

Una propuesta de distancia entre procesos cuánticos

Universidad Nacional de Córdoba
Facultad de Matemática Astronomía y Física

Walter León Boyajian
Director: Pedro Walter Lamberti
2011

Una propuesta de distancia entre procesos cuánticos

Autor:

Walter L. Boyajian

Director:

Pedro W. Lamberti

Facultad de Matemática Astronomía y Física
Trabajo final de Licenciatura en Física

Marzo de 2011

Agradecimientos:

A mis padres, por su apoyo durante toda mi vida.

A mi director del trabajo especial y seminario, por el esfuerzo que hizo para que yo terminara mi carrera, por todos los conocimientos que compartió conmigo y por estar siempre presente cuando requería de sus consejos.

A mis profesores que siempre estuvieron dispuestos a saciar mi deseo de conocimiento.

A mis amigos, por su compañía y apoyo desinteresados.

A todo el personal de la facultad, por su dedicación y por su increíble buena voluntad para prestar su ayuda en todo momento.

A la Nación Argentina, por la educación gratuita que me permitió estudiar con gusto y sin preocupaciones y por el Plan Nacional de Becas Universitarias que me apoyó durante toda la carrera económicamente.

Resumen

La propia naturaleza estadística de la Mecánica Cuántica, hace que dos estados cuánticos no sean perfectamente distinguibles, salvo en el caso particular en que ambos estados son ortogonales. Esto hace que del problema de distinguibilidad entre estados uno de gran importancia conceptual y de difícil resolución. Desde el punto de vista formal, el problema de distinguibilidad se enuncia por medio de distancias entre estados, tanto entre estados puros como entre estados mezclas. Un problema cercano al anterior es el problema de distinguibilidad de procesos cuánticos, es decir entre operaciones que actúan sobre los estados. Físicamente, estas operaciones pueden representar la evolución temporal del sistema, un proceso de medición o la aplicación de una compuerta lógica. En este trabajo estudiamos el problema de distinguibilidad de procesos por medio de la extensión de la definición de una distancia entre estados cuánticos - conocida como la divergencia de Jensen-Shannon cuántica- al marco de los procesos. Se investigan sus principales propiedades matemáticas y las posibles interpretaciones físicas asociadas.

Palabras Clave: Distancias entre procesos cuánticos. Distancias entre estados cuánticos. Métricas en el espacio de Hilbert. Divergencia de Jensen-Shannon

Clasificación PACS:

- 03.65.-w Quantum mechanics.
- 03.67.Lx Quantum computation architectures and implementations.
- 03.67.-a Quantum information.

Índice general

1. Introducción	2
1.1. Distancias usuales entre estados cuánticos	3
1.1.1. Distancia de Wootters	4
1.1.2. Distancia de Hellinger	5
1.1.3. Distancia traza	5
1.1.4. Distancias basadas en la fidelidad	6
1.1.5. Divergencia de Kullback-Leibler y entropía relativa	6
1.1.6. Distancia basada en la entropía	7
1.1.7. Divergencia basada en la divergencia de Jensen-Shannon	8
2. Desigualdades	10
3. Propiedades de la Divergencia de Jensen-Shannon	13
3.1. Relación con la fidelidad	15
4. Aplicaciones a la discriminación de procesos	17
4.1. Descripción matemática de procesos cuánticos	17
4.2. Distinguibilidad entre procesos	18
4.2.1. ¿Porqué es importante distinguir procesos?	18
4.2.2. Criterio para determinar una buena medida de distinguibilidad entre procesos	19
4.2.3. Distancia propuesta utilizando la DJS y el isomorfismo de Ja- miołkowski	20
4.3. Propiedades de la distancia entre procesos basada en la DJS	21
4.3.1. Métrica, facilidad de cálculo y medición	21
4.3.2. Estabilidad	22
4.3.3. Encadenado	22
4.3.4. Interpretación física	23
5. Conclusiones	26
6. Apéndices	27

Capítulo 1

Introducción

Todas las teorías física modernas se formulan en términos de espacios abstractos, sobre los cuales se definen diferentes estructuras geométricas. Tal es el caso de la teoría de la Relatividad General, para la cual el espacio-tiempo se representa como una variedad Riemanniana de dimensión cuatro. La métrica de la variedad está dada por la distribución de materia-energía a través de las ecuaciones de campo y de las condiciones de contorno adecuadas al sistema físico que se estudia. En Mecánica Cuántica el espacio abstracto subyacente es un espacio de Hilbert de dimensión finita o infinita y los estados se representan por vectores en ese espacio de Hilbert (estados puros) o por operadores que actúan sobre ese espacio (estados mezclas). A diferencia de la Relatividad General, en el caso de la Mecánica Cuántica no hay una prescripción dinámica de una métrica para medir la distancia entre los estados cuánticos. No obstante es interesante mencionar que la naturaleza estadística de la teoría define una distancia entre estados tal como lo describe Wootters en un trabajo pionero de 1981 [1]. Más interesante aún es el hecho que esa distancia se puede expresar en términos del producto interno definido en el espacio de Hilbert, es decir en término de la estructura geométrica natural del espacio de Hilbert.

La noción de distancia entre estados cuánticos es de fundamental importancia, tanto desde un punto de vista básico como desde el de las aplicaciones. El primero tiene que ver con el hecho de que la geometría del espacio de los estados está fuertemente relacionada con la noción de distancia. El segundo está asociado al hecho de que la noción de distancia entre estados cuánticos permite estudiar conceptos tales como el de distinguibilidad entre estados, definir el grado de entrelazamiento o definir la tasa de transmisión de información entre canales cuánticos, todas estas, cuestiones de suma importancia en información y computación cuántica.

Como se indicó más arriba, la noción de distancia entre estados cuánticos está relacionada con el problema de distinguibilidad entre estados. Antes de avanzar en el tema específico del presente trabajo, describiremos brevemente el problema de la distinguibilidad de estados cuánticos. En cierto sentido, este problema pone en relieve

la naturaleza de la temática que tratamos en el presente trabajo: la distinguibilidad entre *procesos cuánticos*.

El criterio de distinguibilidad de Wootters puede ser establecido, en el marco de la teoría de probabilidades, del siguiente modo: dos distribuciones de probabilidad, $P = (P_1, P_2, \dots, P_N)$ y $Q = (Q_1, Q_2, \dots, Q_N)$ son distinguibles luego de L pruebas ($L \rightarrow \infty$) si y solo si

$$\frac{\sqrt{L}}{2} \left[\sum_{i=1}^N \frac{(P_i - Q_i)^2}{P_i} \right]^{1/2} > 1 \quad (1.1)$$

Este criterio de distinguibilidad se formula en término de una distancia definida sobre el espacio de distribuciones de probabilidad

$$d_S^2(P, Q) = \frac{1}{2} \sum_{i=1}^N \frac{(P_i - Q_i)^2}{P_i}. \quad (1.2)$$

Esta distancia, en el ámbito estadístico, es conocida como la métrica de Fisher.

La noción de distinguibilidad entre estados en un espacio de Hilbert juega un rol importante en una diversidad de circunstancias. Es un concepto básico para la manipulación de información, en el sentido de que siendo capaz de discernir diferentes estados físicos de un sistema dado, le permite a uno determinar cuánta información puede ser codificada en dicho sistema, construyéndose así un puente entre la teoría cuántica y la teoría de la información.

Cuando se realizan tareas de procesamiento de la información en la vida real, los resultados obtenidos en general no coincidirán con los esperados desde un punto de vista teórico, más aún, es posible que distintas implementaciones experimentales den resultados distintos. Es por eso que surge la necesidad de distinguir las implementaciones, ya sea respecto al caso teórico ideal, o entre distintos experimentos que realizan la misma tarea entre sí. Este hecho ha motivado a algunos autores a preocuparse por la distinguibilidad entre procesos cuánticos [3]. Este problema no ha sido hasta ahora tratado en gran detalle y el objetivo de este trabajo es extender una definición de distancia entre estados de un espacio de Hilbert al marco de los procesos cuánticos.

1.1. Distancias usuales entre estados cuánticos

Desde un punto de vista formal, la *medición* en mecánica cuántica está descripta por medio de un operador hermitiano. Si los autoestados de dicho operador son $|\phi_k\rangle$ y el estado del sistema a medirse es $|\Psi\rangle = \sum_k c_k |\phi_k\rangle$, entonces, de acuerdo a los axiomas de la teoría cuántica, el resultado de la medición será el correspondiente autovalor a_k , con probabilidad $|c_k|^2$, representado físicamente por un estado apropiado del dispositivo de medición \mathcal{A} . Nuestra preocupación aquí es cómo discernir dos estados $|\Psi^{(1)}\rangle$ y $|\Psi^{(2)}\rangle$ de un dado sistema físico, utilizando un dispositivo de medición \mathcal{A} . Para resolver este

problema se han definido una gran cantidad de distancias, tanto entre estados puros, descritos por vectores como los anteriores, o entre estados mezclas, descritos por matrices densidad.

En las secciones siguientes listamos algunas de las definiciones más importantes de distancia y métrica, pero antes consideramos importante recordar las definiciones formales de lo que ello significa.

Si denotamos como \mathbb{X} a un espacio abstracto, la función

$$d : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R} \tag{1.3}$$

es una distancia definida sobre el espacio \mathbb{X} si para cada $x, y \in \mathbb{X}$ se satisfacen las siguientes propiedades:

1. $d(x, y) > 0$ para $x \neq y$ y $d(x, x) = 0$,
2. $d(x, y) = d(y, x)$
3. Si para cada $x, y, z \in \mathbb{X}$, la función d también satisface la desigualdad triangular:

$$d(x, y) + d(y, z) - d(x, z) \geq 0, \tag{1.4}$$

se dice que d es una métrica del espacio \mathbb{X} .

1.1.1. Distancia de Wootters

La distancia de Wootters entre dos distribuciones de probabilidad $P = \{p_i\}_i$ y $Q = \{q_i\}_i$ se define como

$$d_W(P, Q) = \arccos \left(\sum_i \sqrt{p_i q_i} \right). \tag{1.5}$$

Esta expresión corresponde a la distancia geodésica asociada con la métrica de Fisher (1.2).

Una correspondencia entre distribuciones de probabilidad y estados cuánticos puede realizarse de la siguiente manera. A cada estado normalizado $|\Psi\rangle$ en el espacio de Hilbert \mathcal{H} de interés, le asignamos un elemento $\{p_i\}$ del conjunto de distribuciones de probabilidad, de forma tal que $p_i = |\langle \phi_i | \Psi \rangle|^2$, con $|\phi_i\rangle$ los autoestados de un operador \mathcal{A} .

Aplicando esta correspondencia a (1.5) se puede escribir

$$d_W^{\mathcal{A}}(|\psi\rangle \parallel |\varphi\rangle) = \arccos \left(\sum_i |\langle \phi_i | \varphi \rangle| |\langle \phi_i | \psi \rangle| \right), \tag{1.6}$$

donde el superíndice hace referencia que se está utilizando el dispositivo de medición \mathcal{A} . Sin embargo puede verse que (1.6) es maximizada por la cantidad [4]

$$d_W(|\psi\rangle \parallel |\varphi\rangle) := \max_{\mathcal{A}} d_W^{\mathcal{A}}(|\psi\rangle \parallel |\varphi\rangle) = \arccos(|\langle\psi|\varphi\rangle|), \quad (1.7)$$

donde el máximo se toma sobre todos los operadores \mathcal{A} . La interpretación geométrica de esta cantidad es obvia: es el ángulo que forman los estados $|\psi\rangle$ y $|\varphi\rangle$. Esta expresión define la distancia entre estados del espacio de Hilbert asociada a la distancia de Wootters (1.2).

1.1.2. Distancia de Hellinger

El cuadrado de esta distancia entre dos distribuciones de probabilidad P y Q se define como

$$d_H^2(P, Q) = \frac{1}{2} \sum_i |\sqrt{p_i} - \sqrt{q_i}|^2. \quad (1.8)$$

Su extensión a la mecánica cuántica resulta en la distancia

$$[d_H^{\mathcal{A}}(|\psi\rangle \parallel |\varphi\rangle)]^2 = \frac{1}{2} \sum_i [|\langle\phi_i|\psi\rangle| - |\langle\phi_i|\varphi\rangle|]^2, \quad (1.9)$$

que depende del dispositivo de medición \mathcal{A} , pero al igual que en el caso de la distancia de Wootters, puede verse que es maximizada por la cantidad [4]

$$d_H^2(|\psi\rangle \parallel |\varphi\rangle) = 1 - |\langle\psi|\varphi\rangle|. \quad (1.10)$$

1.1.3. Distancia traza

Tengamos presente que los estados mezclas se representan por operadores hermitianos ρ , definidos positivos y de traza unidad.

