilities, then analogous to Equation (45), we define

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

and use it throughout the paper.
4 OUR RESULTS

4.2 Optimal Interaction to Maximize Information Gain ($G$) - a Generic Form of Optimal $|\xi_i\rangle, |\zeta_j\rangle$

We use the following fact to find an expression of $|\xi_i\rangle, |\zeta_j\rangle$ for optimal interaction:

**Lemma 1.** Optimality conditions for $G_{xy}$ ensures that each $G^*_\lambda$ is equal to $G^*_{xy}$ and the corresponding optimal value is given by

$$G^*_{xy} = 2\sqrt{D_{uv}(1 - D_{uv})} = G^*_\lambda, \forall \lambda$$  (47)

**Note 2.** Since we consider equal prior probabilities, we use the following working formula of $G_\lambda$ in the process of the derivation,

$$G_\lambda := \frac{|Q_{x\lambda} - Q_{y\lambda}|}{P_{\lambda x} + P_{\lambda y}}$$  (48)

As a first step, 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 5.** For a POVM $\{E_\lambda\}_{\lambda \in \{0,1,2,3\}}$

$$P_{\lambda x} = (1 - D_{xy})|\xi_x\rangle\langle E_\lambda| + D_{xy}|\zeta_x\rangle\langle E_\lambda|$$  (49)

$$P_{\lambda y} = (1 - D_{xy})|\xi_y\rangle\langle E_\lambda| + D_{xy}|\zeta_y\rangle\langle E_\lambda|$$  (50)

**Proof.** Using Equation (30) in Equation (8), we get,

$$\rho_x := \text{Tr}_A (|X\rangle\langle X|) = (1 - D_{xy})\hat{\xi}_x + D_{xy}\hat{\zeta}_x$$  (51)

where

$$\hat{\xi}_x := |\xi_x\rangle\langle \xi_x|, \quad \hat{\zeta}_x := |\zeta_x\rangle\langle \zeta_x|$$  (52)

By Equation (12),

$$P_{\lambda x} = \text{Tr} (\rho_x E_\lambda)$$

$$= (1 - D_{xy})\text{Tr} (\hat{\xi}_x E_\lambda) + D_{xy}\text{Tr} (\hat{\zeta}_x E_\lambda)$$

$$= (1 - D_{xy})|\xi_x\rangle\langle E_\lambda| + D_{xy}|\zeta_x\rangle\langle E_\lambda|$$

Similarly, we can derive an expression for $P_{\lambda y}$.  

Now, consider an arbitrary orthonormal eigenbasis of $\Gamma_{xy}$ (which corresponds to an optimal POVM that leads to optimal $G$). We derive the general form of $|\xi_i\rangle, |\zeta_j\rangle$, described in that eigenbasis, for optimal interaction.
Theorem 6. Let \( \{ |E_\lambda \rangle \} \) be an orthonormal eigenbasis of \( \Gamma_{xy} \) (\( \{ E_\lambda \} \) be the corresponding eigenprojector). Then for optimal interaction, the general form of \( |\xi_i\rangle, |\zeta_j\rangle \), described in that eigenbasis becomes

\[
|\xi_x\rangle = D_{uv} |E_0\rangle + \overline{D}_{uv} |E_1\rangle \\
|\xi_y\rangle = \overline{D}_{uv} |E_0\rangle + D_{uv} |E_1\rangle \\
|\zeta_x\rangle = D_{uv} |E_2\rangle + \overline{D}_{uv} |E_3\rangle \\
|\zeta_y\rangle = \overline{D}_{uv} |E_2\rangle + D_{uv} |E_3\rangle
\]  

(53)

where \( D_{uv}, \overline{D}_{uv} \) are as defined in Equation (41).

