

---

Exploración de correlaciones cuánticas en sistemas multipartidos, basada en el principio de máxima entropía.

---

Sergio Daniel López Castro

UNIVERSIDAD DEL VALLE DE GUATEMALA  
Facultad de Ciencias y Humanidades



**Exploración de correlaciones cuánticas en sistemas  
multipartidos, basada en el principio de máxima  
entropía.**

Trabajo de graduación en modalidad de Tesis presentado por  
Sergio Daniel López Castro  
para optar al grado académico de Licenciado en Física

Guatemala  
2021



UNIVERSIDAD DEL VALLE DE GUATEMALA  
Facultad de Ciencias y Humanidades



**Exploración de correlaciones cuánticas en sistemas  
multipartidos, basada en el principio de máxima  
entropía.**


Trabajo de graduación en modalidad de Tesis presentado por  
Sergio Daniel López Castro  
para optar al grado académico de Licenciado en Física

Guatemala  
2021

Vo.Bo.:

(f)   
MSc. Mariano Lemus

Tribunal Examinador:

(f)   
MSc. Mariano Lemus

(f)   
MSc. Eduardo Álvarez

(f)   
MSc. Irene Aguilar

Fecha de aprobación: Guatemala, 3 de junio de 2021.

La elaboración del presente trabajo de graduación fue motivada por mi asesor y profesor, el MSc. Mariano Lemus, quien me introdujo en el área de Información y Computación Cuántica. Estas disciplinas capturaron mi atención y debo agradecerle a Mariano por su apoyo, conocimiento, amistad y por tomarme bajo su ala e introducirme al trabajo académico. También el estudio teórico de las correlaciones cuánticas fue llevado por él y su equipo de trabajo. Mi objetivo y papel fue proporcionar las herramientas computacionales para continuar en el estudio de las propiedades de las correlaciones.

Aprovecho este espacio para agradecer a la Fundación Juan Bautista Gutiérrez por el apoyo proporcionado a lo largo de la Licenciatura y ser como una segunda familia; y al Centro Universitario Ciudad Vieja por proporcionarme un hogar. Agradezco también a mis padres, mi hermana y demás familia por ser mi respaldo y motivación en todo momento. Por último, agradezco a mis amigos, porque todos ellos han influido en mi crecimiento personal y profesional.

Esta tesis está organizada de la siguiente manera. Primero, se presentan las secciones introductorias al tema. Luego, el capítulo IV es el Marco Teórico, en el que se hace un resumen de los temas básicos de la Teoría de Información y de la Mecánica Cuántica para luego presentar las correlaciones y las estructuras que se utilizarán para el desarrollo del trabajo. Después, en el capítulo V se habla sobre los trabajos precedentes a este y que son motivadores del mismo. A continuación en el capítulo VII, se presentan los algoritmos con su fundamentación teórica, su método de optimización, la eficiencia lograda y una exploración de los estados GHZ y W. En el capítulo VIII se realiza la implementación de los algoritmos para refutar o validar cuatro hipótesis sobre las propiedades de las correlaciones. Por último se realiza una discusión de los resultados obtenidos tanto en la implementación como de las subrutinas propuestas.

---

## Dedicatoria

---

A Dios, a mí y a todo aquel que me dio una oportunidad.

|   |             |
|---|-------------|
| <b>Prefacio</b>   | <b>III</b>  |
| <b>Dedicatoria</b>  | <b>IV</b>   |
| <b>Lista de figuras</b>                                   | <b>VII</b>  |
| <b>Lista de cuadros</b>                                   | <b>VIII</b> |
| <b>Resumen</b>  | <b>IX</b>   |
| <b>I. Introducción</b>                                    | <b>1</b>    |
| <b>II. Objetivos</b>                                      | <b>3</b>    |
| A. Objetivo general . . . . .                             | 3           |
| B. Objetivos específicos . . . . .                        | 3           |
| <b>III. Justificación</b>                                 | <b>4</b>    |
| <b>IV. Marco teórico</b>                                  | <b>5</b>    |
| A. Introducción a la mecánica cuántica . . . . .          | 5           |
| 1. Espacios de Hilbert . . . . .                          | 5           |
| 2. Representación de estados cuánticos y qubits . . . . . | 7           |
| 3. Postulados de la Mecánica cuántica . . . . .           | 8           |
| 4. Canales cuánticos . . . . .                            | 9           |
| B. Contexto histórico . . . . .                           | 9           |
| 1. Paradoja de Einstein-Podolsky-Rosen . . . . .          | 9           |
| 2. Experimento de Bell . . . . .                          | 10          |
| C. Entropía de Shannon . . . . .                          | 11          |
| 1. Propiedades . . . . .                                  | 13          |
| 2. Entropía relativa . . . . .                            | 13          |
| 3. Información mutua . . . . .                            | 14          |
| 4. Teoremas de codificación de Shannon . . . . .          | 14          |
| D. Entropía de von Neumann . . . . .                      | 15          |
| 1. Propiedades . . . . .                                  | 16          |
| 2. Entropía relativa . . . . .                            | 17          |
| 3. Información mutua . . . . .                            | 18          |
| E. Correlaciones clásicas y cuánticas . . . . .           | 19          |
| 1. Entrelazamiento cuántico . . . . .                     | 19          |



|  |           |
|--|-----------|
| 2. <i>Discord</i> cuántico . . . . .   | 19        |
| 3. <i>Discord</i> simétrico . . . . .  | 21        |
| F. Método de máxima entropía . . . . .   | 22        |
| G. Estructuras de correlación . . . . .  | 24        |
| 1. En sistemas clásicos . . . . .  | 24        |
| 2. En sistemas cuánticos . . . . .   | 24        |
| H. Correlaciones en sistemas cuánticos multipartidos . . . . .                       | 25        |
| 1. Estructuras de correlación y correlaciones multipartidas . . . . .                | 25        |
| 2. Correlaciones irreducibles . . . . .  | 25        |
| 3. Correlaciones clásicas y cuánticas . . . . .                                      | 26        |
| <b>V. Antecedentes</b>   | <b>27</b> |
| <b>VI. Descripción del problema: Correlaciones cuánticas <math>k</math>-partidas</b> | <b>28</b> |
| A. Hipótesis . . . . .   | 29        |
| <b>VII. Algoritmo 1: selección de matrices de máxima entropía</b>                    | <b>30</b> |
| A. Fundamentación teórica . . . . .  | 30        |
| 1. Caracterización de estados de máxima entropía . . . . .                           | 30        |
| 2. <i>Gradient descent</i> . . . . .   | 31        |
| B. Algoritmo . . . . .   | 32        |
| 1. Preparación . . . . .   | 32        |
| 2. Optimización . . . . .  | 32        |
| 3. Diagrama de flujo . . . . .   | 32        |
| C. Eficiencia . . . . .  | 33        |
| D. Exploración de estados GHZ y W . . . . .  | 34        |
| <b>VIII. Algoritmo 2: selección de operadores locales</b>                            | <b>35</b> |
| A. Fundamentación teórica . . . . .  | 35        |
| 1. Variación de operadores locales . . . . .   | 35        |
| 2. <i>Differential evolution</i> . . . . .   | 36        |
| B. Algoritmo . . . . .   | 36        |
| 1. Diagrama de flujo . . . . .   | 37        |
| C. Eficiencia . . . . .  | 38        |
| D. Exploración de los estados GHZ y W . . . . .                                      | 38        |
| <b>IX. Implementación</b>  | <b>40</b> |
| A. Simulaciones . . . . .  | 40        |
| 1. Hipótesis 1 y 2 . . . . .   | 40        |
| 2. Hipótesis 3 . . . . .   | 41        |
| 3. Hipótesis 4 . . . . .   | 41        |
| B. Resultados . . . . .  | 41        |
| <b>X. Discusión de resultados</b>  | <b>44</b> |
| <b>XI. Conclusiones</b>  | <b>46</b> |
| <b>XII. Recomendaciones</b>  | <b>47</b> |
| <b>XIII. Referencias</b>   | <b>48</b> |
| <b>XIV. Anexos</b>   | <b>51</b> |
| A. Especificaciones de hardware y software . . . . .                                 | 51        |

---

## Lista de figuras

---

|   |    |
|---|----|
| VII.1. Diagrama de flujo de Algoritmo 1. . . . .  | 33 |
| VII.2. Eficiencia del Algoritmo 1. $T$ es el tiempo, en segundos, que tarda el código en resolver el Algoritmo 1 con un sistema de $N$ qubits. $T_0 = 0.0016$ s. Los puntos azules son las medias obtenidas por ejecución y la línea punteada naranja es la regresión por mínimos cuadrados ponderados. . . . . | 34 |
| VIII.1. Diagrama de flujo de Algoritmo 2. . . . .   | 37 |
| VIII.2. Eficiencia del Algoritmo 2. $T$ es el tiempo, en segundos, que tarda el código en resolver el Algoritmo 2 con un sistema de $N$ qubits. $T_0 = 0.987$ s. La línea punteada corresponde a la regresión lineal obtenida. . . . .  | 38 |

|  |    |
|--|----|
| VII.1 Entropía de las matrices reducidas para $\Phi^{(2)}$ y $N = 3$ . . . . .   | 34 |
| VII.2 Entropía de las matrices reducidas para $\Phi^{(3)}$ y $N = 4$ . . . . .   | 34 |
| VIII.1. Comparación de métodos de optimización. . . . .                          | 38 |
| VIII.2. Correlaciones de los estados GHZ y W con $N = 3$ . . . . .               | 39 |
| VIII.3. Correlaciones de los estados GHZ y W con $N = 4$ . . . . .               | 39 |
| IX.1. Correlaciones $k$ -partidas en el estado GHZ. . . . .                      | 42 |
| IX.2. Correlaciones $k$ -partidas en el estado W. . . . .                        | 42 |
| IX.3. Resultados de las simulaciones realizadas para hipótesis 1, 2 y 3. . . . . | 42 |
| IX.4. Mapeos de canal cuántico con $r = 2$ sobre GHZ. . . . .                    | 42 |
| IX.5. Mapeos de canal cuántico con $r = 3$ sobre GHZ. . . . .                    | 42 |
| IX.6. Mapeos de canal cuántico con $r = 2$ sobre W. . . . .                      | 43 |
| IX.7. Mapeos de canal cuántico con $r = 3$ sobre W. . . . .                      | 43 |

En el desarrollo de tecnologías cuánticas frecuentemente se utilizan como recurso las correlaciones entre los componentes de sistemas multipartidos. Las correlaciones en dichos sistemas son diferentes a las que presentan los sistemas clásicos, y esto permite que se pueda tomar ventaja de ellas. Sin embargo, el progreso en el entendimiento de correlaciones cuánticas para sistemas bipartidos no se ha visto reflejado para el caso de  $n$ -partidos, con  $n > 2$ . Con el propósito de aportar al entendimiento de estos casos, y utilizando los conocimientos de mecánica e información cuántica, programación y optimización adquiridos durante el programa de licenciatura, este trabajo pretende proporcionar un método para estudiar las características de correlaciones entre los componentes de un sistema cuántico multipartido. Para este propósito, en este trabajo se desarrollan dos subrutinas principales en Python: la primera, dedicada a encontrar la matriz de densidad que maximiza la entropía de von Neumann, dado un conjunto de restricciones en términos de sus matrices reducidas; la segunda, dada una matriz de densidad, se encarga de encontrar los observables locales que maximizan la información mutua entre los resultados de las mediciones asociadas. Las subrutinas utilizan los métodos de optimización *Gradient Descent* y *Differential Evolution*, respectivamente, para resolver el problema. Las subrutinas son utilizadas para verificar un conjunto de hipótesis sobre las propiedades de correlaciones en sistemas con características específicas.

Se concluyó a partir de las simulaciones realizadas que los operadores locales que maximizan las correlaciones de un sistema no siempre maximizan las correlaciones de sus matrices reducidas y que las correlaciones cuánticas  $k$ -partidas no se conservan ante la aplicación de canales cuánticos locales. Además, no se encontró evidencia para negar que todas las correlaciones cuánticas son no-negativas ni para negar que las operaciones unitarias locales modifican las correlaciones de un sistema.

Se espera que los resultados de este trabajo permitan que personas dedicadas al estudio de la teoría de información cuántica tengan una herramienta para simular sistemas aleatorios con los que se puedan estudiar nuevas propiedades de las correlaciones que están albergadas en ellos, con la posibilidad de ser adaptado a sistemas con características específicas.

Desde los inicios de la mecánica cuántica, se ha observado que los sistemas estudiados por la misma tienen características que no son usuales para la forma clásica de pensamiento a la que las personas están acostumbradas. Una muestra de la inquietud que genera el comportamiento de los sistemas cuánticos es la discusión que plantearon Albert Einstein, Boris Podolsky y Nathan Rosen en su famoso artículo sobre la completitud de la teoría (Einstein, Podolsky, y Rosen, 1935). Este *paper*, sin embargo, despertó un gran interés en estas características que hacen tan extraña a la mecánica cuántica.

J.S. Bell argumentó en una posición diferente a Einstein y sus colegas (Bell, 1964), contribuyendo al estudio del entrelazamiento cuántico. Además, se han propuesto experimentos mentales (*e.g.* (Bojić, 2013; Clauser, Horne, Shimony, y Holt, 1969)) y físicos (*e.g.* (Kumar, Nirala, y Marino, 2020)) que comprueban los resultados de Bell. A partir de estos resultados, se ha dado paso a explorar las correlaciones y sus características desconocidas. Analizar y aprender sobre el comportamiento y la clasificación de los tipos de correlaciones en los sistemas cuánticos nos permite comprender mejor la forma en que los sistemas cuánticos almacenan información y cómo ésta se comporta al momento de ser transmitida. Por lo tanto, el desarrollo de estos temas puede influir en la creación de tecnología con base en información cuántica y en el desarrollo de protocolos de seguridad a través de la criptografía cuántica.

Las correlaciones en sistemas cuánticos se clasifican, generalmente, por su naturaleza o por la forma en que se distribuyen. En los últimos 20 años, se han realizado diversos acercamientos en el entendimiento de las correlaciones, tanto desde el punto de vista de las correlaciones totales como de las correlaciones cuánticas (Modi, Brodutch, Cable, Paterek, y Vedral, 2012). Una de las definiciones más aceptadas para las correlaciones cuánticas, junto al entrelazamiento que es bien conocida, es la del *discord* (Henderson y Vedral, 2001; Ollivier y Zurek, 2001) y su versión simétrica (Loura, s.f.; Rulli y Sarandy, 2011). Sin embargo, a pesar de los avances en el entendimiento de las correlaciones en sistemas bipartidos, las correlaciones en sistemas multipartidos no son tan comprendidas.

En un esfuerzo por clasificar las correlaciones multipartidas, se han propuesto estructuras jerárquicas para medir las correlaciones  $k$ -partidas en un sistema  $n$ -partido ( $n, k \in \mathbb{N}$  tales que  $n \geq k \geq 1$ ) (Girolami, Tufarelli, y Susa, 2017) considerando criterios de producto y separabilidad (Szalay, 2018). Además, para cuantificar las correlaciones de los sistemas  $n$ -partidos, se ha desarrollado una forma de definir las correlaciones  $k$ -partidas basándose en el método de máxima entropía (Amari, 2001; Krippendorff, 1986; Schneidman, Still, Berry, Bialek, y cols., 2003). Estos avances son la base para el desarrollo del presente trabajo, en el cual se pretende combinar los conceptos de correlaciones

$k$ -partidas y *discord* global para obtener una clasificación más completa y de mayor generalidad. Se espera que los resultados provean una herramienta para futuras investigaciones en las que se necesite realizar simulaciones para verificar nuevas hipótesis.

### **A. Objetivo general**

Obtener un método para resolver el problema de selección de máxima entropía y aplicarlo para verificar hipótesis específicas relacionadas con las características de correlaciones de sistemas multipartidos.

### **B. Objetivos específicos**

- Explorar la relación entre el principio de máxima entropía y las correlaciones cuánticas.
- Implementar, en Python, un algoritmo capaz de resolver un problema de selección de matrices de densidad en sistemas multipartidos utilizando el método de máxima entropía bajo restricciones.
- Implementar, en Python, un algoritmo capaz de encontrar los observables locales que maximizan la información mutua entre los resultados de las mediciones asociadas.
- Verificar un conjunto de hipótesis sobre las propiedades de correlaciones en sistemas con características específicas.

En la teoría de información, tanto cuántica como clásica, uno de los objetos de estudio es la relación entre las variables de un sistema. Tener amplio conocimiento de las relaciones permite aprovecharse de ellas para encontrar formas de transmitir y guardar información con alta seguridad y eficiencia. Sin embargo, en el área de información cuántica, el estudio de estas relaciones es reciente y actual, aunque sí se ha logrado una gran cantidad de avances importantes (para un tratamiento más extensivo sobre dichos avances ver (Adesso, Bromley, y Cianciaruso, 2016)), especialmente utilizando estructuras que tienen analogías en la contraparte clásica.

En concreto, se sabe que las correlaciones en sistemas de varias partes se pueden clasificar según su naturaleza y según la cantidad de variables que involucran. También se sabe que la distribución de las correlaciones de un sistema en estas clasificaciones permite obtener estados específicos y con características interesantes. De estas nociones surgen preguntas sobre la forma correcta de medir estas correlaciones. Varios autores han intentado responder esta pregunta y han formulado propuestas que obtienen resultados (Linden, Popescu, y Wootters, 2002; Szalay, 2018; Zhou, 2008, 2009). Aún así, existe la posibilidad de encontrar un método que describa mejor el comportamiento de las correlaciones en sistemas cuánticos multipartidos y para esto es útil tener herramientas que ayuden a comprobar las propuestas o sugerir propiedades de estas. Este trabajo busca proporcionar estas herramientas a través de la programación, y a luz de las ideas y métodos propuestos por (Mariano Lemus y Paunković, s.f.).



## A. Introducción a la mecánica cuántica

La Mecánica Cuántica es una teoría fundamental de la Física. esta se planteó y desarrolló en el periodo de 1900 a 1930 (Sakurai y Napolitano, 2011), logrando explicar con éxito varios fenómenos que la mecánica clásica no podía explicar, como el comportamiento de partículas de escalas similares o incluso mucho menores que las de los átomos, o fenómenos como el efecto fotoeléctrico (Serway, Moses, y Moyer, 2004). Es por esto que se convirtió en una parte importante de la Física Moderna, y del mundo moderno, en general, a pesar de las críticas fuertes que le han realizado.

La Mecánica Cuántica es un marco matemático, o bien, un conjunto de reglas para la construcción de teorías físicas. Algunos ejemplos de estas teorías físicas son la Electrodinámica Cuántica o la Información Cuántica (Nielsen y Chuang, 2011). Ambas son construidas dentro del marco de la Mecánica cuántica, pero cada una tiene sus propias reglas y objetos que no son necesariamente determinados por la Mecánica Cuántica. Un sistema cuántico generalmente presenta las siguientes características:

1. La existencia de un límite en la precisión con la que una cantidad física puede medirse, *i.e.*, el principio de incertidumbre de Heisenberg.
2. El sistema se puede representar por estados, que representan la cuantización de las cantidades físicas del sistema, y por la superposición de estos estados.
3. Los estados, mediciones y resultados tienen una distribución de probabilidad asociada.
4. El marco matemático está construido sobre los espacios de Hilbert y se respalda axiomáticamente en los 4 postulados de la Mecánica Cuántica.