Se define la distancia traza entre dos estados cuánticos ρ y σ como

$$d_{Tr}(\rho \parallel \sigma) \equiv \frac{1}{2} \text{tr} |\rho - \sigma|. \quad (1.11)$$

Cuando ρ y σ son estados puros, i.e. $\rho = |\psi\rangle\langle\psi|$ y $\sigma = |\phi\rangle\langle\phi|$, esta distancia puede escribirse en términos del producto escalar entre los vectores $|\psi\rangle$ y $|\phi\rangle$. Si definimos $x := |\langle\psi|\phi\rangle|$ entonces puede calcularse fácilmente que

$$d_{Tr}(|\psi\rangle\langle\psi| \parallel |\phi\rangle\langle\phi|) = \sqrt{1 - x^2}. \quad (1.12)$$

1.1.4. Distancias basadas en la fidelidad

La fidelidad entre matrices densidad ρ y σ se define como:

$$F(\rho \parallel \sigma) \equiv \left(\text{tr} \sqrt{\sqrt{\rho} \sigma \sqrt{\rho}} \right)^2. \quad (1.13)$$

Cuando ρ es un estado puro dado por $\rho = |\psi\rangle \langle \psi|$, la fidelidad se reduce a

$$F(\rho \parallel \sigma) = \langle \psi | \sigma | \psi \rangle, \quad (1.14)$$

es decir, la *proyección* de σ sobre $|\psi\rangle$.

La fidelidad es una cantidad acotada : $0 \leq F(\rho \parallel \sigma) \leq 1$, con la segunda igualdad dada si y solo si $\rho = \sigma$. Este último dato por si solo muestra que la fidelidad no es una distancia, sin embargo puede usarse fácilmente para definir distancias. Dos métricas usualmente utilizadas en la literatura, basadas en la fidelidad, son la *métrica de Bures* y del *ángulo*, definidas respectivamente por

$$d_B(\rho \parallel \sigma) := \sqrt{2 - 2\sqrt{F(\rho \parallel \sigma)}}, \quad (1.15)$$

$$d_A(\rho \parallel \sigma) := \arccos \sqrt{F(\rho \parallel \sigma)}. \quad (1.16)$$

Además de estas dos distancias, en la Ref. [3] se introduce una tercera métrica basada en la fidelidad, definida por la expresión

$$d_C(\rho \parallel \sigma) := \sqrt{1 - F(\rho \parallel \sigma)}. \quad (1.17)$$

1.1.5. Divergencia de Kullback-Leibler y entropía relativa

A diferencia de las anteriores distancias, la descrita en este punto se origina en Teoría de la Información.

La divergencia de Kullback-Leibler se define para distribuciones de probabilidad P y Q como

$$KL(P, Q) = \sum_i p_i \log_2 \left(\frac{p_i}{q_i} \right). \quad (1.18)$$

Esta cantidad no es una distancia en un sentido formal, pero en determinados contextos puede usarse como tal. Para un listado de las principales propiedades de la divergencia de Kullback-Leibler véase la Ref. [15]. Esta divergencia puede ser extendida a la mecánica cuántica, obteniéndose la distancia conocida como entropía relativa.

La entropía relativa de un operador ρ con respecto a un operador σ esta dada por

$$S(\rho \parallel \sigma) = \text{tr} [\rho (\log_2 \rho - \log_2 \sigma)], \quad (1.19)$$

la cual es una cantidad no negativa y vale cero si y solo si $\rho = \sigma$. Es asimétrica y no está acotada. En particular, la entropía relativa está bien definida solamente cuando el soporte de σ es igual o mayor que el de ρ . En otro caso, se define como $+\infty$. Éste es un requerimiento muy restrictivo que es violado en ciertas situaciones físicas relevantes, como por ejemplo cuando σ es un estado puro.

1.1.6. Distancia basada en la entropía

La entropía es una noción fundamental en teoría de la información clásica y cuántica. En el caso clásico, la mayoría de los resultados concernientes al problema de codificación pueden ser expresados en términos de la entropía de Shannon

$$H_S(P) := - \sum_i p_i \log_2(p_i), \quad (1.20)$$

donde $P = \{p_i\}_i$ es una distribución de probabilidad (discreta). En el contexto cuántico, la expresión para la entropía difiere de la entropía de Shannon. Para un estado mezcla descrito por un operador densidad ρ , von Neumann definió su entropía como

$$H_N(\rho) = -\text{tr}(\rho \log_2 \rho) = - \sum_i \lambda_i \log_2 \lambda_i, \quad (1.21)$$

donde $\{\lambda_i\}$ es el conjunto de autovectores del operador ρ .

Esta cantidad tiene varias interpretaciones físicas interesantes. A partir de su definición se puede ver que para un estado puro la entropía de von Neumann se anula, i.e.

$$H_N(|\psi\rangle\langle\psi|) = 0 \quad (1.22)$$

Si $|\psi\rangle$ y $|\varphi\rangle$ son dos estados puros en un espacio de Hilbert \mathcal{H} , puede definirse la distancia entre esos estados usando la entropía de von Neumann mediante la expresión [5]

$$d_N(|\psi\rangle \parallel |\varphi\rangle) := \sqrt{H_N\left(\frac{|\psi\rangle\langle\psi| + |\varphi\rangle\langle\varphi|}{2}\right)}, \quad (1.23)$$

mientras que para estados mezcla arbitrarios se extiende la definición de métrica anterior a [5]

$$d_N(\rho \parallel \sigma) := \min_{|\psi\rangle, |\varphi\rangle} \sqrt{H_N\left(\frac{|\psi\rangle\langle\psi| + |\varphi\rangle\langle\varphi|}{2}\right)}, \quad (1.24)$$

donde, la minimización se toma sobre todas las *purificaciones*¹ $|\psi\rangle$ de ρ y todas las

¹Si ρ es un estado mezcla actuando en un espacio de Hilbert \mathcal{H} , la acción de purificar este estado consiste en agregar un espacio de Hilbert auxiliar \mathcal{H}_{aux} y un estado σ en éste espacio de forma de poder escribir un estado puro $|\psi\rangle\langle\psi| := \rho \otimes \sigma$ en $\mathcal{H}' := \mathcal{H} \otimes \mathcal{H}_{aux}$. En este caso decimos que $|\psi\rangle$ es una purificación de ρ .

Puede verse [9] que siempre es posible hallar una purificación de cualquier estado mezcla y en particular también puede extenderse un estado puro de \mathcal{H} a un espacio mayor \mathcal{H}' .

purificaciones $|\varphi\rangle$ de σ . Esta expresión es a su vez equivalente a una en función de la fidelidad

$$d_N(\rho \parallel \sigma) = \sqrt{H_S \left\{ \frac{1}{2} \left[1 - \sqrt{F(\rho \parallel \sigma)} \right] \right\}}. \quad (1.25)$$

Como se verá más adelante esta cantidad es una verdadera métrica ente estados puros.

1.1.7. Divergencia basada en la divergencia de Jensen-Shannon

La divergencia de Kullback-Leibler tiene ciertos inconvenientes como el hecho de que no es simétrica ni acotada y no siempre esta bien definida. Para solucionar estos inconvenientes, Rao y Lin introdujeron una versión simetrizada de la divergencia de Kullback-Leibler conocida como la divergencia de Jensen-Shannon (DJS), que se define según la expresión

$$JS(P, Q) := H_S \left(\frac{P + Q}{2} \right) - \frac{1}{2} H_S(P) - \frac{1}{2} H_S(Q), \quad (1.26)$$

donde $H(P)$ es la entropía de Shannon definida antes.

En Ref. [6], Lin propuso una generalización de la Ec. (1.26) al caso de tener un conjunto de n distribuciones de probabilidad P_1, \dots, P_n , con la colección de pesos π_1, \dots, π_n respectivamente. La *divergencia de Jensen-Shannon general* se define entonces según la siguiente relación

$$JS^{(\pi_1, \dots, \pi_n)}(P_1, \dots, P_k) := H_S \left(\sum_{i=1}^n \pi_i P_i \right) - \sum_{i=1}^n \pi_i H_S(P_i). \quad (1.27)$$

Por otra parte, la divergencia de Jensen-Shannon se extiende a la mecánica cuántica definiendo

$$JS(\rho \parallel \sigma) := H_N \left(\frac{\rho + \sigma}{2} \right) - \frac{1}{2} H_N(\rho) - \frac{1}{2} H_N(\sigma). \quad (1.28)$$

que es una cantidad que está siempre bien definida y acotada.

Cuando ρ y σ son estados puros, i.e. $\rho = |\psi\rangle\langle\psi|$ y $\sigma = |\phi\rangle\langle\phi|$, la DJS puede escribirse en términos únicamente de la variable $x := |\langle\psi|\phi\rangle|$, es decir del modulo del producto escalar entre los vectores $|\psi\rangle$ y $|\phi\rangle$. En ese caso queda

$$JS(|\psi\rangle\langle\psi| \parallel |\phi\rangle\langle\phi|) = -\frac{(1-x)}{2} \log_2 \left(\frac{1-x}{2} \right) - \frac{(1+x)}{2} \log_2 \left(\frac{1+x}{2} \right). \quad (1.29)$$

La DJS por si sola es una distancia, pero además en las Refs. [7, 8] se mostró que la raíz de la divergencia de Jensen-Shannon es también una métrica (i.e. su raíz

cuadrada es simétrica, nula solo cuando las distribuciones de probabilidad coinciden y verifica la desigualdad del triángulo). Esto motiva a definir la siguiente métrica entre estados cuánticos:

$$d_{JS}(\rho \parallel \sigma) := \sqrt{JS(\rho \parallel \sigma)}. \quad (1.30)$$

Otras propiedades importantes de la DJS se verán con mayor detalle en el capítulo siguiente.

Capítulo 2

Desigualdades

En diversos contextos es de gran utilidad e importancia conceptual disponer de desigualdades entre cantidades relacionadas con algunas de las distancias definidas anteriormente. Por ejemplo se puede demostrar la siguiente desigualdad entre la distancia de Hellinger y la divergencia de Kullback-Leibler.

$$KL(Q, P) \geq \frac{2}{\ln 2} d_H(P, Q). \quad (2.1)$$

Algunas de estas desigualdades tienen un importante significado físico. Ese es el caso de la desigualdad de Holevo que veremos durante este capítulo.

Antes de presentar la desigualdad de Holevo, introducimos algunos conceptos básicos de teoría de la información¹

Un concepto clave en teoría de la información cuántica es la *entropía*. Ésta mide cuánta incerteza existe en un estado del sistema. Si X e Y son dos variables aleatorias, la información contenida en X se relaciona con la información contenida en Y mediante las cantidades de *entropía condicional* e *información mutua*. La información mutua de X y Y mide cuánta información en común tienen estas dos variables aleatorias [9].

Asumamos que una fuente clásica emite mensajes escritos en un alfabeto X que son enviados a un receptor usando un dispositivo cuántico. Cada letra a_i es codificada en un estado cuántico, i.e. en una matriz densidad ρ_i . Obtenemos de esta manera un ensamble cuántico $\{q_i, \rho_i\}$, donde q_i es la probabilidad de que la fuente emita la letra a_i . El receptor realiza una medición y lee información clásica escrita en un alfabeto Y . La información mutua es

$$H_S(X : Y) = H_S(X) + H_S(Y) - H_S(X, Y),$$

donde $H_S(X, Y)$ es la entropía de Shannon de la distribución de probabilidad conjunta de enviar en a_i y leer en b_j .

¹Aquí nombraremos brevemente un par de conceptos de teoría de la información. El tema se trata con mucho más detalle y de manera muy didáctica en Ref. [9].

Una cota fundamental para la información mutua en términos del conjunto $\{q_i, \rho_i\}$ fue provista por Holevo en forma de un importante teorema [10]. Holevo probó que, si un sistema se prepara en un estado X elegido de un conjunto de operadores densidad $\{\rho_i\}_{i=1,\dots,n}$ con probabilidad p_i , entonces la información mutua $I(X : Y)$ que puede ser obtenida sobre la identidad del estado X por un una medición con salida Y está acotada de acuerdo a la desigualdad

$$I(X : Y) \leq H_N \left(\sum_i q_i \rho_i \right) - \sum_i q_i H_N(\rho_i) =: \chi(\{q_i, \rho_i\}). \quad (2.2)$$

Aquí, $H_N(\rho)$ es la entropía de von Neumann de ρ definida por la identidad (1.21). La cantidad χ se conoce como cantidad de Holevo, y es precisamente la extensión de la expresión (1.27) a operadores densidad con los pesos π_i reemplazados por las probabilidades q_i , esto es

$$\chi(\{q_i, \rho_i\}) \equiv JS^{(q_1, \dots, q_n)}(\rho_1 \parallel \dots \parallel \rho_n). \quad (2.3)$$

Esta cantidad está acotada superiormente por la entropía de Shannon del vector de probabilidad $Q = \{q_i\}_i$ [9],

$$\chi(\{q_i, \rho_i\}) \leq H_S(Q)$$

En un trabajo reciente, Roga, Fannes y Życzkowski [11] encontraron una cota más ajustada para χ que como corolario permite encontrar una relación entre las métricas (1.30) y (1.24) para el caso de distribuciones de probabilidades clásicas.