Proof. First we need to fix an orthonormal basis to describe \( |\xi_i\rangle, |\zeta_j\rangle \) following restriction (31). For that purpose, there is no harm to choose the above orthonormal basis. For that purpose, there is no harm to choose the above orthonormal basis. So the general form of \( |\xi_i\rangle, |\zeta_j\rangle \) becomes

\[
|\xi_x\rangle = \sqrt{\alpha} |E_0\rangle + \sqrt{1 - \alpha} |E_1\rangle \\
|\xi_y\rangle = \sqrt{\beta} |E_0\rangle + \sqrt{1 - \beta} |E_1\rangle \\
|\zeta_x\rangle = \sqrt{\mu} |E_2\rangle + \sqrt{1 - \mu} |E_3\rangle \\
|\zeta_y\rangle = \sqrt{\nu} |E_2\rangle + \sqrt{1 - \nu} |E_3\rangle
\]  

(54)

Using this form of \( |\xi_i\rangle, |\zeta_j\rangle \) in Equation (49) we find values of \( G_\lambda \) as shown in Table I below:

| \( \lambda \) | \( P_{\lambda x} \) | \( P_{\lambda y} \) | \( G_\lambda = \frac{|P_{\lambda x} - P_{\lambda y}|}{P_{\lambda x} + P_{\lambda y}} \) |
|---|---|---|---|
| 0 | \( (1 - D_{xy}) |\xi_x|E_0\rangle|^2 = (1 - D_{xy}) \alpha \) | \( (1 - D_{xy}) |\xi_y|E_0\rangle|^2 = (1 - D_{xy}) \beta \) | \( \frac{|\alpha - \beta|}{\alpha + \beta} \) |
| 1 | \( (1 - D_{xy}) |\xi_x|E_1\rangle|^2 = (1 - D_{xy}) (1 - \alpha) \) | \( (1 - D_{xy}) |\xi_y|E_1\rangle|^2 = (1 - D_{xy}) (1 - \beta) \) | \( \frac{|\alpha - \beta|}{\alpha + 1 - \beta} \) |
| 2 | \( (D_{xy}) |\zeta_x|E_2\rangle|^2 = (D_{xy}) \mu \) | \( (D_{xy}) |\zeta_y|E_2\rangle|^2 = (D_{xy}) \nu \) | \( \frac{|\mu - \nu|}{\mu + \nu} \) |
| 3 | \( (D_{xy}) |\zeta_x|E_3\rangle|^2 = (D_{xy}) (1 - \mu) \) | \( (D_{xy}) |\zeta_y|E_3\rangle|^2 = (D_{xy}) (1 - \nu) \) | \( \frac{|\mu - \nu|}{1 - \mu - \nu} \) |

By Lemma 1, for optimal \( G_{xy}, G_\lambda \)'s are equal. Equating \( G_0, G_1 \) in Table I we get,

\[ \alpha + \beta = 1, \quad G_0 = G_1 = |(2\alpha - 1)| \]

Similarly, equating \( G_2, G_3 \) in Table I we get,

\[ \mu + \nu = 1, \quad G_2 = G_3 = |(2\mu - 1)| \]
Together, Equating $G_0, G_2$, we get,

$$
\mu = \alpha, \quad \nu = \beta = 1 - \alpha
$$

(55)

Thus,

$$
G_0^* = G_{uv}^2 - \bar{G}_{uv}^2 = 2G_{uv}^2 - 1 = |2\alpha - 1|
$$

gives rise to

$$
\sqrt{\alpha} = D_{uv}, \quad \sqrt{1 - \alpha} = \bar{D}_{uv}
$$

(56)

Using Equations (56) and (55) in Equation (54), we get a generic form for optimal $|\xi_i\rangle, |\zeta_j\rangle$ as in Equation (53).

Remark 2. For equal error rates, $D_{uv}, \bar{D}_{uv}$ will be replaced by $D, \bar{D}$ respectively in Equation (53).

Remark 3. We can rewrite Equation (53) as below

\begin{align*}
|\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{align*}

(57)

where

\begin{align*}
|\tilde{E}_0\rangle &= \frac{1}{\sqrt{2}} \left( |E_0\rangle + |E_1\rangle \right), \quad |\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), \quad |\tilde{E}_3\rangle = \frac{1}{\sqrt{2}} \left( |E_2\rangle - |E_3\rangle \right)
\end{align*}

(58)

is another orthonormal basis (called, Bell basis), written in terms of an optimal eigenbasis $\{E_{\lambda}\}$. Clearly, these form to describe $|\xi_i\rangle, |\zeta_j\rangle$ is analogous to Equation (50), (51) in [3].