### 1. Espacios de Hilbert

La matemática de la Mecánica Cuántica se ha construido sobre la teoría del Espacio de Hilbert debido a las características útiles e interesantes que posee. La siguiente sección se tomó de (Ponciano, 2018). Un espacio de Hilbert se representa por la letra  $\mathcal{H}$  y posee 4 características importantes:

### Espacio vectorial lineal.

El espacio de Hilbert  $\mathcal{H}$  es un espacio vectorial lineal sobre el campo de los complejos, con las operaciones:

- Suma:  $|v\rangle + |w\rangle$ .
- Producto por escalares:  $a|v\rangle$ .

**Nota:** El espacio dual del espacio de Hilbert se representa como  $\mathcal{H}^*$  y a este pertenecen los vectores duales  $\langle v_i|$  tales que  $|v_j\rangle \in \mathcal{H}$  y  $\langle v_i|(|v_j\rangle) = \langle v_i|v_j\rangle \in \mathbb{C}$ .

### Producto interno.

El espacio de Hilbert está dotado de producto interno. El producto interno de  $|v\rangle$  con  $|w\rangle$  está definido por

$$(\cdot, \cdot) : \mathcal{H} \times \mathcal{H} \longrightarrow \mathbb{C} \ni (|v\rangle, |w\rangle) = \langle v|(|w\rangle) = \langle v|w\rangle.$$

El producto interno cumple con las siguientes propiedades:

1. Es lineal respecto del segundo argumento:  $\langle v|(a|w\rangle + b|z\rangle) = a\langle v|w\rangle + b\langle v|z\rangle$
2. Satisface hermiticidad:  $\langle v|w\rangle^* = \langle w|v\rangle$
3. Es no-negativo:  $\langle v|v\rangle \geq 0$ , con igualdad si y solo si  $|v\rangle = |0\rangle$ .

Al estar dotado de producto interno, la norma de  $|v\rangle$  se define como

$$\|v\| \equiv \sqrt{\langle v|v\rangle} \tag{IV.1}$$

Por lo tanto, el espacio de Hilbert es un espacio normado y, consecuentemente, métrico. También son válidas la desigualdad triangular y la desigualdad de Cauchy-Schwartz. Por otra parte, el producto interno muestra ortogonalidad: dos estados  $|v\rangle, |w\rangle$  son ortogonales si y solo si su producto interno es igual a cero.

### $\mathcal{H}$ es completo.

Un espacio es completo si toda sucesión de Cauchy del espacio converge a un límite que está en el espacio. Esto permite que se utilicen técnicas de cálculo.

### $\mathcal{H}$ es separable.

Un espacio  $\mathcal{H}$  es separable si existe un conjunto de vectores denso en  $\mathcal{H}$  y es contable. Es equivalente decir que  $\mathcal{H}$  es separable si existe un sistema contable completo en el espacio, o bien, si contiene una base ortonormal contable. Por lo tanto, la dimensión de  $\mathcal{H}$  es la cardinalidad de la base completa que genera  $\mathcal{H}$ .

## 2. Representación de estados cuánticos y qubits

Los estados cuánticos son representaciones matemáticas abstractas de el sistema que se está estudiando. Para esta sección se utilizó como referencia a (Nielsen y Chuang, 2011). Estos estados cuánticos, dependiendo de la formulación que se utilice, se pueden representar por los **vectores estado** o por **matrices de densidad**, donde las últimas proveen una descripción más completa. Utilizando la notación de Dirac (también llamada “notación bra-ket”), se representará un vector estado  $v$  con el símbolo  $|v\rangle$ , llamado **ket**, y tiene una representación isomorfa en  $\mathbb{C}^n$  si pertenece a un espacio de dimensión finita. Así, la transpuesta conjugada de  $|v\rangle$  se representa por el símbolo **bra**  $\langle v|$ . Al ser vectores, una combinación lineal se puede escribir como  $\sum_i a_i |v_i\rangle$ , donde  $a_i$  es la amplitud de probabilidad de cada estado.

Por ejemplo, considérese el sistema de un qubit (un bit cuántico). Un qubit tiene dos posibles estados:  $|0\rangle$  y  $|1\rangle$ , que corresponden a los estados clásicos 0 y 1 en un bit. Es decir, pertenece a un espacio vectorial complejo de dimensión 2. Pero un qubit puede estar en superposición de estos dos estados:  $|\psi\rangle = a|0\rangle + b|1\rangle$ , con  $a, b \in \mathbb{C}$ . Al medir el estado del sistema, se puede obtener 0 con probabilidad  $|a|^2$  o bien 1 con probabilidad  $|b|^2$ .

Por otra parte, una matriz  $n \times n$  se puede escribir en notación de Dirac como el operador  $A = \sum_{i,j} a_{ij} |\psi_i\rangle \langle \psi_j|$ , donde  $|\psi_i\rangle$  son todos los estados en los que actúa la matriz. Entonces al aplicar el operador a un estado  $|\psi\rangle \in V$ , se escribe como  $A|\psi\rangle \in V$ . Ahora, se procede a describir los operadores de densidad. El operador de densidad provee un lenguaje conveniente para describir sistemas cuánticos cuyo estado no es completamente conocido. Supóngase un *ensamble de estados puros*  $\{p_i, |\psi_i\rangle\}$ , con los posibles estados  $|\psi_i\rangle$  en los que el sistema cuántico se puede encontrar con probabilidades  $p_i$ , respectivamente. El operador densidad del sistema se define como:

$$\rho \equiv \sum_i p_i |\psi_i\rangle \langle \psi_i| \quad (\text{IV.2})$$

donde  $\rho$  es hermítico. También se puede describir un sistema como una mezcla de estados  $\rho_i$ :

$$\rho = \sum_i p_i \rho_i \quad (\text{IV.3})$$

donde

$$\rho_i = \sum_j p_{ij} |\psi_{ij}\rangle \langle \psi_{ij}| \quad (\text{IV.4})$$

**Caracterización:** Un operador  $\rho$  es un operador de densidad asociado a  $\{p_i, |\psi_i\rangle\}$  si y solo si satisface

1. **Condición de la traza:**  $\text{Tr}(\rho) = 1$
2. **Condición de positividad:**  $\rho$  es un operador positivo semi-definido, es decir, las mediciones sobre los estados son positivas.

$$\langle \phi | \rho | \phi \rangle = \sum_i p_i |\langle \phi | \psi_i \rangle|^2 \geq 0. \quad (\text{IV.5})$$

Luego, un estado se puede reconocer como un **estado puro** si la traza del operador densidad al cuadrado es igual a uno ( $\text{tr}(\rho^2) = 1$ ) y se reconoce como un **estado mezcla** cuando la traza del operador al cuadrado es menor a uno ( $\text{tr}(\rho^2) < 1$ ).

**Nota:** Un operador de densidad no está asociado a un único ensamble de estados puros, lo cual se conoce como “Libertad unitaria de ensamble”. Véase el siguiente teorema:

**Teorema 1.** Los conjuntos  $\{|\bar{\psi}_i\rangle\}$  y  $\{|\bar{\phi}_j\rangle\}$  (estados no normalizados,  $|\bar{\psi}_i\rangle = \sqrt{p_i}|\psi_i\rangle$ ) generan la misma matriz de densidad si y solo si

$$|\bar{\psi}_i\rangle = \sum_j u_{ij} |\bar{\phi}_j\rangle \quad (\text{IV.6})$$

donde  $u_{ij}$  es una matriz unitaria en los complejos.

### 3. Postulados de la Mecánica cuántica

Además de definir el espacio de Hilbert, para comprender el marco matemático de la Mecánica Cuántica y sus interpretaciones es necesario introducir los postulados, permitiéndonos tener una perspectiva completa de esta. A continuación se presentan los postulados de la Mecánica Cuántica, utilizando operadores de densidad, que es equivalente a si se utilizase la descripción en vectores. Se utilizó (Nielsen y Chuang, 2011) como referencia en esta sección. El primer postulado describe qué tipo de objetos matemáticos se utilizarán:

**Postulado 1.** Un sistema físico aislado está asociado con un espacio de Hilbert, el cual es conocido como **espacio estado** del sistema. El sistema es completamente descrito por su **vector estado** (operador de densidad)  $\rho$ , el cual es positivo definido y tiene  $\text{tr}(\rho) = 1$ . Un sistema cuántico también se puede escribir como la mezcla de los estados  $\rho_i$  con probabilidades  $p_i$ :

$$\rho = \sum_i p_i \rho_i \quad (\text{IV.7})$$

Nótese que, como  $\mathcal{H}$  es separable, cada vector estado se puede escribir como una combinación lineal de los vectores de la base. El segundo postulado da una descripción de cómo cambian los estados a lo largo del tiempo.

**Postulado 2.** La evolución de un sistema cuántico cerrado está descrito por una transformación unitaria. Es decir, el estado  $\rho$  de un sistema en el tiempo  $t_1$  está relacionado con el estado  $\rho'$  del sistema en el tiempo  $t_2$  a través del operador  $U$ , el cual depende de los tiempos  $t_1$  y  $t_2$ :

$$\rho' = U\rho U^\dagger = \sum_i p_i U |\psi_i\rangle\langle\psi_i| U^\dagger \quad (\text{IV.8})$$

Luego, es necesario explicar qué sucede cuando se interactúa con un sistema y, por lo tanto, ya no es cerrado. El tercer postulado explica cómo reacciona un sistema ante las mediciones.

**Postulado 3.** Las mediciones cuánticas están descritas por un conjunto  $\{M_m\}$  de operadores de medición, que actúan en el estado espacio del sistema.  $m$  se refiere al resultado de la medición que pueda ocurrir en el experimento. La probabilidad de obtener  $m$  es

$$p(m) = \text{Tr}(M_m^\dagger M_m \rho), \quad (\text{IV.9})$$

el estado post-medición es

$$\rho_m = \frac{M_m \rho M_m^\dagger}{\text{Tr}(M_m^\dagger M_m \rho)} \quad (\text{IV.10})$$

Y los operadores de medición satisfacen con la condición de completitud:

$$\sum_m M_m^\dagger M_m = \mathbb{I} \quad (\text{IV.11})$$

Por último, ya que generalmente se trabaja con sistemas compuestos por dos o más subsistemas, el cuarto postulado explica cómo se describen dichos sistemas.

**Postulado 4.** *El espacio estado de un sistema físico compuesto es el producto tensorial de los espacios estados de los subsistemas físicos. Es más, si cada subsistema  $i$  ( $i = 1, 2, \dots, n$ ) está en el estado  $\rho_i$ , el estado conjunto del sistema es*

$$\rho = \rho_1 \otimes \rho_2 \otimes \dots \rho_n. \quad (\text{IV.12})$$

donde  $\otimes$  representa el producto tensorial.

Este último postulado también permite definir uno de los fenómenos más interesantes de la Mecánica Cuántica: el entrelazamiento cuántico.

## 4. Canales cuánticos

**Definición:** Un mapeo lineal  $\Lambda: M_{d_1}(\mathbb{C}) \rightarrow M_{d_2}(\mathbb{C})$  es llamado canal cuántico si es:

1. completamente positivo: mapea matrices semi-definidas positivas a matrices semi-definidas positivas.
2. traza-preservante: para todo  $X \in M_{d_1}(\mathbb{C})$ ,  $\text{Tr}\{\Lambda(X)\} = \text{Tr}\{X\}$ .

En (Nielsen y Chuang, 2011) se muestra un teorema que establece que un canal cuántico admite la descomposición de Kraus:

$$\forall X \in M_{d_1}(\mathbb{C}), \Lambda(X) = \sum_{i=1}^r A_i X A_i^\dagger \quad (\text{IV.13})$$

para matrices  $A_1, \dots, A_r \in M_{d_2 \times d_1}(\mathbb{C})$  llamadas operadores de Kraus. Los operadores de Kraus cumplen con  $\sum_{i=1}^r A_i^\dagger A_i = I_{d_1}$ .

## B. Contexto histórico

### 1. Paradoja de Einstein-Podolsky-Rosen

En la década de 1920, para muchos físicos era claro que la mecánica clásica no podía describir enteramente el mundo atómico y subatómico. Era necesario reconstruir la física y esto llevó a los inicios de la mecánica cuántica. Werner Heisenberg y Niels Bohr, entre otros, habiendo contribuido en gran parte a la creación de esta teoría, insistían en que no hay forma con sentido suficiente para explicar comportamientos atómicos, por ejemplo, uno no puede predecir el momento preciso en el que un átomo liberará un fotón. Esto le molestaba mucho a Albert Einstein y nunca pudo aceptar esta aleatoriedad (*'God does not play dice'*, declaró) (Giuliani, 2020; of Photonic Sciences, s.f.).

En 1935, Einstein, Podolsky y Rosen publicaron su famoso *paper*, en el cual introdujeron un experimento mental que argumentaba que la mecánica cuántica no es una teoría física completa, a través de establecer que una teoría física completa debe cumplir con la *condición de completitud*: “todos los elementos de una realidad física deben tener una contraparte en la teoría física”. Actualmente se conoce como ‘La paradoja EPR’ y su objetivo era demostrar las dificultades conceptuales intrínsecas en la mecánica cuántica (Giuliani, 2020; Einstein y cols., 1935).

Los autores argumentaron que los elementos de la realidad física no se determinan por consideraciones filosóficas *a priori*, sino que deben apoyarse de resultados de experimentos y mediciones. Con esto proponen el siguiente criterio: ‘si, sin perturbar de alguna forma el sistema, podemos predecir con certeza (probabilidad igual a uno) el valor de una cantidad física, entonces existe un elemento de

realidad física correspondiente a su cantidad física'. Luego, los autores conducen a que el principio de incertidumbre de Heisenberg, uno de los principios más importantes de la mecánica cuántica, contiene una contradicción importante al aplicarlo a los operadores de posición y momentum. De acuerdo al criterio que Einstein, Podolsky y Rosen propusieron, el principio de incertidumbre de Heisenberg nos dice que el momentum y la posición (y cualquier par de variables complementarias) no son elementos de realidad simultáneamente. Finalmente, concluyen que la mecánica cuántica no es una teoría física completa, ya que variables que se pueden medir simultáneamente en la mecánica clásica no se pueden medir simultáneamente en la mecánica cuántica (Einstein y cols., 1935).

Por otra parte, hay otra cualidad que molestaba a Einstein: el entrelazamiento. Esta cualidad permite que dos sistemas cuánticos que han interactuado en el pasado sean afectados instantáneamente por mediciones en uno u otro sistema. Esto parece violar uno de los principios de la relatividad: la información no se puede transmitir más rápido que la velocidad de la luz, porque esto violaría la causalidad. Es por esto que también propusieron una alternativa: la teoría de la mecánica cuántica tiene variables adicionales escondidas, i.e, hay variables que no se tomaron en cuenta al desarrollar esta teoría (Brassard y Méthot, 2006; Giuliani, 2020; Einstein y cols., 1935).

## 2. Experimento de Bell

En 1964, John S. Bell escribió un *paper* en respuesta a los argumentos de la paradoja EPR. Bell toma, inicialmente, el planteamiento de Bohm y Aharonov: un par de partículas de espín 1/2 que están entrelazadas y se mueven libremente en direcciones opuestas. Luego, desarrolla la demostración por contradicción de que la mecánica cuántica no se puede explicar por variables locales escondidas y que, en efecto, es una teoría no local. (Bell, 1964)

Para comprobar lo que Bell dijo en su *paper* y entenderlo de mejor manera, se ha realizado una variedad de experimentos, tanto mentales como físicos. Hasta ahora, los experimentos realizados han obtenido resultados a favor del, actualmente, teorema de Bell (Bell y Aspect, 2004; Bohm y Aharonov, 1957). Uno de los experimentos importantes, debido a que muestra una ventaja de la información cuántica y además tiene resultados interesantes, es el experimento CHSH.

### Experimento CHSH

También conocido como el juego CHSH, fue planteado por John Clauser, Michael Horne, Abner Shimony y Richard Holt (Clauser y cols., 1969). El juego consiste en lo siguiente: Alice y Bob, que están separados espacialmente, recibirán aleatoriamente un bit cada uno  $(x, y)$  y cada uno regresará un bit de resultado  $(a, b)$  sin comunicarse entre ellos. Alice y Bob pueden utilizar cualquier estrategia clásica o cuántica, pero para ganar deben cumplir que

$$x \wedge y = a \oplus b \tag{IV.14}$$

es decir, ganan cuando los resultados son diferentes y los bits originales son iguales a 1 o ganan cuando los resultados son iguales para cualquier otro par de bits iniciales.

Un análisis para el caso en que Alice y Bob están limitados a compartir solamente información clásica previo a iniciar el juego muestra que la máxima probabilidad de ganar es del 75% (por ejemplo cuando ambos eligen 0). Definiendo

$$S = 4(P_{ganar} - P_{perder}) \tag{IV.15}$$

y obteniendo la cota máxima, se puede observar la Desigualdad CHSH:

$$||S|| \leq 2. \tag{IV.16}$$

Todas las estrategias que corresponden a teorías de variables locales escondidas tienen esta cota superior. Por otra parte, Alice y Bob pueden utilizar una estrategia cuántica. Antes de iniciar el juego, tanto Alice como Bob tienen un qubit del estado entrelazado

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \quad (\text{IV.17})$$

entonces Alice y Bob realizarán una medición del estado  $\psi$  en una base ortogonal diferente para cada bit recibido. Estos son los casos:

- si  $x = 0$ , Alice mide en la base estándar.
- si  $x = 1$ , Alice mide en una base rotada  $\pi/4$ .
- si  $y = 0$ , Bob mide en una base rotada  $\pi/8$ .
- si  $y = 1$ , Bob mide en una base rotada  $-\pi/8$ .

Ahora, la probabilidad de que los resultados sean iguales depende del ángulo entre las bases ( $\theta$ ) en las que se midió, así

$$P_c(\theta) = \cos^2 \theta \quad (\text{IV.18})$$

Nótese que para las entradas 00, 01 y 10,  $\theta = \pi/8$  por lo que la probabilidad de que las mediciones coincidan es  $\cos^2(\pi/8)$ . En cambio, para la entrada 11,  $\theta = 3\pi/8$ , por lo que la probabilidad de que **no** coincidan es también  $\cos^2(\pi/8)$ . Por lo tanto, la probabilidad de ganar es siempre

$$\cos^2(\pi/8) = \frac{2 + \sqrt{2}}{4} \approx 0.85 \quad (\text{IV.19})$$

Obsérvese que la probabilidad de ganar supera un 10% a la cota clásica. Ahora, comparando la cantidad  $S$ , la desigualdad CHSH en la mecánica cuántica resulta:

$$\|S\| \leq 2\sqrt{2} \quad (\text{IV.20})$$

que es una cota mayor que la obtenida de forma clásica. Por lo tanto, se concluye que la estrategia cuántica es mejor para este juego y que las teorías de variables locales escondidas no pueden describir completamente los fenómenos de la mecánica cuántica (Bojić, 2013; Clauser y cols., 1969).

## C. Entropía de Shannon

La Teoría de Información es un campo de estudio que se ocupa de cuantificar la información para la comunicación. Es una subárea de la Matemática y estudia tópicos como compresión de información y los límites de procesamiento de señales. Claude Shannon realizó grandes contribuciones a la Teoría de Información, entre ellos fue la definición de entropía y el establecimiento de ‘Cantidad de información’ en eventos o variables aleatorias. Este es un concepto fundamental que ha tenido otras definiciones, pero que en este contexto es más útil utilizar la definición de Shannon. Esta sección, incluyendo sus subsecciones, recopilan definiciones, propiedades e interpretaciones de (Cover y Thomas, 2006; Gray, 2011; Klüver y Klüver, 2007; MacKay y Mac Kay, 2003; Shannon y Weaver, 1963; Vajapeyam, 2014; Witten, 2020).

