



UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

Optimización estocástica de desigualdades no locales en redes de Bell

Profesor Guía: **Dr. Aldo Patricio Delgado Hidalgo**
Departamento de Física
Facultad de Ciencias Físicas y Matemáticas
Universidad de Concepción

Tesis presentada en la Facultad de Ciencias Físicas y Matemáticas para
optar al grado de Doctor en Ciencias Físicas de la Universidad de
Concepción

Código Programa: 4226

por

Jean Franco Cortés Vega
CONCEPCIÓN - CHILE
Septiembre — 2026

Comisión examinadora : Dr. Aldo Patricio Delgado Hidalgo

Dr. Esteban Sepúlveda Gómez

Dr. Pablo Solano Palma

Dr. Dardo Goyeneche

Agradecimientos

Esta tesis es la llave que me permite cerrar uno de las etapas más complejas que he vivido. Durante el desarrollo de este trabajo he enfrentado un sin fin de dificultades que me han permitido crecer no solo como profesional, sino que también como una persona.

Estos agradecimientos serán escritos en función del rol que ha jugado cada persona en esta travesía, ya que cada uno de ustedes ha aportado desde sus respectivas veredas a que pueda concluir esta etapa.

En primer lugar quiero agradecer el incondicional apoyo de mis padres Marisol y Ricardo. Sus palabras de aliento son el dulce bálsamo que permite mantener la cabeza en alto a pesar de las tormentas. Sin ellos el camino que elegí en esta vida sería muy diferente. No puedo dejar de lado a mi hermana Andrea, quien también ha jugado un papel importante en mi vida. Ella siempre escoge las palabras correctas para poder reconectar mis ideas y emociones cuando se acerca el caos.

A mi psicóloga y terapeuta Daniella Rossi. Eres una tremenda profesional y agradezco la aleatoriedad de la vida que me permitió ser tu paciente. Gracias a ti he podido aprender mucho sobre mi mismo, conocer y reconocer mis capacidades, a enfrentar mis miedos y a superarlos. A planificar y ordenar la casa cuando se avecina la catástrofe. Sin tu apoyo terapéutico es altamente probable que esta etapa no hubiese llegado a puerto.

Gracias Francisca. Tu has tenido las palabras de apoyo, la eterna paciencia, el inmenso cariño de aceptarme tal cual soy. De ti aprendí a no bajar los brazos, a hacer los sacrificios necesarios para concluir los objetivos y a tener la resiliencia de levantarse las veces que sean necesarias para cumplir las metas y sueños que uno se propone. Te quiero.

Adicionalmente quiero agradecer a mi supervisor Dr. Aldo Delgado quien ha supervisado tanto mi tesis de magíster como esta etapa de doctorado. Su apoyo durante este proceso fue incondicional. Muy fiel a su estilo ha tenido la paciencia de apoyarme durante todo este proceso, a pesar de las dificultades que he tenido que afrontar. Estaré eternamente agradecido de siempre tener las puertas abiertas para

poder regresar a continuar nuestros trabajos. También agradecer las innumerables oportunidades que me brindo para poder continuar con mis estudios al director del programa de Doctorado en Ciencias Físicas, el Dr. Esteban Sepúlveda. Infinitas gracias por comprender y confiar que todas estas oportunidades concluirían en este trabajo.

También quiero agradecer a mi amigo, compañero y colega Luciano Pereira. El también ha sido un pilar fundamental en este proceso. No solamente me ha apoyado desde lo técnico, sino que también desde lo personal y desde el bienestar. Estoy eternamente agradecido de tener una amistad tan genuina y profunda.

No puedo dejar de lado a los colegas y amistades que he forjado en mi trabajo como profesor en la Universidad San Sebastián. Este proceso hubiese sido tan diferente sin su apoyo diario principalmente a Anita Zapata, Paola Meza, Cristian Quinzácarra, Nicole Carrasco, Sebastián Pons, Patricio Aguilera, Francisco Avilés, Fernando Izaurieta, Daniel Insunza, Astrid Haengsen, Luis Angulo, Cristian Reyes. Lamento si se me queda a alguien en el tintero, no es nada personal.

¿Qué sería de la vida sin los amigos? Gracias por estar siempre Vanesa, Alexis, Gerald, Luciano (si, el mismo de arriba), Sebastián (también se repite), Nayda, Nicole, Patrick y Claudia. Todos ustedes han estado en diferentes etapas de mi vida y de este proceso. Gracias por las conversaciones profundas Morin, Fabio y Joaquín. Ustedes han estado presente sobretodo en último año de este proceso. Me han ayudado a no bajar los brazos y me han acompañado desde la paciencia y el cariño.

Resumen

El propósito de esta tesis es desarrollar y evaluar un protocolo autoguiado para certificar correlaciones no clásicas en la red de Bell triangular. En este escenario, tres fuentes independientes distribuyen sistemas bipartitos entre tres nodos conectados cíclicamente, cada uno de los cuales recibe dos subsistemas y realiza una medición conjunta. A diferencia del escenario bipartito convencional, un modelo clásico debe reproducir las correlaciones observadas respetando simultáneamente la topología causal de la red y la independencia de las fuentes. Estas condiciones dan lugar a un conjunto de correlaciones no convexo sujeto a restricciones no lineales. Por ello, la presencia de entrelazamiento en los estados distribuidos no garantiza por sí sola que las mediciones generen una distribución incompatible con el modelo clásico triangular.

El marco teórico se construye a partir del teorema de Bell, el intercambio de entrelazamiento y la bilocalidad, para luego introducir la red triangular y la distribución de Fritz como ejemplo de no localidad en una red sin entradas externas. La incompatibilidad con el modelo clásico se evalúa mediante el funcional Wagon-Wheel I_{w-w} , una expresión no lineal construida a partir de probabilidades conjuntas y marginales de los resultados observables de la red. Este funcional fue obtenido por Fraser y Wolfe mediante la técnica de inflación y actúa como un testigo de compatibilidad causal. Toda distribución generada por tres fuentes clásicas independientes satisface $I_{w-w} \leq 0$, por lo que un valor positivo demuestra que las correlaciones observadas no admiten dicho modelo y certifica su no localidad en la red. Sin embargo, un resultado no positivo solo indica que este funcional no permitió obtener la certificación.

Sobre esta base, se propone un protocolo que utiliza Complex Simultaneous Perturbation Stochastic Approximation (CSPSA) para encontrar las mediciones conjuntas que maximizan directamente I_{w-w} . CSPSA es un método de optimización estocástica que estima la dirección de actualización a partir de dos evaluaciones de I_{w-w} , obtenidas al perturbar simultánea y aleatoriamente todos los parámetros complejos. Los estados distribuidos permanecen fijos, mientras que las bases de medición de los tres nodos se parametrizan mediante unitarias pertenecientes a $U(4)$. El algoritmo explora este espacio a partir de inicializaciones aleatorias e independientes,

sin imponer una distribución objetivo ni utilizar como punto de partida las mediciones asociadas a la distribución de Fritz. El protocolo se estudia para fuentes que distribuyen estados maximalmente entrelazados, estados puros en descomposición de Schmidt y estados de Werner, además de configuraciones en las que una de las fuentes distribuye un estado separable.

En una primera etapa, I_{w-w} se evalúa de forma exacta a partir de las probabilidades calculadas mediante la regla de Born. Para tres estados maximalmente entrelazados, el protocolo encuentra mediciones que violan la desigualdad Wagon-Wheel y se aproxima al valor de referencia de Fritz. El estudio de las familias de Schmidt y Werner permite identificar cómo cambia la capacidad de detección frente al entrelazamiento parcial y al ruido. En las configuraciones con una fuente separable no se encuentran violaciones dentro del dominio explorado. Estos resultados establecen la factibilidad numérica del método y muestran que la detección depende conjuntamente de los estados distribuidos, las mediciones y la estructura causal de la red.

En una segunda etapa, el análisis se extiende al escenario en que I_{w-w} debe estimarse a partir de un número finito de mediciones. Para ello, se incorpora Classical Shadows, un método que permite estimar propiedades del estado cuántico mediante mediciones aleatorias locales de Pauli, sin reconstruir completamente su matriz de densidad. Cada ronda de medición genera una snapshot o instantánea, que corresponde a un estimador clásico del estado. El tamaño del conjunto fijo de snapshots, denotado por N_{sh} , define el presupuesto de shadows y, por tanto, el número de copias medidas. Estos datos se reutilizan durante la optimización para estimar I_{w-w} en las distintas configuraciones exploradas por CSPA. Las simulaciones muestran que un presupuesto insuficiente puede llevar al algoritmo a seleccionar mediciones favorecidas por las fluctuaciones estadísticas, produciendo valores positivos que no corresponden a una violación al evaluar exactamente esas mediciones. Al aumentar N_{sh} , esta discrepancia disminuye y las mediciones encontradas permiten obtener violaciones consistentes para las configuraciones más robustas estudiadas. En consecuencia, el presupuesto de shadows y el número de iteraciones deben elegirse conjuntamente, mientras que una certificación experimental requeriría evaluar finalmente las mediciones seleccionadas mediante datos independientes.