Since the expression (53) of $|\xi_i\rangle, |\zeta_j\rangle$ corresponds to optimal $G$, therefore, by Theorem 4, $\Gamma_{xy}$ should become a diagonal matrix for this form of $|\xi_i\rangle, |\zeta_j\rangle$ in its eigenbasis. We perform this routine task now, while we figure out the eigenvalues.

Theorem 7. For an optimal POVM $\{E_{\lambda}\}$,

$$
\Gamma_{xy} = \left( G_{uv}^2 - \bar{G}_{uv}^2 \right) \left[ (1 - D_{xy}) (E_{00} - E_{11}) + D_{xy} (E_{22} - E_{33}) \right]
$$

(59)
where
\[ E_{ij} := |E_i \rangle \langle E_j | \] (60)

Clearly, the eigenbasis diagonalize \( \Gamma_{xy} \) with eigenvalues
\[ \gamma_0 = \left( D^2_{uv} - \overline{D}^2_{uv} \right) (1 - D_{xy}), \quad \gamma_1 = -\gamma_0 \]
\[ \gamma_2 = \left( D^2_{uv} - \overline{D}^2_{uv} \right) D_{xy}, \quad \gamma_3 = -\gamma_2 \] (61)

Proof. Using expressions of \( |\xi_i \rangle, |\zeta_j \rangle \) in Equation (53), we get,
\[ \hat{\xi}_x := |\xi_x \rangle \langle \xi_x | = \mathcal{D}^2_{uv} E_{00} + \overline{\mathcal{D}}^2_{uv} E_{11} + 2 \mathcal{D}_{uv} \cdot \overline{\mathcal{D}}_{uv} (E_{01} + E_{10}) \]
\[ \hat{\xi}_y := |\xi_y \rangle \langle \xi_y | = \mathcal{D}^2_{uv} E_{22} + \overline{\mathcal{D}}^2_{uv} E_{33} + 2 \mathcal{D}_{uv} \cdot \overline{\mathcal{D}}_{uv} (E_{23} + E_{32}) \]

So,
\[ \rho_x = (1 - D_{xy}) \hat{\xi}_x + D_{xy} \hat{\xi}_y \]
\[ = \begin{bmatrix} 1 - D_{xy} \\ 1 - D_{xy} \\ D_{xy} \\ D_{xy} \end{bmatrix} \oplus \begin{bmatrix} \mathcal{D}^2_{uv} \\ \overline{\mathcal{D}}^2_{uv} \\ \mathcal{D}^2_{uv} \\ \overline{\mathcal{D}}^2_{uv} \end{bmatrix} + (1 - 2D_{uv}) \begin{bmatrix} E_{00} \\ E_{10} \\ E_{01} \\ E_{11} \end{bmatrix} \] (62)

where the symbol \( \oplus \) denotes componentwise multiplication, known as Hadamard product.

Similarly,
\[ \hat{\xi}_y := |\xi_y \rangle \langle \xi_y | = \mathcal{D}^2_{uv} E_{00} + \overline{\mathcal{D}}^2_{uv} E_{11} + 2 \mathcal{D}_{uv} \cdot \overline{\mathcal{D}}_{uv} (E_{01} + E_{10}) \]
\[ \hat{\xi}_y := |\xi_y \rangle \langle \xi_y | = \mathcal{D}^2_{uv} E_{22} + \overline{\mathcal{D}}^2_{uv} E_{33} + 2 \mathcal{D}_{uv} \cdot \overline{\mathcal{D}}_{uv} (E_{23} + E_{32}) \]

\[ \rho_y = (1 - D_{xy}) \hat{\xi}_y + D_{xy} \hat{\xi}_y \]
\[ = \begin{bmatrix} 1 - D_{xy} \\ 1 - D_{xy} \\ D_{xy} \\ D_{xy} \end{bmatrix} \oplus \begin{bmatrix} \mathcal{D}^2_{uv} \\ \overline{\mathcal{D}}^2_{uv} \\ \mathcal{D}^2_{uv} \\ \overline{\mathcal{D}}^2_{uv} \end{bmatrix} + (1 - 2D_{uv}) \begin{bmatrix} E_{00} \\ E_{10} \\ E_{01} \\ E_{11} \end{bmatrix} \] (63)