Conceptualmente, **la información** puede pensarse como almacenada o transmitida en *variables* que pueden tener diferentes valores. Una *variable* puede entenderse como una unidad de almacenaje que puede guardar uno de diferentes valores específicos a través de algún proceso. Para calcular la *cantidad de información* o la *sorpresa* de una variable/evento, utilizamos la *Información de Shannon* o *información propia*:

$$I(x) = -\log_2(P(x)) \quad (\text{IV.21})$$

donde  $x$  representa al evento y  $P(x)$  es la probabilidad de que suceda. Esta cantidad está medida en bits o Shannons y siempre es positiva. Esto se puede interpretar como qué tan difícil es adivinar esa información sin observar la variable. La información tiene propiedades fundamentales:

- $I(p)$  es monótona decreciente en  $p$ , es decir, un aumento en la probabilidad del evento disminuye la información obtenida de observar el evento.
- La información es una cantidad no negativa:

$$I(p) \geq 0 \quad (\text{IV.22})$$

- Los eventos que ocurren siempre no transmiten información:

$$I(p = 1) = 0 \quad (\text{IV.23})$$

- La información obtenida de eventos independientes es la suma de la información obtenida de cada evento:

$$I(p_1, p_2) = I(p_1) + I(p_2). \quad (\text{IV.24})$$

La **entropía** de una variable es la *cantidad de información* en promedio contenida dentro de la variable. Según Shannon, la entropía  $H$  de una variable aleatoria discreta  $X$  con posibles valores  $\{x_1, \dots, x_n\}$  y función de probabilidad  $P(X)$  se define como

$$H(X) = E[I(X)] = E[-\log_2(P(X))] = -\sum_{i=1}^n P(x_i) \log_2 P(x_i) \quad (\text{IV.25})$$

donde  $E$  es el operador de valor esperado. En el caso de que  $P(x_i) = 0$  se define que  $P(x_i) \log_2 P(x_i) \equiv 0$ , es decir, no aporta a la entropía de la variable. Como se establece en el **teorema de codificación de Shannon**, la entropía es una medida directa del número óptimo de bits que se necesitan para almacenar la información en una variable. La entropía también se puede interpretar como la cantidad de incertidumbre en una distribución de probabilidad.

Es conveniente definir la **entropía conjunta** de un par de variables aleatorias  $X$  y  $Y$ :

$$H(X, Y) \equiv -\sum_{x, y} p(x, y) \log_2 p(x, y), \quad (\text{IV.26})$$

la cual es la información obtenida de medir dos variables simultáneamente. También se define la **entropía condicional** de la variable  $X$  respecto de la variable  $Y$  (con posibles valores  $x$  y  $y$ , respectivamente) como

$$H(X|Y) \equiv H(X, Y) - H(Y). \quad (\text{IV.27})$$

La entropía condicional cuantifica la cantidad de información necesaria para describir el resultado de la variable aleatoria  $X$  dado que se sabe el valor de la variable aleatoria  $Y$ . Al restar las ecuaciones anteriores, se puede reescribir para obtener la regla de Bayes:

$$H(Y|X) = H(X|Y) - H(X) + H(Y) \quad (\text{IV.28})$$

La entropía condicional también se puede escribir como

$$H(X|Y) = -\sum_y p(y) H(X|y). \quad (\text{IV.29})$$



## 1. Propiedades

Algunas propiedades que tiene la entropía de Shannon son:

- Añadir o remover un evento con cero probabilidad no afecta a la entropía:

$$H_{n+1}(p_1, \dots, p_n, 0) = H_n(p_1, \dots, p_n). \quad (\text{IV.30})$$

- La entropía de un evento  $X$ , con  $n$  posibles valores, tiene un máximo, el cual se obtiene cuando todos los posibles valores son equiprobables:

$$H(x) \leq \log_2(n) \quad (\text{IV.31})$$

- Si  $X$  y  $Y$  son variables independientes aleatorias, entonces saber el valor de alguna no afecta al conocer el valor de la otra:

$$H(X|Y) = H(X) \quad (\text{IV.32})$$

$$H(Y|X) = H(Y) \quad (\text{IV.33})$$

- **Subaditividad:** La entropía conjunta de dos variables aleatorias no es mayor que la suma de las entropías de cada variable individual:

$$H(X, Y) \leq H(X) + H(Y). \quad (\text{IV.34})$$

- **Subaditividad fuerte:** es la desigualdad

$$H(X, Y, Z) + H(Y) \leq H(X, Y) + H(Y, Z) \quad (\text{IV.35})$$

donde se cumple la igualdad si y solo si  $Z \rightarrow Y \rightarrow X$  forma una cadena de Markov.

- **Regla de la cadena:** Sean  $X_1, \dots, X_n$  y  $Y$  un conjunto cualquiera de variables aleatorias. Entonces

$$H(X_1, \dots, X_n|Y) = \sum_{i=1}^n H(X_i|Y, X_1, \dots, X_{i-1}) \quad (\text{IV.36})$$

## 2. Entropía relativa

La **entropía relativa**, también conocida como **divergencia de Kullback-Leibler**, es otra medida útil, la cual mide la cercanía de dos distribuciones de probabilidad,  $p(x)$  y  $q(x)$ , sobre el mismo conjunto índice  $x$ . La entropía relativa de  $p(x)$  a  $q(x)$  se define como

$$H(p(x)||q(x)) \equiv \sum_x p(x) \log_2 \frac{p(x)}{q(x)} \equiv -H(p(x)) - \sum_x p(x) \log_2 q(x) \quad (\text{IV.37})$$

donde se define  $0 \log_2 0 \equiv 0$  y  $p(x) \log_2 0 \equiv +\infty$ , si  $p(x) > 0$ . Nótese que la entropía relativa es una función convexa y asimétrica en la pareja  $(p(x), q(x))$ . Además existe el teorema de la no-negatividad de la entropía relativa, es decir,

$$H(p(x)||q(x)) \geq 0 \quad (\text{IV.38})$$

donde

$$H(p(x)||q(x)) = 0 \Leftrightarrow p(x) = q(x) \quad (\text{IV.39})$$

Es decir, mientras más similares son las distribuciones de probabilidad, la entropía relativa es más cercana a cero. Ahora, suponga que  $p(x)$  es una distribución de probabilidad para la variable  $X$ , con

$d$  posibles resultados, y que  $q(x) = 1/d$  es una distribución uniforme de probabilidad con los mismos posibles resultados. Entonces

$$H(p(x)||1/d) = -H(X) - \sum_x p(x) \log_2(1/d) \quad (\text{IV.40})$$

$$= \log_2 d - H(X) \quad (\text{IV.41})$$

Luego, por la propiedad de no-negatividad, encontramos el valor máximo de la entropía de una variable  $X$  con  $d$  posibles resultados:

$$\log_2 d \geq H(X). \quad (\text{IV.42})$$

### 3. Información mutua

La **información mutua** de dos variables aleatorias es una medida de la dependencia entre ambas variables. En otras palabras, mide qué tanta información tienen  $X$  y  $Y$  en común. Se obtiene la información mutua al sustraer la entropía conjunta de  $(X, Y)$  a la suma de las entropías independientes de  $X$  y  $Y$ :

$$H(X : Y) \equiv H(X) + H(Y) - H(X, Y) \quad (\text{IV.43})$$

La información mutua también se escribe como  $I(X; Y)$ . Utilizando la entropía condicional, se puede escribir la igualdad

$$H(X : Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \quad (\text{IV.44})$$

Además, la información mutua también se puede expresar en términos de la entropía relativa:

$$H(X : Y) = H(p(x, y)||p(x)p(y)) \quad (\text{IV.45})$$

Otra forma de interpretar la información mutua es la siguiente: “(la información mutua) mide una reducción promedio en la incerteza de  $X$  que resulta de saber el valor de  $Y$ ; o bien, la cantidad promedio de información que  $Y$  transmite sobre  $X$ ”. Este concepto es utilizado como una forma general de una medida de correlación.

### 4. Teoremas de codificación de Shannon

La teoría de información y la teoría de codificación están fuertemente relacionadas en el sentido de que la teoría de información establece lo que es posible hacer con la información, mientras que la teoría de codificación establece formas de lograr estos resultados (Santos, 2016). Ahora, la forma principal que utiliza la teoría de información para establecer estos límites es a través de los teoremas de codificación que estableció Claude E. Shannon en 1948 y que fueron publicados, en colaboración con Warren Weaver, en el libro “*The Mathematical Theory of Communication*” (Shannon y Weaver, 1963).

En un sistema de comunicación, la información procede de una **fuentes** y es enviado y procesado por un **codificador** de manera que tenga un formato adecuado para la transmisión a través de un medio físico de comunicaciones, llamado **canal**. Algunas veces se presentan perturbaciones, llamadas **ruido**, en el canal e introducen errores a la información que se transmite. La información enviada es, finalmente, procesada por un **decodificador** para obtener el formato original para que lo pueda leer el usuario final, llamado **receptor**. Un canal es llamado **ruidoso** si la salida del canal no es igual a la entrada del canal.

Un codificador puede ejecutar dos procesos básicos: **codificación de fuente** y **codificación de canal**. El primero tiene como objetivo eliminar redundancia, *i.e.*, comprimir la información lo mejor posible sin que se pierda información del mensaje original. El segundo, en cambio, tiene como objetivo introducir redundancia adicional para tomar en cuenta posibles ruidos que puedan afectar la transmisión de la información (Santos, 2016).

### Teorema de codificación de canales sin ruido.

Este teorema cuantifica el grado al cual se puede comprimir la información producida por una fuente de información clásica. Suponga una fuente de información independiente e idénticamente distribuida (i.i.d) que produce los bits  $X_1, X_2, \dots$ , cada uno puede ser cero con probabilidad  $p$  o puede ser uno con probabilidad  $1 - p$ . Sean  $x_1, x_2, \dots, x_n$  las posibles secuencias de valores para las variables aleatorias  $X_1, X_2, \dots, X_n$ . Si una secuencia tiene alta probabilidad de ocurrir se conoce como **secuencia típica** o, de lo contrario, **secuencias atípicas**. Supóngase que se obtiene una secuencia de  $n$  valores lo suficientemente larga, entonces, en una secuencia típica, se espera con gran probabilidad que contenga una fracción  $p$  de ceros y el resto  $(1 - p)$  de unos. La probabilidad de obtener esta secuencia es

$$p(x_1, \dots, x_n) = p(x_1)p(x_2)\dots p(x_n) \approx p^{np}(1 - p)^{(1-p)n}. \quad (\text{IV.46})$$

Al tomar, el logaritmo a ambos lados se obtiene

$$-\log_2 p(x_1, \dots, x_n) \approx -np \log_2 p - n(1 - p) \log_2(1 - p) = nH(X) \quad (\text{IV.47})$$

donde  $X$  es una variable aleatoria distribuida de acuerdo a la fuente de distribución y  $H(X)$  es la entropía de la fuente de distribución. Finalmente, se obtiene que

$$p(x_1, \dots, x_n) \approx 2^{-nH(X)} \quad (\text{IV.48})$$

por lo que existen, como máximo,  $2^{nH(X)}$  secuencias típicas. Como consecuencia, sólo se requieren  $nH(X)$  bits para codificar de forma única cada secuencia típica (Santos, 2016; Shannon y Weaver, 1963).

### Teorema de codificación de canales ruidosos

Este teorema establece un máximo sobre la información que se puede transmitir a través de un canal con ruido casi completamente sin errores. Un canal se caracteriza por la distribución de probabilidad condicional  $p(Y|X)$  y es lo que diferencia un canal de otro. La **capacidad**  $C$  del canal ruidoso  $\mathcal{N}$  se define como

$$C(\mathcal{N}) = \max_{p(x)} H(X : Y) \quad (\text{IV.49})$$

donde el máximo se toma sobre todas las distribuciones de entrada  $p(x)$  para  $X$  e  $Y$  es la variable aleatoria correspondiente a la salida del canal (Santos, 2016; Shannon y Weaver, 1963).

## D. Entropía de von Neumann

John von Neumann realizó contribuciones importantes a la formulación matemática de la mecánica cuántica a través de su libro “*Mathematische Grundlagen der Quantenmechanik*” (1932), en el que introdujo las matrices de densidad. Las matrices de densidad fueron pensadas por von Neumann para tener un rol similar a las distribuciones de probabilidad y las funciones de partición en el marco clásico. Estas matrices de densidad permiten calcular todas las cantidades cuánticas promedio de manera similar a lo antes conocido y representar los estados cuánticos de forma completa. La siguiente sección recopila definiciones, propiedades e interpretaciones de (Nielsen y Chuang, 2011; Ohya y Petz, 2004; Petz, 2001; Preskill, 2018; Wilde, 2013). La entropía de von Neumann de un estado cuántico  $\rho$  se define como

$$S(\rho) = -\text{tr}(\rho \log_2 \rho), \quad (\text{IV.50})$$

donde  $\text{tr}(\cdot)$  es la traza. Sean  $\lambda(\rho) = \{\lambda_x\}$  los eigenvalores de  $\rho$  de forma en que  $\sum_x \lambda_x = 1$ , entonces la entropía de von Neumann se puede reescribir como

$$S(\rho) = - \sum_x \lambda_x \log_2 \lambda_x = H(\lambda(\rho)) \quad (\text{IV.51})$$

en donde, al igual que la entropía de Shannon, se define  $0 \log_2 0 \equiv 0$ . La entropía de von Neumann mide tanto la información cuántica contenida por letra del ensamble de estados puros (el número mínimo de qubits por letra que se necesitan para codificar de forma segura la información), como también su información clásica (la cantidad máxima de información por letra, en bits, que se obtiene). Obsérvese el siguiente teorema.

**Teorema 2.** *Considérese el estado mezcla  $\rho = \sum_i p_i \rho_i$ , donde  $\{p_i\}$  es una distribución de probabilidad y  $\rho_i$  son matrices de densidad. Entonces*

$$S(\rho) \leq H(p_i) + \sum_i p_i S(\rho_i) \quad (\text{IV.52})$$

con igualdad si y solo si los estados  $\rho_i$  tienen soporte en subespacios ortogonales.

Este teorema muestra que la entropía de Shannon de los eigenvalores de  $\rho$  es la entropía mínima que se puede obtener de  $\rho$ , ya que se puede tener en cuenta cualquier combinación lineal con la que se pueda escribir  $\rho$  y la entropía de la distribución de probabilidad  $p_i$ .

**Nota.** Sea  $\rho$  una matriz de densidad y  $\lambda_i$  sus eigenvalores, y sea  $M_j$  un operador de medición para el resultado  $j$  con eigenvalores  $\gamma_j$ . En el caso en que cada  $\gamma_j$  es un múltiplo de  $\lambda_i$ , es decir,  $\rho$  conmuta con  $M_j$ , los resultados de la medición están determinados por la distribución de probabilidad de los eigenvalores de  $\rho$ . A estos operadores de medición se les llama Medición proyección-valuada (PVM, por sus siglas en inglés).

La **entropía conjunta**  $S(A, B)$  de un sistema compuesto por dos componentes  $A$  y  $B$  está definida por

$$S(A, B) \equiv - \text{tr}(\rho^{AB} \log_2(\rho^{AB})) \quad (\text{IV.53})$$

donde  $\rho^{AB}$  es la matriz de densidad del sistema compuesto  $AB$ . La **entropía condicional** se define por

$$S(A|B) \equiv S(A, B) - S(B) \quad (\text{IV.54})$$

y es la información cuántica obtenida de  $A$  dado que se midió antes en  $B$ , o viceversa. Al restar las ecuaciones anteriores, se obtiene la regla de Bayes:

$$S(A|B) = S(B|A) + S(A) - S(B) \quad (\text{IV.55})$$

Suponga un estado puro  $|AB\rangle$  de un sistema compuesto perteneciente a Alice y Bob. Entonces  $|AB\rangle$  está entrelazado si y solo si  $S(B|A) < 0$ .

## 1. Propiedades

La entropía de von Neumann tiene las siguientes propiedades:

- Es no-negativa. La entropía es cero si y solo si el estado es puro.

$$S(\rho) \geq 0 \quad (\text{IV.56})$$

- En un espacio de Hilbert  $n$ -dimensional, la entropía es como máximo  $\log_2 n$ . La entropía es igual a  $\log_2 n$  si y solo si el sistema está en un estado completamente mezclado  $I/n$ .