A continuación presentamos ambos resultados, para ello introduciremos superficialmente el concepto de *operación cuántica*, que se trata en mayor detalle en la sección 4.1 y apéndices citados allí.

Desde un punto de vista matemático una operación cuántica Φ general es un mapa que lleva un estado ρ que actúa en un espacio de Hilbert \mathcal{H} en otro estado τ que actúa en \mathcal{H}' . En particular podemos encontrar operaciones que llevan un estado de \mathcal{H}_A a uno de un espacio extendido $\mathcal{H}_{AB} := \mathcal{H}_A \otimes \mathcal{H}_B$. En ese caso siempre podemos definir las operaciones Φ' que lleva estados de \mathcal{H}_A en sí mismo, y $\tilde{\Phi}$ que lleva estados de \mathcal{H}_A en estados de \mathcal{H}_B tomando traza parcial de la siguiente manera:

$$\rho' = \Phi'(\rho) = \text{Tr}_B[\Phi(\rho)] \quad (2.4)$$

$$\sigma = \tilde{\Phi}(\rho) = \text{Tr}_A[\Phi(\rho)], \quad (2.5)$$

donde la matriz σ recibe el nombre de matriz correlación. Además una operación cuántica siempre puede descomponerse en la llamada descomposición de Kraus, que en particular permite descomponer a Φ' en término de operaciones elementales

$$\Phi(\rho) = \sum_i K_i \rho K_i^\dagger \quad (2.6)$$

Con estas definiciones podemos enunciar el siguiente teorema demostrado en Ref. [11]:

Teorema: Se considera un estado ρ y una operación cuántica Φ , y se definen ρ' y σ como en las ecuaciones (2.4) y (2.5). Además se define una distribución de probabilidad $Q = \{q_i\}_i$ con elementos $q_i := \text{Tr}(\rho K_i^\dagger K_i)$ y estados cuánticos $\rho'_i := K_i \rho K_i^\dagger / q_i$ de modo que $\rho' = \sum_i q_i \rho'_i$. Entonces:

1. La cantidad de Holevo χ está acotada por la entropía de intercambio

$$\chi(\{q_i, \rho'_i\}) \leq H_N(\sigma) \leq H_S(Q). \quad (2.7)$$

2. La entropía promedio está acotada por la entropía del estado inicial

$$\sum_i q_i H_N(\rho'_i) \leq H_N(\rho). \quad (2.8)$$

Si se consideran dos matrices densidad ρ_1 y ρ_2 que ocurren con probabilidades $\frac{1}{2}$ puede encontrarse [11] una expresión explícita para la entropía de intercambio $H_N(\sigma)$ y por lo tanto también para la cota superior de χ en la ecuación (2.7)

$$\chi \leq H_N^{(2)} \left\{ \frac{1}{2} \left[1 - \sqrt{F(\rho_1 \parallel \rho_2)} \right] \right\}, \quad (2.9)$$

donde $H_2(x) := -x \ln x - (1-x) \ln(1-x)$ es la entropía de Shannon de un vector de probabilidades de dos componentes. De esta desigualdad y de la Ec. (1.25) que en el caso clásico de tener dos distribuciones de probabilidad P y Q se cumple que

$$d_{JS}(P, Q) = \sqrt{JS(P, Q)} \leq d_N(P, Q). \quad (2.10)$$

Capítulo 3

Propiedades de la Divergencia de Jensen-Shannon

La divergencia de Jensen-Shannon (DJS) puede escribirse en términos de la entropía relativa como

$$JS(\rho \parallel \sigma) = \frac{1}{2} \left[S\left(\rho \parallel \frac{\rho + \sigma}{2}\right) + S\left(\sigma \parallel \frac{\rho + \sigma}{2}\right) \right]. \quad (3.1)$$

Si ρ y σ son operadores densidad con un conjunto completo de autovectores $\{|r_i\rangle\}$ y $\{|s_i\rangle\}$ tales que $\rho = \sum_i r_i |r_i\rangle \langle r_i|$ y $\sigma = \sum_j s_j |s_j\rangle \langle s_j|$, entonces la DJS entre ρ y σ se expresa como

$$JS(\rho \parallel \sigma) = \frac{1}{2} \left[\sum_{k,i} |\langle t_k | r_i \rangle|^2 r_i \log\left(\frac{2r_i}{\tau_k}\right) + \sum_{k,j} |\langle t_k | s_j \rangle|^2 s_j \log\left(\frac{2s_j}{\tau_k}\right) \right], \quad (3.2)$$

donde $\tau_k = \sum_i r_i |\langle t_k | r_i \rangle|^2 + \sum_j s_j |\langle t_k | s_j \rangle|^2$ y $\langle t_k |$ es un conjunto completo de autovectores de $\rho + \sigma$.

Esta cantidad es positiva, nula si y solo si $\rho = \sigma$, simétrica y siempre bien definida. De hecho, la restricción impuesta sobre los soportes de ρ y σ para la entropía relativa dejan de ser necesarios para la DJS.

Si ρ y σ conmutan, entonces pueden diagonalizarse en la misma base, i.e

$$\rho = \sum_i r_i |i\rangle \langle i|, \quad \sigma = \sum_i s_i |i\rangle \langle i|, \quad (3.3)$$

donde $|i\rangle$ es una base ortogonal. Entonces se tiene que

$$JS(\rho \parallel \sigma) = JS(\{r_i\}, \{s_i\}). \quad (3.4)$$

La divergencia de Jensen-Shannon está acotada. Esta propiedad puede demostrarse de la siguiente forma. La entropía de von Neumann satisface la siguiente propiedad:

Si $\rho = \sum_i p_i \rho_i$ es una mezcla de estados cuánticos ρ_i , mientras que $\{p_i\}$ es un conjunto de números reales y positivos, tales que $\sum_i p_i = 1$, entonces

$$H_N \left(\sum_i p_i \rho_i \right) \leq \sum_i p_i H_N(\rho_i) + H_S(\{p_i\}), \quad (3.5)$$

donde la igualdad se obtiene si y solo si los estados ρ_i tienen soporte en subespacios ortogonales. Tomando $p_1 = p_2 = \frac{1}{2}$, $\rho_1 = \rho$ y $\rho_2 = \sigma$ en la desigualdad anterior se obtiene que

$$0 \leq JS(\rho \parallel \sigma) \leq 1, \quad (3.6)$$

pues $H_S(\{\frac{1}{2}, \frac{1}{2}\}) = 1$.

Otras propiedades de la divergencia de Jensen-Shannon son:

(1) La DJS es invariante ante transformaciones unitarias, esto es

$$JS(U\rho U^\dagger \parallel U\sigma U^\dagger) = JS(\rho \parallel \sigma) \quad (3.7)$$

para cualquier operador unitario U . Esta es una identidad bastante natural para ser satisfecha por una distancia, dado que una transformación unitaria representa una rotación en el espacio de Hilbert, y una distancia entre dos estados debe ser invariante ante rotaciones de los estados.

(2) Teorema H generalizado: Para cualquier mapa Φ completamente positivo¹ (\mathcal{CP}) y que preserve la traza, se tiene que

$$JS(\Phi\rho \parallel \Phi\sigma) \leq JS(\rho \parallel \sigma) \quad (3.8)$$

Esta propiedad es un resultado significativo, pues el modo más general en el que un sistema cuántico abierto evoluciona está representado matemáticamente por un mapa \mathcal{CP} . Por lo tanto el significado de esta identidad es que una transformación no unitaria disminuye la distinguibilidad entre estados². Podemos tomar un mapa \mathcal{CP} particular dado por $\sum_i P_i \sigma P_i$ con P_i siendo un conjunto completo de proyectores ortogonales ($P_i^\dagger = P_i$ y $P_i^2 = P_i$). Entonces

$$JS(P_i \rho P_i \parallel P_i \sigma P_i) \leq JS(\rho \parallel \sigma) \quad (3.9)$$

(3) La DJS es conjuntamente convexa:

$$JS \left(\sum_i \lambda_i \rho^{(i)} \parallel \sum_i \lambda_i \sigma^{(i)} \right) \leq \sum_i \lambda_i JS(\rho^{(i)} \parallel \sigma^{(i)}), \quad (3.10)$$

¹Más adelante se tratan con detalle los mapas completamente positivos al introducirlos como un requerimiento para aquellos mapas que describen un proceso físico.

²el caso unitario esta incluido en esta desigualdad, en cuyo caso se cumple la igualdad de acuerdo con la identidad anterior

donde los coeficiente λ_i son números reales positivos tales que $\sum_i \lambda_i = 1$.

(4) Si definimos ρ^{AB} y σ^{AB} como operadores densidad en el estado compuesto AB (representado por el espacio de Hilbert $\mathcal{H}_A \otimes \mathcal{H}_B$), y sus respectivas trazas parciales $\rho^A = \text{Tr}_B(\rho^{AB})$ y $\sigma^A = \text{Tr}_B(\sigma^{AB})$, entonces:

$$JS(\rho^A \parallel \sigma^A) \leq JS(\rho^{AB} \parallel \sigma^{AB}) \quad (3.11)$$

$$JS(\rho^A \parallel \sigma^A) \leq JS(\rho^{AB} \parallel \sigma^{AB}) \quad (3.12)$$

Estas desigualdades poseen una interpretación natural muy clara: Tomar la traza sobre una parte del sistema lleva a una pérdida de información y por lo tanto la distinguibilidad entre los dos estados del sistema es menor.

(5) La aditividad no se satisface para la DJS pero si la “aditividad restringida”:

$$JS(\rho_1 \otimes \rho_2 \parallel \sigma_1 \otimes \rho_2) = JS(\rho_1 \parallel \sigma_1), \quad (3.13)$$

con $\rho_1, \sigma_1 \in \mathcal{S}_A$, y $\rho_2 \in \mathcal{S}_B$.

La DJS también puede relacionarse con las métricas de Bures y de Wootters. En Ref. [12] los autores hallaron que:

$$d_B(\rho \parallel \rho + d\rho) \approx \sqrt{2JS(\rho \parallel \rho + d\rho)}, \quad (3.14)$$

donde ρ y $\rho + d\rho$ son dos operadores densidad vecinos. Mientras que para estados puros,

$$d_W(|\psi\rangle \parallel |\varphi\rangle) \approx \sqrt{2JS(|\psi\rangle \parallel |\varphi\rangle)}. \quad (3.15)$$

Lo último que diremos respecto a las propiedades de la DJS es que ésta también puede relacionarse con el criterio de distinguibilidad de Wootters mencionado al comienzo del trabajo, en la sección (1.1). En [4] los autores encontraron una relación entre la distancia de Wootters y la DJS que les permitió reescribir el criterio anterior de la siguiente forma: dos distribuciones de probabilidad $P^{(1)}$ y $P^{(2)}$ son distinguibles luego de L pruebas ($L \rightarrow \infty$) si y solo si

$$[d_{JS}(P^{(1)}, P^{(2)})]^{1/2} > \frac{1}{\sqrt{2L}}. \quad (3.16)$$

3.1. Relación con la fidelidad

La DJS también permite definir una fidelidad alternativa a la dada en (1.13). Dicha definición es [12]:

$$F_{JS}(\rho \parallel \sigma) \equiv [1 - JS(\rho \parallel \sigma)]^2. \quad (3.17)$$

Puede verse que F_{JS} está acotada entre 0 y 1. En particular cuando $\rho = \sigma$ se tiene que $F_{JS} = 1$ mientras que $F_{JS} = 0$ si y solo si ρ y σ tienen soporte en subespacios ortogonales.