Thus,
\[ \Gamma_{xy} := \rho_x - \rho_y = \left( \mathcal{D}^2_{uv} - \overline{\mathcal{D}}^2_{uv} \right) \begin{bmatrix} 1 - D_{xy} \\ 1 - D_{xy} \\ D_{xy} \\ D_{xy} \end{bmatrix} \oplus \begin{bmatrix} +E_{00} \\ -E_{11} \\ +E_{22} \\ -E_{33} \end{bmatrix} \]

which leads to the desired form of \( \Gamma_{xy} \).
Before discussing the general form of an optimal POVM, we consider a special case - in the next section we derive an optimal POVM for the optimal interaction chosen in [4] for equal error rates.

4.3 Optimal POVM for the Specific Interaction for Equal Error Rates ($D_{xy} = D_{uv} = D$) by Fuchs et al. [4]

For equal error rates, i.e., $D_{xy} = D_{uv} = D$, [4] describes a choice of $|\xi_i\rangle$, $|\zeta_j\rangle$ that optimize $I$ (and therefore $G$). For this optimal $|\xi_i\rangle$, $|\zeta_j\rangle$ as described in Equation (44), we now attempt to figure out the optimal POVM.

**Theorem 8.** Consider a canonical basis for Eve as (35). For the optimal interactions (44), the optimal POVM $\{E_\lambda\}$ could be given as

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

where

$$|E_0\rangle = \mathcal{D}|\mathcal{E}_0\rangle - \mathcal{D}|\mathcal{E}_1\rangle, \quad |E_1\rangle = \mathcal{D}|\mathcal{E}_0\rangle + \mathcal{D}|\mathcal{E}_1\rangle$$

$$|E_2\rangle = \mathcal{D}|\mathcal{E}_2\rangle - \mathcal{D}|\mathcal{E}_3\rangle, \quad |E_3\rangle = \mathcal{D}|\mathcal{E}_2\rangle + \mathcal{D}|\mathcal{E}_3\rangle$$

**Proof.** Comparing a special form of $|\xi_x\rangle$, $|\xi_y\rangle$ given by Equation (44) and the general form of $|\xi_x\rangle$, $|\xi_y\rangle$ described in Equation (53) but for equal error rates, we get

$$\mathcal{D}|E_0\rangle + \mathcal{D}|E_1\rangle = |E_0\rangle$$

$$\mathcal{D}|E_0\rangle + \mathcal{D}|E_1\rangle = 2\mathcal{D} \cdot \mathcal{D}|\mathcal{E}_0\rangle + \left(\mathcal{D}^2 - \mathcal{D}^2\right)|\mathcal{E}_1\rangle$$

Subtracting $\mathcal{D}$ times Equation (67) from $\mathcal{D}$ times Equation (66), we get,

$$\left(\mathcal{D}^2 - \mathcal{D}^2\right)|E_0\rangle = \mathcal{D}\left(1 - 2\mathcal{D}^2\right)|\mathcal{E}_0\rangle - \mathcal{D}\left(\mathcal{D}^2 - \mathcal{D}^2\right)|\mathcal{E}_1\rangle$$

$$= \mathcal{D}\left(\mathcal{D}^2 - \mathcal{D}^2\right)|\mathcal{E}_0\rangle - \mathcal{D}\left(\mathcal{D}^2 - \mathcal{D}^2\right)|\mathcal{E}_1\rangle$$

$$\Rightarrow |E_0\rangle = \mathcal{D}|\mathcal{E}_0\rangle - \mathcal{D}|\mathcal{E}_1\rangle$$