- La entropía es invariante ante transformaciones unitarias de cambio de base, ya que depende de los eigenvalores de  $\rho$ :

$$S(U\rho U^{-1}) = S(\rho). \quad (\text{IV.57})$$

- La entropía es una función cóncava. Es decir, sean  $p_i$  probabilidades correspondientes a los operadores de densidad  $\rho_i$ , la entropía satisface que

$$S\left(\sum_i p_i \rho_i\right) \geq \sum_i p_i S(\rho_i). \quad (\text{IV.58})$$

- Si el sistema compuesto  $AB$  está en un estado puro, entonces  $S(A) = S(B)$ .
- **Teorema de entropía conjunta:** Sean  $p_i$  probabilidades;  $|i\rangle$ , estados ortogonales para un sistema  $A$ ; y  $\rho_i$ , un conjunto cualquiera de operadores de densidad para un sistema  $B$ . Entonces

$$S\left(\sum_i p_i |i\rangle\langle i| \otimes \rho_i\right) = H(p_i) + \sum_i p_i S(\rho_i). \quad (\text{IV.59})$$

- La entropía del producto tensorial de dos matrices de densidad es igual a la suma de la entropía de cada matriz de densidad:

$$S(\rho \otimes \sigma) = S(\rho) + S(\sigma). \quad (\text{IV.60})$$

- Suponga que  $\{P_i\}$  es un conjunto completo de proyectores ortogonales y  $\rho$  un operador de densidad. Entonces la entropía del estado  $\rho' \equiv \sum_i P_i \rho P_i$  del sistema antes de la medición es, como mínimo, tan grande como la entropía original:

$$S(\rho') \geq S(\rho) \quad (\text{IV.61})$$

- La entropía de von Neumann tiene la siguiente propiedad de **subaditividad** (también conocida como desigualdad del triángulo):

$$|S(A) - S(B)| \leq S(A, B) \leq S(A) + S(B). \quad (\text{IV.62})$$

- **Subaditividad fuerte:** Para cualquier trío de sistemas cuánticos,  $A, B, C$ , se cumplen las desigualdades:

$$S(A, B, C) + S(B) \leq S(A, B) + S(B, C) \quad (\text{IV.63})$$

## 2. Entropía relativa

Así mismo, la **entropía relativa** tiene una versión cuántica. Esta cantidad ha sido especialmente estudiada (Audenaert, 2013; Vanslette, 2017a; Vedral, 2002) debido a que posee características útiles. Sean  $\rho$  y  $\sigma$  dos operadores de densidad, la entropía relativa de  $\rho$  a  $\sigma$  es

$$S(\rho||\sigma) \equiv \text{tr}(\rho \log_2 \rho) - \text{tr}(\rho \log_2 \sigma) = -S(\rho) - \text{tr}(\rho \log_2 \sigma) \quad (\text{IV.64})$$

donde la entropía relativa es infinita positiva si el kernel de  $\sigma$  tiene intersección no trivial con el soporte de  $\rho$ . La entropía relativa es conjuntamente convexa. La condición de no-negatividad en la entropía relativa cuántica se conoce como la desigualdad de Klein:

$$S(\rho||\sigma) \geq 0 \quad (\text{IV.65})$$

donde

$$S(\rho||\sigma) = 0 \Leftrightarrow \rho = \sigma. \quad (\text{IV.66})$$

La entropía relativa, que suele ser utilizada para determinar la cercanía de dos estados, debe tratarse con precaución, ya que, por ejemplo, los estados ortogonales (la forma en la que los estados son más diferentes) tienen entropía relativa divergente, pero no todos los estados que tienen entropía relativa divergente son ortogonales. Además, la entropía relativa no es simétrica, por lo que no se puede utilizar como una métrica. Por otra parte, considérese el estado de mezcla máxima  $I/n$ , donde  $I$  es la identidad en un sistema de dimensión  $n$ , y el estado  $\rho$  en el mismo sistema. Entonces se cumple que

$$S(\rho||I/n) = \log_2 n - S(\rho), \quad (\text{IV.67})$$

ya que la traza de una matriz de densidad es igual a uno. Nótese que, por la desigualdad de Klein, se obtiene una cota superior para la entropía de von Neumann

$$\log_2 n \geq S(\rho). \quad (\text{IV.68})$$

La entropía relativa cuántica tiene otra propiedad: **la monotonicidad de la entropía relativa de von Neumann**. Esta es una desigualdad que muestra que la entropía relativa decrece monótonamente bajo operaciones  $\mathcal{N}$  que son completamente positivas y preservan la traza de las matrices de densidad:

$$S(\mathcal{N}(\rho)||\mathcal{N}(\sigma)) \leq S(\rho||\sigma). \quad (\text{IV.69})$$

Además, la monotonicidad también decrece cuando una parte del sistema es ignorada:

$$S(\rho_A||\sigma_A) \leq S(\rho_{AB}||\sigma_{AB}) \quad (\text{IV.70})$$

### 3. Información mutua

La **información mutua** en información cuántica tiene una interpretación similar a su versión clásica, así como también tiene la misma formulación. Se define por

$$S(A : B) \equiv S(A) + S(B) - S(A, B) \quad (\text{IV.71})$$

También se puede reescribir utilizando la entropía condicional:

$$S(A : B) = S(A) - S(A|B) = S(B) - S(B|A); \quad (\text{IV.72})$$

y, además, se puede escribir en términos de la entropía relativa:

$$S(A : B) = S(\rho_{AB}||\rho_A \otimes \rho_B) \quad (\text{IV.73})$$

donde  $\rho_X$  es la matriz de densidad en el sistema  $X$ . Es posible observar que la información mutua cuántica es no-negativa:

$$S(A : B) \geq 0 \quad (\text{IV.74})$$

donde se cumple que

$$S(A : B) = 0 \Leftrightarrow \rho_{AB} = \rho_A \otimes \rho_B \quad (\text{IV.75})$$

En un trío de sistemas cuánticos,  $A, B, C$ , la información mutua obedece la siguiente desigualdad

$$S(A : B) + S(A : C) \leq 2S(A) \quad (\text{IV.76})$$

que se puede interpretar de forma que la suma de las informaciones mutuas entre una parte del sistema con cada una de las otras partes no puede ser mayor que dos veces la entropía del primer sistema. Por otra parte, también se sabe que descartar un sistema de un sistema cuántico compuesto no incrementa la información mutua:

$$S(A : B) \leq S(A : B, C). \quad (\text{IV.77})$$

Supóngase el sistema cuántico compuesto  $AB$  y sea  $\varepsilon$  una operación cuántica que preserva la traza en el sistema  $B$ . Sea  $S(A : B)$  la información mutua entre los sistemas antes de aplicar  $\varepsilon$  al sistema  $B$ , y  $S(A' : B')$  la información mutua después de que  $\varepsilon$  es aplicado al sistema  $B$ . Entonces

$$S(A' : B') \leq S(A : B), \quad (\text{IV.78})$$

es decir, las operaciones cuánticas locales no incrementan la información mutua en sistemas separados.

## E. Correlaciones clásicas y cuánticas

### 1. Entrelazamiento cuántico

Matemáticamente, se dice que dos sistemas cuánticos están entrelazados si su estado compuesto no se puede factorizar como el producto tensorial de los estados individuales, es decir, los estados de cada subsistema no se pueden especificar sin el otro (Nielsen y Chuang, 2011). Por ejemplo, considérese el estado de dos qubits

$$\rho = \frac{1}{2}(|00\rangle + |11\rangle)(\langle 00| + \langle 11|) = \frac{1}{2}(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11| + |11\rangle\langle 00|).$$

Se podrá notar que no hay qubits individuales  $\rho_a, \rho_b$  tales que se pueda escribir  $\rho$  de la forma  $\rho = \rho_a \otimes \rho_b$ , por lo tanto, están entrelazados. Esta propiedad de los sistemas cuánticos compuestos se obtiene cuando han interactuado entre ellos y se mantiene a pesar de la separación física de los sistemas. Que dos sistemas estén entrelazados implica que las mediciones que se realicen en uno de los sistemas, afecta directa e instantáneamente al otro sistema (Nielsen y Chuang, 2011). En Información Cuántica, el entrelazamiento es considerado un recurso, el cual se puede cuantificar, y se vincula con el grado de correlación entre los sistemas. Además de ser una de las características de los sistemas cuánticos, resulta ser una de las cualidades que brindan eficiencia a la computación cuántica y seguridad a la criptografía cuántica (Wilde, 2013).

### 2. *Discord* cuántico

En la sección de entropía, hemos observado la “cuantización” (es decir, cambiar los objetos clásicos por los objetos cuánticos:  $H \rightarrow S$  y  $P(X) \rightarrow \rho \in \mathcal{D}_X$ ) de dicha cantidad, tanto para un sistema simple como para un sistema compuesto. Nótese lo siguiente: la entropía condicional clásica se definió como

$$H(X|Y) = - \sum_y p(y) H(X|Y = y), \quad (\text{IV.79})$$

donde  $H(X|Y = y)$  es la entropía de la distribución de probabilidad condicional  $p(X|Y = y)$ . Pero también se definió como

$$H(X|Y) = H(X, Y) - H(Y). \quad (\text{IV.80})$$

Clásicamente, estas dos ecuaciones son equivalentes y de ellas se obtienen resultados iguales. También se ha realizado la observación de que la versión cuántica de la entropía condicional puede tener valores negativos, lo cual es anti-intuitivo y difícil de interpretar. Por otra parte, cuando se busca cuantizar la primera, se tiene un problema: se encontró que no existe una matriz de densidad equivalente a una distribución de probabilidad condicional, lo cual impide realizar el procedimiento deseado. Es por esto que varios autores han propuesto una forma diferente de adaptar la primera definición de la entropía condicional a la teoría cuántica. Esta sección se basa en los trabajos de (Modi y cols., 2012; Henderson y Vedral, 2001; Ollivier y Zurek, 2001).

Supóngase el sistema bipartido  $AB$  y el conjunto de operadores de medición positivo-valuados en  $A$  (POVM, en inglés)  $\{E_a\}$ , donde  $E_a = M_a^\dagger M_a$  y  $M_a$  son operadores de medición con resultado clásico  $a$ . El estado inicial  $\rho_{AB}$  es transformado por la medición:

$$\rho_{AB} \rightarrow \rho'_{AB} = \sum_a M_a \rho_{AB} M_a^\dagger \quad (\text{IV.81})$$

donde el sistema  $A$  observa el resultado  $a$  con probabilidad  $p(a) = \text{tr}(E_a \rho_{AB})$  y  $B$  tiene el estado condicional  $\rho_{B|a} = \text{tr}_A(E_a \rho_{AB})/p(a)$ . Entonces se define la **entropía condicional clásico-cuántica** como

$$S(B|\{E_a\}) \equiv \sum_a p(a) S(\rho_{B|a}) \quad (\text{IV.82})$$

Finalmente, la **entropía condicional cuántica medida** se define escogiendo los operadores  $E_a$  que minimizan la entropía:

$$S_q(B|A) = \min_{E_a} S(B|\{E_a\}) = \min_{E_a} \left( \sum_a p(a) S(\rho_{B|a}) \right). \quad (\text{IV.83})$$

Ahora que se presentan dos expresiones para la entropía condicional, que no son equivalentes en el mundo cuántico, es posible notar que la información mutua también tiene dos expresiones diferentes al momento de cuantizarla. En efecto, ambas expresiones miden diferentes cantidades, por lo que se denotará

$$I(A : B) = S(A) + S(B) - S(AB) \quad (\text{IV.84})$$

como la información mutua cuántica y mide la cantidad total de correlaciones de  $\rho_{AB}$ . Alternativamente, se escribirá la **información mutua lateral** como

$$J(B|A) = S(B) - S_q(B|A) = S(B) - \min_{E_a} \left( \sum_a p(a) S(\rho_{B|a}) \right) \quad (\text{IV.85})$$

como el cuantificador de correlaciones clásicas de  $\rho_{AB}$ . Nótese que  $J(B|A)$  es un valor máximo, ya que a la entropía de  $B$  se le está restando el mínimo de todos los posibles  $S(B|\{E_a\})$ . Es natural entonces pensar en la diferencia entre la cantidad total de correlaciones y la cantidad de correlaciones clásicas de un sistema, dado que son valores que no siempre van a ser iguales. Efectivamente, esta diferencia se conoce como **Discord cuántico** y se define como

$$\delta(B|A) \equiv I(A : B) - J(B|A) \quad (\text{IV.86})$$

$$= S(A) - S(AB) + \min_{E_a} \left( \sum_a p(a) S(\rho_{B|a}) \right). \quad (\text{IV.87})$$

Nótese que esta cantidad es también la resta de las versiones de entropía condicional que clásicamente son equivalentes:

$$\delta(B|A) = S_q(B|A) - S(B|A) \quad (\text{IV.88})$$

Esta cantidad puede ser interpretada como la cantidad de “cuanticidad” que hay en las correlaciones el sistema: si  $I(A : B)$  es el total de correlaciones, incluidas las correlaciones cuánticas, y  $J(B|A)$  es la cantidad de correlaciones clásicas, entonces la resta, o sea el *discord*, es la cantidad de correlaciones cuánticas. El *discord* cuántico tiene las siguientes propiedades:

- No es simétrico:  $\delta(B|A) \neq \delta(A|B)$ .
- Es positivo semi-definido ( $\delta \geq 0$ ), con igualdad si y solo si existe una medición PVM  $\{\Pi_j^B\}$  tal que

$$\rho_{AB} = \sum_j \Pi_j^B \rho_{AB} \Pi_j^B. \quad (\text{IV.89})$$



Esto implica que

$$I(A : B) \geq \text{máx}\{J(B|A), J(A|B)\}. \quad (\text{IV.90})$$

- Es invariante bajo transformaciones unitarias locales.
- Es menor que las entropías marginales de cada subsistema

$$\delta(B|A) \leq \text{mín}\{S(A), S(B)\}. \quad (\text{IV.91})$$

Hay una situación que envuelve al *discord* y se debe a la asimetría de este: el problema de utilizar esta cantidad es que un sistema compuesto debería de mostrar la misma “cuantidad” en sus correlaciones independientemente del subsistema sobre el cual se mida. Con el fin de encontrar una cantidad similar, pero de mejor interpretación, Loura (Loura, s.f.) y otros autores (Rulli y Sarandy, 2011; Ollivier y Zurek, 2001; Henderson y Vedral, 2001) han propuesto el *discord* global o *discord* simétrico.

### 3. *Discord* simétrico

Dado el estado  $\rho_{AB}$  del sistema compuesto  $AB$ , y los conjuntos de mediciones PVM  $\{\Pi_i^A\}$  y  $\{\Pi_j^B\}$  en los subsistemas  $A$  y  $B$ , respectivamente. Se define la **información mutua bilateral** como

$$J^\Delta(A : B) = \text{máx}_{\Pi_i^A \otimes \Pi_j^B} \{H(\{p_{ij}\}) - H(\{p_i\}) - H(\{p_j\})\}, \quad (\text{IV.92})$$

donde  $\{p_{ij}\}$ ,  $\{p_i\}$  y  $\{p_j\}$  son distribuciones categóricas inducidas para la medición conjunta, la medición en  $A$ , y la medición en  $B$ , respectivamente. Esta cantidad está maximizada respecto a todas las posibles mediciones, ya que se busca tener un valor invariante. Luego, se define el ***discord* simétrico**, o ***discord* global**, como

$$\delta^\Delta(A : B) = I(A : B) - J^\Delta(A : B). \quad (\text{IV.93})$$

Algunas de las ventajas inmediatas de esta cantidad son la simetría y la facilidad de generalización a sistemas multipartidos. A continuación se presentan algunos resultados obtenidos por R. Loura en (Loura, s.f.), en forma de propiedades y relaciones con el *discord* usual:

- Sea  $\rho_{AB}$  un estado bipartido cualquiera. Entonces

$$\delta^\Delta(A : B) \geq \delta(A|B). \quad (\text{IV.94})$$

- Dado un estado bipartido  $\rho_{AB}$ , se da la igualdad

$$\delta^\Delta(A : B) = \delta(A|B) \quad (\text{IV.95})$$

si y solo si existe alguna medición PVM  $\{\Pi_j^B\}$  y alguna medición PVM  $\{\Pi_i^A\}$ , que maximizan tanto  $J(A|B)$  como  $J^\Delta(A : B)$ . Esto implica el criterio de conmutatividad adicional:

$$\rho_{A|B=j} \rho_{A|B=j'} = \rho_{A|B=j'} \rho_{A|B=j} \quad (\text{IV.96})$$

para todos  $j, j'$ .

- Sea  $\rho_{AB}$  un estado bipartido. Entonces las siguientes proposiciones son equivalentes:

1.  $\delta^\Delta(A : B) = 0$ .
2.  $\delta(A|B) = \delta(B|A) = 0$ .

3. Existen PVM's  $\{\Pi_i^A\}$  y  $\{\Pi_j^B\}$  en los sistemas  $A$  y  $B$ , respectivamente, tales que

$$\rho_{AB} = \sum_{i,j} (\Pi_i^A \otimes \Pi_j^B) \rho_{AB} (\Pi_i^A \otimes \Pi_j^B). \quad (\text{IV.97})$$

4. El estado  $\rho_{AB}$  es diagonalizable por una matriz unitaria de la forma  $U_A \otimes U_B$ , donde  $U_A$  y  $U_B$  son operadores unitarios en los espacios de Hilbert de los sistemas  $A$  y  $B$ , respectivamente.

R. Loura también muestra que, para que un sistema sea clásico según el *discord* simétrico, la cantidad que importa es  $H(\{p_{ij}\})$ . Consecuentemente, define una nueva cantidad: el **discord reducido**. Dado el sistema bipartido  $AB$ , en el estado  $\rho_{AB}$ , el *discord* reducido es

$$\delta^*(AB) = \min_{\{\Pi_i^A\}, \{\Pi_j^B\}} H(\{p_{ij}\}) - S(AB) \quad (\text{IV.98})$$

donde  $\{\Pi_i^A\}$  y  $\{\Pi_j^B\}$  son PVM's en cada uno de los subsistemas separados  $A$  y  $B$ , respectivamente, y  $\{p_{ij}\}$  la distribución categórica inducida por la PVM  $\{\Pi_i^A \otimes \Pi_j^B\}$  en el sistema completo. Además, demuestra la siguiente relación:

$$\delta^*(A : B) \geq \delta^\Delta(A : B) \quad (\text{IV.99})$$

donde  $\delta^*(A : B) = 0$  ssi  $\delta^\Delta(A : B) = 0$ .

## F. Método de máxima entropía

La inferencia bayesiana es un método de inferencia estadística que permite actualizar la distribución de probabilidad a partir de que se obtiene evidencia o información acerca de la distribución. Es una técnica estadística de gran utilidad debido a su amplio rango de aplicaciones. El objetivo de este método es, dado un conjunto de proposiciones  $x \in X$  con distribución de probabilidad *a priori*  $\phi(x)$ , tanto determinar si es posible actualizar la distribución de probabilidad como encontrar la distribución de probabilidad posterior  $\rho(x)$  a la que debemos actualizar  $\phi(x)$ . Para ello, la información se define operacionalmente como la razón que causa que una distribución de probabilidad cambie. La inferencia inductiva está sustentada por el siguiente principio:

**Principio de Actualización Mínima (PAM):** *Una distribución de probabilidad solo debe actualizarse en la medida requerida por la nueva información.*

Esto indica que si se tiene nueva información sobre la variable, la distribución de probabilidad debe permanecer igual. Ahora, es necesaria una herramienta que permita determinar si la nueva información puede actualizar la distribución previa a una nueva. En (Vanslette, 2017a), Vanslette propone dos criterios que debe tener este funcional, basándose en el PAM:

- DC1.** *Independencia de subdominios:* Cuando se recibe información sobre un conjunto de proposiciones  $x$ , no debe afectar o cambiar el estado de conocimiento de las otras proposiciones. De lo contrario, significa que la información recibida también era sobre esas otras proposiciones.
- DC2.** *Independencia de subsistemas:* Cuando se cree *a priori* que dos sistemas son independientes y sólo se recibe información sobre uno de ellos, entonces el estado de conocimiento del otro sistema permanece sin cambios.

Es importante notar que estos criterios instruyen en qué casos **no** actualizar, más que indicar cómo actualizar. Luego, Vanslette muestra que, tanto para el caso clásico como para el cuántico, la entropía relativa es el funcional que cumple con los criterios y lleva a inferencias demostrablemente

correctas, pero también menciona que la justificación para escogerla es que los demás candidatos fallan demostrablemente a desempeñarse como se desea (Vanslette, 2017a). La entropía relativa resultante para el caso clásico, aplicando condiciones de normalización, es

$$S(\rho(x)||\phi(x)) = -|A| \int \rho(x) \ln \left( \frac{\rho(x)}{\phi(x)} \right) dx \quad (\text{IV.100})$$

donde  $|A|$  es una constante positiva arbitraria. Ahora que se tiene una forma de comparar la distribución previa con las distribuciones candidatas a ser la posterior, el método de máxima entropía (que fue creado por E.T. Jaynes) consiste en comparar cada posible  $\rho(x)$  con  $\phi(x)$  a través de la entropía relativa y escoger la distribución que maximiza la entropía, respecto a un conjunto de restricciones de valores esperados  $\langle A_i \rangle = \int \rho(x) A_i(x) dx$  y bajo normalización. En (Vanslette, 2017a) y (Vanslette, 2017b), Vanslette encuentra la solución general utilizando el método de multiplicadores de Lagrange:

$$\rho(x) = \frac{\phi(x)}{Z} \exp \left( \sum_i \alpha_i A_i(x) \right) \quad (\text{IV.101})$$

donde  $\alpha_i$  son los multiplicadores de Lagrange de cada observable  $A_i(x)$  y  $Z$  es la constante de normalización. Vanslette resuelve los multiplicadores de Lagrange evaluando el valor esperado correspondiente, utilizando los métodos estándar de Mecánica Estadística  $\langle A_i \rangle = \frac{\partial}{\partial \alpha_i} \ln(Z)$ . Además, se muestra que este método es coherente con la regla de Bayes.