La definición de esta identidad está motivada por la relación entre la métrica de Bures y la fidelidad (1.15), y por la expresión (3.14). Otra motivación puede verse de la siguiente manera. La fidelidad entre dos estados ρ y σ puede ser evaluada en términos de purificaciones de dichos estados como

$$F(\rho \parallel \sigma) = \max_{|\varphi\rangle} |\langle \psi | \varphi \rangle|^2, \quad (3.18)$$

donde $|y\rangle$ es cualquier purificación fija de ρ , y la maximización se realiza sobre todas las purificaciones de σ [9]. De la expresión para la distancia de Wootters entre estados puros (1.7), queda claro que la fidelidad puede expresarse en términos de dicha distancia como

$$\begin{aligned} F(\rho \parallel \sigma) &= \max_{|\varphi\rangle} \cos^2 [W(|\psi\rangle \parallel |\varphi\rangle)] \\ &= \max_{|\varphi\rangle} [1 - W(|\psi\rangle \parallel |\varphi\rangle) + \dots], \end{aligned} \quad (3.19)$$

que a su vez, por la relación entre la distancia de Wootters con la DJS (3.15) puede aproximarse como

$$F(\rho \parallel \sigma) \approx \max_{|\varphi\rangle} [1 - JS(|\psi\rangle \langle \psi | \parallel |\varphi\rangle \langle \varphi |)]^2 \quad (3.20)$$

Capítulo 4

Aplicaciones a la discriminación de procesos

4.1. Descripción matemática de procesos cuánticos

El formalismo de operaciones cuánticas es una herramienta general para describir la evolución de sistemas cuánticos en una gran variedad de circunstancias. Los estados cuánticos están descriptos por matrices densidad ρ , que se transforman como

$$\rho' = \mathcal{E}(\rho). \quad (4.1)$$

El mapa \mathcal{E} en esta ecuación es una *operación cuántica*, ρ es el estado inicial antes del proceso, y $\mathcal{E}(\rho)$ es el estado final luego de que el proceso ocurra. Las operaciones cuánticas *describen* los procesos físicos más generales que pueden ocurrir en un sistema cuántico [9], incluyendo evolución unitaria, medición, ruido y decoherencia. Un ejemplo simple de operación cuántica es una transformación unitaria, en la que $\mathcal{E}(\rho) = U\rho U^\dagger$, donde U es una matriz unitaria.

Toda operación cuántica puede ser descripta mediante la *representación de suma de operadores* que relaciona un estado de entrada ρ_{ent} con otro de salida ρ_{sal} a través de

$$\rho_{sal} = \mathcal{E}(\rho_{ent}) = \sum_j E_j \rho_{ent} E_j^\dagger, \quad (4.2)$$

donde los operadores E_j son conocidos como operaciones elementales y obedecen la condición de que $\sum_j E_j^\dagger E_j = \mathbb{I}^1$. En esta representación, los operadores $\{E_j\}$ describen

¹En Ref. [3] se considera que una operación cuántica general incluye los casos en los que $\sum_j E_j^\dagger E_j \leq \mathbb{I}$, el caso de igualdad físicamente corresponde a procesos físicos sin post-selección. En Ref. [13], donde se describen las condiciones que un mapa debe cumplir para representar una operación física, se requieren operaciones que preserven la traza, por lo que incluyen solo el caso en que se verifica la igualdad.

completamente el efecto del proceso. En Ref. [13], se muestra que estos son los únicos requerimientos que debe cumplir un mapa para representar una operación física. Una demostración de este hecho se incluye en el apéndice B.

El inconveniente de la representación de suma de operadores es la no unicidad, pues existe una libertad en la elección de las operaciones elementales. Esto lleva a inconvenientes a la hora de comparar dos procesos. Para solucionar esto, se fija una cierta base $\{A_j\}$ para el espacio de operadores, que por conveniencia es mejor tomarla ortogonal bajo el producto interno de Hilbert-Schmidt, i.e. se requiere que $\text{Tr}(A_j^\dagger A_k) = \delta_{jk}$. Entonces, utilizando esta base para expandir las operaciones elementales, $E_j = \sum_m a_{jm} A_m$, es posible reescribir la ecuación anterior como

$$\mathcal{E}(\rho) = \sum_{mn} (\chi_{\mathcal{E}})_{mn} A_m \rho A_n^\dagger, \quad (4.3)$$

donde $(\chi_{\mathcal{E}}) \equiv \sum_j a_{jm} a_{jn}^*$ son los elementos de la matriz del proceso $\chi_{\mathcal{E}}$. Esta ecuación muestra que la matriz del proceso describe completamente la acción del proceso cuántico. La gran ventaja de la representación mediante la matriz del proceso es que, a diferencia de la representación de suma de operadores, una vez que la base $\{A_j\}$ está fijada, la matriz del proceso resulta única ².

4.2. Distinguibilidad entre procesos

4.2.1. ¿Porqué es importante distinguir procesos?

Al realizar experimentalmente tareas de procesamiento cuántico en el mundo real es cuando aparecen muchas imperfecciones en el procesamiento. Estas pueden aparecer tanto en la creación como en la edición de estados cuánticos, así como en la manipulación de dichos estados mediante ciertos procesos cuánticos. Es por ello que es importante medir y caracterizar cuantitativamente estas imperfecciones en un modo que tenga sentido teórico y que sea práctico desde el punto de vista experimental [3].

En las secciones previas, al estudiar distancias, hemos considerado estados cuánticos. Estos estados pueden determinarse completamente utilizando tomografía de estados cuánticos para luego compararse mediante las distancias que ya fueron definidas. Esta técnica permite reconstruir un estado cuántico de una fuente de sistemas cuánticos mediante mediciones de los sistemas que provienen de la fuente, donde por fuente se entiende cualquier sistema que prepara estados cuánticos ya sea en estados puros e estados mezcla. Para que el estado quede unívocamente identificado, la medición debe ser tomográficamente completa, i.e. los operadores medidos deben formar una base

²La matriz del proceso aparece frecuentemente en la literatura por caminos diversos. La forma anterior es una manera de presentarla. En el Apéndice A se presenta de un modo diferente junto con otras definiciones importantes.

de operadores en el espacio de Hilbert del sistema. De esta forma se obtiene toda la información del estado.

Los procesos por su parte pueden medirse utilizando un proceso similar al de tomografía de estados cuánticos, (denominado tomografía de procesos cuánticos), aunque el problema de desarrollar medidas cuánticas para comparar procesos cuánticos reales e idealizados no ha sido obtenido de manera clara [3].

Según los autores de ese trabajo, debería existir una buena medida (única) que permita una comparación precisa entre diferentes implementaciones de procesamiento de información cuántica en la que además, estén de acuerdo tanto teóricos como experimentales. Esta medida de distancia debería ser útil tanto al comparar un experimento respecto al experimento teórico ideal, como para comparar experimentos que intentan realizar la misma tarea.

4.2.2. Criterio para determinar una buena medida de distinguibilidad entre procesos

La propuesta de Nielsen et. al. [3] se basa en una lista de 6 criterios simples y motivados físicamente que deban ser satisfechos por una buena definición de distancia entre procesos cuánticos, de forma de poder eliminar la mayoría de las definiciones de medidas de distancia que a priori pueden parecer plausibles.

Una medida entre dos procesos cuánticos $\Delta(\mathcal{E}, \mathcal{F})$ debe ser pensada físicamente de dos maneras: como una medida de error en el procesamiento de información cuántica cuando uno quiere realizar el proceso ideal \mathcal{F} pero en su lugar realiza el proceso \mathcal{E} ; o como distinguibilidad entre dos procesos \mathcal{E} y \mathcal{F} . Las siguientes seis propiedades que una buena medida debería satisfacer son las presentadas en [3].

1. *Métrica*: Δ debe ser una métrica.
2. *Fácil de calcular*: debe ser posible evaluar Δ de manera directa.
3. *Fácil de medir*: debe haber una manera clara y procedimientos experimentales posibles para determinar el valor de Δ .
4. *Interpretación física*: Δ debe tener una interpretación física bien motivada.
5. *Estabilidad*: $\Delta(\mathcal{I} \otimes \mathcal{E}, \mathcal{I} \otimes \mathcal{F}) = \Delta(\mathcal{E}, \mathcal{F})$, donde \mathcal{I} representa la operación identidad en un sistema cuántico auxiliar.
6. *Regla de la cadena o encadenado*: $\Delta(\mathcal{E}_2 \circ \mathcal{E}_1, \mathcal{F}_2 \circ \mathcal{F}_1) \leq \Delta(\mathcal{E}_1, \mathcal{F}_1) + \Delta(\mathcal{E}_2, \mathcal{F}_2)$, de modo que para un proceso compuesto de muchos pasos menores, el error total será menor que la suma de los errores individuales de cada paso.

La regla de la cadena y la estabilidad son criterios claves para estimar el error en una tarea de procesamiento cuántico compleja. Esto es así porque las tareas de procesamiento de información generalmente se parten en una secuencia de operaciones simples, de modo que una cota en el error total puede hallarse simplemente analizando las componentes individuales. Esto es crítico en el marco de las aplicaciones como son la computación cuántica, en donde los procesos completos de tomografía en una computación de n -qubits requieren muchas mediciones. De modo que es imposible realizarlos de esta manera. La regla de la cadena y la estabilidad permiten en cambio separar los procesos constitutivos involucrados en la computación, los cuales pueden ser utilizados para inferir si la computación completa es robusta.

Otra deducción que puede tomarse de esos criterios es por ejemplo, que de la métrica y encadenado puede verse que $\Delta(\mathcal{R} \circ \mathcal{E}, \mathcal{R} \circ \mathcal{F}) \leq \Delta(\mathcal{E}, \mathcal{F})$ donde \mathcal{R} es cualquier operación cuántica. Esta desigualdad corresponde al requerimiento de que post-procesamiento mediante \mathcal{R} no puede aumentar la distinguibilidad entre dos procesos \mathcal{E} y \mathcal{F} . Otra consecuencia elemental de la métrica y el encadenado es la invariancia unitaria, i.e. $\Delta(\mathcal{U} \circ \mathcal{E} \circ \mathcal{V}, \mathcal{U} \circ \mathcal{F} \circ \mathcal{V}) = \Delta(\mathcal{E}, \mathcal{F})$ donde \mathcal{U} y \mathcal{V} son operaciones unitarias.

4.2.3. Distancia propuesta utilizando la DJS y el isomorfismo de Jamiołkowski

Hasta ahora, las distancias en la sección 1.1 se habían definido entre estados cuánticos, los cuales están representados por vectores y por matrices densidad. Definir una distancia entre procesos presenta como primer inconveniente, el hecho de que las operaciones cuánticas son mapas. Una opción es utilizar la matriz del proceso (cf. Apéndice A), pues representa completamente a una operación \mathcal{E} en forma de matriz. Sin embargo, esta matriz en general no será una matriz densidad, por lo que las definiciones de distancia dadas antes no pueden aplicarse a matrices de procesos.

Existe una manera directa de definir una distancia entre procesos cuánticos utilizando las distancias definidas para estados cuánticos, y es utilizando el *isomorfismo de Jamiołkowski*. Éste relaciona una operación cuántica \mathcal{E} con un estado cuántico $\rho_{\mathcal{E}}$ ³:

$$\rho_{\mathcal{E}} = [\mathcal{E} \otimes \mathcal{I}] (|\Phi\rangle \langle \Phi|), \quad (4.4)$$

donde $|\Phi\rangle = \sum_j |j\rangle |j\rangle / \sqrt{d}$ es un estado máximamente entrelazado en el sistema (d -dimensional) con otra copia de sí mismo, y $\{|j\rangle\}$ es alguna base ortonormal.

El mapa $\mathcal{E} \rightarrow \rho_{\mathcal{E}}$ es invertible, esto es, el conocimiento de $\rho_{\mathcal{E}}$ es equivalente al conocimiento de \mathcal{E} . Además, $\rho_{\mathcal{E}}$ tiene las mismas propiedades matemáticas que una

³En algunas fuentes se suele definir el estado de *Jamiołkowski* como $\rho_{\mathcal{E}} = [\mathcal{I} \otimes \mathcal{E}] (|\Phi\rangle \langle \Phi|)$. Ambos estados son equivalentes en el sentido de que ambos son isomorfismos del operador \mathcal{E} , el uso de uno u otro depende del orden en el que los autores colocan el espacio de Hilbert auxiliar, en el que actúa \mathcal{I} . Véase 4.1

matriz densidad de dimensión d (véase Apéndice A) y posee otras propiedades que serán útiles a la hora de demostrar propiedades en este trabajo. Esto nos motiva a definir como distancia entre procesos a la raíz cuadrada de la divergencia de Jensen-Shannon entre sus respectivos isomorfismos de Jamiołkowski:

$$\Delta(\mathcal{E}, \mathcal{F}) := d_{JS}(\rho_{\mathcal{E}} \parallel \rho_{\mathcal{F}}) = \sqrt{JS(\rho_{\mathcal{E}} \parallel \rho_{\mathcal{F}})}. \quad (4.5)$$

La prueba de las propiedades de la sección anterior utilizando esta definición de distancia constituye el aporte central del presente trabajo. En la próxima sección estudiaremos las propiedades de esta métrica entre procesos.