Using Equation (68) in Equation (66), we get,

$$|E_1\rangle = \mathcal{D}|\mathcal{E}_0\rangle + \mathcal{D}|\mathcal{E}_1\rangle$$

Similarly, comparing the expressions for $|\zeta_x\rangle$, $|\zeta_y\rangle$ in Equations (44),(53) we can establish that

$$|E_2\rangle = \mathcal{D}|\mathcal{E}_2\rangle - \mathcal{D}|\mathcal{E}_3\rangle$$

$$|E_3\rangle = \mathcal{D}|\mathcal{E}_2\rangle + \mathcal{D}|\mathcal{E}_3\rangle$$

Hence proved.
For convenience, we can easily check that $\Gamma_{xy}$ is diagonalized by the eigenbasis $\{E_\lambda\}_{\lambda \in \{0,1,2,3\}}$ given by Equation (65). In this process we can figure out the eigenvalues of $\Gamma_{xy}$ as well.

**Theorem 9.** For $\Gamma_{xy}$ described in Equation (59), we prove that $\Gamma_{xy}|E_0\rangle = \gamma_0|E_0\rangle$, where $\gamma_0 := 2\sqrt{D(1-D)(1-D)} = (D^2 - \overline{D}^2)(1-D)$.

**Proof.** For equal error rates, expression of $\Gamma_{xy}$ in Equation (59) becomes

$$\Gamma_{xy} = (D^2 - \overline{D}^2)(1-D)(E_{00} - E_{11}) + D(E_{22} - E_{33})$$

which gives

$$\Gamma_{xy}|E_0\rangle = (D^2 - \overline{D}^2)(1-D)(E_{00} - E_{11})|E_0\rangle$$

Considering $E_{ij} := |E_i\rangle\langle E_j|$, Equation (65) gives rise to

- $E_{00} = D^2E_{00} + \overline{D}^2E_{11} - \overline{D} \underbrace{(E_{01} + E_{10})}_{E_{11}}$
- $E_{11} = \overline{D}^2E_{00} + D^2E_{11} + D \underbrace{(E_{01} + E_{10})}_{E_{11}}$
- $E_{00} - E_{11} = (D^2 - \overline{D}^2)(E_{00} - E_{11}) - 2D \overline{D} (E_{01} + E_{10})$

For $|E_0\rangle = D|E_0\rangle - \overline{D}|E_1\rangle$,

$$\langle E_{00} - E_{11}|E_0\rangle = \left[ (D^2 - \overline{D}^2)(E_{00} - E_{11}) - 2D \overline{D} (E_{01} + E_{10}) \right] \left( D|E_0\rangle - \overline{D}|E_1\rangle \right)$$

$$= D \left[ (D^2 - \overline{D}^2)|E_0\rangle - 2D \overline{D} |E_1\rangle \right] - \overline{D} \left[ -(D^2 - \overline{D}^2)|E_1\rangle - 2D \overline{D} |E_0\rangle \right]$$

which leads to

$$\Gamma_{xy}|E_0\rangle = (D^2 - \overline{D}^2)(1-D)|E_0\rangle = \gamma_0|E_0\rangle$$

and the proof is completed for $\lambda = 0$. □

**Remark 4.** One can calculate and check that the eigenvalues of $\Gamma_{xy}$ described in eigenbasis (65) matches those as in Equation (61) calculated for the generic form of optimal $|\xi_i\rangle, |\zeta_j\rangle$ described by Equation (53).

This section has provided us an insight of a possible generic form of the optimal POVM. We continue with the task to find a generic form of the optimal POVM in the next section.
4 OUR RESULTS

4.4 Optimal POVM to maximize Information Gain (G) - a Generic Form

Now we consider a generic form of the eigenbasis and prove that it indeed forms an orthonormal eigenbasis of $\Gamma_{xy}$.