La investigación distingue finalmente entre la búsqueda de mediciones y la certificación de no localidad. Classical Shadows orienta la optimización, mientras que en una implementación experimental las mediciones encontradas deben evaluarse mediante frecuencias obtenidas de datos independientes. Esta evaluación final es la que certifica la incompatibilidad con el modelo clásico triangular. En conjunto, el trabajo establece un procedimiento constructivo para optimizar directamente un testigo no lineal bajo condiciones ideales y con información estadística finita. Sus resultados abren la posibilidad de extender el protocolo a redes con más nodos y explorar

su aplicación en arquitecturas cuánticas distribuidas, un futuro internet cuántico y protocolos de comunicación cuántica segura.

Abstract

The purpose of this thesis is to develop and evaluate a self-guided protocol for certifying nonclassical correlations in the triangle Bell network. In this scenario, three independent sources distribute bipartite systems among three cyclically connected nodes, each of which receives two subsystems and performs a joint measurement. Unlike the conventional bipartite scenario, a classical model must reproduce the observed correlations while simultaneously respecting the causal topology of the network and the independence of the sources. These conditions give rise to a nonconvex set of correlations subject to nonlinear constraints. Therefore, the presence of entanglement in the distributed states does not by itself guarantee that the measurements generate a distribution incompatible with the classical triangle model.

The theoretical framework builds on Bell’s theorem, entanglement swapping, and bilocality, before introducing the triangle network and the Fritz distribution as an example of nonlocality in a network without external inputs. Incompatibility with the classical model is assessed using the Wagon-Wheel functional I_{w-w} , a nonlinear expression constructed from joint and marginal probabilities of the observable outcomes in the network. This functional was derived by Fraser and Wolfe using the inflation technique and acts as a witness of causal compatibility. Every distribution generated by three independent classical sources satisfies $I_{w-w} \leq 0$. Therefore, a positive value demonstrates that the observed correlations do not admit such a model and certifies their network nonlocality. However, a nonpositive result only indicates that this functional did not provide such a certification.

On this basis, a protocol based on Complex Simultaneous Perturbation Stochastic Approximation (CSPSA) is proposed to find the joint measurements that directly maximize I_{w-w} . CSPSA is a stochastic optimization method that estimates the update direction from two evaluations of I_{w-w} , obtained by simultaneously and randomly perturbing all complex parameters. The distributed states are held fixed, while the measurement bases of the three nodes are parameterized by unitary operators in $U(4)$. The algorithm explores this space from random and independent initializations, without imposing a target distribution or using the measurements associated with the Fritz distribution as starting points. The protocol is studied for sources

distributing maximally entangled states, pure states in Schmidt form, and Werner states, as well as for configurations in which one source distributes a separable state.

In the first stage, I_{w-w} is evaluated exactly from the probabilities calculated using the Born rule. For three maximally entangled states, the protocol finds measurements that violate the Wagon-Wheel inequality and yield values approaching the Fritz reference value. The study of the Schmidt and Werner families reveals how the detection capability changes under partial entanglement and noise. No violations are found within the explored domain for configurations containing one separable source. These results establish the numerical feasibility of the method and show that detection depends jointly on the distributed states, the measurements, and the causal structure of the network.

In the second stage, the analysis is extended to the scenario in which I_{w-w} must be estimated from a finite number of measurements. For this purpose, the Classical Shadows method is incorporated to estimate properties of the quantum state through random local Pauli measurements without fully reconstructing its density matrix. Each measurement round generates a snapshot, which is a classical estimator of the state. The size of the fixed set of snapshots, denoted by N_{sh} , defines the shadow budget and therefore the number of measured copies. These data are reused during the optimization to estimate I_{w-w} for the different configurations explored by CSPSA. The simulations show that an insufficient budget can lead the algorithm to select measurements favored by statistical fluctuations, producing positive values that do not correspond to a violation when those measurements are evaluated exactly. As N_{sh} increases, this discrepancy decreases, and the measurements found yield consistent violations for the most robust configurations studied. Consequently, the shadow budget and the number of iterations must be chosen jointly, while experimental certification would require a final evaluation of the selected measurements using independent data.

Finally, this research distinguishes between the search for measurements and the certification of nonlocality. Classical Shadows guides the optimization, whereas, in an experimental implementation, the measurements found must be evaluated using frequencies obtained from independent data. This final evaluation certifies incompatibility with the classical triangle model. Overall, this work establishes a constructive procedure for directly optimizing a nonlinear witness under both ideal conditions and finite statistical data. Its results open the possibility of extending the protocol to networks with more nodes and exploring its application to distributed quantum architectures, a future quantum internet, and secure quantum communication protocols.

Tabla de contenido

Agradecimientos	III
Índice de tablas	XII
Índice de figuras	XIII
Resumen	1
Introducción	1
1. Fundamentos de la teoría de información cuántica	5
1.1. Elementos de probabilidad y estadística	5
1.1.1. Correlaciones Condicionales y la Frontera Clásica	8
1.2. Espacios de Hilbert	9
1.2.1. Espacios vectoriales	10
1.2.2. Producto interior	10
1.2.3. Completitud y espacios de Hilbert	11
1.2.4. Producto tensorial	12
1.2.5. Descomposición de Schmidt	13
1.3. Postulados de la mecánica cuántica	14
1.4. Operadores lineales en $\mathcal{H}$	17
1.5. Introducción a teoría de información cuántica	19
1.5.1. Qubits y qudits	20
1.6. Mediciones proyectivas de qubit	23
1.6.1. Matrices de Pauli	23
1.7. La matriz densidad	25
1.7.1. Ensemble de estados cuánticos	25
1.7.2. Matriz densidad reducida	28
2. Correlaciones Cuánticas	31
2.1. Estados cuánticos y correlaciones clásicas	32
2.2. Estados entrelazados	33

2.3.	La crítica EPR a la completitud de la mecánica cuántica	37
2.4.	El teorema de Bell y las correlaciones cuánticas	39
2.5.	La desigualdad CHSH y su alcance experimental	42
2.6.	El intercambio de entrelazamiento entre sistemas independientes . . .	46
2.6.1.	Ignorancia del resultado y pérdida aparente de entrelazamiento	48
2.7.	Bilocalidad	50
2.8.	Red triangular	53
2.8.1.	Estructura de la red triangular	54
2.8.2.	Correlaciones clásicas y cuánticas en la red triangular	55
2.8.3.	La distribución de Fritz	57
2.8.4.	Algunas estrategias de optimización para el estudio de la red triangular	59
3.	Optimización variacional de la desigualdad Wagon-Wheel en la red triangular	62
3.1.	El funcional Wagon-Wheel	63
3.1.1.	CSPSA como método de optimización estocástica	68
3.1.2.	Aplicación a la maximización de I_{w-w}	70
3.1.3.	Protocolo de Classical Shadows	73
3.1.4.	Protocolo de Classical Shadows	73
3.1.5.	Classical Shadows con Pauli local	76
3.1.6.	Estimación de I_{w-w} mediante Classical Shadows	77
4.	Resultados numéricos de la maximización de la desigualdad Wagon-Wheel	81
4.1.	Aplicación del protocolo al problema de optimización	81
4.2.	Evaluación exacta del funcional wagon-wheel	82
4.2.1.	Configuraciones con una fuente separable	82
4.2.2.	Configuraciones con tres fuentes equivalentes	88
4.3.	Estimación de I_{w-w} mediante Classical Shadows	96
5.	Conclusiones	106
	Bibliografía	110

Índice de tablas

1.1. Propiedades de las matrices de Pauli.	24
--	----

Índice de figuras

1.1. Esfera de Bloch	20
2.1. Escenario de la desigualdad CHSH	43
2.2. Escenario de intercambio de entrelazamiento	47
2.3. Estructura de la red triangular	55
3.1. Estructura de la red triangular	63
3.2. Inflación Wagon-Wheel de la red triangular	66
4.1. Evolución de I_{w-w} para dos fuentes maximalmente entrelazadas y una separable	84
4.2. Distribución acumulada del máximo de I_{w-w} por repetición	85
4.3. Evolución de I_{w-w} para dos fuentes con estados en descomposición de Schmidt y una fuente separable	86
4.4. Convergencia de I_{w-w} para dos estados de Werner y uno separable	87
4.5. Evolución de I_{w-w} para tres fuentes maximalmente entrelazadas	89
4.6. Aproximación de I_{w-w} al valor de Fritz	90
4.7. Convergencia de la mediana de I_{w-w} para distintos coeficientes de Schmidt	92
4.8. Convergencia de la mediana de I_{w-w} para valores intermedios de α	93
4.9. Convergencia de I_{w-w} para tres estados de Werner	94
4.10. Convergencia de I_{w-w} para estados de Werner próximos al singlete	95
4.11. Evaluación exacta de las mediciones obtenidas mediante Classical Shadows	98
4.12. Evaluación exacta de I_{w-w} para distintos valores de N_{sh}	99
4.13. Distancia de la mediana de I_{w-w}^{ex} a I_{w-w}^{Fritz}	100
4.14. Comparación entre Full Classical Shadows y la evaluación exacta	101
4.15. Comparación de los procedimientos para estados en descomposición de Schmidt	102
4.16. Comparación de los procedimientos para estados en descomposición de Schmidt	102
4.17. Full Classical Shadows para estados en descomposición de Schmidt	103