4.3. Propiedades de la distancia entre procesos basada en la DJS

4.3.1. Métrica, facilidad de cálculo y medición

La definición de una nueva distancia mediante la ecuación (4.5) hereda automáticamente la propiedad de métrica de d_{JS} , a su vez, dado que $d_{JS}(\cdot \parallel \cdot)$ es fácil de calcular pues tiene una definición muy clara que requiere de evaluaciones directas de operadores, lleva a que $\Delta(\cdot, \cdot)$ sea a su vez fácil de calcular. Más aún, dado que \mathcal{E} puede ser determinado experimentalmente usando tomografía de estados cuánticos, Δ también puede ser determinado experimentalmente, al menos en principio.

La propiedad de métrica de Δ puede verse fácilmente utilizando las propiedades de d_{JS} . A continuación probamos que la definición de Δ dada por la Ec. (4.5) satisface las propiedades que definen una métrica que presentamos en la sección 1.1, recordando que la distancia $d_{JS}(\cdot \parallel \cdot)$ es de por sí una métrica.

Las propiedades 1 y 2 se cumplen trivialmente, pues dado que existe una relación unívoca entre un proceso \mathcal{E} y el estado de Jamiołkowski $\rho_{\mathcal{E}}$, entonces si $\mathcal{E} \neq \mathcal{F}$ entonces se cumple que $\rho_{\mathcal{E}} \neq \rho_{\mathcal{F}}$. Luego, como d_{JS} satisface las propiedades 1 y 2, se sigue que

$$\Delta(\mathcal{E}, \mathcal{F}) = d_{JS}(\rho_{\mathcal{E}} \parallel \rho_{\mathcal{F}}) > 0 \quad \text{para } \mathcal{E} \neq \mathcal{F}, \quad (4.6)$$

$$\Delta(\mathcal{E}, \mathcal{E}) = d_{JS}(\rho_{\mathcal{E}} \parallel \rho_{\mathcal{E}}) = 0, \quad (4.7)$$

que corresponde a la prueba de la propiedad 1 para $\Delta(\cdot, \cdot)$. Además

$$\Delta(\mathcal{E}, \mathcal{F}) = d_{JS}(\rho_{\mathcal{E}} \parallel \rho_{\mathcal{F}}) = d_{JS}(\rho_{\mathcal{F}} \parallel \rho_{\mathcal{E}}) = \Delta(\mathcal{F}, \mathcal{E}), \quad (4.8)$$

que es la demostración de la propiedad 2.

La propiedad 3 se muestra del mismo modo haciendo

$$\begin{aligned} \Delta(\mathcal{E}, \mathcal{F}) + \Delta(\mathcal{F}, \mathcal{G}) - \Delta(\mathcal{E}, \mathcal{G}) &= d_{JS}(\rho_{\mathcal{E}} \parallel \rho_{\mathcal{F}}) \\ &\quad + d_{JS}(\rho_{\mathcal{F}} \parallel \rho_{\mathcal{G}}) - d_{JS}(\rho_{\mathcal{E}} \parallel \rho_{\mathcal{G}}) \geq 0. \end{aligned} \quad (4.9)$$

4.3.2. Estabilidad

Queremos ver que $\Delta(\mathcal{I} \otimes \mathcal{E}, \mathcal{I} \otimes \mathcal{F}) = \Delta(\mathcal{E}, \mathcal{F})$. De acuerdo a la definición que dimos de Δ en (4.5) la estabilidad queda probada viendo que $JS(\rho_{\mathcal{I} \otimes \mathcal{E}} \parallel \rho_{\mathcal{I} \otimes \mathcal{F}}) = JS(\rho_{\mathcal{E}} \parallel \rho_{\mathcal{F}})$.

Utilizando la propiedad $\rho_{\mathcal{A} \otimes \mathcal{B}} = \rho_{\mathcal{A}} \otimes \rho_{\mathcal{B}}$ [3] podemos escribir,

$$JS(\rho_{\mathcal{I} \otimes \mathcal{E}} \parallel \rho_{\mathcal{I} \otimes \mathcal{F}}) = JS(\rho_{\mathcal{I}} \otimes \rho_{\mathcal{E}} \parallel \rho_{\mathcal{I}} \otimes \rho_{\mathcal{F}}). \quad (4.10)$$

A su vez el lado derecho de esta igualdad puede reescribirse utilizando la propiedad de aditividad restringida (3.13), de modo de obtener la igualdad.

$$JS(\rho_{\mathcal{I} \otimes \mathcal{E}} \parallel \rho_{\mathcal{I} \otimes \mathcal{F}}) = JS(\rho_{\mathcal{I}} \otimes \rho_{\mathcal{E}} \parallel \rho_{\mathcal{I}} \otimes \rho_{\mathcal{F}}) = JS(\rho_{\mathcal{E}} \parallel \rho_{\mathcal{F}}), \quad (4.11)$$

que es lo que queríamos ver.

4.3.3. Encadenado

Para probar la propiedad de encadenado, seguimos los pasos utilizados en [3]. En dicha prueba, Nielsen et. al. utilizan como hipótesis fundamental que la distancia definida sobre los estados de Jamiołkowski cumpla con la propiedad de contractividad, por lo que nosotros vamos a necesitar corroborar que la métrica d_{JS} definida según (1.30) sea contractiva, i.e.

$$d_{JS}(\mathcal{E}(\sigma_1) \parallel \mathcal{E}(\sigma_2)) \leq d_{JS}(\sigma_1 \parallel \sigma_2), \quad (4.12)$$

para operaciones \mathcal{E} que preserven la traza. La DJS por si sola posee la propiedad de contractividad, que fue presentada en el capítulo 3 en la Ec. (3.8). Esta propiedad es fácilmente extendible a la d_{JS} de modo que de (3.8) podemos escribir

$$d_{JS}(\mathcal{E}(\sigma_1) \parallel \mathcal{E}(\sigma_2)) = \sqrt{JS(\mathcal{E}(\sigma_1) \parallel \mathcal{E}(\sigma_2))} \leq \sqrt{JS(\sigma_1 \parallel \sigma_2)} = d_{JS}(\sigma_1 \parallel \sigma_2), \quad (4.13)$$

que es la propiedad deseada para la métrica.

A partir de esta propiedad de contractividad, podemos probar a continuación que $\Delta(\cdot, \cdot)$ satisface la propiedad de encadenado, es decir que

$$\Delta(\mathcal{E}_2 \circ \mathcal{E}_1, \mathcal{F}_2 \circ \mathcal{F}_1) \leq \Delta(\mathcal{E}_2, \mathcal{F}_2) + \Delta(\mathcal{E}_1, \mathcal{F}_1), \quad (4.14)$$

con la suposición adicional de que \mathcal{F}_1 es *doble estocástica*, i.e. \mathcal{F}_1 preserva la traza y satisface que $\mathcal{F}_1(I) = I$, suposición que se utiliza en cierto punto de la demostración. Esta suposición puede parecer muy restrictiva pues hay procesos físicos que no lo cumplen. Sin embargo en información cuántica son de interés típico los casos en los que \mathcal{F}_1 y \mathcal{F}_2 son procesos unitarios ideales y se utiliza Δ como medio para comparar la composición de esos dos procesos ideales con los realizados experimentalmente $\mathcal{E}_2 \circ$

\mathcal{E}_1 [3]. Dado que los procesos unitarios son automáticamente doble estocásticos, en el caso de interés usual se cumplirá el encadenado.

Comenzamos la demostración aplicando la desigualdad triangular de modo que

$$\begin{aligned} \Delta(\mathcal{E}_2 \circ \mathcal{E}_1, \mathcal{F}_1 \circ \mathcal{F}_2) &= d_{JS}(\rho_{\mathcal{E}_2 \circ \mathcal{E}_1} \parallel \rho_{\mathcal{F}_2 \circ \mathcal{F}_1}) \\ &\leq d_{JS}(\rho_{\mathcal{E}_2 \circ \mathcal{E}_1} \parallel \rho_{\mathcal{E}_2 \circ \mathcal{F}_1}) + d_{JS}(\rho_{\mathcal{E}_2 \circ \mathcal{F}_1} \parallel \rho_{\mathcal{F}_2 \circ \mathcal{F}_1}). \end{aligned} \quad (4.15)$$

Luego utilizamos la identidad $\rho_{\mathcal{E} \circ \mathcal{F}} = (\mathcal{E} \otimes \mathcal{F}^T)(\Phi)$ ⁴, donde Φ es un estado máximamente entrelazado definido antes. Definimos el mapa completamente positivo $\mathcal{F}^T(\rho) \equiv \sum_j F_j^T \rho F_j^*$, y F_j son las operaciones elementales de \mathcal{F} . Aplicando esta identidad a ambas matrices densidad en el segundo término del lado derecho de la ecuación anterior obtenemos

$$\begin{aligned} \Delta(\mathcal{E}_2 \circ \mathcal{E}_1, \mathcal{F}_2 \circ \mathcal{F}_1) &\leq d_{JS}(\rho_{\mathcal{E}_2 \circ \mathcal{E}_1} \parallel \rho_{\mathcal{E}_2 \circ \mathcal{F}_1}) \\ &\quad + d_{JS}[(\mathcal{E}_2 \otimes \mathcal{F}_1^T)(\Phi) \parallel (\mathcal{F}_2 \otimes \mathcal{F}_1^T)(\Phi)]. \end{aligned} \quad (4.16)$$

Como \mathcal{F}_1 es doble estocástica entonces \mathcal{F}_1^T es una operación que preserva la traza. Luego podemos aplicar contractividad a ambos términos del lado derecho de la ecuación anterior, de modo que

$$\begin{aligned} \Delta(\mathcal{E}_2 \circ \mathcal{E}_1, \mathcal{F}_2 \circ \mathcal{F}_1) &\leq d_{JS}(\rho_{\mathcal{E}_1} \parallel \rho_{\mathcal{F}_1}) + d_{JS}(\rho_{\mathcal{E}_2} \parallel \rho_{\mathcal{F}_2}) \\ &= \Delta(\mathcal{E}_1, \mathcal{F}_1) + \Delta(\mathcal{E}_2, \mathcal{F}_2), \end{aligned} \quad (4.17)$$

que es el resultado que se buscaba.

4.3.4. Interpretación física

En Ref. [3], los autores se han centrado en obtener interpretaciones físicas para las distancias utilizadas en el contexto de la computación cuántica, encontrando propiedades de las distancias que permiten dar una interpretación clara como mediciones del error entre un proceso ideal y uno real.

Con la distancia definida de acuerdo a (4.5), no nos fue posible encontrar una expresión completamente equivalente a las halladas en [3], sin embargo hemos probado varias expresiones parcialmente equivalentes utilizando la divergencia de Jensen-Shannon.

En el cálculo de funciones mediante computación cuántica, el objetivo es calcular una función f , exactamente o con alta probabilidad de éxito. Más precisamente, el objetivo es tomar como entrada a un estado x , del problema, y producir un estado final ρ_x de la computadora que sea o bien igual a $|f(x)\rangle$, o suficientemente cercano, de modo que cuando se realice una medición en la base computacional la salida sea $f(x)$ con alta probabilidad.

⁴Una demostración de esta identidad se ofrece en el apéndice C.2

Supongamos que intentamos calcular $f(x)$ mediante computación cuántica, para ello, se diseña un algoritmo que será representado como un proceso \mathcal{F} que actúa sobre una entrada $|x\rangle$, donde x representa la instancia del problema a ser resuelto, por ejemplo un número que debe ser factorizado. Este proceso no necesariamente obtiene como salida el estado $|f(x)\rangle$ con probabilidad 1, pues los algoritmos en computación cuántica se basan, como se dijo más arriba, en que luego de una medición el resultado sea $f(x)$ con alta probabilidad. Surge entonces naturalmente la noción de probabilidad de éxito⁵ del proceso. Evidentemente esta probabilidad es la proyección de $\mathcal{F}(|x\rangle\langle x|)$ sobre $|f(x)\rangle$, i.e.

$$p^{id} = \langle f(x) | \mathcal{F}(|x\rangle\langle x|) |f(x)\rangle, \quad (4.18)$$

donde el superíndice id indica que este es el error *ideal*, por tratarse de la implementación del algoritmo teórico. En la realidad, se realizará una operación no ideal \mathcal{E} , de modo que definimos igualmente la probabilidad de éxito de este proceso como

$$p = \langle f(x) | \mathcal{E}(|x\rangle\langle x|) |f(x)\rangle. \quad (4.19)$$

Es posible que estemos interesados no en el error sobre un valor de entrada x particular sino en el error *promedio*, donde el promedio se toma con respecto a una distribución uniforme sobre instancias x . Por lo tanto definimos estas cantidades promedio como

$$\bar{p}^{id} = \frac{1}{d} \sum_y \langle f(y) | \mathcal{F}(|y\rangle\langle y|) |f(y)\rangle, \quad (4.20)$$

mientras que \bar{p}_e se define de modo equivalente.