**Theorem 10.** Consider a canonical basis $\{|E_\lambda\rangle\}$ for Eve as (35). Let

$$
|E_0\rangle = \sqrt{a}|E_0\rangle - \sqrt{1-a}|E_1\rangle, \quad |E_1\rangle = \sqrt{1-a}|E_0\rangle + \sqrt{a}|E_1\rangle \\
|E_2\rangle = \sqrt{a}|E_2\rangle - \sqrt{1-a}|E_3\rangle, \quad |E_3\rangle = \sqrt{1-a}|E_2\rangle + \sqrt{a}|E_3\rangle
$$

We prove that $\{E_\lambda\}_{\lambda \in \{0,1,2,3\}}$ forms an orthonormal eigenbasis for $\Gamma_{xy}$.

**Proof.** For equal error rates, expression of $\Gamma_{xy}$ in Equation (59) becomes

$$
\Gamma_{xy} = (D^2 - \bar{D}^2) [(1 - D) (E_{00} - E_{11}) + D (E_{22} - E_{33})]
$$

which gives

$$
\Gamma_{xy}|E_0\rangle = \left(\mathcal{D}^2 - \mathcal{D}^2\right) (1 - D) (E_{00} - E_{11}) |E_0\rangle
$$

Considering $E_{ij} := |E_i\rangle\langle E_j|$, Equation (72) gives rise to

$$
E_{00} = aE_{00} + (1 - a)E_{11} - \sqrt{a(1-a)} (E_{01} + E_{10}) \\
E_{11} = (1 - a)E_{00} + aE_{11} + \sqrt{a(1-a)} (E_{01} + E_{10}) \\
E_{00} - E_{11} = (2a - 1) (E_{00} - E_{11}) - 2\sqrt{a(1-a)} (E_{01} + E_{10})
$$

For $|E_0\rangle = \sqrt{a}|E_0\rangle - \sqrt{1-a}|E_1\rangle$,

$$
\langle E_{00} - E_{11} | E_0\rangle \\
= \left[(2a - 1) (E_{00} - E_{11}) - 2\sqrt{a(1-a)} (E_{01} + E_{10})\right] \left(\sqrt{a}|E_0\rangle - \sqrt{1-a}|E_1\rangle\right) \\
= \sqrt{a} \left[(2a - 1) |E_0\rangle - 2\sqrt{a(1-a)} |E_1\rangle\right] \\
- \sqrt{1-a} \left[- (2a - 1) |E_1\rangle - 2\sqrt{a(1-a)} |E_0\rangle\right] \\
= \sqrt{a}|E_0\rangle - \sqrt{1-a}|E_1\rangle = |E_0\rangle
$$

which leads to

$$
\Gamma_{xy}|E_0\rangle = \left(\mathcal{D}^2 - \mathcal{D}^2\right) (1 - D) |E_0\rangle = \gamma_0 |E_0\rangle
$$

and the proof is completed for $\lambda = 0$.

Similarly, for unequal error rates $\mathcal{D}$ will be replaced by $\mathcal{D}_{uv}$. \qed
4.5 The Interaction is Optimal

Now we show that the interaction given by Equation (53) and the POVM given by Equation (72) leads to optimal G. To do so, we need to prove that they satisfy the necessary and sufficient conditions given by Proposition 3. The initial task is to find an expression for $|\xi_u\rangle, |\xi_v\rangle, |\zeta_u\rangle, |\zeta_v\rangle$ using Equation (53).

**Theorem 11.** Using Equations (57) and (58) in Equation (33), we can derive,

\[
\begin{align*}
|\xi_u\rangle &= \sqrt{1-D_{xy}} |E_0\rangle + \sqrt{D_{xy}} |E_2\rangle \\
|\xi_v\rangle &= \sqrt{1-D_{xy}} |E_0\rangle - \sqrt{D_{xy}} |E_2\rangle \\
|\zeta_u\rangle &= \sqrt{1-D_{xy}} |E_1\rangle - \sqrt{D_{xy}} |E_3\rangle \\
|\zeta_v\rangle &= \sqrt{1-D_{xy}} |E_1\rangle + \sqrt{D_{xy}} |E_3\rangle
\end{align*}
\]

where the basis $\{|E_\lambda\rangle\}$ is as described in Equation (58).