4.18. Full Classical Shadows para estados de Werner	105
---	-----

Introducción

Desde su fundación, la mecánica cuántica representó un desafío intelectual sin precedentes al romper con el determinismo clásico que había gobernado la física durante siglos. Fue una transición de la certeza absoluta a la probabilidad estadística que no fue aceptada universalmente. Figuras clave como Albert Einstein se mostraron profundamente escépticos, argumentando que una teoría que dependiera del azar intrínseco no podía ser la descripción final de la naturaleza. Para él, se trataba apenas de una aproximación incompleta de una realidad más profunda. Esta postura se formalizó en 1935, junto a Podolsky y Rosen, mediante el artículo *¿Puede considerarse completa la descripción mecánico-cuántica de la realidad física?*, que cuestionó la consistencia de la interpretación de Copenhague. El artículo mostró que la posibilidad de inferir el estado de una partícula lejana mediante la medición de su par entrelazado implicaba la existencia simultánea de valores definidos para magnitudes que el formalismo cuántico excluye medir conjuntamente. De ahí concluyeron que la teoría ofrecía entonces una descripción incompleta de la realidad física [21]. Esta discusión permaneció en el terreno filosófico hasta 1964, cuando John Bell transformó el argumento en una predicción comprobable. Bell demostró que ninguna teoría de variables ocultas locales podía reproducir la totalidad de las predicciones de la mecánica cuántica, estableciendo así una cota matemática para las correlaciones entre partículas distantes [3]. Este resultado fue generalizado en 1969 por Clauser, Horne, Shimony y Holt mediante la desigualdad CHSH, que permitió su verificación en experimentos de laboratorio sin exigir las correlaciones perfectas del planteamiento original [14]. La relevancia crítica de este criterio radica en que la violación de los límites clásicos no se manifiesta en cualquier configuración arbitraria de los aparatos. Para evidenciar el carácter no local de un estado cuántico resulta indispensable que dicho estado sea generado por la fuente como un estado entrelazado, que se identifiquen con precisión los observables involucrados y que se seleccionen ángulos específicos entre los detectores que maximicen la correlación cuántica y expongan la ruptura con el realismo local. Esta predicción fue confirmada experimentalmente en 1982 por Alain Aspect y colaboradores [1], consolidando a la desigualdad CHSH como una herramienta operacional para excluir teorías de variables ocultas locales mediante observaciones controladas, y no solo como una consecuencia formal del teorema

de Bell. La trascendencia de comprender la naturaleza no local del entrelazamiento fue reconocida en 2022 con el Premio Nobel de Física, otorgado conjuntamente a Alain Aspect, John Clauser y Anton Zeilinger [46] por sus experimentos con fotones entrelazados, que establecieron la violación de las desigualdades de Bell y sentaron las bases de la ciencia de la información cuántica.

El primer paso natural en esta dirección es el intercambio de entrelazamiento (*entanglement swapping*), un esquema en el que dos fuentes independientes, en lugar de una única fuente compartida, distribuyen pares entrelazados hacia un nodo intermedio. Una medición conjunta sobre los dos sistemas que llegan a ese nodo proyecta a los dos sistemas extremos sobre un estado entrelazado, pese a que estos nunca han interactuado ni comparten un origen común [35, 54]. La correlación resultante ya no admite una única variable oculta compartida por todo el sistema, sino que depende de la independencia entre las dos fuentes y del resultado de la medición intermedia comunicado clásicamente. Al permitir además que el nodo intermedio elija entre distintas mediciones, se obtiene la noción de bilocalidad. Una correlación es bilocal si admite una descripción clásica compatible con la independencia de las dos fuentes [7]. A diferencia del escenario bipartito, el conjunto de correlaciones bilocales no es convexo, de modo que sus testigos ya no pueden formularse, en general, como desigualdades lineales [7].

El intercambio de entrelazamiento sugiere de inmediato una pregunta más amplia. ¿Qué correlaciones son posibles cuando se admite un número arbitrario de fuentes independientes y de nodos que realizan mediciones locales? Esta generalización da lugar a lo que se conoce como una red de Bell, una estructura causal representada por un grafo cuya topología determina directamente qué correlaciones no clásicas resultan alcanzables. Un caso particularmente interesante surge al cerrar esta red en un ciclo. Es la red triangular, denotada usualmente C_3 , que extiende el intercambio de entrelazamiento a tres fuentes independientes conectadas entre tres observadores, pero sin el nodo intermedio distinguido que articulaba la correlación en aquel esquema, de modo que ningún mecanismo clásico único puede explicar las correlaciones conjuntas. Fritz [25] mostró que, precisamente por carecer de un nodo privilegiado, las correlaciones clásicamente explicables en esta red quedan sujetas a condiciones no lineales, en contraste con las desigualdades lineales del escenario bipartito estándar. Además, ilustró esta ruptura mediante una distribución explícita, construida a partir de una elección concreta de estados y observables en cada nodo, incompatible con cualquier modelo causal clásico de la topología triangular. Esta distribución, conocida desde entonces como la distribución de Fritz, se ha convertido en un ejemplo de referencia para el estudio de la no clasicidad en la red triangular. Además plantea una pregunta más general. Hasta dónde puede persistir la no localidad en esta estructura cuando los criterios de clasicidad ya no admiten una formulación lineal simple.

Dentro del espectro de las redes de Bell, la red triangular constituye su extensión cíclica más simple. El estudio de esta red ha sido abordado desde direcciones muy distintas. Entre ellas destacan las redes neuronales entrenadas como oráculos de clasicidad [32], la parametrización numérica directa de modelos locales [18], la caracterización analítica del sector simétrico del conjunto local [10, 19], y la búsqueda de realizaciones cuánticas explícitas a través del formalismo de testers cuánticos [20]. Sin embargo, ninguno de estos enfoques resuelve directamente el problema de optimizar una figura de mérito sobre el espacio completo de estados y mediciones, sin fijar de antemano una distribución objetivo. CSPSA ha demostrado versatilidad en escenarios muy diversos de información cuántica, desde la estimación de estados puros en alta dimensión hasta la discriminación de estados cuánticos desconocidos y el entrelazamiento geométrico [16, 33, 53], lo que la vuelve una herramienta natural para abordar un funcional no lineal como el Wagon-Wheel, poco habitual en la literatura sobre optimización en escenarios de Bell. Precisamente en esta línea, un trabajo previo empleó CSPSA para optimizar el conjunto de mediciones que maximiza la violación de la desigualdad CHSH sobre estados de prueba [17]. Esto permitió clasificar la naturaleza de estados cuya estructura no se conoce a priori sin recurrir a una tomografía completa. Abordar la red triangular con la misma herramienta constituye, en ese sentido, una continuación natural de esa línea de trabajo, ahora sobre un escenario de mayor complejidad.

Esta investigación tiene como objetivo establecer un protocolo autoguiado que permita testificar la presencia de entrelazamiento en una red de Bell. En particular, se busca que el protocolo determine los observables que maximizan el funcional I_{w-w} sin fijar de antemano una distribución objetivo. De este modo, cuando el valor optimizado supera la cota clásica, la violación obtenida permite certificar que las correlaciones generadas por el sistema poseen una naturaleza no local y que, por tanto, no pueden explicarse mediante un modelo causal clásico de la red triangular. La búsqueda de estos observables se realiza mediante CSPSA y se desarrolla en dos niveles progresivos. En el primero se supone acceso exacto a las probabilidades $P_{ABC}(a, b, c)$ que definen I_{w-w} , lo que permite estudiar la factibilidad del método en un entorno idealizado. En el segundo se considera un escenario más cercano a una implementación experimental, en el que dichas probabilidades deben estimarse a partir de un número finito de mediciones sobre el estado físico. Para este propósito se incorpora la técnica de Classical Shadows, que permite estimar el funcional sin reconstruir el estado cuántico completo ni realizar nuevas mediciones para cada punto examinado durante la optimización. En ambos niveles se estudian tres familias de estados distribuidos por las fuentes de la red triangular, correspondientes a estados maximalmente entrelazados, estados en descomposición de Schmidt con un parámetro común y estados de Werner con un parámetro común. Esta exploración permite identificar configuraciones cuyas correlaciones violan efectivamente la desigualdad de Wagon-Wheel y, por consiguiente, son incompatibles con cualquier modelo causal clásico de

la red triangular, pese a la complejidad física y numérica asociada con la optimización conjunta de las mediciones sobre seis qubits. La principal contribución metodológica de este trabajo consiste en mostrar que CSPA, combinado con Classical Shadows, proporciona una estrategia viable para abordar problemas de optimización no lineal y de alta dimensionalidad en redes de Bell, especialmente en situaciones donde la programación semidefinida y sus relajaciones jerárquicas resultan difíciles de aplicar directamente. Este enfoque abre una vía para estudiar topologías con un mayor número de fuentes y nodos, cuya complejidad aumenta rápidamente con el tamaño de la red.