Evidentemente, la fidelidad y la distancia traza poseen una relación privilegiada con estas cantidades sobre cualquier otra definición de distancia. Esto queda claro al notar que las definiciones de p y de p^{id} tienen la misma forma que (1.14), de modo que la probabilidad de éxito está relacionada con la fidelidad mediante:

$$p = F[\mathcal{E}(|x\rangle\langle x|), |f(x)\rangle\langle f(x)|], \quad (4.21)$$

$$p^{id} = F[\mathcal{F}(|x\rangle\langle x|), |f(x)\rangle\langle f(x)|]. \quad (4.22)$$

En [3] Nielsen et. al. hallaron que

$$\bar{p}_e \leq \bar{p}_e^{id} + \Delta_{Tr}(\mathcal{E}, \mathcal{F}), \quad (4.23)$$

donde el subíndice e hace referencia a que estas probabilidades de error (i.e. $\bar{p}_e := 1 - \bar{p}$ y $\bar{p}_e^{id} = 1 - \bar{p}^{id}$) y la distancia Δ_{Tr} está definida al igual que definimos Δ en (4.5), pero utilizando la distancia Traza (1.11), es decir que

$$\Delta_{Tr}(\mathcal{E}, \mathcal{F}) := d_{Tr}(\rho_{\mathcal{E}} \| \rho_{\mathcal{F}}). \quad (4.24)$$

⁵Equivalentemente se puede definir la probabilidad de fracaso o de error, que será simplemente 1 menos la probabilidad de éxito. El uso de una o de la otra es puramente arbitrario y dependerá de la comodidad para trabajar las ecuaciones en cada caso.

Como se dijo antes, no nos fue posible encontrar una relación equivalente a (4.23) en el caso general que utilice nuestra definición de Δ , sin embargo pueden encontrarse ciertas relaciones semejantes que utilicen la divergencia de Jensen-Shannon.

Lo primero que podemos notar es que así como las probabilidades de éxito pueden relacionarse con la fidelidad mediante (4.21) y (4.22), podemos utilizar (6.45) para extender esa relación a la DJS. Utilizando las relaciones mencionadas, podemos escribir:

$$p \lesssim [1 - JS(|\psi\rangle\langle\psi| \parallel |\tau\rangle\langle\tau|)]^2, \quad (4.25)$$

donde $|\psi\rangle\langle\psi|$ y $|\tau\rangle\langle\tau|$ son purificaciones arbitrarias de $\mathcal{E}(|x\rangle\langle x|)$ y $|f(x)\rangle\langle f(x)|$ respectivamente en el espacio $\mathcal{H} \otimes \mathcal{H}_{\text{aux}}$. Como $|f(x)\rangle\langle f(x)|$ ya es un estado puro, entonces $|\tau\rangle\langle\tau| = |f(x)\rangle\langle f(x)| \otimes \mathbb{I}_{\text{aux}}$, donde \mathbb{I}_{aux} es el operador identidad en el espacio auxiliar \mathcal{H}_{aux} .

Además es posible hallar una relación entre las probabilidades de error y la DJS. En el apéndice C.3 mostramos que:

$$\sqrt{p_e} \lesssim \sqrt{p_e^{id}} + \sqrt{1 - [1 - JS(|\psi\rangle\langle\psi| \parallel |\varphi\rangle\langle\varphi|)]^2}, \quad (4.26)$$

donde $|\psi\rangle\langle\psi|$ y $|\varphi\rangle\langle\varphi|$ son purificaciones arbitrarias de $\mathcal{E}(|x\rangle\langle x|)$ y $\mathcal{F}(|x\rangle\langle x|)$ respectivamente.

Otro resultado puede obtenerse por un camino diferente, al reescribir la identidad (4.23) para la DJS. En el apéndice C.4, mostramos que si $\rho = |\psi\rangle\langle\psi|$ y $\sigma = |\varphi\rangle\langle\varphi|$ son dos estados puros, entonces $d_{JS}(\rho \parallel \sigma) \geq d_{Tr}(\rho \parallel \sigma)$. Por lo tanto si asumimos que $\rho_{\mathcal{E}}$ y $\rho_{\mathcal{F}}$ son estados puros⁶, podemos escribir, utilizando (4.23) y (4.24):

$$\begin{aligned} \bar{p}_e &\leq \bar{p}_e^{id} + \Delta_{Tr}(\mathcal{E}, \mathcal{F}) \\ &= \bar{p}_e^{id} + d_{Tr}(\rho_{\mathcal{E}} \parallel \rho_{\mathcal{F}}) \\ &\leq \bar{p}_e^{id} + d_{JS}(\rho_{\mathcal{E}} \parallel \rho_{\mathcal{F}}) \\ &\leq \bar{p}_e^{id} + \Delta(\mathcal{E}, \mathcal{F}). \end{aligned} \quad (4.27)$$

Que es equivalente a (4.23) cuando $\rho_{\mathcal{E}}$ y $\rho_{\mathcal{F}}$ son estados puros.

Lo último que diremos con respecto al sentido físico de la DJS, es acerca de su interpretación en la teoría de la información. Como se dijo al principio de esta sección, las distancia traza y la fidelidad tienen una relación privilegiada respecto a las definiciones de probabilidades de éxito o fracaso en computación cuántica. La divergencia de Jensen-Shannon por su parte es una cantidad originada en la teoría de la información, y posee varias interpretaciones interesantes en el marco de dicha teoría, como la relación con la información mutua. Un ejemplo de esto es la cota de Holevo presentada en la sección 2.

⁶Esta suposición es sumamente restrictiva, sin embargo creemos que el hecho de que la identidad pueda probarse en este caso particular, puede ser un punto de partida junto con las otras identidades que están definidas a su vez en base a purificaciones, para encontrar una relación más general en el futuro.

Capítulo 5

Conclusiones

En este trabajo se estudió el problema de distinguibilidad de procesos cuánticos. Dado que la definición de distinguibilidad involucra el concepto geométrico de distancia, se revisaron las principales distancias entre estados cuánticos. Entre ellas se trabajó con una conocida como la Divergencia de Jensen-Shannon. De ésta se estudiaron sus propiedades, como así también los distintos contextos físicos que permiten una adecuada interpretación. Se investigó su aplicabilidad como distancia entre procesos. Dentro de esta problemática mostramos que la extensión de la definición de distancia entre estados, utilizando el isomorfismo de Jamiołkowski, satisface una lista de propiedades usualmente requeridas por otros autores. Uno de los principales resultados de esta monografía es que la distancia inducida por la DJS satisface los criterios antes expuestos. Creemos que en el futuro puede obtenerse una interpretación física más clara o más general en el marco de la computación de funciones utilizando la DJS. Es por ello que se han presentado varias propiedades de ésta divergencia que se relacionan con la computación de funciones (algunas en mayor medida que otras) con el fin de presentar un punto de partida amplio para investigaciones futuras.

En este trabajo se procuró presentar, con un grado de formalismo acorde al contexto del mismo, los conceptos de operaciones cuánticas y mapas \mathcal{CP} , procurándose una descripción matemática clara y sencilla. Sin embargo puede darse un enfoque distinto tanto en las definiciones de distancia, como en los mapas en general (donde por ejemplo pueden incluirse mapas que no sean \mathcal{CP} o que no preserven la traza) e indagar en la relación con la geometría del espacio de Hilbert.

Capítulo 6

Apéndices

Apéndice A: Matriz dinámica y su relación con el isomorfismo de Jamiołkowski

Notación De ahora en adelante denotaremos por $\mathbb{M}_{m,n}$ al conjunto de todas las matrices $m \times n$ sobre \mathbb{C} . El conjunto de matrices cuadradas $n \times n$ se denota por \mathbb{M}_n . El conjunto de matrices densidad n -dimensionales se define por \mathcal{H}_n .

El conjunto \mathbb{M}_n tiene la estructura del espacio de Hilbert con producto escalar definido por la expresión

$$(A, B) = \text{Tr } A^\dagger B. \quad (6.1)$$

Este espacio de Hilbert particular se conoce como espacio de Hilbert-Schmidt de operadores actuando en \mathbb{C}^n y lo denotaremos por \mathcal{H}_{HS}

Reforma

Sea $A \in \mathbb{M}_{m,n}(\mathbb{C})$ cuyos elementos se definen $a_{ij} := [A]_{i,j}$. Entonces, definimos la *reforma* de A como

$$\mathbf{res}(A) = (a_{11}, a_{12}, \dots, a_{1n}, a_{21}, a_{22}, \dots, \dots, a_{2n}, \dots, a_{m1}, a_{m2}, \dots, a_{mm})^T. \quad (6.2)$$

De modo que la operación de reforma coloca fila tras fila los elementos de una matriz $m \times n$ en un vector de tamaño mn . De manera inversa, cualquier vector de dimensión mn puede ser llevado a una matriz rectangular.

Un ejemplo sencillo permite comprender mejor la definición de más arriba:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \longleftrightarrow \mathbf{res}(A) = (A_{11}, A_{12}, A_{21}, A_{22}). \quad (6.3)$$

Matriz asociada al mapa

Como se dijo en la sección 4.1, una operación cuántica se define como un mapa $\Phi(\rho_{ent}) = \rho_{sal}$ con ciertas propiedades. Mediante la definición de la operación de reforma podemos expresar Φ mediante una matriz. Para ello vamos a recordar que existe un isomorfismo entre \mathbb{M}_{mn} y $\mathbb{C}^{m^2n^2}$. Todo mapa lineal $\Phi : \mathcal{H}_n \mapsto \mathcal{H}_n$ puede escribirse como una matriz $M_\Phi \in \mathbb{M}_{n^2}$ [14]:

$$\mathbf{res}(\rho_{sal}) = M_\Phi \mathbf{res}(\rho_{ent}), \quad (6.4)$$

donde

$$\begin{aligned} [M_\Phi]_{k,l} &= (\epsilon_k, \Phi(\epsilon_l)) \\ &= \text{Tr} \left[\epsilon_k^\dagger \Phi(\epsilon_l) \right], \end{aligned} \quad (6.5)$$

donde $k, l = 1, \dots, n^2$, de modo que la matriz tiene n^4 elementos. Aquí hemos usado la base canónica $\{\epsilon_k\}_{k=1, \dots, n^2}$ en \mathbb{M}_{n^2} .

Matriz dinámica

Podemos ahora definir la *matriz dinámica* de Φ como la matriz $D_\Phi \in \mathbb{M}_{n^2}$ cuyos elementos vienen dados por

$$[D_\Phi]_{i,j} = \text{Tr} [(\epsilon_i \otimes \epsilon_j) M_\Phi], \quad (6.6)$$

donde $i, j = 1, \dots, n$ mientras que $\{\epsilon_i\}_{i=1, \dots, n}$ es una base canónica en \mathbb{M}_n .

Notación de cuatro índices

Una notación mucho más compacta y que es de gran utilidad en ciertas ocasiones es la notación de cuatro índices para los elementos de la matriz de procesos y la matriz dinámica. Antes de introducir esta notación veamos que la estructura del espacio de Hilbert-Schmidt sin embargo permite relacionar el producto escalar entre matrices con el de vectores haciendo

$$(A, B) = \text{Tr} A^\dagger B = \mathbf{res}(A)^* \mathbf{res}(B) = (\mathbf{res}(A), \mathbf{res}(B)), \quad (6.7)$$

de modo que el producto escalar en el espacio de Hilbert Schmidt de matrices de tamaño N toma la forma de producto ordinario entre dos vectores de tamaño N^2 . A su vez esta identidad sugiere que puede utilizarse sin peligro de ambigüedad la notación habitual más compacta de bra-kets tanto para matrices como para los vectores que surgen de reformar. Así, si definimos

$$|a\rangle \equiv \mathbf{res}(A), \quad (6.8)$$

y el producto escalar de Hilbert-Schmidt (6.1) como

$$\langle A|B\rangle \equiv (A, B), \quad (6.9)$$

entonces por (6.7) se tiene

$$\langle A|B\rangle = \langle a|b\rangle. \quad (6.10)$$

De esta forma es posible escribir de forma más compacta (6.4) como

$$|\rho_{sal}\rangle = M_\Phi |\rho_{ent}\rangle. \quad (6.11)$$

De esta forma además queda claro que los elementos de matriz de M_Φ pueden obtenerse mediante

$$\begin{aligned} [M_\Phi]_{kl} &= \langle \epsilon_k | \Phi(\epsilon_l) \rangle \\ &= \langle \epsilon_k | M_\Phi | \epsilon_l \rangle, \end{aligned} \quad (6.12)$$

que es una definición completamente equivalente a (6.5).