**Remark 5.** To get expressions of $|\xi_i\rangle, |\zeta_i\rangle$ in u-v basis symmetric to those in x-y basis, e.g., like [4, Equation (52)], one must consider the canonical basis in the order $|E_0\rangle = |x\rangle|x\rangle, |E_1\rangle = |y\rangle|y\rangle, |E_2\rangle = |x\rangle|y\rangle, |E_3\rangle = |y\rangle|x\rangle$, compatible with [4].

**Theorem 12.** The interaction given by Equation (53) and the POVM given by Equation (72) satisfy the necessary and sufficient conditions given by Proposition 3 and therefore attains optimal information gain.

**Proof.** We prove it for $\lambda = 1$ only. Other cases are similar.

\[
\begin{align*}
|U_{1u}\rangle &= B_u \otimes \sqrt{E_1} |U\rangle = B_u \otimes E_1 |U\rangle \\
&= B_u \otimes E_1 \left( \sqrt{1-D_{uv}} |u\rangle|\xi_u\rangle + \sqrt{D_{uv}} |v\rangle|\zeta_u\rangle \right), \text{ by Eq. (32)} \\
&= \sqrt{1-D_{uv}} \left( B_u |u\rangle \otimes (E_1 |\xi_u\rangle) + \sqrt{D_{uv}} (B_u |v\rangle) \otimes (E_1 |\zeta_u\rangle) \right) \\
&= \sqrt{1-D_{uv}} \langle E_1 |\xi_u\rangle |u\rangle |E_1\rangle, \\
&\text{since } B_u |u\rangle = |u\rangle, B_u |v\rangle = 0; \text{ and } E_1 |\xi_u\rangle = \langle E_1 |\xi_u\rangle |E_1\rangle \\
&= \frac{1}{\sqrt{2}} \sqrt{1-D_{uv}} \sqrt{1-D_{xy}} |u\rangle |E_1\rangle,
\end{align*}
\]

since, by Eq. (73), $\langle E_1 |\xi_u\rangle = \frac{1}{\sqrt{2}} \sqrt{1-D_{xy}}$.

Similarly,

\[
\begin{align*}
|V_{1u}\rangle &= B_u \otimes \sqrt{E_1} |V\rangle = B_u \otimes E_1 |V\rangle \\
&= \sqrt{D_{uv}} \langle E_1 |\zeta_u\rangle |u\rangle |E_1\rangle \\
&= -\frac{1}{\sqrt{2}} \sqrt{D_{uv}} \sqrt{1-D_{xy}} |u\rangle |E_1\rangle.
\end{align*}
\]
Therefore,

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

with $\varepsilon_1 = -1$

It completes the proof for $\lambda = 1$.

**Note 3.** It could be shown that

$$\varepsilon_0 = +1, \quad \varepsilon_1 = -1, \quad \varepsilon_2 = +1, \quad \varepsilon_3 = -1$$

which completes the proof of the theorem. \(\square\)

Further, we take the opportunity to establish a direct relation between the sign parameter $\varepsilon_\lambda$ and the sign of eigenvalues $\gamma_\lambda$.

**Lemma 2.** For optimal $G$,

$$\varepsilon_\lambda = \text{sgn } \gamma_\lambda$$

**Proof.** For optimal $G$, $\Gamma$ is a diagonal matrix with diagonal entries $\gamma_\lambda$. Thus,

$$\gamma_\lambda = \text{tr} (\Gamma E_\lambda) = \text{tr} (\rho_x E_\lambda) - \text{tr} (\rho_y E_\lambda) = P_{\lambda x} - P_{\lambda y}$$

By Equation (28),

$$\varepsilon_\lambda = \text{sgn} \left( Q_{x\lambda} - Q_{y\lambda} \right) = \text{sgn} \left( P_{\lambda x} - P_{\lambda y} \right) = \text{sgn} \gamma_\lambda$$

which establishes the relation. \(\square\)

**Remark 6.** By Lemma 2, a telltale indication for optimality is that Equation (74) should match with the sign of the eigenvalues $\gamma_\lambda$ of $\Gamma_{xy}$ as in Equation (61), which indeed happens here. Therefore optimality is achieved for the interaction given by Equation (53) and the POVM given by Equation (72).