La presente tesis se organiza de la siguiente manera. En el Capítulo 2 se introduce el escenario triangular y se presentan las descripciones clásica y cuántica de sus correlaciones. Además, se desarrolla la construcción de la distribución de Fritz y se explica cómo esta permite demostrar la existencia de correlaciones no locales en una red sin entradas externas. En el Capítulo 3 se presenta la desigualdad de Wagon-Wheel como criterio para evaluar la compatibilidad causal de las correlaciones observadas en la red triangular. También se describen los métodos de optimización empleados para buscar violaciones de esta desigualdad y se introduce el protocolo de Classical Shadows para estimar el funcional a partir de un número finito de mediciones. En el Capítulo 4 se exponen y analizan los resultados numéricos obtenidos para distintas configuraciones de estados cuánticos. Asimismo, se estudia el desempeño del procedimiento que combina CSPA con Classical Shadows y se compara con la evaluación exacta del funcional. Finalmente, en el Capítulo 5 se presentan las conclusiones generales de este trabajo y las posibles líneas de investigación futura.

Capítulo 1

Fundamentos de la teoría de información cuántica

Este capítulo establece el marco probabilístico y matemático que se utilizará para describir estados cuánticos, mediciones y correlaciones entre subsistemas. Su propósito no es desarrollar una exposición exhaustiva de la mecánica cuántica, sino fijar las herramientas necesarias para formular con precisión los escenarios de Bell y las redes cuánticas que serán estudiadas en los capítulos posteriores. Desde esta perspectiva, las nociones de independencia estadística, producto tensorial, medición proyectiva y matriz densidad aparecen como elementos estructurales para analizar correlaciones no clásicas.

1.1. Elementos de probabilidad y estadística

Definición 1.1 (Espacio muestral). El conjunto Ω de todos los posibles resultados de un experimento aleatorio es llamado *espacio muestral*.

Definición 1.2 (Evento). Un *evento* o *suceso* es cualquier colección de los posibles resultados de un experimento, esto es, cualquier subconjunto de Ω , incluido el conjunto Ω .

Si $A \subseteq \Omega$, diremos que el evento A ocurre cuando el resultado del experimento aleatorio pertenece a A . Una variable aleatoria es una función que asigna a cada resultado posible de un experimento aleatorio un valor en un conjunto determinado [28]. En lo que sigue, estas variables permitirán describir los resultados observados por distintos agentes y estudiar si dichos resultados presentan dependencia estadística. En física, particularmente en información cuántica, resulta importante analizar cómo se relacionan los resultados obtenidos por observadores espacialmente separados.

Consideremos un escenario bipartito en el que una fuente común prepara, en cada corrida experimental, dos sistemas físicos y envía uno a Alice y otro a Bob. Cada observador realiza una medición local y registra un resultado, descrito por las variables aleatorias X e Y , respectivamente. Si la preparación de la fuente introduce una relación común entre ambos sistemas, entonces los resultados observados pueden presentar dependencia estadística: conocer el resultado obtenido por Alice puede modificar la información disponible sobre el resultado obtenido por Bob. Sin embargo, la presencia de una fuente común no implica por sí sola que las correlaciones sean cuánticas o no clásicas; únicamente establece un mecanismo compartido que puede generar dependencia entre los resultados.

En contraste, si Alice y Bob reciben sistemas preparados de manera independiente, sin una fuente común ni un mecanismo que vincule ambas preparaciones, los resultados locales se describen mediante variables aleatorias independientes. En ese caso, conocer el valor de X no modifica la distribución de probabilidad asociada a Y . Esta diferencia entre una preparación común y preparaciones independientes será esencial más adelante, cuando se estudien correlaciones en escenarios de Bell y redes cuánticas con múltiples fuentes.

Para describir formalmente este tipo de situaciones, no basta con especificar las probabilidades asociadas a cada observador por separado. Es necesario introducir la distribución conjunta de los resultados, sus distribuciones marginales y las distribuciones condicionales, las cuales permiten expresar cómo cambia la descripción probabilística de una variable cuando se conoce el valor de la otra.

Definición 1.3 (Distribución conjunta). Dadas dos variables aleatorias discretas X e Y , su *distribución conjunta* $p_{XY}(x, y)$ asigna a cada par de valores (x, y) la probabilidad de que X tome el valor x y, simultáneamente, Y tome el valor y [28]:

$$p_{XY}(x, y) = \mathbb{P}(X = x, Y = y). \quad (1.1)$$

⊗

Definición 1.4 (Distribuciones marginales). Dada una distribución conjunta $p_{XY}(x, y)$, las *distribuciones marginales* de X e Y se obtienen sumando sobre la otra variable:

$$p_X(x) = \sum_y p_{XY}(x, y), \quad p_Y(y) = \sum_x p_{XY}(x, y). \quad (1.2)$$

Cada marginal describe el comportamiento estadístico de una variable por separado, ignorando la información sobre la otra. ⊗

Definición 1.5 (Distribución condicional). Dada una distribución conjunta $p_{XY}(x, y)$ con $p_X(x) > 0$, la *distribución condicional* de Y dado que X toma el valor x se define como

$$p_{Y|X}(y|x) = \frac{p_{XY}(x, y)}{p_X(x)}. \quad (1.3)$$

Esta distribución describe cómo se actualiza la descripción probabilística de Y una vez conocido el valor de X .

Con estas nociones, la independencia estadística puede formularse de manera precisa: dos variables son independientes cuando la descripción conjunta no contiene información adicional respecto de las descripciones marginales por separado.

Definición 1.6 (Independencia). Dos variables aleatorias X e Y son *independientes* si su distribución conjunta factoriza como el producto de sus marginales [28]:

$$p_{XY}(x, y) = p_X(x) p_Y(y) \quad \forall (x, y). \quad (1.4)$$

De manera equivalente, X e Y son independientes cuando conocer el valor de una no modifica la descripción probabilística de la otra, es decir,

$$p_{Y|X}(y|x) = p_Y(y) \quad \text{para todo } x \text{ con } p_X(x) > 0. \quad (1.5)$$

⊛

La independencia estadística describe el caso en que las variables no presentan dependencia entre sí. Por tanto, una forma general de introducir la noción de correlación consiste en definirla como la falla de esta factorización.

Definición 1.7 (Correlación). Dos variables aleatorias X e Y están *correlacionadas* cuando no son independientes, es decir, cuando existe al menos un par (x, y) para el cual

$$p_{XY}(x, y) \neq p_X(x) p_Y(y). \quad (1.6)$$

⊛

Una vez establecida la correlación como ausencia de independencia estadística, resulta útil introducir cantidades que permitan caracterizar parcialmente la relación entre dos variables aleatorias. Entre las más utilizadas se encuentran la covarianza y el coeficiente de correlación, los cuales describen el sentido y la intensidad de una dependencia lineal.

Definición 1.8 (Covarianza). La *covarianza* de X e Y es el número definido por

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)]. \quad (1.7)$$

donde $\mathbb{E}$ denota el operador de valor esperado, mientras que $\mu_X = \mathbb{E}[X]$ y $\mu_Y = \mathbb{E}[Y]$ corresponden a las medias de X e Y , respectivamente. Equivalentemente,

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X] \mathbb{E}[Y]. \quad (1.8)$$

La covarianza indica cómo varían conjuntamente dos variables aleatorias con respecto a sus medias [28]. ⊛

La covarianza es positiva cuando las desviaciones de X e Y respecto a sus medias tienden a tener el mismo signo, y es negativa cuando dichas desviaciones tienden a tener signos opuestos. Por tanto, su signo describe el sentido de la asociación entre las variables. Sin embargo, su magnitud depende de las escalas en que se miden X e Y , por lo que no permite comparar directamente la intensidad de la relación entre experimentos distintos. Para ello conviene introducir una versión normalizada: el coeficiente de correlación.

Definición 1.9 (Coeficiente de correlación). El *coeficiente de correlación* de X e Y se define como

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}. \quad (1.9)$$

donde

$$\sigma_X = \sqrt{\mathbb{E}[(X - \mu_X)^2]}, \quad \sigma_Y = \sqrt{\mathbb{E}[(Y - \mu_Y)^2]}$$

son las desviaciones estándar de X e Y , respectivamente, y se supone que $\sigma_X > 0$ y $\sigma_Y > 0$. Esta cantidad corresponde a una versión normalizada de la covarianza y satisface $-1 \leq \rho_{XY} \leq 1$ [28]. *

El coeficiente de correlación está definido entre los valores -1 y 1 , donde:

- Si $\rho_{XY} = 1$ entonces existe una dependencia lineal total entre las dos variables, con relación de pendiente positiva: cuando una aumenta, la otra también aumenta en proporción constante. Usualmente se llama *correlación positiva perfecta*.
- Si $\rho_{XY} = -1$ entonces existe una dependencia lineal total entre las dos variables, con relación de pendiente negativa: cuando una aumenta, la otra disminuye en proporción constante. Usualmente se llama *correlación negativa perfecta*.