Si se piensa que el mapa Φ actúa en un sistema compuesto $\mathcal{H} = \mathcal{H}_n \otimes \mathcal{H}_m$, entonces el producto tensorial de cualesquiera dos bases en ambos factores provee una base en \mathcal{H} . Luego las definiciones anteriores deben ser reemplazadas por la definición más general de cuatro índices [13]

$$[M_\Phi]_{\substack{m\mu \\ n\nu}} = \langle e_m \otimes f_\mu | \Phi(e_n \otimes f_\nu) \rangle \quad (6.13)$$

$$= \langle e_m \otimes f_\mu | M_\Phi | e_n \otimes f_\nu \rangle, \quad (6.14)$$

donde los índices latinos hacen referencia al subsistema $\mathcal{H}_A = \mathcal{H}_n$ mientras que los índices griegos a $\mathcal{H}_B = \mathcal{H}_m$.

Remezclado

Con la notación de cuatro índices es fácil definir una operación importante de reordenamiento de elementos matriciales. Se define la operación de *remezclado* mediante la expresión

$$X_{\substack{m\mu \\ n\nu}}^R \equiv X_{\substack{mn \\ \mu\nu}}. \quad (6.15)$$

Esta operación permite relacionar, para un dado mapa Φ , sus matrices dinámica y de proceso [13, 14] mediante

$$D_\Phi = M_\Phi^R. \quad (6.16)$$

Isomorfismo de Jamiołkowski

En el apéndice B se muestra que si Φ es un mapa que representa una operación física, entonces debe ser un mapa \mathcal{CP} que preserve la traza. Esta condición se refleja en la exigencia de que su matriz dinámica D_Φ sea positiva, hermitiana y su traza sea siempre igual a n . Por todas estas condiciones la matriz normalizada

$$\rho_\Phi = \frac{D_\Phi}{N}, \quad (6.17)$$

representa efectivamente un estado mixto en \mathcal{H}_{n^2} . Este estado se conoce como *estado de Jamiołkowski* o más comúnmente como *isomorfismo de Jamiołkowski*, precisamente porque establece un isomorfismo entre un mapa \mathcal{CP} que preserve la traza y un subespacio de los estados en \mathcal{H}_{n^2} ¹.

Por último veremos que este estado puede ser expresado de la forma

$$\rho_\Phi = [\Phi \otimes \mathbb{I}] (\rho^\phi), \quad (6.18)$$

donde

$$\rho^\phi = |\phi\rangle \langle \phi| = \frac{1}{n} \sum_{i,j} (|i\rangle \otimes |i\rangle) (\langle j| \otimes \langle j|) \quad (6.19)$$

es un estado máximamente entrelazado en el espacio de Hilbert $\mathcal{H} = \mathcal{H}_n \otimes \mathcal{H}_n$.

Las matrices dinámicas de ρ^ϕ y del operador identidad pueden obtenerse fácilmente utilizando la definición (6.14). Sus elementos resultan ser

$$[M_{\mathbb{I}}]_{n\nu}^{m\mu} = \delta_{mn} \delta_{\mu\nu}, \quad (6.20)$$

$$[M_{\rho^\phi}]_{n\nu}^{m\mu} = \frac{1}{n} \delta_{m\mu} \delta_{n\nu}. \quad (6.21)$$

Con estas matrices podemos escribir el operador $\rho_\Phi = [\Phi \otimes \mathbb{I}] (\rho^\phi)$ como una matriz de ocho índices actuando sobre el estado ρ^ϕ :

$$[\rho_\Phi]_{n\nu}^{m\mu} = [M_\Phi]_{m'n'}^{mn} [M_{\mathbb{I}}]_{\mu'\nu'}^{\mu\nu} [M_{\rho^\phi}]_{n'\nu'}^{m'\mu'} = \frac{1}{n} [M_\Phi]_{\mu\nu}^{mn} = \frac{1}{n} [D_\Phi]_{n\nu}^{m\mu}, \quad (6.22)$$

demostrándose que las dos definiciones son equivalentes.

¹En efecto no todo estado del espacio de Hilbert \mathcal{H}_{n^2} puede ser expresado como una matriz dinámica porque existe una restricción adicional sobre la matriz dinámica D_Φ de un mapa que representa una operación física (cf. (6.27) en el apéndice B)

Apéndice B: Propiedades de mapas que representan una operación física

Sea ρ una matriz densidad actuando sobre un espacio de Hilbert N-dimensional. cuáles son las condiciones que debe cumplir un mapa $\Phi : \mathcal{M}^{(N)} \rightarrow \mathcal{M}^{(N)}$ si debe representar una operación física?

Los mapas que se admitirán son aquellos que pueden tomar la forma

$$\rho \rightarrow \rho' = \sum_i A_i \rho A_i^\dagger, \quad \sum_i A_i^\dagger A_i = \mathbb{I} \quad (6.23)$$

Aquí seguiremos los pasos dados en [13] para mostrar que no es necesario ningún otro tipo de mapa².

El primer requerimiento es que el mapa sea lineal. La linealidad es requerida pues se espera que la imagen de ρ no dependa del modo en que ρ es preparado como mezcla de estados puros. Por lo tanto postulamos la existencia de un *superoperador lineal* Φ ,

$$\rho' = \Phi(\rho) \text{ o } \rho'_{m\mu} = [M_\Phi]_{m\mu}^{n\nu} \rho_{n\nu}. \quad (6.24)$$

Notemos que los mapas inhomogéneos están contenidos automáticamente, pues

$$[M_\Phi]_{m\mu}^{n\nu} \rho_{n\nu} + \sigma_{m\mu} = \left([M_\Phi]_{m\mu}^{n\nu} + \sigma_{m\mu} \delta_{n\nu} \right) \rho_{n\nu} = [M_{\Phi'}]_{m\mu}^{n\nu} \rho_{n\nu} \quad (6.25)$$

dado que $\text{tr}(\rho) = 1$. Entonces, podemos tratar directamente con mapas afines de matrices densidad.

Además el mapa debe llevar matrices densidad en operadores densidad (aunque no necesariamente operadores que actúen en el mismo espacio de Hilbert). Esto significa que siempre que ρ sea hermitiana, de traza unidad y positiva, su imagen ρ' debe satisfacer a su vez esas tres propiedades. Estas tres condiciones imponen las siguientes restricciones sobre el matriz de Φ :

1.

$$\rho' = (\rho')^\dagger \iff [M_\Phi]_{m\mu}^{n\nu} = [M_\Phi]_{\nu n}^{*\mu m}$$

2.

$$\text{Tr}(\rho') = 1 \iff [M_\Phi]_{m\mu}^{m\mu} = \delta_{m\mu}$$

3.

$$\rho' \geq 0 \iff [M_\Phi]_{m\mu}^{n\nu} \rho_{n\nu} \geq 0 \text{ cuando } \rho > 0.$$

²En este apéndice utilizaremos fuertemente las nociones y notación introducidas en el Apéndice A.

Estas condiciones toman una interpretación más clara si las escribimos para la matriz dinámica de $\Phi, D_\Phi \equiv M_\Phi^R$, de modo que

$$[D_\Phi]_{\mu\nu}^{mn} = [M_\Phi]_{n\nu}^{m\mu} \quad (6.26)$$

En términos de la matriz dinámica, las tres condiciones anteriores quedan:

1. $\rho' = (\rho')^\dagger \Leftrightarrow [D_\Phi]_{\mu\nu}^{mn} = [D_\Phi]_{\mu\nu}^{\dagger mn}$ de modo que $D_\Phi = D_\Phi^\dagger$
2. $\text{Tr}(\rho') = 1 \Leftrightarrow [D_\Phi]_{m\nu}^{mn} = \delta_{n\nu}$
3. $\rho' \geq 0 \Leftrightarrow [D_\Phi]_{\mu\nu}^{mn} \rho_{n\nu} \geq 0$ cuando $\rho > 0$.

La condición (1) se cumple si y solo si D_Φ es Hermitiana. La condición (2) también toma una forma familiar: la traza parcial con respecto al primer subsistema es el operador identidad para el segundo subsistema, i.e.

$$[D_\Phi]_{m\nu}^{mn} = \delta_{n\nu} \Leftrightarrow \text{Tr}_A D_\Phi = \mathbb{I} \quad (6.27)$$

Se dice de un mapa que es un *mapa positivo* si lleva matrices positivas en matrices positivas. Es decir que la propiedad (3) enunciada para Φ es en rigor la exigencia de que Φ sea un mapa positivo. Sin embargo dicha condición enunciada para D_Φ requiere de más trabajo para interpretar. Supongamos momentáneamente que la matriz densidad original sobre la que actúa Φ es un estado puro. Entonces podemos escribir $\rho_{n\nu} = z_n z_\nu^*$. La imagen va a ser positiva si y solo si, para todos los vectores x_m

$$x_m \rho' x_\mu^* = x_m z_n [D_\Phi]_{\mu\nu}^{mn} x_\mu^* z_\nu^* \geq 0. \quad (6.28)$$

Esto significa que la propia matriz dinámica debe ser positiva cuando actúa sobre estados producto en \mathcal{H}_{n^2} . Esta propiedad se llama *positividad por bloques*. Notemos que como (3) se cumple para todo estado inicial ρ , entonces aunque (6.28) haya sido obtenida suponiendo estados puros como estados iniciales, la condición de positividad por bloques para D_Φ se debe cumplir siempre.

Sin embargo la condición (3) en rigor no es suficiente desde un punto de vista físico. Cualquier estado cuántico ρ podría ser extendido por un conjunto auxiliar en un estado $\rho \otimes \sigma$ de un sistema compuesto más grande. La posibilidad de que dicho espacio auxiliar sea agregado, requiere corroborar ya no solo que Φ sea un mapa positivo, sino que $\Phi \otimes \mathbb{I}$ también lo sea. Esta condición se enuncia precisamente mediante la siguiente definición: Se dice que un mapa Φ es *completamente positivo* si y solo si para una extensión K -dimensional arbitraria $\mathcal{H} = \mathcal{H}_N \otimes \mathcal{H}_K$, el mapa

$$\Phi \otimes \mathbb{I}_K \tag{6.29}$$

es positivo. Esta es la última condición sobre nuestro mapa que generaliza la condición (3). Dos teoremas importantes permiten a partir de estas tres condiciones encontrar una expresión explícita para el mapa.

El teorema de Choi se enuncia como: Un mapa lineal Φ es completamente positivo si y solo si la correspondiente matriz dinámica D_Φ es positiva. Es decir que por este teorema, incrementar la condición sobre Φ de ser positivo a completamente positivo, llevo a la restricción sobre D_Φ de ser positiva por bloques a ser positiva.

Finalmente el teorema de Representación por suma de operadores permite escribir un mapa completamente positivo de una forma particular: Un mapa lineal Φ es completamente positivo si y solo si es de la forma

$$\rho \rightarrow \rho' = \sum_i A_i \rho A_i^\dagger \tag{6.30}$$

La condición (2) de preservación de la traza puede entonces ser incorporada al exigir que además

$$\sum_i A_i A_i^\dagger = \mathbb{I} \tag{6.31}$$

La condición restante (1) que exigía la hermitianicidad de D_Φ esta incluida en el teorema de Choi, de modo que todas las condiciones que impusimos sobre un mapa para que éste represente una operación lineal quedan reducidas a las condiciones (6.30) y (6.31) que es lo que queríamos ver.

Apéndice C: Demostración de propiedades adicionales

C.1

Queremos ver que

$$\rho_{\mathcal{E} \circ \mathcal{F}} = (\mathcal{E} \otimes \mathcal{F}^T) (\rho^\phi), \tag{6.32}$$

para ello calcularemos los elementos de matriz de cada lado de la igualdad y corroboramos que son iguales.

Comenzamos calculando los elementos de matriz del lado izquierdo. Por la definición del estado de Jamiołkowski (4.4), podemos escribir

$$\rho_{\mathcal{E} \circ \mathcal{F}} = [(\mathcal{E} \circ \mathcal{F}) \otimes \mathcal{I}] (\rho^\phi), \tag{6.33}$$

donde ρ^ϕ es un estado máximamente entrelazado en el espacio compuesto $\mathcal{H} \otimes \mathcal{H}_{\text{aux}}$.