En el caso cuántico, se busca actualizar la matriz de densidad previa  $\hat{\phi}$  a una matriz de densidad posterior  $\hat{\rho}$ , mediante un proceso similar. En (Vanslette, 2017a), la entropía relativa cuántica, bajo normalización, es

$$S(\hat{\rho}||\hat{\phi}) = -\text{Tr} \left( \hat{\rho} \log \hat{\rho} - \hat{\rho} \log \hat{\phi} \right). \quad (\text{IV.102})$$

Al maximizar la entropía relativa respecto a un conjunto de valores esperados  $\langle \hat{A}_i \rangle$  de los operadores hermíticos  $\hat{A}_i$ , *i.e.*  $\text{Tr}(\hat{A}_i \hat{\rho}) = \langle \hat{A}_i \rangle$ , y bajo normalización, se obtiene la forma general de la densidad posterior:

$$\hat{\rho} = \frac{1}{Z} \exp \left( \sum_i \alpha_i \hat{A}_i + \ln \hat{\phi} \right), \quad (\text{IV.103})$$

donde  $Z$  es una constante de normalización y  $\alpha_i$  son los multiplicadores de Lagrange (Vanslette, 2017b). Los multiplicadores también son encontrados utilizando los métodos de Mecánica Estadística,

$$\langle \hat{A}_i \rangle = \frac{\partial}{\partial \alpha_i} \ln(Z) \quad (\text{IV.104})$$

e invirtiendo para encontrar  $\alpha_i = \alpha_i(\langle \hat{A}_i \rangle)$ . La constante de normalización  $Z$  se encuentra al diagonalizar la matriz  $\hat{C} = \sum_i \alpha_i \hat{A}_i + \ln \hat{\phi}$  de forma que  $\hat{C} \rightarrow \hat{\Lambda} = \sum_\lambda \lambda |\lambda\rangle\langle\lambda|$ , entonces

$$Z = \text{Tr} \left( \exp(\hat{C}) \right) = \text{Tr} \left( \exp(\hat{\Lambda}) \right) = \sum_\lambda e^\lambda. \quad (\text{IV.105})$$

Hay matrices de densidad que no se pueden actualizar por este método, al igual que hay distribuciones de probabilidad que no se pueden actualizar: las sesgadas. Una matriz de densidad previa que no genera todo el espacio de Hilbert es llamada sesgada de orden  $M$ . Vanslette establece en (Vanslette, 2017b) el siguiente teorema:

**Teorema 3.** *Una matriz de densidad previa que es sesgada de orden  $M$  sólo puede ser actualizada inferencialmente en el eigenspacio que esta genera. Las regiones no generadas por dicha matriz de densidad permanecen en cero.*

Este teorema indica que la actualización entrópica solo puede causar cambios inferenciales y epistémicos a  $\hat{\rho}$ . Estas matrices solo pueden cambiar si son rotadas físicamente aplicando operadores unitarios.

## G. Estructuras de correlación

### 1. En sistemas clásicos

En un sistema clásico de  $n$  variables aleatorias,  $S = \{X_i\}_{i=1}^n$ , se define un subsistema  $R$  de  $S$  como un subconjunto  $R \subseteq S$  tal que las distribuciones de probabilidad marginal asociadas son

$$P_R(x_{k_1}, x_{k_2}, \dots, x_{k_m}) = \sum_{x_i \notin R} P_S(x_1, x_2, \dots, x_n), \quad (\text{IV.106})$$

donde  $R = \{X_{k_i}\}_{i=1}^m$ . A cualquier conjunto de subsistemas de  $S$  se le denota como  $\Phi = \{R_i\}_{i=1}^u$ .

Una **estructura de correlación**  $\{\Phi, P_\Phi\}$  asociada a  $S$  es un conjunto de subsistemas  $\Phi = \{R_i\}_{i=1}^u$  y una distribución de probabilidad asociada  $P_\Phi$  tal que:

1.  $\Phi$  es una cubierta de  $S$ , *i.e.*, para todo  $X_i \in S$ , existe  $R_j \in \Phi$  tal que  $X_i \in R_j$ .
2.  $P_\Phi$  es la solución al problema de optimización sobre el conjunto  $\mathcal{P}$  de distribuciones para  $S$  dado por

$$P_\Phi = \arg \max_{P \in \mathcal{P}} \{H(P)\} \quad (\text{IV.107})$$

y sujeto a las restricciones

$$\sum_{x_i \notin R_j} P_\Phi(x_1, x_2, \dots, x_n) = \sum_{x_i \notin R_j} P_S(x_1, x_2, \dots, x_n), \forall R_j \in \Phi \quad (\text{IV.108})$$

**Propiedad:** Dada una cubierta  $\Phi$ ,  $P_\Phi$  es única.

Defínase  $Red(\Phi)$  como el conjunto de todos los subsistemas  $R_i \in \Phi$  para los cuales existe  $R_j \in \Phi$ , tales que  $R_i \subset R_j$ . Se dice que dos relaciones estructurales  $\Phi_1, \Phi_2$  en  $S$  son equivalentes en correlaciones multivariantes si

$$\Phi_1 - Red(\Phi_1) = \Phi_2 - Red(\Phi_2). \quad (\text{IV.109})$$

En efecto, esta es una relación de equivalencia, ya que la distribución asociada de máxima entropía es la misma para todas las estructuras equivalentes según (IV.109). Entonces, para cualquier  $\Phi$ , se puede obtener la clase de equivalencia

$$Class(\Phi) = \Phi - Red(\Phi), \quad (\text{IV.110})$$

donde  $\Phi - Red(\Phi) = \{R_1, R_2, \dots, R_k\}$  es la menor estructura de correlación equivalente a  $\Phi$  (Mariano Lemus y Paunković, s.f.).

### 2. En sistemas cuánticos

Sea  $S = \{\mathcal{H}_{x_i}\}_{i=1}^n$  tal que  $\mathcal{H}_S = \otimes_i \mathcal{H}_{x_i}$  son espacios de Hilbert. Denotamos el conjunto de operadores de densidad que actúan sobre  $\mathcal{H}_S$  como  $\mathcal{D}_S$ . Un sistema cuántico multipartido  $\{S, \rho_S\}$  es un conjunto  $S$  de espacios de Hilbert junto a un operador de densidad  $\rho_S \in \mathcal{D}_S$ . Además, por simpleza, se hará referencia a cada espacio de Hilbert  $\mathcal{H}_{X_i}$  como  $X_i$  y se utilizará la notación de sistemas clásicos.

Una estructura de correlación  $\{\Phi, \rho_\Phi\}$  asociada a un sistema cuántico multipartido  $S$  es un conjunto de subsistemas  $\Phi = \{R_j\}$  y un operador de densidad  $\rho_\Phi$  tal que:

1.  $\Phi$  es una cubierta de  $S$ .
2.  $\rho_\Phi$  es la solución al problema de optimización sobre el conjunto  $\mathcal{D}_S$  dado por

$$\rho_\Phi = \arg \max_{\rho \in \mathcal{D}_S} \{S(\rho)\}, \quad (\text{IV.111})$$

sujeto a las restricciones

$$\text{Tr}_{\bar{R}_j}(\rho_\Phi) = \text{Tr}_{\bar{R}_j}(\rho_S). \quad (\text{IV.112})$$

En sistemas cuánticos también se da la propiedad mencionada anteriormente (cambiando  $P_\Phi$  por  $\rho_\Phi$ ) y además se conservan la relación de equivalencia y las clases de equivalencia, debido a que la matriz de densidad de máxima entropía asociada es la misma para todas las estructuras equivalentes (Mariano Lemus y Paunković, s.f.; Zhou, 2009).

## H. Correlaciones en sistemas cuánticos multipartidos

### 1. Estructuras de correlación y correlaciones multipartidas

Para entender mejor el rol de estructuras de correlación en el contexto de estudiar la manera en que los diferentes subsistemas de un sistema cuántico están correlacionados, considere las correlaciones totales

$$I(\rho_S) = \sum_i S(\rho_i) - S(\rho_S). \quad (\text{IV.113})$$

Nótese que las matrices de densidad asociadas a las diferentes estructuras de un sistema  $S$  son matrices de densidad que actúan sobre el espacio de Hilbert  $\mathcal{H}_S = \bigotimes_i^n \mathcal{H}_i$  y están definidas de manera que mantienen todas las correlaciones dentro de los subsistemas que componen la estructura pero no más de eso. Considere ahora la diferencia entre la entropía de  $\rho_S$  y la de  $\rho_\Phi$  asociada a una estructura dada  $\Phi$ :

$$S(\rho_\Phi) - S(\rho_S) = \left( S(\rho_\Phi) - \sum_{i=1}^n S(\rho_i) \right) - \left( S(\rho_S) - \sum_{i=1}^n S(\rho_i) \right) = I(\rho_S) - I(\rho_\Phi). \quad (\text{IV.114})$$

Esta cantidad es siempre no-negativa debido a que  $\rho_\Phi$  tiene máxima entropía. Además, si  $S(\rho_\Phi) - S(\rho_S) \neq 0$  entonces existen correlaciones en  $S$  que no están en  $\rho_\Phi$ . En conclusión, maximizar la entropía de  $\rho_S$  (IV.111) garantiza que  $\rho_\Phi$  sea la matriz de densidad menos correlacionada que todavía mantiene las correlaciones dentro de sus subsistemas (Mariano Lemus y Paunković, s.f.). Esto motiva a estudiar el comportamiento de las correlaciones que permanecen en estructuras que contienen todos los subsistemas de  $k$  variables.

### 2. Correlaciones irreducibles

Sea  $\{S, \rho_S\}$  un sistema cuántico  $n$ -partido y denótese  $\Phi^{(k)}$ , con  $1 < k \leq n$ , la estructura de correlación compuesta por todos los subsistemas de  $S$  que tienen exactamente  $k$  elementos. Las correlaciones irreducibles  $k$ -partidas en  $S$  se definen como

$$I_{irr}^{(k)} = I(\rho_{\Phi^{(k)}}) - I(\rho_{\Phi^{(k-1)}}) = S(\rho_{\Phi^{(k-1)}}) - S(\rho_{\Phi^{(k)}}). \quad (\text{IV.115})$$

Además, la suma de todas las correlaciones irreducibles de un sistema es igual a las correlaciones totales del sistema. Las correlaciones irreducibles de orden  $k$  muestran las correlaciones intrínsecas que existen entre todos los subsistemas de  $k$  variables, y que existen por la naturaleza del sistema físico (Kato, Furrer, y Murao, 2016; Liu, Zeng, y Zhou, 2016).

### 3. Correlaciones clásicas y cuánticas

Para los propósitos de este trabajo, se realizará una generalización del *discord* simétrico, cantidad que se discutió en la sección 3. Sea  $\{S = \{X_i\}_{i=1}^n, \rho_S\}$  un sistema multipartido y sea  $LM(S)$  el conjunto de operadores locales de medición en  $\mathcal{H}_S$ :

$$LM(S) = \{\Pi = \{\bigotimes_{i=1}^n \Pi_{X_i}^{k_i}\}_{k_1, k_2, \dots, k_n} \mid \{\Pi_{X_i}^{k_i}\}_{k_i} \text{ es un conjunto completo de proyectores para } X_i\}. \quad (\text{IV.116})$$

La correlaciones clásicas totales en  $S$  se definen como

$$C(\rho_S) = \max_{\{\Pi \in LM(S)\}} \left\{ \sum_{i=1}^n S(\rho_{\{\Pi_i^{k_i}\}}) - S(\rho_\Pi) \right\} \quad (\text{IV.117})$$

y las correlaciones cuánticas en  $S$  son

$$Q(\rho_S) = I(\rho_S) - C(\rho_S). \quad (\text{IV.118})$$

Inicialmente, en la Teoría de Información Clásica, se han construido estructuras jerárquicas que han permitido el estudio de las correlaciones en variables aleatorias con distribuciones de probabilidad específicas. S. Amari (Amari, 2001) utilizó Geometría de la Información para describir el comportamiento de las jerarquías de las distribuciones de probabilidad. Esto lo hizo a través de descomposiciones ortogonales de los sistemas que estudiaba (familias exponenciales o familias mezclas) y logra describir las dependencias entre varias componentes del sistema a través de ‘dependencias estocásticas’.

Desde el punto de estudio cuántico, uno de los resultados más interesantes es el de N. Linden, S. Popescu y W.K. Wootters (Linden y cols., 2002). Ellos introducen una caracterización de las correlaciones irreducibles  $n$ -partidas, basándose en la medición de información en el estado cuántico de  $n$  partes que no está incluida en el conjunto de estados reducidos de  $n-1$  partes. En este artículo, se concluye que casi todos los estados puros de tres qubits son determinados por sus matrices reducidas de dos componentes y que los estados que no tienen esta propiedad son equivalentes bajo rotaciones locales a estados de la forma  $a|000\rangle + b|111\rangle$ .

Por otra parte, D.L. Zhou (Zhou, 2008) realizó avances interesantes en el estudio de las correlaciones irreducibles  $k$ -partidas en sistemas de  $n$  qubits, siguiendo el método que Linden, Popescu y Wootters propusieron y formulando una generalización. En este se denota la importancia de la forma exponencial de una matriz de densidad para caracterizar las correlaciones irreducibles multipartidas, ya que permite acceder a estados maximales con mayor facilidad. Sin embargo, cuando Zhou (Zhou, 2009) implementa una versión cuántica del trabajo de Amari (estructuras de correlación) con el objetivo de formular cantidades equivalentes a las discutidas anteriormente, se encuentra con un resultado contra-intuitivo: se pueden crear correlaciones irreducibles tri-partidas a partir de un estado tri-partido que sólo tiene correlaciones irreducibles bipartidas. Las implicaciones de esta conclusión no son deseables para la medición de las correlaciones, por lo que sirven de motivación para proponer una forma diferente de medir las correlaciones irreducibles, tomando de referencia el desarrollo logrado en cuanto a las estructuras de correlación.

---

Descripción del problema: Correlaciones cuánticas  $k$ -partidas

---

Con el objetivo de estudiar cómo se distribuyen las correlaciones en un sistema multipartido, (Mariano Lemus y Paunković, s.f.) propone las siguientes definiciones:

Dado un sistema cuántico multipartido  $\{S, \rho_S\}$  de  $n$  variables y las estructuras de correlación  $\{\Phi^{(k)}, \rho_{\phi^{(k)}}\}$ , se escriben las correlaciones clásicas  $k$ -partidas como

$$I_c^{(k)} = I(\rho_{\phi^{(k)}}^c) - I(\rho_{\phi^{(k-1)}}^c), \quad (\text{VI.1})$$

donde  $\rho_{\phi^{(k)}}^c$  es la matriz post medición resultante de aplicar las mediciones locales que maximizan sus correlaciones a  $\rho_{\phi^{(k)}}$ .

**Propiedad:** La suma de las correlaciones clásicas  $k$ -partidas es igual a las correlaciones clásicas de  $\rho_S$ :

$$\sum_{k=2}^n I_c^{(k)} = C(\rho_S). \quad (\text{VI.2})$$

Luego, se definen las correlaciones cuánticas  $k$ -partidas como la diferencia entre de las correlaciones irreducibles de orden  $k$  y las correlaciones clásicas  $k$ -partidas:

$$I_Q^{(k)} = I_{irr}^{(k)} - I_c^{(k)}. \quad (\text{VI.3})$$

**Propiedad:** La suma de las correlaciones cuánticas  $k$ -partidas es igual a las correlaciones cuánticas de  $\rho_S$ :

$$\sum_{k=2}^n I_Q^{(k)} = Q(\rho_S). \quad (\text{VI.4})$$

Es deseable que estas cantidades cumplan con ciertas propiedades y que estas sean demostrables. Como son cantidades que se están proponiendo, es necesario tener una noción de su comportamiento en ciertas situaciones, de forma que sugieran el camino a las demostraciones correspondientes. Se formulan a continuación las siguientes hipótesis:



## A. Hipótesis

**Hipótesis 1.** Las correlaciones cuánticas  $k$ -partidas siempre son no-negativas.

**Hipótesis 2.** Los operadores locales que maximizan  $C(\rho_S)$  también maximizan  $C(\rho_{\Phi}^{(k)})$ .

**Hipótesis 3.** Las cantidades  $I(\rho_S)$ ,  $C(\rho_S)$ , y  $Q(\rho_S)$  se conservan ante la aplicación de operadores unitarios locales.

**Hipótesis 4.** Las correlaciones cuánticas  $k$ -partidas no se incrementan ante la aplicación de canales cuánticos locales.

---

Algoritmo 1: selección de matrices de máxima entropía

---

El objetivo de este algoritmo, dado un sistema cuántico multipartido  $\{S, \rho_S\}$  y una cubierta  $\Phi$ , es resolver el problema de optimización dado por la ecuación (IV.111) y las restricciones (IV.112).

## A. Fundamentación teórica

### 1. Caracterización de estados de máxima entropía

Primero, recordamos que una matriz de densidad de  $n$  qubits se puede representar como

$$\rho = \frac{1}{2^n} \left( I + \sum_{\vec{k}} \alpha_{\vec{k}} \sigma_{\vec{k}} \right) \quad (\text{VII.1})$$

donde  $\vec{k} = (k_1, k_2, \dots, k_n)$ ,  $k_i \in \{0, 1, 2, 3\}$ ,  $\sigma_{\vec{k}} = \bigotimes_{i=1}^n \sigma_{k_i}$  son las matrices de Pauli junto a la identidad y todos los  $\alpha_{\vec{k}}$  son coeficientes reales tales que  $\sum_{\vec{k}} \alpha_{\vec{k}}^2 = 1$ . Además, una matriz de densidad también se puede escribir en su forma exponencial:

$$\rho = \exp \left( \sum_{\vec{k}} \theta_{\vec{k}} \sigma_{\vec{k}} \right), \quad (\text{VII.2})$$

donde  $\theta_{\vec{k}}$  son coeficientes reales y  $\theta_{\vec{0}}$  es una constante de normalización. De esta forma es más sencillo escoger los coeficientes que representen un estado específico.

Sea  $TC(\{S, \rho_S\}, \Phi)$  el conjunto de operadores de densidad que cumplen con las Condiciones de Traza (IV.112), aunque no necesariamente cumplen (IV.111), para la cubierta  $\Phi = \{R_j\}$ . El estado de máxima entropía en el conjunto  $TC(\{S, \rho_S\}, \Phi)$  se denotará como  $\rho_{\Phi}$ . El objetivo de este algoritmo será encontrar  $\rho_{\Phi}$ .