Teorema 1.1. Si X e Y son variables aleatorias independientes, entonces $\text{Cov}(X, Y) = 0$ y $\rho_{XY} = 0$.

Es importante notar que el recíproco no es válido en general: la condición $\text{Cov}(X, Y) = 0$ no implica necesariamente que X e Y sean independientes. La covarianza y el coeficiente de correlación detectan dependencia lineal, pero no agotan todas las formas posibles de dependencia estadística.

1.1.1. Correlaciones Condicionales y la Frontera Clásica

En el estudio de los escenarios de Bell, el objeto central de análisis no son las variables aleatorias aisladas, sino las distribuciones de probabilidad condicionales

conjuntas, denotadas como $p(a, b|x, y)$ [9]. Estas representan la probabilidad de obtener los resultados a y b (en los sistemas de Alice y Bob) dado que se seleccionaron las configuraciones de medición x y y , respectivamente.

Desde una perspectiva estrictamente clásica, cualquier correlación entre observadores distantes que comparten una historia común o una fuente de información se modela asumiendo la existencia de una variable oculta local λ , distribuida según una densidad de probabilidad $\rho(\lambda)$ que satisface $\int \rho(\lambda)d\lambda = 1$ [3]. Bajo esta hipótesis, la distribución conjunta se factoriza de forma local:

$$p(a, b|x, y) = \int \Lambda(\lambda)p(a|x, \lambda)p(b|y, \lambda)d\lambda \quad (1.10)$$

El conjunto de todas las tablas de probabilidad $p(a, b|x, y)$ que admiten esta descomposición define un espacio geométrico convexo conocido como el polítopo local [9]. En el escenario bipartito con dos elecciones de medición y dos resultados posibles en cada nodo, el teorema de Fine [22] establece que una familia de distribuciones observables $p(a, b|x, y)$ admite un modelo de variables ocultas locales si y solo si existe una distribución conjunta global $p(a_0, a_1, b_0, b_1)$ para los resultados potenciales de todas las mediciones. Aunque las mediciones correspondientes a distintas elecciones en un mismo nodo sean incompatibles y no puedan realizarse simultáneamente, las marginales de esta distribución global deben reproducir cada una de las distribuciones observables $p(a, b|x, y)$. Esta condición también es equivalente al cumplimiento de todas las desigualdades CHSH.

Por lo tanto, la formalización de la probabilidad clásica en esta sección no constituye un mero formalismo matemático, sino que delimita geoméricamente el conjunto de correlaciones compatibles con un modelo clásico local. La violación de una desigualdad de Bell, eje central de esta investigación, implica que la distribución observada se encuentra fuera del polítopo local y, en consecuencia, no admite una descripción mediante variables ocultas locales [9].

1.2. Espacios de Hilbert

Un espacio de Hilbert es un espacio vectorial complejo dotado de un producto interior y completo respecto a la norma inducida por dicho producto. Esta estructura constituye el marco matemático básico de la mecánica cuántica: los estados físicos se representan mediante vectores o rayos en un espacio de Hilbert, los observables mediante operadores que actúan sobre él y las probabilidades de medición se calculan a partir del producto interior [34]. En esta tesis se trabajará principalmente con espacios de Hilbert de dimensión finita, isomorfos a $\mathbb{C}^d$, que son los espacios relevantes para describir qubits, qudits y sistemas compuestos finitos.

1.2.1. Espacios vectoriales

Definición 1.10 (Espacio vectorial). Sea $\mathbb{K}$ un campo y sea V un conjunto no vacío equipado con una operación de suma y una multiplicación por escalares. Se dice que V es un *espacio vectorial* sobre $\mathbb{K}$ si dichas operaciones satisfacen los axiomas usuales de asociatividad y conmutatividad de la suma, existencia de neutro e inverso aditivos, distributividad y compatibilidad con el producto por escalares [2]. Los elementos de V se denominan *vectores*. $\circledast$

Definición 1.11 (Base). Un conjunto $\{|e_i\rangle\}_{i=1}^n \subset V$ es una *base* de V si es linealmente independiente y genera todo el espacio. Esto significa que todo vector $|v\rangle \in V$ puede escribirse de manera única como

$$|v\rangle = \sum_{i=1}^n \alpha_i |e_i\rangle, \quad (1.11)$$

con $\alpha_i \in \mathbb{K}$. Si V admite una base finita de n elementos, se dice que tiene *dimensión* n [2]. $\circledast$

1.2.2. Producto interior

Definición 1.12 (Producto interior). Sea V un espacio vectorial sobre $\mathbb{C}$. Un *producto interior* sobre V es una función $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ que satisface, para todo $|a\rangle, |b\rangle, |c\rangle \in V$ y $\alpha \in \mathbb{C}$ [2]:

- (i) $\langle a|b\rangle = \overline{\langle b|a\rangle}$ simetría conjugada.
- (ii) $\langle a|\alpha b + c\rangle = \alpha \langle a|b\rangle + \langle a|c\rangle$ linealidad en la segunda entrada.
- (iii) $\langle a|a\rangle \geq 0$, con igualdad si y solo si $|a\rangle = |0\rangle$ positividad definida.

Un espacio vectorial dotado de un producto interior se denomina *espacio prehilbertiano*. $\circledast$

De las propiedades anteriores se sigue que el producto interior es antilineal en la primera entrada:

$$\langle \alpha a + b|c\rangle = \bar{\alpha} \langle a|c\rangle + \langle b|c\rangle. \quad (1.12)$$

Por tanto, $\langle \cdot, \cdot \rangle$ define una forma sesquilineal positiva. Esta estructura es la que permite calcular amplitudes, probabilidades y valores de expectación en mecánica cuántica.

El producto interior induce una norma

$$\| |a\rangle \| := \sqrt{\langle a|a\rangle}, \quad (1.13)$$

y, a partir de ella, una distancia $d(|a\rangle, |b\rangle) = \||a\rangle - |b\rangle\|$. Una desigualdad fundamental asociada a esta estructura es la desigualdad de Cauchy-Schwarz.

Teorema 1.2 (Cauchy-Schwarz). *Para todo $|a\rangle, |b\rangle \in V$,*

$$|\langle a|b\rangle|^2 \leq \langle a|a\rangle \langle b|b\rangle, \quad (1.14)$$

o equivalentemente,

$$|\langle a|b\rangle| \leq \||a\rangle\| \||b\rangle\|. \quad (1.15)$$

La igualdad se cumple si y solo si $|a\rangle$ y $|b\rangle$ son linealmente dependientes [2].

Ejemplo 1.1. *En $\mathbb{C}^n$, el producto interior estándar está definido por*

$$\langle v|w\rangle = \sum_{k=1}^n v_k^* w_k, \quad (1.16)$$

donde $|v\rangle = (v_1, \dots, v_n)^\top$ y $|w\rangle = (w_1, \dots, w_n)^\top$.

1.2.3. Completitud y espacios de Hilbert

Definición 1.13 (Sucesión de Cauchy). Una sucesión $\{|v_n\rangle\}_{n \geq 1}$ en un espacio normado V es una *sucesión de Cauchy* si para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que

$$\||v_n\rangle - |v_m\rangle\| < \varepsilon \quad (1.17)$$

para todo $n, m > N$ [2]. *

Definición 1.14 (Espacio completo). Un espacio normado V es *completo* si toda sucesión de Cauchy en V converge a un elemento de V . *

Toda sucesión convergente es de Cauchy, pero el recíproco no es válido en espacios normados arbitrarios. La completitud garantiza que los límites de sucesiones de Cauchy permanecen dentro del espacio. En dimensión finita, todo espacio prehilbertiano es completo, por lo que esta condición queda automáticamente satisfecha en los espacios $\mathbb{C}^d$ utilizados en información cuántica de dimensión finita.

Definición 1.15 (Espacio de Hilbert). Un *espacio de Hilbert* $\mathcal{H}$ es un espacio vectorial sobre $\mathbb{C}$ dotado de un producto interior y completo respecto a la norma inducida por dicho producto [34]. *

En información cuántica, el caso central corresponde a espacios de Hilbert de dimensión finita, $\mathcal{H} = \mathbb{C}^d$. En este contexto, los estados puros se representan mediante vectores normalizados de $\mathcal{H}$, mientras que las mediciones y observables se representan mediante operadores que actúan sobre dicho espacio [34].

Definición 1.16 (Base ortonormal). Un conjunto $\{|e_i\rangle\}_{i=1}^d \subset \mathcal{H}$ es una *base ortonormal* de $\mathcal{H}$ si

$$\langle e_i | e_j \rangle = \delta_{ij}, \quad (1.18)$$

y todo $|v\rangle \in \mathcal{H}$ se escribe de manera única como

$$|v\rangle = \sum_{i=1}^d \langle e_i | v \rangle |e_i\rangle. \quad (1.19)$$

⊛

La existencia de una base ortonormal permite escribir la relación de completitud

$$\sum_{i=1}^d |e_i\rangle\langle e_i| = \mathbb{1}, \quad (1.20)$$

la cual será utilizada de manera recurrente para descomponer estados, operadores y probabilidades de medición.