Utilizando la notación de cuatro índices para un mapa introducida en el apéndice A³, podemos escribir la matriz dinámica del mapa $(\mathcal{E} \circ \mathcal{F}) \otimes \mathcal{I}$ como un operador de ocho índices de la siguiente manera:

$$(\rho_{\mathcal{E} \circ \mathcal{F}})_{n\nu}^{m\mu} = \mathcal{E} \circ \mathcal{F}_{m'n'}^{mn} \mathcal{I}_{\mu'\nu'}^{\mu\nu} \rho_{n'\nu'}^{\phi}_{m'\mu'}. \quad (6.34)$$

Ahora notamos que la matriz dinámica del operador identidad es

$$\mathcal{I}_{\mu'\nu'}^{\mu\nu} = \delta_{\mu\mu'} \delta_{\nu\nu'}, \quad (6.35)$$

mientras que la del estado máximamente entrelazado a su vez es

$$\rho_{n'\nu'}^{\phi}_{m'\mu'} = \frac{1}{N} \delta_{m'\mu'} \delta_{n'\nu'}, \quad (6.36)$$

identidades que pueden ser obtenidas por álgebra directa utilizando la Ec. (6.5). Además, para encontrar la matriz dinámica del mapa compuesto $(\mathcal{E} \circ \mathcal{F})$, conviene introducir una matriz densidad σ y pensar a dicho mapa como dos mapas que actúan de la siguiente forma:

$$\mathcal{F}(\sigma) = \sigma' \quad \text{y} \quad (\mathcal{E} \circ \mathcal{F})(\sigma) = \mathcal{E}(\mathcal{F}(\sigma)) = \mathcal{E}(\sigma') = \sigma''. \quad (6.37)$$

Entonces tenemos que

$$\begin{aligned} (\sigma'')_{mn} &= \mathcal{E}_{op}^{mn} (\sigma')_{op} \\ &= \mathcal{E}_{op}^{mn} \mathcal{F}_{m'n'}^{op} \sigma_{m'n'} \\ &= (\mathcal{E} \circ \mathcal{F})_{m'n'}^{mn} \sigma_{m'n'} \end{aligned} \quad (6.38)$$

De las últimas dos igualdades concluimos que la matriz dinámica del mapa compuesto $(\mathcal{E} \circ \mathcal{F})$ está relacionado con las matrices dinámicas de los mapas \mathcal{E} y \mathcal{F} mediante

$$(\mathcal{E} \circ \mathcal{F})_{m'n'}^{mn} = \mathcal{E}_{op}^{mn} \mathcal{F}_{m'n'}^{op}. \quad (6.39)$$

Si reemplazamos las matrices dinámicas (6.35), (6.36) y (6.39) en (6.34) obtenemos:

$$\begin{aligned} (\rho_{\mathcal{E} \circ \mathcal{F}})_{n\nu}^{m\mu} &= \mathcal{E}_{op}^{mn} \mathcal{F}_{m'n'}^{op} \delta_{\mu\mu'} \delta_{\nu\nu'} \frac{1}{N} \delta_{m'\mu'} \delta_{n'\nu'} \\ &= \frac{1}{N} \mathcal{E}_{op}^{mn} \mathcal{F}_{\mu\nu}^{op}. \end{aligned} \quad (6.40)$$

³Por comodidad conviene reducir un poco la notación de los elementos matriciales de un mapa CP introducida en dicho apéndice a costa de un poco de claridad. Confiamos en que la diferencia entre un mapa y su matriz densidad ha quedado suficientemente clara en dicho apéndice y denotaremos aquí a los elementos de la matriz de un proceso ε simplemente como $\varepsilon_{n\nu}^{m\mu}$ (en vez de $[M_\varepsilon]_{n\nu}^{m\mu}$)

Ahora vamos a calcular los elementos de matriz del lado derecho de (6.32). Estos elementos pueden calcularse haciendo

$$[(\mathcal{E} \otimes \mathcal{F}^T) (\rho^\Phi)]_{m\mu, n\nu} = \mathcal{E}_{m'n'}^{mn} (\mathcal{F}^T)_{\mu'\nu'}^{\mu\nu} \rho_{n'\nu'}^{\Phi, m'\mu'} \quad (6.41)$$

Los elementos de la operación transpuesta están relacionados por $(\mathcal{F}^T)_{\mu'\nu'}^{\mu\nu} = \mathcal{F}_{\mu\nu}^{\mu'\nu'}$, de modo que reemplazando esta igualdad en (6.41), junco con (6.36) obtenemos

$$\begin{aligned} [(\mathcal{E} \otimes \mathcal{F}^T) (\rho^\Phi)]_{m\mu, n\nu} &= \mathcal{E}_{m'n'}^{mn} \mathcal{F}_{\mu\nu}^{\mu'\nu'} \frac{1}{N} \delta_{m'\mu'} \delta_{n'\nu'} \\ &= \frac{1}{N} \mathcal{E}_{m'n'}^{mn} \mathcal{F}_{\mu\nu}^{m'n'}, \end{aligned} \quad (6.42)$$

que es igual al resultado obtenido para el lado izquierdo (6.40) luego de renombrar algunas variables, con lo que queda demostrada la igualdad (6.32).

C.2

Partiendo de las relaciones (4.21), y motivados por la definición de distancia dada en (1.17), escribimos:

$$\sqrt{p_e} = \sqrt{1-p} = \sqrt{1 - F[\mathcal{E}(|x\rangle\langle x|) \parallel |f(x)\rangle\langle f(x)|]}. \quad (6.43)$$

Como $d_C(\rho \parallel \sigma) = \sqrt{1 - F(\rho \parallel \sigma)}$ es una métrica, satisface en particular la desigualdad triangular, por lo que podemos reescribir (6.43) como

$$\begin{aligned} \sqrt{p_e} &= d_C(\mathcal{E}(|x\rangle\langle x|) \parallel |f(x)\rangle\langle f(x)|) \\ &\leq d_C(\mathcal{F}(|x\rangle\langle x|) \parallel |f(x)\rangle\langle f(x)|) + d_C(\mathcal{E}(|x\rangle\langle x|) \parallel \mathcal{F}(|x\rangle\langle x|)) \\ &= \sqrt{p_e^{id}} + d_C(\mathcal{E}(|x\rangle\langle x|) \parallel \mathcal{F}(|x\rangle\langle x|)) \\ &= \sqrt{p_e^{id}} + \sqrt{1 - F(\mathcal{E}(|x\rangle\langle x|) \parallel \mathcal{F}(|x\rangle\langle x|))} \end{aligned} \quad (6.44)$$

Pero, en la sección 3.1 habíamos mostrado la relación entre la fidelidad y la divergencia de Jensen-Shannon (cf. Ec. (3.20)) que puede reescribirse como

$$F(\rho \parallel \sigma) \gtrsim [1 - JS(|\psi\rangle\langle\psi| \parallel |\varphi\rangle\langle\varphi|)]^2, \quad (6.45)$$

donde $|\psi\rangle$ y $|\varphi\rangle$ son purificaciones arbitrarias de $\mathcal{E}(|x\rangle\langle x|)$ y $\mathcal{F}(|x\rangle\langle x|)$ respectivamente. Luego, (6.44) puede escribirse en términos de la DJS como

$$\sqrt{p_e} \lesssim \sqrt{p_e^{id}} + \sqrt{1 - [1 - JS(|\psi\rangle\langle\psi| \parallel |\varphi\rangle\langle\varphi|)]^2}, \quad (6.46)$$

que es el resultado que queríamos probar.

C.3

En las secciones 1.1.3 y 1.1.7 se citaron expresiones para la distancia traza y para la divergencia de Jensen-Shannon el caso particular en el que $\rho = |\psi\rangle\langle\psi|$ y $\sigma = |\phi\rangle\langle\phi|$ son estados puros. Ambas expresiones, pueden en este caso expresarse como función únicamente del producto escalar entre los vectores $|\psi\rangle$ y $|\phi\rangle$. Habíamos dicho que si definíamos la variable $x := |\langle\psi|\phi\rangle|$ entonces la distancia traza y la DJS están dadas respectivamente por las ecuaciones (1.12) y (1.29):

$$d_{tr}(|\psi\rangle\langle\psi| \parallel |\phi\rangle\langle\phi|) = \sqrt{1 - x^2}, \quad (6.47)$$

$$JS(|\psi\rangle\langle\psi| \parallel |\phi\rangle\langle\phi|) = -\frac{(1-x)}{2} \log_2\left(\frac{1-x}{2}\right) - \frac{(1+x)}{2} \log_2\left(\frac{1+x}{2}\right). \quad (6.48)$$

Entonces recordando la definición de la distancia basada en la DJS (1.30) y las identidades anteriores podemos escribir la diferencia entre las distancias como una función únicamente de x :

$$f(x) := d_{JS}(|\psi\rangle\langle\psi| \parallel |\phi\rangle\langle\phi|) - d_{Tr}(|\psi\rangle\langle\psi| \parallel |\phi\rangle\langle\phi|), \quad (6.49)$$

con

$$f(x) = \sqrt{-\frac{(1-x)}{2} \log_2\left(\frac{1-x}{2}\right) - \frac{(1+x)}{2} \log_2\left(\frac{1+x}{2}\right) - \sqrt{1-x^2}}. \quad (6.50)$$

Un gráfico de $f(x)$ en el intervalo $[0,1]$ se muestra en la Figura 1. En éste se ve claramente que la función siempre toma valores positivos, de modo que para cualquier par de estados puros $|\psi\rangle$ y $|\phi\rangle$, la distancia definida a partir de la DJS es mayor que la distancia traza.

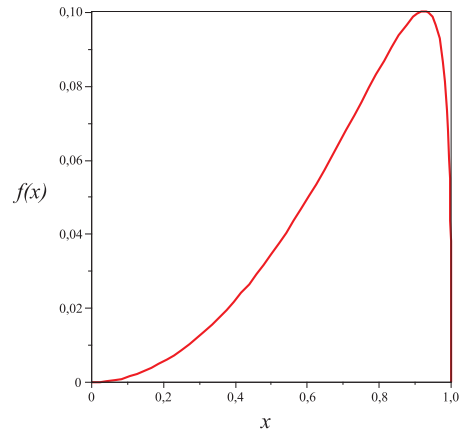


Figura 1: Gráfico de la función $f(x)$ definida en (6.50) que representa la diferencia entre la distancia definida a partir de la DJS y la distancia traza entre los estados puros $|\psi\rangle$ y $|\varphi\rangle$. La variable x está definida como $x := |\langle\psi|\varphi\rangle|$. Éste gráfico muestra claramente que la distancia en base a la DJS es siempre mayor que la distancia traza para estados puros.

Bibliografía

- [1] W. K. Wootters, *Statistical distance and Hilbert space*, Phys. Rev. D, **23**, 357 (1981)
- [2] S. Braunstein and C. Caves, *Statistical distance and the geometry of quantum states*, Phys. Rev. Lett. **72**, 3439 (1994)
- [3] A. Gilchrist, N. K. Langford and M. A. Nielsen, *Distance measures to compare real and ideal quantum processes*, Phys. Rev. A **71**, 062310 (2005)
- [4] A. Majtey, P. W. Lamberti, M. T. Martin, and A. Plastino, *Wootters' distance revisited: a new distinguishability criterium*, Eur. Phys. J. D. **32**, 413 (2005)
- [5] P. W. Lamberti, M. Portesi, and J. Sparacino, *Natural metric for quantum information theory*, Int. J. Quantum Inf. **7**, 1009 (2009)
- [6] J. Lin, *Divergence measures based on the Shannon entropy*, IEEE Trans. Inf. Theory, **37**, 145 (1991)
- [7] P. W. Lamberti, A. P. Majtey, A. Borras, M. Casas, and A. Plastino, *Metric character of the quantum Jensen-Shannon divergence*, Phys. Rev. A **77**, 052311 (2008)
- [8] J. Briët and P. Harremoës, *Properties of classical and quantum Jensen-Shannon divergence*, Phys. Rev. A **79**, 052311 (2009)
- [9] M. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000)
- [10] A. S. Holevo, *Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel*, Probl. Inf. Transm. **9**, 177 (1973)
- [11] W. Roga, M. Fannes, and K. Życzkowski, *Universal Bounds for the Holevo Quantity, Coherent Information, and the Jensen-Shannon Divergence*, Phys. Rev. Lett. **105**, 040505 (2010)

- [12] A. P. Majtey, P. W. Lamberti, and D. P. Prato, *Jensen-Shannon divergence as a measure of distinguishability between mixed quantum states*, Phys. Rev. A **72**, 052310 (2005)
- [13] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States* (Cambridge University Press, Cambridge, UK, 2006)
- [14] J. A. Miszczak, *Singular values decomposition and matrix reorderings in quantum information theory*, arXiv:1011.1585v1
- [15] T.M. Cover and J. A. Thomas, *Elements of Information Theory*, (Wiley Interscience, Hoboken, USA, 2006)