### 4.6 A Generic form of Optimal Interaction in Canonical Basis

Summarizing our results: an optimal interaction has a general form given by Equation (53) and a general form of the optimal POVM described in a canonical basis (35) is given by Equation (72). We combine these results in a theorem below.
Theorem 13. An optimal interaction expressed in the canonical basis \((35)\) is a one-parameter family as described below:

\[
\begin{align*}
|\xi_x\rangle &= (D_{uv}\sqrt{a} + D_{uv}\sqrt{1-a}) |E_0\rangle + (D_{uv}\sqrt{a} - D_{uv}\sqrt{1-a}) |E_1\rangle \\
|\xi_y\rangle &= (D_{uv}\sqrt{a} + D_{uv}\sqrt{1-a}) |E_0\rangle + (D_{uv}\sqrt{a} - D_{uv}\sqrt{1-a}) |E_1\rangle \\
|\zeta_x\rangle &= (D_{uv}\sqrt{a} + D_{uv}\sqrt{1-a}) |E_2\rangle + (D_{uv}\sqrt{a} - D_{uv}\sqrt{1-a}) |E_3\rangle \\
|\zeta_y\rangle &= (D_{uv}\sqrt{a} + D_{uv}\sqrt{1-a}) |E_2\rangle + (D_{uv}\sqrt{a} - D_{uv}\sqrt{1-a}) |E_3\rangle
\end{align*}
\]

\((76)\)

5 Connecting Our Results with Fuchs et al.

Finally, we connect our results with those in \([4]\). We establish that the optimal interaction chosen in \([4]\) is a special case of the general form of optimal interaction provided by us.

For unequal error rates, Equation \((37)\) is a special case (apart from a permutation of the canonical basis) with \(a = 1\) in Equation \((76)\). Similarly, for equal error rates \((D_{xy} = D_{uv} = D)\), Equation \((13)\) is a special case with \(a = D^2\) in Equation \((76)\). One may consider an innumerable such optimal interaction (and corresponding optimal POVM) by tuning rotation parameter \(a\) in the range \([0, 1]\). One such example is given below for unequal error rates.

Example 1. Let \(a = \frac{1}{2}\). Thus the optimal interaction in Equation \((76)\) becomes

\[
\begin{align*}
|\xi_x\rangle &= \sqrt{1 - D_{uv}} |E_0\rangle - \sqrt{D_{uv}} |E_1\rangle \\
|\xi_y\rangle &= \sqrt{1 - D_{uv}} |E_0\rangle + \sqrt{D_{uv}} |E_1\rangle \\
|\zeta_x\rangle &= \sqrt{1 - D_{uv}} |E_2\rangle - \sqrt{D_{uv}} |E_3\rangle \\
|\zeta_y\rangle &= \sqrt{1 - D_{uv}} |E_2\rangle + \sqrt{D_{uv}} |E_3\rangle
\end{align*}
\]

and the corresponding optimal POVM becomes

\[
\begin{align*}
|E_0\rangle &= \frac{1}{\sqrt{2}} (|E_0\rangle - |E_1\rangle), \quad |E_1\rangle = \frac{1}{\sqrt{2}} (|E_0\rangle + |E_1\rangle) \\
|E_2\rangle &= \frac{1}{\sqrt{2}} (|E_2\rangle - |E_3\rangle), \quad |E_3\rangle = \frac{1}{\sqrt{2}} (|E_2\rangle + |E_3\rangle)
\end{align*}
\]

Clearly, the general form of optimal interaction provided in this paper yields different choices of those in \([4]\). Moreover, it’s implementation is independent of equal or unequal error rates.
6 Conclusion

We have established a unique form describing an optimal interaction (followed by corresponding optimal measurement) in the eigenbasis of $\Gamma_{xy}$. We also have derived the optimal POVM for the chosen interaction in [4] for equal error rates. We have shown that the choice of optimal interaction in [4], for equal as well as unequal error rates, is a special case of the form provided by us.

References