1.2.4. Producto tensorial

Los sistemas cuánticos compuestos requieren una operación que combine los espacios de Hilbert de sus subsistemas en un único espacio de estados. Esta operación es el producto tensorial. Su relevancia es central, ya que permite describir sistemas multipartitos, mediciones locales, correladores y estados entrelazados [34].

Definición 1.17 (Producto tensorial de espacios de Hilbert). Sean $\mathcal{H}_A$ y $\mathcal{H}_B$ dos espacios de Hilbert de dimensiones d_A y d_B , con bases ortonormales $\{|i\rangle_A\}_{i=1}^{d_A}$ y $\{|j\rangle_B\}_{j=1}^{d_B}$. El *producto tensorial* $\mathcal{H}_A \otimes \mathcal{H}_B$ es el espacio de Hilbert de dimensión $d_A d_B$ cuya base ortonormal está formada por los vectores

$$\{|i\rangle_A \otimes |j\rangle_B\}_{i=1,\dots,d_A; j=1,\dots,d_B}. \quad (1.21)$$

Un estado normalizado general en $\mathcal{H}_A \otimes \mathcal{H}_B$ puede escribirse como

$$|\Psi\rangle_{AB} = \sum_{i=1}^{d_A} \sum_{j=1}^{d_B} c_{ij} |i\rangle_A \otimes |j\rangle_B, \quad \sum_{i,j} |c_{ij}|^2 = 1. \quad (1.22)$$

⊛

Un caso particular es el estado producto

$$|\Psi\rangle_{AB} = |\psi\rangle_A \otimes |\phi\rangle_B, \quad (1.23)$$

con $|\psi\rangle_A \in \mathcal{H}_A$ y $|\phi\rangle_B \in \mathcal{H}_B$. En este caso, el estado global factoriza en estados asociados a cada subsistema. En general, no todo estado de $\mathcal{H}_A \otimes \mathcal{H}_B$ admite esta factorización; los estados que no la admiten se denominan entrelazados y serán estudiados con mayor detalle en el Capítulo 2 [34].

Definición 1.18 (Producto tensorial de operadores). Dados un operador $\hat{A}$ sobre $\mathcal{H}_A$ y un operador $\hat{B}$ sobre $\mathcal{H}_B$, el producto tensorial $\hat{A} \otimes \hat{B}$ es el operador sobre $\mathcal{H}_A \otimes \mathcal{H}_B$ definido por su acción sobre estados producto:

$$(\hat{A} \otimes \hat{B})(|\psi\rangle_A \otimes |\phi\rangle_B) = (\hat{A}|\psi\rangle_A) \otimes (\hat{B}|\phi\rangle_B), \quad (1.24)$$

y extendido por linealidad a todo $\mathcal{H}_A \otimes \mathcal{H}_B$ [34]. ⊗

Los observables locales, que actúan sobre un único subsistema, se escriben como $\hat{A} \otimes \mathbb{1}_B$ o $\mathbb{1}_A \otimes \hat{B}$. En escenarios de Bell, los correladores bipartitos toman la forma

$$\langle \hat{A}_i \otimes \hat{B}_j \rangle = \langle \Psi | \hat{A}_i \otimes \hat{B}_j | \Psi \rangle, \quad (1.25)$$

lo que permite describir correlaciones estadísticas entre mediciones locales realizadas sobre subsistemas distintos.

Esta construcción se extiende de manera natural a n subsistemas. Si cada subsistema tiene espacio de Hilbert $\mathcal{H}_k$, el espacio de estados compuesto es

$$\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_n, \quad (1.26)$$

con dimensión $\prod_{k=1}^n d_k$ y base formada por los vectores $|i_1\rangle_1 \otimes \cdots \otimes |i_n\rangle_n$. En el caso de la red triangular estudiada en el Capítulo 3, el espacio de estados conjunto se construye mediante productos tensoriales de los espacios asociados a los sistemas distribuidos por las fuentes independientes. Esta estructura tensorial describe el espacio total de estados, mientras que la independencia de las fuentes corresponde a una condición adicional sobre la preparación física del sistema.

1.2.5. Descomposición de Schmidt

La construcción del producto tensorial permite escribir cualquier estado bipartito como una combinación lineal general de vectores base, pero no toda esa generalidad es igualmente relevante. Para caracterizar cuándo un estado puro admite o no una descripción producto, conviene disponer de una forma canónica que reduzca la doble suma sobre i, j a una única suma. Esta forma canónica es la descomposición de Schmidt.

Teorema 1.3 (Descomposición de Schmidt). *Sea $|\Psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ un estado normalizado, con $\dim \mathcal{H}_A = d_A$ y $\dim \mathcal{H}_B = d_B$. Existen bases ortonormales $\{|i\rangle_A\}_{i=1}^{d_A}$ de $\mathcal{H}_A$ y $\{|i\rangle_B\}_{i=1}^{d_B}$ de $\mathcal{H}_B$, y coeficientes reales no negativos $\alpha_1, \dots, \alpha_d$, con $d = \min(d_A, d_B)$, tales que*

$$|\Psi\rangle_{AB} = \sum_{i=1}^d \sqrt{\alpha_i} |i\rangle_A \otimes |i\rangle_B, \quad \sum_{i=1}^d \alpha_i = 1. \quad (1.27)$$

Los coeficientes $\{\alpha_i\}$, llamados coeficientes de Schmidt, quedan determinados de forma única por $|\Psi\rangle_{AB}$ (salvo su orden). Las bases $\{|i\rangle_A\}$ y $\{|i\rangle_B\}$, llamadas bases de Schmidt, no son en general únicas cuando algún α_i está degenerado [34].

La demostración se obtiene aplicando la descomposición en valores singulares a la matriz de coeficientes c_{ij} que expresa $|\Psi\rangle_{AB}$ en una base producto arbitraria. Los valores singulares de dicha matriz son las raíces $\sqrt{\alpha_i}$, mientras que las bases de Schmidt corresponden a las bases en las que dicha matriz se diagonaliza [4, 34]. El número de coeficientes α_i estrictamente positivos se denomina rango de Schmidt, y no depende de la base producto elegida para calcularlo.

El rango de Schmidt provee un criterio directo de separabilidad para estados puros. $|\Psi\rangle_{AB}$ admite la factorización $|\psi\rangle_A \otimes |\phi\rangle_B$ si y solo si su rango de Schmidt es igual a 1, en cuyo caso $\alpha_1 = 1$. Asimismo, la descomposición de Schmidt entrega directamente los estados reducidos de cada subsistema. Trazando $|\Psi\rangle_{AB} \langle\Psi|_{AB}$ sobre B se obtiene

$$\rho_A = \text{Tr}_B (|\Psi\rangle\langle\Psi|) = \sum_{i=1}^d \alpha_i |i\rangle\langle i|_A, \quad (1.28)$$

y de forma análoga para ρ_B , de modo que ambos estados reducidos comparten el mismo espectro $\{\alpha_i\}$, aunque actúen sobre espacios de Hilbert distintos.

1.3. Postulados de la mecánica cuántica

La mecánica cuántica puede formularse a partir de un conjunto de postulados que establecen cómo se representan los estados físicos, qué objetos matemáticos describen las cantidades medibles y cómo se calculan las probabilidades asociadas a los resultados de medición [15, 34]. En esta sección se presenta una versión sintética de estos postulados, utilizando el lenguaje de espacios de Hilbert introducido previamente. El objetivo no es desarrollar una discusión exhaustiva de sus fundamentos, sino fijar el marco operacional que permitirá describir estados compuestos, mediciones locales y correlaciones cuánticas en los capítulos posteriores.

Primer postulado: A tiempo fijo t_0 , un estado puro de un sistema físico se representa mediante un vector de estado $|\Psi(t_0)\rangle$, perteneciente a un espacio de Hilbert llamado *espacio de estados* $\mathcal{H}$. Este vector debe estar normalizado, por lo que satisface

$$\langle \Psi(t_0) | \Psi(t_0) \rangle = 1. \quad (1.29)$$

Más generalmente, un estado cuántico se describe mediante un operador densidad $\rho(t_0)$ que actúa sobre $\mathcal{H}$, es semidefinido positivo y tiene traza unitaria,

$$\rho(t_0) \geq 0, \quad \text{Tr}[\rho(t_0)] = 1. \quad (1.30)$$

Un estado es puro cuando su operador densidad puede escribirse como

$$\rho(t_0) = |\Psi(t_0)\rangle \langle \Psi(t_0)|. \quad (1.31)$$

En caso contrario, se denomina *estado mixto*. A partir de estas representaciones se calculan las probabilidades asociadas a los posibles resultados de medición permitidos por la teoría [34, 39].