Sea  $K = \{0, 1, 2, 3\}^n$  el conjunto de todos los  $\vec{k}$  posibles de  $n$  componentes y defínase la función

$A : K \times \Phi \rightarrow \{0, 1\}$  tal que

$$A(\vec{k}, R) = \begin{cases} 0 & \text{si } \exists i \leq n \text{ tal que } k_i \neq 0 \text{ y } x_i \notin R \\ 1 & \text{en cualquier otro caso} \end{cases} \quad (\text{VII.3})$$

Entonces construimos  $K_\Phi = \{\vec{k} \in K \ni \exists R \in \Phi \text{ y } A(\vec{k}, R) = 1\}$ . Este conjunto fue construido con el objetivo de permitirnos movernos a través de los índices asociados a los subsistemas que aparecen en alguna de las condiciones de traza, ya que los índices muestran qué matriz de Pauli actúa sobre cuál sistema. El resultado de esto es que, al escribir el estado original  $\rho$  y el estado de máxima entropía  $\rho_\Phi$  en su forma de la esfera de Bloch, obtenemos la siguiente igualdad para  $\vec{k} \in K_\Phi$ :

$$\alpha_{\vec{k}}^\Phi = \alpha_{\vec{k}}. \quad (\text{VII.4})$$

El problema de optimización se reduce a encontrar los  $\alpha_{\vec{k}}$  tales que  $\vec{k} \notin K_\Phi$  que maximizan la entropía. El resultado se encuentra utilizando el siguiente teorema:

**Teorema 4.** *Sea  $\rho = \exp(\sum_{\vec{k}} \theta_{\vec{k}} \sigma_{\vec{k}})$  el estado (de rank máximo) en el espacio de Hilbert  $\mathcal{H}_S$  de  $n$ -qubits que cumple con las condiciones de traza de  $\Phi$  para el problema de máxima entropía. Entonces*

$$\theta_{\vec{k}} = 0, \text{ para } \vec{k} \notin K_\Phi. \quad (\text{VII.5})$$

Este teorema resuelve el problema de optimización hasta el punto de que se sabe que los coeficientes  $\theta_{\vec{k}} = 0$  para los  $\vec{k} \notin K_\Phi$  en la forma exponencial de la matriz, y se sabe el valor de los coeficientes  $\alpha_{\vec{k}}$  para el resto de  $\vec{k}$ 's. Entonces para obtener la matriz completa hay que resolver la igualdad entre la matriz exponencial y la matriz en la forma de la esfera de Bloch con los coeficientes que se conocen, obteniendo los coeficientes restantes.

## 2. Gradient descent

El *Gradient descent*, o descenso en gradiente, es un método numérico de optimización iterativa de primer orden que encuentra un mínimo local de una función diferenciable. El método es atribuido a Cauchy, quien fue motivado por cálculos astronómicos, ya que para encontrar la órbita de un objeto masivo necesitaba resolver tanto ecuaciones diferenciales como algebraicas (Lemaréchal, 2012). La idea del método es dar pasos repetidos en dirección opuesta al gradiente de la función, hasta llegar a un mínimo local. Análogamente, avanzar en dirección al gradiente es conocido como *gradient ascent*.

Uno de los problemas de este método es que puede atascarse en valles. Por esta razón se utiliza una extensión del método conocida como **momentum**, por su analogía con el momentum en mecánica. El momentum permite que la búsqueda tenga cierta 'inercia' en el espacio de búsqueda, lo cual sobrepasa las oscilaciones de gradientes ruidosos y permite avanzar en los valles. Además, permite acelerar el proceso de optimización (Kochenderfer y Wheeler, 2019).

Resumidamente, el *gradient descent* consiste en los siguientes pasos:

1. Se inicializan las variables  $\vec{x}_i$  y se calcula  $F(\vec{x}_i)$ .
2. Se calcula el gradiente de la función objetivo  $\nabla F(\vec{x}_i)$ .
3. Se actualizan las variables  $x_{i+1} = \vec{x}_i - \eta \nabla F(\vec{x}_i) + \mu \vec{x}_i$ , donde  $\eta \in (0, 1]$  es la tasa de aprendizaje y  $\mu \in [0, 1]$  es el 'momentum'.
4. Se calcula  $F(x_{i+1})$  y se repite el proceso hasta que  $F(x_{i+1}) - F(\vec{x}_i) < \varepsilon$ , donde  $\varepsilon$  es el error máximo que se desea.

## B. Algoritmo

En este algoritmo se utilizaron los paquetes *NumPy*, *QuTiP*, *SciPy*, *Itertools* y *Random*. El algoritmo recibe como entradas el número  $N$  de qubits del Sistema, la cubierta  $\Phi$  y la matriz de densidad  $\rho$ . El programa interpreta el sistema como la lista de enteros  $S = [0, 1, \dots, N]$  donde cada entero representa un espacio de Hilbert. De igual forma, la cubierta es interpretada como una lista que contiene sublistas de  $S$ . Este algoritmo se puede separar en dos etapas:

### 1. Preparación

Como primer paso, luego de recibir los datos, el programa verifica que la cubierta ingresada sea, en efecto, una cubierta y elimina los subconjuntos redundantes. Después, genera el conjunto  $K$ , donde cada vector  $\vec{k}$  está representado por una lista con las componentes del mismo, y utilizando la función  $A(\vec{k}, R)$  descrita en (VII.3) selecciona el conjunto  $K_\Phi$ .

Por último, el programa determina los coeficientes  $\alpha_{\vec{k}}$  de  $\rho$  en su forma de la esfera de Bloch. A través del isomorfismo que reordena las matrices  $2^N \times 2^N$  en vectores en  $\mathbb{C}^{4^N}$ , transforma el problema  $\rho = \sum_{\vec{k}} \alpha_{\vec{k}} \sigma_{\vec{k}}$  a la forma  $\vec{\rho} = M_\sigma \vec{\alpha}$ , donde  $\vec{\rho}$  es el vector isomorfo a  $\rho$ ,  $M_\sigma$  es una matriz cuyos vectores columna son isomorfos a cada  $\sigma_{\vec{k}}$  y  $\vec{\alpha}$  es el vector con componentes  $\alpha_{\vec{k}}$ . La solución es determinada por medio de mínimos cuadrados de *SciPy*.

### 2. Optimización

Para la optimización, se utiliza el *Gradient Descent* con un máximo de  $10^5$  iteraciones, error máximo de  $10^{-6}$ , tasa de aprendizaje de 0.45 y momentum de 0.5. Los parámetros que se varían son los coeficientes  $\theta_{\vec{k}}$  tales que  $\vec{k} \in K_\Phi$  y se inicializan en cero.

La función objetivo funciona de la siguiente manera: recibe los  $\theta_{\vec{k}}$  y genera una matriz de la forma  $\exp(\sum_{\vec{k}} \theta_{\vec{k}} \sigma_{\vec{k}})$ . Luego, de la misma forma que en la preparación, resuelve para los coeficientes  $\alpha_{\vec{k}}^{it}$  y los guarda en un archivo de texto para su utilización posterior. Por último, selecciona los  $\alpha_{\vec{k}}^{it}$  y los  $\alpha_{\vec{k}}$  de la matriz original relacionados por  $\vec{k} \in K_\Phi$  y encuentra  $\sum_{\vec{k} \in K_\Phi} (\alpha_{\vec{k}}^{it} - \alpha_{\vec{k}})^2$ , el cual es el valor que buscamos minimizar. La función gradiente se definió como  $\delta\theta_{\vec{k}} = 2^{-N} (\alpha_{\vec{k}}^{it} - \alpha_{\vec{k}})$ , de forma que si la función objetivo está cercana al mínimo el gradiente disminuye y viceversa.

### 3. Diagrama de flujo

El presente diagrama de flujo muestra el funcionamiento del algoritmo expuesto en este capítulo.

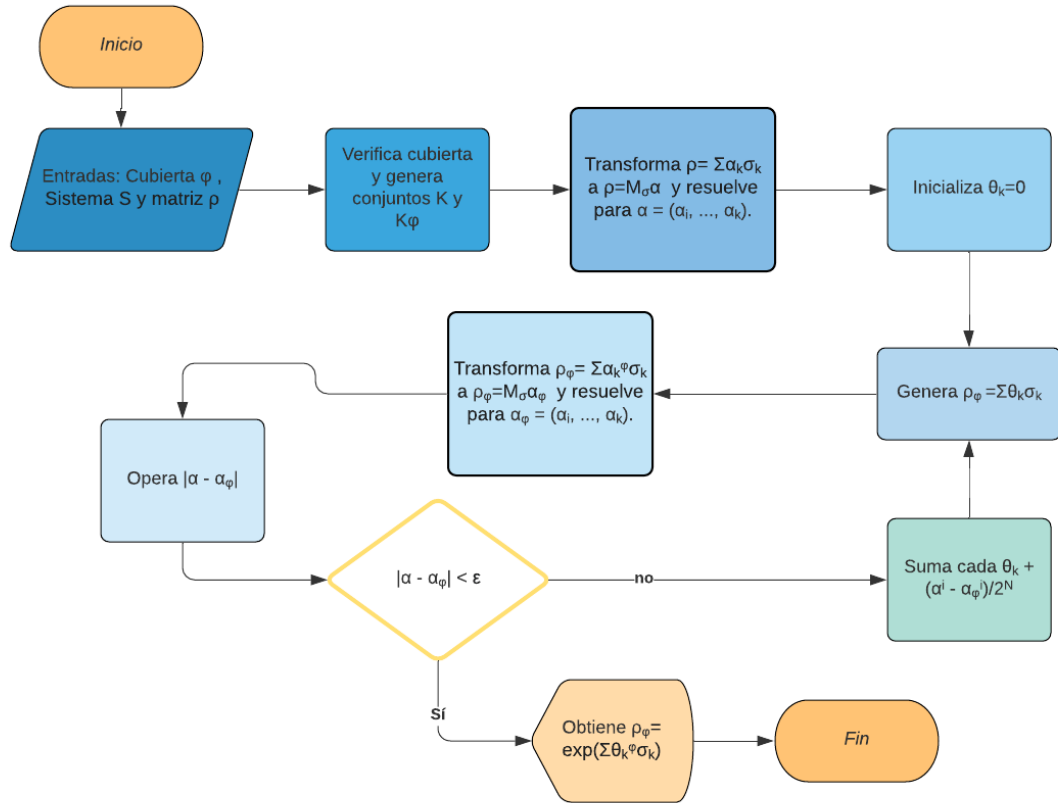


Figura VII.1: Diagrama de flujo de Algoritmo 1.

### C. Eficiencia

Se utilizó la herramienta *iPython Magic* para medir los tiempos de ejecución (en segundos) para sistemas de  $N = 2, 3, 4, 5$  qubits. Se ejecutaron 100 ciclos con matrices de densidad aleatorias para  $N = 2$  y  $N = 3$ , pero 30 ciclos para  $N = 4$  y  $N = 5$ . En la Figura VII.2 se muestra el logaritmo del tiempo de corrida contra el tamaño  $N$  del sistema. Se realizó una regresión por mínimos cuadrados ponderados, la cual muestra que el tiempo aumenta de forma exponencial en función del tamaño del sistema. Nótese que no es necesario medir para sistemas de  $N = 1$  porque ya representan al sistema más simple.

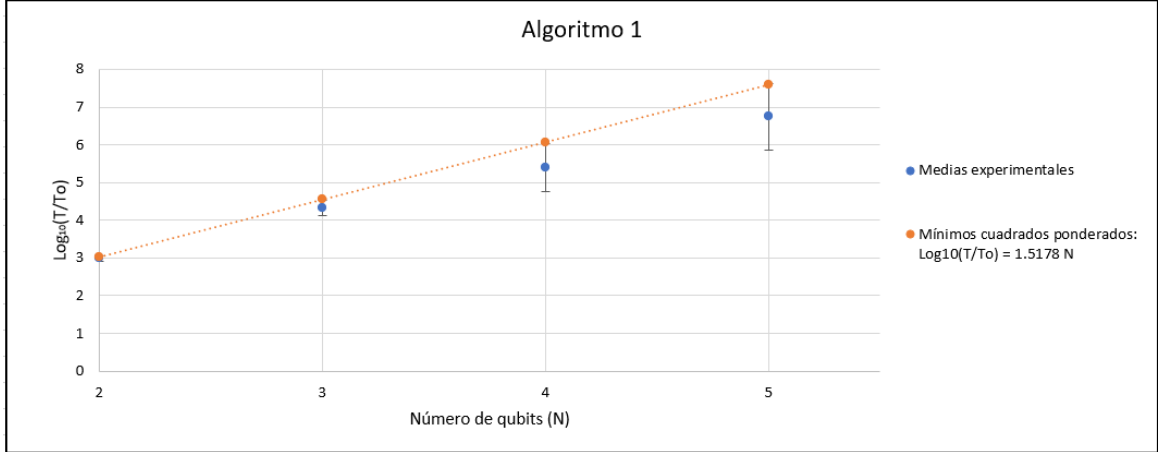


Figura VII.2: Eficiencia del Algoritmo 1.  $T$  es el tiempo, en segundos, que tarda el código en resolver el Algoritmo 1 con un sistema de  $N$  qubits.  $T_0 = 0.0016$  s. Los puntos azules son las medias obtenidas por ejecución y la línea punteada naranja es la regresión por mínimos cuadrados ponderados.

## D. Exploración de estados GHZ y W

Para verificar la precisión del algoritmo, se implementó en los estados GHZ y W de 3 y 4 qubits para obtener las matrices reducidas en las cubiertas  $\Phi^{(N-1)}$  y se les midió la entropía de von Neumann. Obsérvese que las matrices originales tienen entropía cero, mientras que las reducidas tienen entropía mayor. Se observa un error porcentual máximo de 7% entre el valor teórico, el cual fue calculado analíticamente, y el obtenido.

| Estado | Entropía de $\rho_S$ | Entropía de $\rho_\Phi^{(2)}$ |                |
|--------|----------------------|-------------------------------|----------------|
|        |                      | Valor teórico                 | Valor obtenido |
| GHZ    | 0.000                | 1.000                         | 1.027          |
| W      | 0.000                | 0.072                         | 0.077          |

Tabla VII.1: Entropía de las matrices reducidas para  $\Phi^{(2)}$  y  $N = 3$ .

| Estado | Entropía de $\rho_S$ | Entropía de $\rho_\Phi^{(3)}$ |                |
|--------|----------------------|-------------------------------|----------------|
|        |                      | Valor teórico                 | Valor obtenido |
| GHZ    | 0.000                | 1.000                         | 1.140          |
| W      | 0.000                | 0.120                         | 0.161          |

Tabla VII.2: Entropía de las matrices reducidas para  $\Phi^{(3)}$  y  $N = 4$ .

---

Algoritmo 2: selección de operadores locales

---

El objetivo de este algoritmo es seleccionar los operadores locales que resuelvan el problema de optimización dado por la ecuación (IV.117) :

$$C(\rho_S) = \max_{\{\Pi \in LM(S)\}} \left\{ \sum_{i=1}^n S(\rho_{\{\Pi_i^{k_i}\}}) - S(\rho_\Pi) \right\}.$$

## A. Fundamentación teórica

### 1. Variación de operadores locales

Un observable  $M$ , debemos recordar, es un operador hermitiano que se puede escribir por descomposición espectral de la forma  $M = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|$ , donde el conjunto de proyectores  $\{|\lambda_i\rangle\langle\lambda_i|\}$  es completo y cada  $|\lambda_i\rangle$  corresponde a un estado ortonormal del sistema (en el caso de un qubit, cero y uno) (Nielsen y Chuang, 2011). Ese conjunto de proyectores es conocido como un conjunto de operadores locales y se denota por  $\Pi$ . Ahora, bien, para todo espacio de Hilbert en  $\mathcal{C}^2$ , es posible escribir una base de estados ortonormales en función de dos parámetros y la base computacional de la forma:

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right) |0\rangle + e^{-i\phi} \sin\left(\frac{\theta}{2}\right) |1\rangle \quad (\text{VIII.1})$$

$$|\psi^\perp\rangle = \sin\left(\frac{\theta}{2}\right) |0\rangle - e^{-i\phi} \cos\left(\frac{\theta}{2}\right) |1\rangle, \quad (\text{VIII.2})$$

donde  $0 \leq \theta \leq \pi$  y  $0 \leq \phi \leq 2\pi$  son coordenadas en la esfera de Bloch (Nielsen y Chuang, 2011; Cover y Thomas, 2006). Entonces el conjunto de operadores locales de la base en función de  $(\theta, \phi)$  es  $\Pi = \{\Pi^0 = |\psi\rangle\langle\psi|, \Pi^1 = |\psi^\perp\rangle\langle\psi^\perp|\}$ . Nótese que cuando  $\theta = 0$  se obtiene la base computacional, independientemente del valor de  $\phi$ . En un sistema cuántico  $S$  de  $n$  variables, por cada variable  $X_i$  se tiene un conjunto de operadores locales  $\Pi_{X_i} = \{\Pi_{X_i}^0, \Pi_{X_i}^1\}$  que depende de un par de parámetros  $(\theta_i, \phi_i)$ . La variación de estos parámetros nos permitirá el movimiento en los operadores locales de  $LM(S)$ .

## 2. *Differential evolution*

El algoritmo de Evolución Diferencial es un algoritmo de optimización estocástico basado en población, el cual ha sido introducido recientemente (Storn y Price, 1997). Algunas ventajas de este algoritmo de optimización son: simplicidad y eficiencia de código, facilidad de uso y rapidez de convergencia.

El *Differential Evolution* trabaja con dos poblaciones: nueva y vieja generación de la misma población. El tamaño de la población es ajustada por el parámetro  $N_p$ . La población consiste en vectores en  $\mathbb{R}^D$ , donde  $D$  es el número de variables de control. La población se inicializa aleatoriamente dentro de los límites de los parámetros.

El proceso de optimización es realizado por tres operaciones principales: mutación, mestizaje y selección. En cada generación, los individuos de la población actual son **vectores objetivo**. Para cada vector objetivo, la operación de mutación produce un **vector mutante** al sumar la diferencia ponderada entre dos vectores elegidos aleatoriamente y un tercer vector aleatorio. La operación de mestizaje genera un nuevo vector, llamado **vector prueba**, mediante mezclar los parámetros del vector mutante con los del vector objetivo. Luego el vector prueba reemplaza el vector objetivo en la nueva generación (Rout, Sahu, y Panda, 2013).

El procedimiento, según (Rout y cols., 2013), es el siguiente:

1. **Inicialización:** Para cada parámetro  $j$  con límite inferior  $L_j$  y superior  $U_j$ , los valores se escogen aleatoriamente mediante una distribución uniforme en el intervalo  $[L_j, U_j]$ .
2. **Mutación:** Para un vector parámetro  $X_{i,G}$ , tres vectores  $(X_{r1,G}, X_{r2,G}, X_{r3,G})$  son escogidos aleatoriamente de forma que los índices  $i, r1, r2, r3$  son distintos. Un vector donante  $V_{i,G+1}$  se crea por la suma de uno de los vectores con la diferencia medida entre los otros dos vectores:

$$V_{i,G+1} = X_{r1,G} + F \cdot (X_{r2,G} - X_{r3,G}), \quad (\text{VIII.3})$$

donde  $F$  es una constante en el intervalo  $(0, 2)$ .

3. **Mestizaje:** El vector prueba  $U_{i,G+1}$  es desarrollado a partir de los elementos del vector objetivo  $X_{i,G}$  y los elementos del vector donante  $V_{i,G+1}$ . Los elementos del vector donante ingresan al vector prueba con probabilidad  $CR$  si:

$$U_{j,i,G+1} = \begin{cases} V_{j,i,G+1} & \text{si } U(0, 1) \leq CR \text{ ó } j = I_{rdm} \\ X_{j,i,G+1} & \text{en cualquier otro caso} \end{cases} \quad (\text{VIII.4})$$

donde  $U(0, 1)$  es una distribución uniforme en  $(0, 1)$ ,  $I_{rdm}$  es un entero aleatorio entre  $(1, D)$ .

4. **Selección:** El vector objetivo  $X_{i,G}$  es comparado con el vector prueba  $U_{i,G+1}$  y es admitido el vector que mejor ajusta la función. La operación de selección se representa por

$$X_{i,G+1} = \begin{cases} U_{i,G+1} & \text{si } f(U_{i,G+1}) < f(X_{i,G}) \\ X_{i,G} & \text{en cualquier otro caso} \end{cases} \quad (\text{VIII.5})$$

donde  $i \in [1, N_p]$ .