La estructura lineal de $\mathcal{H}$ da lugar al principio de superposición. Si $|\Psi_1\rangle$ y $|\Psi_2\rangle$ son vectores de estado, entonces toda combinación lineal no nula $\alpha |\Psi_1\rangle + \beta |\Psi_2\rangle$, una vez normalizada, representa también un posible estado puro. Además, dos vectores $|\Psi\rangle$ y $e^{i\varphi} |\Psi\rangle$, con $\varphi \in \mathbb{R}$, representan el mismo estado físico, ya que todas las probabilidades de medición son invariantes bajo una fase global. Por tanto, un estado puro queda determinado por la clase de equivalencia de $|\Psi\rangle$ módulo fases globales [15, 39].

Segundo postulado: Cada cantidad física medible A es descrita por un operador $\hat{A}$ que actúa sobre el espacio de estados $\mathcal{H}$. Este operador, denominado *observable*, es lineal y hermítico, lo que significa que coincide con su operador adjunto, $\hat{A}^\dagger = \hat{A}$ [39, 41].

Tercer postulado: El espacio de estados de un sistema compuesto por n subsistemas con espacios de Hilbert $\mathcal{H}_1, \dots, \mathcal{H}_n$ es el producto tensorial

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots \otimes \mathcal{H}_n. \quad (1.32)$$

Si cada subsistema k se encuentra en el estado $|\psi_k\rangle \in \mathcal{H}_k$, entonces el estado producto asociado al sistema compuesto es $|\psi_1\rangle \otimes \dots \otimes |\psi_n\rangle$. En general, no todo estado puro de $\mathcal{H}$ admite esta factorización. Los estados puros que no factorizan se denominan *entrelazados* [4, 34]. Los observables que actúan localmente sobre el subsistema k se representan en $\mathcal{H}$ como

$$\mathbb{1}_1 \otimes \dots \otimes \hat{A}_k \otimes \dots \otimes \mathbb{1}_n, \quad (1.33)$$

donde $\hat{A}_k$ actúa sobre $\mathcal{H}_k$ y $\mathbb{1}_j$ denota el operador identidad sobre $\mathcal{H}_j$ para cada $j \neq k$.

Cuarto postulado: Cuando se mide un observable $\hat{A}$ sobre un sistema descrito por el operador densidad ρ , los resultados posibles son los autovalores a_n de $\hat{A}$. Si P_n denota el proyector sobre el subespacio propio asociado a a_n , la probabilidad de obtener dicho resultado es

$$P(a_n) = \text{Tr}(\rho P_n). \quad (1.34)$$

Los proyectores satisfacen $\sum_n P_n = \mathbb{1}$, por lo que $\sum_n P(a_n) = 1$. Para un estado puro $\rho = |\Psi\rangle\langle\Psi|$, esta expresión se reduce a

$$P(a_n) = \langle\Psi| P_n |\Psi\rangle. \quad (1.35)$$

En el caso no degenerado, donde $P_n = |a_n\rangle\langle a_n|$, se obtiene

$$P(a_n) = |\langle a_n|\Psi\rangle|^2. \quad (1.36)$$

Dado que $\hat{A}$ es hermítico, sus autovalores son reales. Su valor de expectación es

$$\langle\hat{A}\rangle = \sum_n a_n P(a_n) = \text{Tr}(\rho\hat{A}). \quad (1.37)$$

Para un estado puro, $\langle\hat{A}\rangle = \langle\Psi|\hat{A}|\Psi\rangle$. Esta cantidad representa el promedio estadístico de los resultados obtenidos al medir un conjunto de sistemas preparados en el mismo estado [36, 39].

Quinto postulado: Si la medición de $\hat{A}$ sobre el estado $|\Psi\rangle$ da como resultado el autovalor a_n , el estado inmediatamente después de la medición queda dado por

$$|\Psi\rangle \mapsto \frac{P_n |\Psi\rangle}{\sqrt{\langle\Psi| P_n |\Psi\rangle}}, \quad (1.38)$$

siempre que $\langle\Psi| P_n |\Psi\rangle \neq 0$. El estado resultante pertenece al subespacio propio asociado al resultado observado [15, 34].

Sexto postulado: La evolución temporal de un estado cuántico $|\Psi(t)\rangle$ está gobernada por la ecuación de Schrödinger

$$i\hbar \frac{d}{dt} |\Psi(t)\rangle = \hat{H}(t) |\Psi(t)\rangle, \quad (1.39)$$

donde $\hbar$ es la constante de Planck reducida y $\hat{H}(t)$ es el observable asociado a la energía total del sistema, llamado *operador Hamiltoniano*. Para un Hamiltoniano independiente del tiempo, la solución formal puede escribirse como

$$|\Psi(t)\rangle = U(t, t_0) |\Psi(t_0)\rangle, \quad U(t, t_0) = e^{-i\hat{H}(t-t_0)/\hbar}, \quad (1.40)$$

donde $U(t, t_0)$ es un operador unitario. Por ello, la evolución temporal preserva la norma del estado [39, 41].

Los postulados anteriores fijan el lenguaje mínimo para describir estados, observables, mediciones y sistemas compuestos. En lo que sigue se introducen con mayor detalle los operadores lineales y sus propiedades algebraicas, ya que serán utilizados para formular valores de expectación, proyectores y correlaciones cuánticas.

1.4. Operadores lineales en $\mathcal{H}$

Los postulados de la mecánica cuántica utilizan operadores lineales para representar observables, transformaciones y procesos de medición. En esta sección se introducen las propiedades algebraicas básicas de estos operadores, con énfasis en aquellas que serán necesarias para formular valores de expectación, proyectores, descomposiciones espectrales y correladores cuánticos.

Definición 1.19 (Operador lineal). Sea $\mathcal{H}$ un espacio de Hilbert de dimensión d . Un *operador lineal* $\hat{A} : \mathcal{H} \rightarrow \mathcal{H}$ es una aplicación que satisface

$$\hat{A}(\alpha |\psi\rangle + \beta |\phi\rangle) = \alpha \hat{A} |\psi\rangle + \beta \hat{A} |\phi\rangle \quad (1.41)$$

para todo $|\psi\rangle, |\phi\rangle \in \mathcal{H}$ y $\alpha, \beta \in \mathbb{C}$. Dada una base ortonormal $\{|e_i\rangle\}_{i=1}^d$, el operador queda determinado por sus elementos de matriz

$$A_{ij} = \langle e_i | \hat{A} | e_j \rangle. \quad (1.42)$$

En consecuencia, su acción sobre un vector arbitrario $|\psi\rangle \in \mathcal{H}$ puede escribirse como

$$\hat{A} |\psi\rangle = \sum_{i,j} A_{ij} \langle e_j | \psi \rangle |e_i\rangle. \quad (1.43)$$

⊗

Definición 1.20 (Operador adjunto). El *operador adjunto* $\hat{A}^\dagger$ de $\hat{A}$ es el único operador que satisface

$$\langle \phi, \hat{A} \psi \rangle = \langle \hat{A}^\dagger \phi, \psi \rangle \quad (1.44)$$

para todo $|\psi\rangle, |\phi\rangle \in \mathcal{H}$. En notación de Dirac, esta relación puede escribirse como

$$\langle \phi | \hat{A} | \psi \rangle = \langle \hat{A}^\dagger \phi | \psi \rangle. \quad (1.45)$$

En términos matriciales, $(A^\dagger)_{ij} = A_{ji}^*$, es decir, $\hat{A}^\dagger$ es la transpuesta conjugada de $\hat{A}$. A partir del operador adjunto se definen tres clases de operadores de especial relevancia en mecánica cuántica [34, 41]:

- **Hermítico:** $\hat{A}^\dagger = \hat{A}$. Sus autovalores son reales y, en dimensión finita, sus autovectores pueden elegirse formando una base ortonormal de $\mathcal{H}$. Los observables del Segundo postulado son operadores de este tipo.
- **Unitario:** $\hat{U}^\dagger \hat{U} = \hat{U} \hat{U}^\dagger = \mathbb{1}$. Estos operadores preservan el producto interior y la norma. La evolución temporal del Sexto postulado genera operadores de este tipo.

- **Proyector:** $\hat{P}^2 = \hat{P} = \hat{P}^\dagger$. Los proyectores representan proyecciones ortogonales sobre subespacios de $\mathcal{H}$. El operador P_n del Quinto postulado es el proyector sobre el subespacio propio asociado al resultado a_n .