## B. Algoritmo

Se utilizaron los paquetes *NumPy*, *QuTiP*, *SciPy*, *Itertools* y *Random*. Las entradas del algoritmo son el sistema de  $N$  qubits (interpretado de la misma forma que el Algoritmo 1) y la matriz de



densidad  $\rho$ . Se utilizó el método *Differential evolution* proporcionado por *SciPy* con los parámetros estándar para realizar la optimización.

Para este algoritmo, en lugar de variar  $\theta$ , utilizamos  $\beta = 2\theta$  para poder utilizar los mismos límites de  $\phi$ . Se inicializan los  $\beta$ 's y  $\phi$ 's en  $\pi$ , recordando que hay un par por cada qubit. La función objetivo recibe cada par  $(\beta_i, \phi_i)$  y genera un conjunto de operadores locales parametrizados según lo visto en la Fundamentación teórica. Luego, encuentra las trazas parciales individuales y le aplica su respectivo operador local. Posteriormente, encuentra los productos tensoriales de los operadores locales de forma que el resultado pertenece a  $LM(S)$  y se aplica a la matriz completa. Finalmente, obtiene las entropías de von Neumann de cada matriz y calcula las correlaciones clásicas.

## 1. Diagrama de flujo

El presente diagrama de flujo muestra el funcionamiento del algoritmo expuesto en este capítulo.

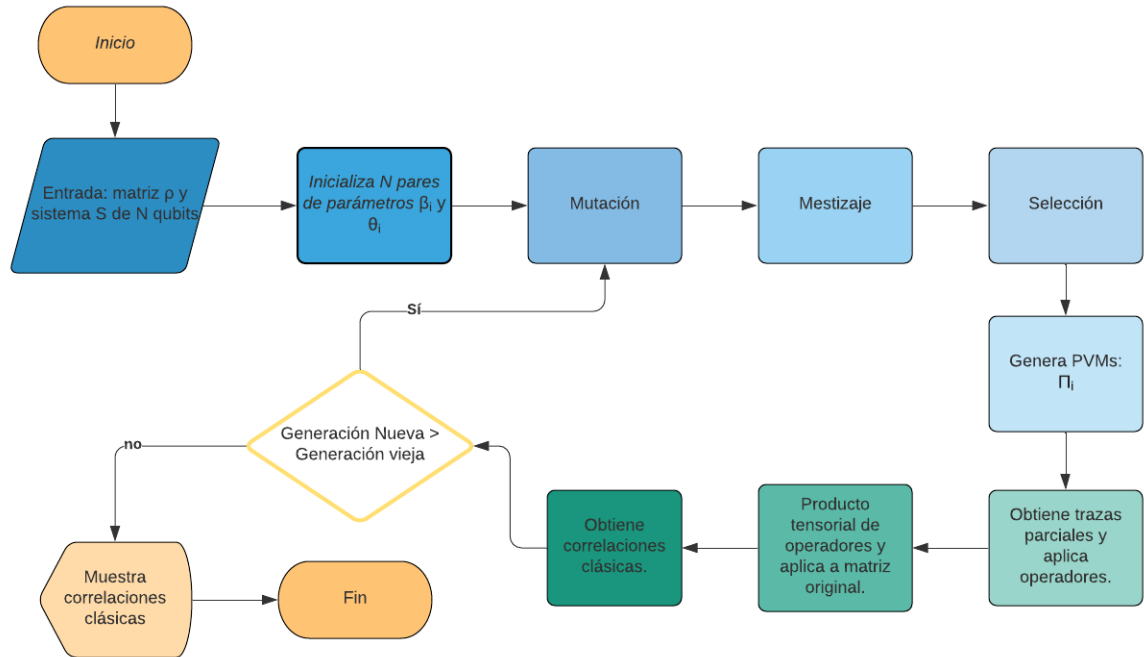


Figura VIII.1: Diagrama de flujo de Algoritmo 2.

## Otros algoritmos de optimización

Es importante mencionar que no todos los algoritmos de optimización logran converger para resolver este problema de optimización. El algoritmo de Evolución Diferencial se escogió luego de probar 5 métodos diferentes y que el ya mencionado obtuviera los resultados más satisfactorios. A continuación se muestra un cuadro que muestra los algoritmos probados junto a la tasa de éxito y el tiempo de solución para el problema (IV.117) con matrices aleatorias de tres qubits.

| Algoritmo              | Tiempo medio | Tasa de éxito |
|------------------------|--------------|---------------|
| Simulated Annealing    | 1 min 10 s   | 20 %          |
| SHG Optimization       | 2 min 27 s   | 50 %          |
| Dual Annealing         | 1 min 42 s   | 80 %          |
| Basin-Hopping          | 4 min 16 s   | 20 %          |
| Differential Evolution | 37 s         | 100 %         |

Tabla VIII.1: Comparación de métodos de optimización.

### C. Eficiencia

Se utilizó la herramienta *iPython Magic* para realizar mediciones de los tiempos de corrida (en segundos) para sistemas de  $N = 2, 3, 4, 5$  qubits. Se ejecutaron 100 ciclos para cada valor de  $N$ , utilizando matrices de densidad aleatorias para cada ciclo. En la Figura VIII.2 se muestra el logaritmo del tiempo de corrida contra el tamaño  $N$  del sistema. La regresión lineal muestra que el crecimiento del tiempo es exponencial en función del tamaño del sistema.  $T_0$  es una constante medida experimentalmente. Nótese que no es necesario medir para  $N = 1$  porque no hay correlaciones en un sistema de 1 qubit.

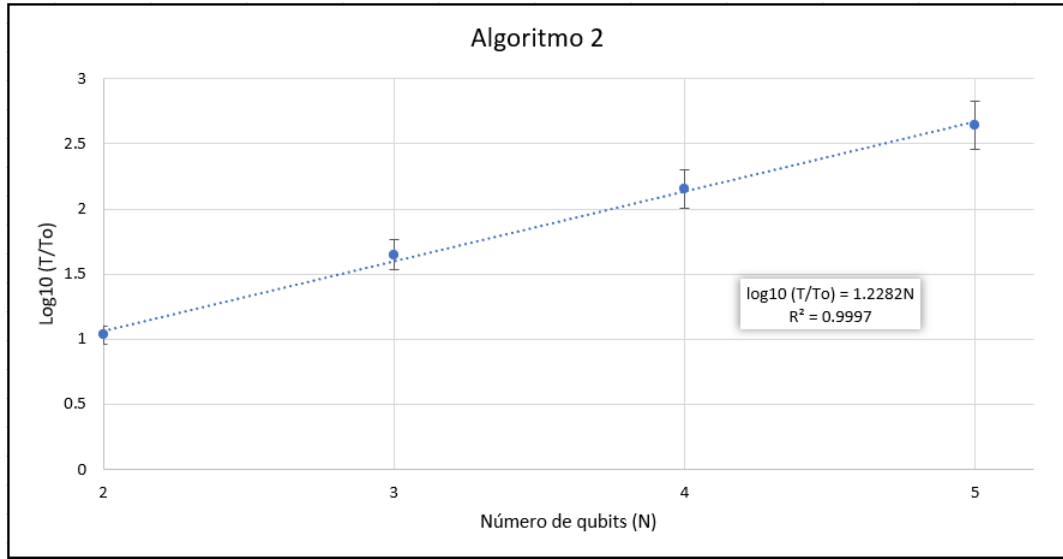


Figura VIII.2: Eficiencia del Algoritmo 2.  $T$  es el tiempo, en segundos, que tarda el código en resolver el Algoritmo 2 con un sistema de  $N$  qubits.  $T_0 = 0.987$  s. La línea punteada corresponde a la regresión lineal obtenida.

### D. Exploración de los estados GHZ y W

Se realizaron las mediciones de correlaciones totales, clásicas y cuánticas a los estados GHZ y W de 3 y 4 qubits. Nótese la diferencia de como se distribuyen las correlaciones que guarda cada estado. Se puede observar que el Algoritmo 2 muestra alta precisión al medir las correlaciones de estos estados específicos, aunque la precisión puede disminuir al procesar matrices con mayor cantidad de entradas no-cero.

| <b>Estado</b> | <b>Valor teórico</b> |          |           | <b>Valor obtenido</b> |          |           |
|---------------|----------------------|----------|-----------|-----------------------|----------|-----------|
|               | Totales              | Clásicas | Cuánticas | Totales               | Clásicas | Cuánticas |
| GHZ           | 3.000                | 2.000    | 1.000     | 2.999                 | 1.999    | 1.000     |
| W             | 2.755                | 1.170    | 1.585     | 2.755                 | 1.169    | 1.586     |

Tabla VIII.2: Correlaciones de los estados GHZ y W con  $N = 3$ .

| <b>Estado</b> | <b>Valor teórico</b> |          |           | <b>Valor obtenido</b> |          |           |
|---------------|----------------------|----------|-----------|-----------------------|----------|-----------|
|               | Totales              | Clásicas | Cuánticas | Totales               | Clásicas | Cuánticas |
| GHZ           | 4.000                | 3.000    | 1.000     | 3.999                 | 3.000    | 0.999     |
| W             | 3.245                | 1.245    | 2.000     | 3.244                 | 1.243    | 2.001     |

Tabla VIII.3: Correlaciones de los estados GHZ y W con  $N = 4$ .

Con el objetivo de estudiar cómo se distribuyen las correlaciones en un sistema multipartido, se ha propuesto el siguiente método (Mariano Lemus y Paunković, s.f.) de obtener las correlaciones clásicas y cuánticas  $k$ -partidas:

1. Se encuentran las estructuras de correlación  $\{\Phi^{(k)}, \rho_{\phi^{(k)}}\}$  que contienen todos los subsistemas de  $S$  que tienen exactamente  $k$  elementos.
2. A partir de esto, se obtienen las correlaciones irreducibles  $I_{irr}^{(k)}$  según la ecuación (IV.115).
3. Para las correlaciones clásicas  $k$ -partidas  $I_c^{(k)}$ , obtenemos los operadores locales que resuelven el problema de correlaciones clásicas (IV.117) y los aplicamos a  $\rho_{\phi^{(k)}}$ , obteniendo la matriz de densidad  $\rho_{\phi^{(k)}}^c$ . Definimos las correlaciones clásicas  $k$ -partidas como

$$I_c^{(k)} = I(\rho_{\phi^{(k)}}^c) - I(\rho_{\phi^{(k-1)}}^c) \quad (\text{IX.1})$$

donde  $\sum_{k=2}^N I_c^{(k)} = C(\rho_S)$ .

4. Finalmente, las correlaciones cuánticas  $k$ -partidas las definimos como

$$I_Q^{(k)} = I_{irr}^{(k)} - I_c^{(k)} \quad (\text{IX.2})$$

donde  $\sum_{k=2}^N I_Q^{(k)} = Q(\rho_S)$ .

## A. Simulaciones

### 1. Hipótesis 1 y 2

Para investigar estas hipótesis, se realizaron 600 iteraciones en las que se generaba una matriz de densidad  $\rho_S$  aleatoria de  $N = 3$  y se realizaba el siguiente proceso:

1. Se utilizó el Algoritmo 2 para obtener las mediciones locales  $\Pi_S$  que maximizan  $C(\rho_S)$ .
2. Se utilizó el Algoritmo 1 para obtener las matrices reducidas  $\rho_{\Phi^{(k)}}$  y luego se obtuvieron las correlaciones irreducibles y las correlaciones totales.
3. A cada matriz  $\rho_{\Phi^{(k)}}$  se le aplicó el Algoritmo 2 para obtener las matrices reducidas post-medición  $\rho_{\Phi^{(k)}}^c$  con los operadores locales  $\Pi_{\Phi^{(k)}}$  que maximizan  $C(\rho_{\Phi^{(k)}})$ . A estas se les midió las correlaciones clásicas  $k$ -partidas y luego las correlaciones cuánticas  $k$ -partidas.
4. Por otra parte, se midió la cantidad  $C(\rho_{\Phi^{(k)}})_S$  con los operadores locales  $\Pi_S$  que maximizan  $C(\rho_S)$ .
5. Se midió el error relativo de  $C(\rho_{\Phi^{(k)}})_S$  respecto a  $C(\rho_{\Phi^{(k)}}^c)$

El medible que permitirá comprobar la hipótesis 1 es el conteo de las veces que se obtuvo correlaciones cuánticas negativas. El medible para la hipótesis 2 es el error relativo mencionado en el último inciso para  $k = 2$ , de forma que un error cercano a cero sugiere que los operadores  $\Pi_S$  sí maximizan las correlaciones clásicas de sus matrices reducidas.

## 2. Hipótesis 3

Se realizaron 25 iteraciones con matrices de densidad de 3 qubits aleatorias en cada iteración. En esta simulación, se realizó lo siguiente:

1. Se generó un operador unitario local de la forma

$$U = U_1 \otimes U_2 \otimes U_3 \quad (\text{IX.3})$$

donde cada  $U_i = e^{i\alpha} R_Z(\beta) R_Y(\gamma) R_Z(\delta)$  es un operador unitario,  $R_Z$  y  $R_Y$  son matrices de rotación en los ejes  $Z$  y  $Y$ , y  $\alpha, \beta, \gamma, \delta$  son reales aleatorios entre 0 y  $2\pi$ .

2. Utilizando los Algoritmos 1 y 2, se obtuvo las correlaciones  $k$ -partidas, clásicas y cuánticas.
3. Se comparó con los resultados de obtener las correlaciones  $k$ -partidas previo a la aplicación del operador unitario.

En este caso, se utilizará como medible el conteo de las rotaciones que causaron diferencias mayores a  $1 \times 10^{-4}$  en las correlaciones totales y clásicas respecto a las correlaciones presentes antes de aplicar los operadores.

## 3. Hipótesis 4

Utilizando el método para generar operadores de Kraus aleatorios mostrado en (Kukulski, Netchita, Pawela, Puchala, y Zyczkowski, 2020), se realizaron 10 iteraciones para cada  $r = 1, 2, 3$  en las que se aplicaba  $\Lambda$  al estado GHZ y al estado W de 3 qubits y se obtuvo todas sus correlaciones  $k$ -partidas.

## B. Resultados

A continuación, en las tablas IX.1 y IX.2 se muestran las correlaciones  $k$ -partidas obtenidas al utilizar los algoritmos presentados en este documento sobre los estados GHZ y W de 3 qubits.

| $k$          | Correlaciones totales | Correlaciones clásicas | Correlaciones cuánticas |
|--------------|-----------------------|------------------------|-------------------------|
| <b>2</b>     | 1.972                 | 1.972                  | 0.000                   |
| <b>3</b>     | 1.028                 | 0.028                  | 1.000                   |
| <b>Total</b> | 3.000                 | 2.000                  | 1.000                   |

Tabla IX.1: Correlaciones  $k$ -partidas en el estado GHZ.

| $k$          | Correlaciones totales | Correlaciones clásicas | Correlaciones cuánticas |
|--------------|-----------------------|------------------------|-------------------------|
| <b>2</b>     | 2.680                 | 1.117                  | 1.563                   |
| <b>3</b>     | 0.074                 | 0.052                  | 0.021                   |
| <b>Total</b> | 2.754                 | 1.170                  | 1.584                   |

Tabla IX.2: Correlaciones  $k$ -partidas en el estado W.

En la Tabla IX.3 se muestran los resultados de las simulaciones realizadas para las hipótesis 1, 2 y 3. El resultado de la hipótesis 2 muestra que, en promedio, las correlaciones clásicas obtenidas de las matrices reducidas son 30 % mayores, en promedio, que las correlaciones obtenidas de las mismas ante la aplicación de las mediciones que maximizan las correlaciones clásicas en su matriz original. También muestra que como mínimo son iguales, pero que pueden ser hasta 60 % mayores.

| Hipótesis | Medible  | Valor           |
|-----------|--|-----------------|
| 1         | Conteo de correlaciones cuánticas negativas                        | 0               |
| 2         | Diferencia relativa promedio                                       | 30 % $\pm$ 30 % |
| 3         | Cambios en las correlaciones $k$ -partidas ante operadores locales | 0               |

Tabla IX.3: Resultados de las simulaciones realizadas para hipótesis 1, 2 y 3.

A continuación se muestran los resultados obtenidos de aplicar mapeos aleatorios de tipo Canal cuántico de  $r$  operadores de Kraus a los estados GHZ y W en 10 iteraciones. En las tablas se utiliza las iniciales de la cantidad para referirse a las mismas (Ej.: 'CQ' corresponde a Correlaciones Cuánticas). Cada tabla tiene cuatro datos resaltados, estos resultados servirán para concluir a favor o en contra de la Hipótesis 4.

| $r = 2$  | CT    |       |       | CC    |       |       | CQ    |       |       |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| <b>k</b> | Min   | Max   | Media | Min   | Max   | Media | Min   | Max   | Media |
| <b>2</b> | 0.027 | 0.516 | 0.253 | 0.026 | 0.477 | 0.219 | 0.002 | 0.085 | 0.035 |
| <b>3</b> | 0.034 | 0.575 | 0.226 | 0.001 | 0.425 | 0.059 | 0.032 | 0.452 | 0.167 |
| <b>T</b> | 0.153 | 0.866 | 0.479 | 0.053 | 0.480 | 0.277 | 0.100 | 0.468 | 0.202 |

Tabla IX.4: Mapeos de canal cuántico con  $r = 2$  sobre GHZ.

| $r = 3$  | CT    |       |       | CC    |       |       | CQ    |       |       |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| <b>k</b> | Min   | Max   | Media | Min   | Max   | Media | Min   | Max   | Media |
| <b>2</b> | 0.016 | 0.344 | 0.148 | 0.010 | 0.326 | 0.123 | 0.003 | 0.071 | 0.026 |
| <b>3</b> | 0.011 | 0.196 | 0.062 | 0.000 | 0.002 | 0.001 | 0.010 | 0.194 | 0.061 |
| <b>T</b> | 0.037 | 0.541 | 0.210 | 0.011 | 0.328 | 0.124 | 0.026 | 0.213 | 0.087 |

Tabla IX.5: Mapeos de canal cuántico con  $r = 3$  sobre GHZ.

| $r = 2$  | CT    |       |       | CC    |       |       | CQ    |       |       |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| <b>k</b> | Min   | Max   | Media | Min   | Max   | Media | Min   | Max   | Media |
| <b>2</b> | 1.019 | 1.425 | 1.215 | 0.310 | 0.844 | 0.487 | 0.581 | 0.869 | 0.728 |
| <b>3</b> | 0.100 | 0.441 | 0.320 | 0.004 | 0.128 | 0.053 | 0.086 | 0.426 | 0.267 |
| <b>T</b> | 1.243 | 1.866 | 1.535 | 0.330 | 0.859 | 0.540 | 0.825 | 1.126 | 0.995 |

Tabla IX.6: Mapeos de canal cuántico con  $r = 2$  sobre W.

| $r = 3$  | CT    |       |       | CC    |       |       | CQ    |       |       |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| <b>k</b> | Min   | Max   | Media | Min   | Max   | Media | Min   | Max   | Media |
| <b>2</b> | 0.655 | 1.167 | 0.872 | 0.225 | 0.650 | 0.386 | 0.328 | 0.593 | 0.485 |
| <b>3</b> | 0.274 | 0.593 | 0.448 | 0.014 | 0.211 | 0.089 | 0.219 | 0.578 | 0.359 |
| <b>T</b> | 1.136 | 1.755 | 1.319 | 0.35  | 0.664 | 0.475 | 0.670 | 1.090 | 0.844 |

Tabla IX.7: Mapeos de canal cuántico con  $r = 3$  sobre W.

---

## Discusión de resultados

---

Una parte importante de los programas desarrollados fue la utilización de los paquetes *QuTiP* y *SciPy*. Esto debido a que el primero proporcionó métodos que permitieron realizar operaciones con objetos cuánticos con gran facilidad y el segundo proporcionó varios métodos de optimización. Es importante notar que lenguajes de programación compiladores (como *C++* o *Fortran*) o lenguajes de programación especializados en operaciones con matrices (como *MATLAB* o *MATA*) podrían llegar a ser más eficientes en la solución de los problemas presentados (Chauhan, 2013). Así pues, se sugiere realizar adaptaciones a los lenguajes mencionados para verificar si hay algún lenguaje que proporcione resultados de forma más eficiente con sistemas más grandes. Se realiza esta sugerencia debido a que las Figuras VII.2 y VIII.2 muestran comportamiento exponencial al aumentar el tamaño del sistema, lo cual es esperado ya que se trabaja con matrices de  $2^N \times 2^N$ .

Para el primer Algoritmo, se seleccionó el algoritmo de Descenso en Gradiente como método de optimización debido a que permite explorar el espacio de parámetros a través de un gradiente, sin ingresar los límites, pues no se tenía idea de los valores que estos parámetros pudiesen tomar. Por otra parte, la función gradiente que se utilizó no es necesariamente el gradiente de la función objetivo. Esto claramente repercute en un aumento en el tiempo de ejecución, sin embargo, la función gradiente que se utilizó sí proporciona resultados con cierta precisión según lo observado anteriormente. Esto se debe a que liga cada parámetro a una constante dependiente de la matriz de densidad ingresada, por lo que el acercamiento a estas constantes es asegurado.

Para el segundo algoritmo, entre los métodos de optimización disponibles se escogió el *Differential Evolution*, porque ante varias pruebas resultó tener una tasa de éxito mayor y una velocidad mayor que los algoritmos probados, con las funciones requeridas (Véase Tabla VIII.1). Además permite variar los parámetros en rangos específicos para cada uno, rangos que ya se conocían. Es importante mencionar que, al inicio, se utilizó *Simulated Annealing* como método de optimización para el Algoritmo 2, pero los resultados no concordaban con los conocimientos previos sobre la teoría, especialmente al simular las rotaciones de la Hipótesis 3.

Obsérvense las tablas IX.1 y IX.2. En estas se encuentran las correlaciones  $k$ -partidas para los estados GHZ y W, respectivamente. Estos resultados son importantes por varias razones. La primera razón es la comparación que se puede realizar entre ambos estados, ya que ambos son estados de alto entrelazamiento aunque cada uno concentra sus correlaciones de formas diferentes: el GHZ tiene



todas sus correlaciones clásicas en los subsistemas de 2 qubits y todas las cuánticas en subsistema de 3 qubits; en cambio, el estado  $W$  concentra todas sus correlaciones, clásicas y cuánticas, en los subsistemas de 2 qubits. Una segunda razón es que permite comprobar que los métodos utilizados funcionan correctamente, ya que los estados GHZ y  $W$  son ampliamente estudiados y utilizados en el ámbito de la computación cuántica y se conoce su comportamiento. Otra razón es que sirven de referencia para comparar el efecto de operaciones aplicadas a estos estados.

Las tablas mencionadas se utilizaron inicialmente para comparar las correlaciones obtenidas en la simulación de la Hipótesis 3 al realizar pruebas preliminares. Sobre los resultados de las pruebas para la Hipótesis 3, Tabla IX.3, se observa que en ningún momento se obtuvo una operación unitaria local que modificara significativamente las correlaciones de los estados aleatorios. Según esta evidencia, no es posible rechazar la Hipótesis 3. Esto es un resultado que le puede dar cierto grado de generalidad al método propuesto.

Por otra parte, obsérvense las tablas de IX.4 a IX.7, las cuales corresponden a los resultados de las pruebas de la Hipótesis 4. Nótese que la Hipótesis 3 es el caso  $r = 1$  de la Hipótesis 4. Lo primero que se debe notar en las tablas mencionadas es que la media de las correlaciones totales disminuye considerablemente conforme  $r$  aumenta. Esto indica que los canales se vuelven más ruidosos conforme más operadores de Kraus utiliza, *i.e.*, destruyen las correlaciones de los estados. Lo segundo a notar es que los estados sufren una redistribución de las correlaciones, tanto clásicas como cuánticas. Con esto vemos, en los datos resaltados de las tablas, el aumento de correlaciones cuánticas  $k$ -partidas. En conclusión, en base a los resultados se rechaza la Hipótesis 4.

Ahora, para las simulaciones de las Hipótesis 1 y 2, se utilizaron matrices de densidad aleatorias, en las que se utilizó un porcentaje de entradas no-cero uniformemente aleatoria para evitar sesgos. Obsérvese en la Tabla IX.3 que no se encontró ninguna correlación cuántica negativa, por lo tanto, no se puede rechazar la Hipótesis 1 a partir de los resultados obtenidos. Por otra parte, el resultado concerniente a la Hipótesis 2 es un porcentaje positivo significativamente alto acompañado de una incertidumbre del mismo tamaño. Esto significa que las mediciones PVM que maximizan las correlaciones de la matriz de densidad del sistema no necesariamente maximizan las correlaciones de las matrices reducidas del sistema, razón por la cual se concluye que el enunciado de la Hipótesis 2 no se satisface.

Utilizando *Python* y principalmente los paquetes *QuTiP* y *SciPy*, se implementó un método para calcular las correlaciones clásicas y cuánticas  $k$ -partidas a partir del desarrollo teórico realizado por (Mariano Lemus y Paunković, s.f.). Este método consiste de dos subrutinas: la selección de matrices reducidas de máxima entropía y la selección de operadores locales que maximizan las correlaciones de un sistema. La primera utiliza *Gradient Descent* para realizar la optimización, mientras que la segunda utiliza *Differential Evolution*. Ambas subrutinas aumentan su tiempo de corrida exponencialmente conforme el tamaño del sistema. Este método se utilizó para realizar pruebas a cuatro hipótesis, a través de simulaciones.

Debido a que no se encontró correlaciones cuánticas negativas ni cambios significativos en las correlaciones ante la aplicación de operadores unitarios locales, las hipótesis que hablan sobre la positividad semi-definida y sobre la conservación de las correlaciones no se pueden rechazar. Las simulaciones mostraron que los operadores locales que maximizan  $C(\rho_S)$  obtienen un 30 % menos correlaciones en  $\rho_{\Phi^{(k)}}$  que los operadores que maximizan  $C(\rho_{\Phi}^{(k)})$ , por lo que se rechaza la hipótesis que propone que las correlaciones obtenidas por ambos conjuntos de operadores son de igual tamaño. Se observó una redistribución de las correlaciones  $k$ -partidas ante la aplicación de canales cuánticos locales en los estados GHZ y W, lo cual permite rechazar la hipótesis relacionada a las aplicaciones mencionadas. Además, se observó una disminución de las correlaciones ante canales cuánticos de mayor ruido.

Se observó que el Algoritmo 1 tiene un tiempo de ejecución bastante alto en comparación con lo deseable para realizar simulaciones de una cantidad grande de iteraciones y de sistemas de más de 4 qubits. Se recomienda encontrar una función gradiente que funcione con mayor eficiencia o, si fuese necesario, realizar una adaptación de los códigos a lenguajes de programación con mayor eficiencia en las operaciones con matrices. En la misma idea de agilizar el Algoritmo 1, una forma de mejorar la eficiencia es simplificar la función objetivo, ya que requiere de mucho espacio en memoria. Por otra parte, se recomienda realizar pruebas con algoritmos de optimización diferentes al *Gradient Descent* con el fin de mejorar la precisión de los resultados.

En la medición de correlaciones cuánticas  $k$ -partidas no es deseable que estas correlaciones aumenten ante la aplicación de canales cuánticos locales, sin embargo, el método propuesto sí sufre estos cambios. Por lo tanto, se recomienda reformular el enfoque con el que se estudia las correlaciones cuánticas, utilizando entropía relativa a través de *Entropic Updating* o incluso a través de diferentes definiciones de entropía como la entropía mínima o la entropía de Rényi.

- Adesso, G., Bromley, T. R., y Cianciaruso, M. (2016). Measures and applications of quantum correlations. *Journal of Physics A: Mathematical and Theoretical*, 49(47), 473001.
- Amari, S.-I. (2001). Information geometry on hierarchy of probability distributions. *IEEE transactions on information theory*, 47(5), 1701–1711.
- Audenaert, K. M. R. (2013, Jul). On the asymmetry of the relative entropy. *Journal of Mathematical Physics*, 54(7), 073506. Descargado de <http://dx.doi.org/10.1063/1.4811856> doi: 10.1063/1.4811856
- Bell, J. S. (1964, Nov). On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1, 195–200. Descargado de <https://link.aps.org/doi/10.1103/PhysicsPhysiqueFizika.1.195> doi: 10.1103/PhysicsPhysiqueFizika.1.195
- Bell, J. S., y Aspect, A. (2004). *Speakable and unspeakable in quantum mechanics: Collected papers on quantum philosophy* (2.<sup>a</sup> ed.). Cambridge University Press. doi: 10.1017/CBO9780511815676
- Bohm, D., y Aharonov, Y. (1957, Nov). Discussion of experimental proof for the paradox of einstein, rosen, and podolsky. *Phys. Rev.*, 108, 1070–1076. Descargado de <https://link.aps.org/doi/10.1103/PhysRev.108.1070> doi: 10.1103/PhysRev.108.1070
- Bojić, A. (2013, 06). A new quantum game based on chsh game. *Journal of Information and Organizational Sciences*, 37, 15-22.
- Brassard, G., y Méthot, A. A. (2006). Can quantum-mechanical description of physical reality be considered incomplete? *International Journal of Quantum Information*, 04(01), 45-54. Descargado de <https://doi.org/10.1142/S0219749906001608> doi: 10.1142/S0219749906001608
- Chauhan, S. (2013). *Programming languages - design and constructs*. University Science Press. Descargado de <https://books.google.com.gt/books?id=qKRnohkDbb0C>
- Clauser, J., Horne, M., Shimony, A., y Holt, R. (1969, 10). Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23, 880-884. doi: 10.1103/PhysRevLett.23.880
- Cover, T. M., y Thomas, J. A. (2006). *Elements of information theory (wiley series in telecommunications and signal processing)*. USA: Wiley-Interscience.
- Einstein, A., Podolsky, B., y Rosen, N. (1935, May). Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47, 777–780. Descargado de <https://link.aps.org/doi/10.1103/PhysRev.47.777> doi: 10.1103/PhysRev.47.777
- Girolami, D., Tufarelli, T., y Susa, C. E. (2017, Oct). Quantifying genuine multipartite correlations and their pattern complexity. *Physical Review Letters*, 119(14). Descargado de <http://dx.doi.org/10.1103/PhysRevLett.119.140505> doi: 10.1103/physrevlett.119.140505
- Giuliani, G. (2020). *On some epr (einstein, podolsky, rosen) issues*.
- Gray, R. M. (2011). *Entropy and information theory* (2nd ed.). Springer Publishing Company, Incorporated.

- Henderson, L., y Vedral, V. (2001, Aug). Classical, quantum and total correlations. *Journal of Physics A: Mathematical and General*, 34(35), 6899–6905. Descargado de <http://dx.doi.org/10.1088/0305-4470/34/35/315> doi: 10.1088/0305-4470/34/35/315
- Kato, K., Furrer, F., y Murao, M. (2016, Feb). Information-theoretical analysis of topological entanglement entropy and multipartite correlations. *Physical Review A*, 93(2). Descargado de <http://dx.doi.org/10.1103/PhysRevA.93.022317> doi: 10.1103/physreva.93.022317
- Klüver, J., y Klüver, C. (2007). *On communication. an interdisciplinary and mathematical approach*. Springer Netherlands. Descargado de <https://books.google.bi/books?id=mwCAQJE-QUEC>
- Kochenderfer, M., y Wheeler, T. (2019). *Algorithms for optimization*. MIT Press. Descargado de <https://books.google.com.gt/books?id=uBSMDwAAQBAJ>
- Krippendorff, K. (1986). *Information theory: structural models for qualitative data* (Vol. 62). Sage.
- Kukulski, R., Nechita, I., Pawela, L., Puchala, Z., y Zyczkowski, K. (2020, 11). Generating random quantum channels.
- Kumar, A., Nirala, G., y Marino, A. M. (2020). *Einstein-podolsky-rosen paradox with position-momentum entangled macroscopic twin beams*.
- Lemaréchal, C. (2012). *Cauchy and the gradient method*.
- Linden, N., Popescu, S., y Wootters, W. K. (2002, Oct). Almost every pure state of three qubits is completely determined by its two-particle reduced density matrices. *Phys. Rev. Lett.*, 89, 207901. Descargado de <https://link.aps.org/doi/10.1103/PhysRevLett.89.207901> doi: 10.1103/PhysRevLett.89.207901
- Liu, Y., Zeng, B., y Zhou, D. L. (2016, Feb). Irreducible many-body correlations in topologically ordered systems. *New Journal of Physics*, 18(2), 023024. Descargado de <http://dx.doi.org/10.1088/1367-2630/18/2/023024> doi: 10.1088/1367-2630/18/2/023024
- Loura, R. (s.f.). *Relationship between quantum discord and global discord*. Instituto Técnico Superior. (Proporcionado por el autor.)
- MacKay, D. J., y Mac Kay, D. J. (2003). *Information theory, inference and learning algorithms*. Cambridge university press.
- Mariano Lemus, P. M., y Paunković, N. (s.f.). *A note on correlations*. Instituto Técnico Superior. (Proporcionado por los autores antes de publicación.)
- Modi, K., Brodutch, A., Cable, H., Paterek, T., y Vedral, V. (2012, Nov). The classical-quantum boundary for correlations: Discord and related measures. *Reviews of Modern Physics*, 84(4), 1655–1707. Descargado de <http://dx.doi.org/10.1103/RevModPhys.84.1655> doi: 10.1103/revmodphys.84.1655
- Nielsen, M. A., y Chuang, I. L. (2011). *Quantum computation and quantum information: 10th anniversary edition* (10th ed.). USA: Cambridge University Press.
- of Photonic Sciences, I.-T. I. (s.f.). *The big bell test: Global physics experiment challenges einstein: Principle of local realism tested with the help of 100,000 volunteers*. ScienceDaily. Descargado 2018, May 9, de [www.sciencedaily.com/releases/2018/05/180509135409.htm](http://www.sciencedaily.com/releases/2018/05/180509135409.htm)
- Ohya, M., y Petz, D. (2004). *Quantum entropy and its use*. Springer Berlin Heidelberg. Descargado de <https://books.google.com.gt/books?id=r2ullNVyESQC>
- Ollivier, H., y Zurek, W. H. (2001, Dec). Quantum discord: A measure of the quantumness of correlations. *Physical Review Letters*, 88(1). Descargado de <http://dx.doi.org/10.1103/PhysRevLett.88.017901> doi: 10.1103/physrevlett.88.017901
- Petz, D. (2001). Entropy, von neumann and the von neumann entropy. En *John von neumann and the foundations of quantum physics* (pp. 83–96). Springer.
- Ponciano, J. (2018). *Introducción a la mecánica cuántica. apuntes de curso*. Descargado de <https://ecfm.usac.edu.gt/japonciano/courses.html>
- Preskill, J. (2018). *Quantum shannon theory*.
- Rout, U. K., Sahu, R. K., y Panda, S. (2013). Design and analysis of differential evolution algorithm based automatic generation control for interconnected power system. *Ain Shams Engineering Journal*, 4(3), 409-421. Descargado de <https://www.sciencedirect.com/science/article/pii/S2090447912000986> doi: <https://doi.org/10.1016/j.asej.2012.10.010>
- Rulli, C. C., y Sarandy, M. S. (2011, Oct). Global quantum discord in multipartite systems. *Physical Review A*, 84(4). Descargado de <http://dx.doi.org/10.1103/PhysRevA.84.042109> doi:

10.1103/physreva.84.042109

- Sakurai, J. J., y Napolitano, J. (2011). *Modern quantum mechanics; 2nd ed.* San Francisco, CA: Addison-Wesley. Descargado de <https://cds.cern.ch/record/1341875>
- Santos, C. (2016). *Shannon's coding theorems* (Tesis de Master no publicada). Saint Mary's College of California.
- Schneidman, E., Still, S., Berry, M. J., Bialek, W., y cols. (2003). Network information and connected correlations. *Physical review letters*, *91*(23), 238701.
- Serway, R., Moses, C., y Moyer, C. (2004). *Modern physics*. Cengage Learning. Descargado de <https://books.google.com.gt/books?id=uTM8AAAAQBAJ>
- Shannon, C. E., y Weaver, W. (1963). *A mathematical theory of communication*. USA: University of Illinois Press.
- Storn, R., y Price, K. (1997). Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, *11*(4), 341–359.
- Szalay, S. (2018, Nov). The classification of multipartite quantum correlation. *Journal of Physics A: Mathematical and Theoretical*, *51*(48), 485302. Descargado de <http://dx.doi.org/10.1088/1751-8121/aae971> doi: 10.1088/1751-8121/aae971
- Vajapeyam, S. (2014). *Understanding shannon's entropy metric for information*.
- Vanslette, K. (2017a, Dec). Entropic updating of probabilities and density matrices. *Entropy*, *19*(12), 664. Descargado de <http://dx.doi.org/10.3390/e19120664> doi: 10.3390/e19120664
- Vanslette, K. (2017b). *The quantum bayes rule and generalizations from the quantum maximum entropy method*.
- Vedral, V. (2002, Mar). The role of relative entropy in quantum information theory. *Reviews of Modern Physics*, *74*(1), 197–234. Descargado de <http://dx.doi.org/10.1103/RevModPhys.74.197> doi: 10.1103/revmodphys.74.197
- Wilde, M. M. (2013). *Quantum information theory*. Cambridge University Press. doi: 10.1017/CBO9781139525343
- Witten, E. (2020, Mar). A mini-introduction to information theory. *La Rivista del Nuovo Cimento*, *43*(4), 187–227. Descargado de <http://dx.doi.org/10.1007/s40766-020-00004-5> doi: 10.1007/s40766-020-00004-5
- Zhou, D. L. (2008, Oct). Irreducible multiparty correlations in quantum states without maximal rank. *Physical Review Letters*, *101*(18). Descargado de <http://dx.doi.org/10.1103/PhysRevLett.101.180505> doi: 10.1103/physrevlett.101.180505
- Zhou, D. L. (2009, Aug). Irreducible multiparty correlations can be created by local operations. *Physical Review A*, *80*(2). Descargado de <http://dx.doi.org/10.1103/PhysRevA.80.022113> doi: 10.1103/physreva.80.022113

## **A. Especificaciones de hardware y software**

A continuación se presentan las características del equipo utilizado para realizar las simulaciones y mediciones.

- Procesador: Intel Core i7-7500U, de 4 núcleos y velocidad de procesamiento 2.70GHz.
- Memoria RAM: 16.0 GB.
- Sistema operativo: Windows 10 Home.

También se presentan los programas y paquetes utilizados, junto a la versión utilizada.

- Python 3.8
  - SciPy 1.6.2
  - NumPy 1.20.1
  - QuTiP 4.6.0
  - Pandas 1.2.4
  - Random 3.8
  - Itertools 8.7.0
  - IPython 7.12.0
- Anaconda 1.10.0
- JupyterLab 1.1.4