⊛

Definición 1.21 (Producto exterior). Dados $|\phi\rangle, |\psi\rangle \in \mathcal{H}$, el *producto exterior* $|\phi\rangle\langle\psi|$ es el operador lineal definido por su acción sobre un vector arbitrario $|\chi\rangle \in \mathcal{H}$:

$$|\phi\rangle\langle\psi||\chi\rangle = \langle\psi|\chi\rangle |\phi\rangle. \quad (1.46)$$

En la representación matricial, sus elementos están dados por

$$(|\phi\rangle\langle\psi|)_{ij} = \phi_i \psi_j^*. \quad (1.47)$$

Este operador permite escribir de forma compacta proyectores, relaciones de completitud y descomposiciones espectrales. En particular, si $\{|e_i\rangle\}_{i=1}^d$ es una base ortonormal de $\mathcal{H}$, entonces

$$\sum_{i=1}^d |e_i\rangle\langle e_i| = \mathbb{1}. \quad (1.48)$$

⊛

Definición 1.22 (Operador normal). Un operador $\hat{A}$ se denomina *normal* si conmuta con su operador adjunto, es decir,

$$\hat{A}^\dagger \hat{A} = \hat{A} \hat{A}^\dagger. \quad (1.49)$$

⊛

Teorema 1.4 (Descomposición espectral). *Todo operador hermítico $\hat{A}$ sobre un espacio de Hilbert complejo de dimensión finita admite la descomposición*

$$\hat{A} = \sum_n a_n |a_n\rangle\langle a_n|, \quad (1.50)$$

donde $\{a_n\}$ son los autovalores reales de $\hat{A}$ y $\{|a_n\rangle\}$ una base ortonormal de autovectores [2, 4]. En el caso no degenerado, los operadores $|a_n\rangle\langle a_n|$ son los proyectores asociados a los posibles resultados de medición. Por tanto, la regla de Born puede escribirse como

$$P(a_n) = \langle\Psi| |a_n\rangle\langle a_n| |\Psi\rangle = |\langle a_n|\Psi\rangle|^2. \quad (1.51)$$

Todo operador que admite una descomposición espectral en una base ortonormal satisface $\hat{A}^\dagger \hat{A} = \hat{A} \hat{A}^\dagger$ y es, por tanto, normal. Recíprocamente, todo operador normal sobre un espacio de Hilbert complejo de dimensión finita admite una descomposición de esta forma.

Definición 1.23 (Traza). La *traza* de un operador $\hat{A}$ se define como la suma de sus elementos diagonales en cualquier base ortonormal $\{|e_i\rangle\}_{i=1}^d$:

$$\text{Tr}[\hat{A}] = \sum_{i=1}^d \langle e_i | \hat{A} | e_i \rangle. \quad (1.52)$$

El resultado es independiente de la base elegida. Entre las propiedades que serán utilizadas en esta tesis se encuentran:

■ **Linealidad:**

$$\text{Tr}[\alpha \hat{A} + \beta \hat{B}] = \alpha \text{Tr}[\hat{A}] + \beta \text{Tr}[\hat{B}]. \quad (1.53)$$

■ **Propiedad cíclica:**

$$\text{Tr}[\hat{A}\hat{B}] = \text{Tr}[\hat{B}\hat{A}]. \quad (1.54)$$

■ **Relación con autovalores:**

$$\text{Tr}[\hat{A}] = \sum_n a_n, \quad (1.55)$$

donde $\{a_n\}$ son los autovalores de $\hat{A}$, contados con su multiplicidad.

En particular, el valor de expectación del Cuarto postulado puede reescribirse como

$$\langle \hat{A} \rangle = \text{Tr} [|\Psi\rangle\langle\Psi| \hat{A}], \quad (1.56)$$

forma que anticipa el formalismo de la matriz densidad introducido más adelante [34, 51]. Esta notación será especialmente útil en los escenarios de Bell, donde los correladores entre mediciones locales sobre sistemas compuestos se expresan mediante trazas. $\otimes$

Las herramientas introducidas en esta sección permiten expresar valores de expectación, proyectores y correladores de forma compacta. En particular, serán fundamentales en los escenarios de Bell, donde las correlaciones entre observadores se escriben mediante operadores locales sobre espacios compuestos, dando lugar a expresiones del tipo $\text{Tr}[\rho(\hat{A} \otimes \hat{B})]$.

1.5. Introducción a teoría de información cuántica

La teoría de información cuántica estudia cómo la información puede ser codificada, procesada y transmitida mediante sistemas que obedecen las leyes de la mecánica cuántica. La unidad básica de información cuántica es el *qubit*, análogo cuántico del bit clásico, cuya estructura permite describir superposiciones, mediciones incompatibles y correlaciones entre sistemas compuestos [34]. En esta sección se introducen los sistemas cuánticos de dos y d niveles, que servirán como bloques locales para formular estados, mediciones y correladores en escenarios de Bell y redes cuánticas.

1.5.1. Qubits y qudits

Qubits

Consideremos un espacio de Hilbert $\mathcal{H}_2$ de dimensión dos. Una base ortonormal estándar para este espacio es la *base computacional*, formada por los vectores

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (1.57)$$

Un estado puro general de un sistema cuántico de dos niveles, o *qubit*, puede escribirse como

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle, \quad (1.58)$$

donde $\alpha, \beta \in \mathbb{C}$ son amplitudes de probabilidad que satisfacen la condición de normalización $|\alpha|^2 + |\beta|^2 = 1$.

Como dos vectores que difieren solo en una fase global representan el mismo estado físico, todo estado puro de un qubit puede parametrizarse como

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right) |0\rangle + e^{i\varphi} \sin\left(\frac{\theta}{2}\right) |1\rangle, \quad (1.59)$$

con $\theta \in [0, \pi]$ y $\varphi \in [0, 2\pi)$. Esta parametrización permite representar cada estado puro como un punto sobre la esfera de Bloch [4, 34].

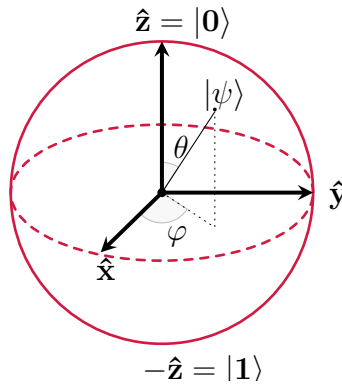


Figura 1.1: Esfera de Bloch. **Fuente:** Elaboración propia.

Esta esfera se denomina *esfera de Bloch* y permite visualizar geoméricamente cualquier estado puro de un qubit. Los puntos especiales de la esfera tienen una interpretación física directa:

- **Polos:** el polo norte ($\theta = 0$) corresponde al estado $|0\rangle$ y el polo sur ($\theta = \pi$) al estado $|1\rangle$, que son los autoestados de σ_z con autovalores $+1$ y -1 respectivamente.

- **Puntos ecuatoriales:** los puntos del ecuador ($\theta = \pi/2$) corresponden a superposiciones balanceadas de $|0\rangle$ y $|1\rangle$. En particular, $|+\rangle$ y $|-\rangle$ (con $\varphi = 0$ y $\varphi = \pi$) son autoestados de σ_x , mientras que $(|0\rangle \pm i|1\rangle)/\sqrt{2}$ (con $\varphi = \pi/2$ y $\varphi = 3\pi/2$) son autoestados de σ_y .

Los estados $|+\rangle$ y $|-\rangle$, que conforman la *base de Hadamard*, tienen la representación explícita:

$$|+\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad (1.60)$$

$$|-\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (1.61)$$

Una propiedad fundamental de la esfera de Bloch es que dos estados antipodales son siempre ortogonales: si $|\psi\rangle$ corresponde al punto (θ, φ) , el punto antipodal $(\pi - \theta, \varphi + \pi)$ corresponde al estado $|\psi^\perp\rangle$, con $\langle\psi|\psi^\perp\rangle = 0$.

A cada punto de la esfera le corresponde un vector unitario

$$\hat{n} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta), \quad (1.62)$$

llamado *vector de Bloch*, cuyos componentes coinciden con los valores de expectación de las matrices de Pauli:

$$\langle\sigma_x\rangle = \sin \theta \cos \varphi, \quad \langle\sigma_y\rangle = \sin \theta \sin \varphi, \quad \langle\sigma_z\rangle = \cos \theta. \quad (1.63)$$

El vector de Bloch $\hat{n}$ codifica los valores de expectación de las matrices de Pauli en el estado $|\psi\rangle$. Por esta razón, elegir una medición proyectiva binaria sobre un qubit equivale a elegir una dirección en la esfera de Bloch. Esta interpretación será utilizada más adelante para parametrizar observables locales en escenarios de Bell. La representación anterior corresponde a estados puros; los estados mixtos, introducidos posteriormente mediante el formalismo de la matriz densidad, se representan por puntos en el interior de la esfera.

La descripción anterior corresponde al caso más simple de un sistema cuántico de dos niveles. Sin embargo, muchos sistemas físicos y protocolos de información cuántica requieren espacios de Hilbert de dimensión mayor. En ese contexto, un *qudit* es la extensión d -dimensional de un qubit [51].

Qudits

Consideremos un espacio de Hilbert $\mathcal{H}_d$. La base computacional de $\mathcal{H}_d$ está formada por d vectores:

Un *qudit* es una extensión d -dimensional de un qubit. Consideremos un espacio de Hilbert $\mathcal{H}_d$. La base computacional de $\mathcal{H}_d$ está formada por d vectores:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad |d-1\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \quad (1.64)$$

Sean $\alpha_i \in \mathbb{C}$, con $i \in \{0, 1, \dots, d-1\}$, se define un estado cuántico llamado *qudit* de d niveles como

$$|\phi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle + \dots + \alpha_{d-1} |d-1\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{d-1} \end{pmatrix} \in \mathbb{C}^d, \quad (1.65)$$

o de manera más compacta,

$$|\phi\rangle = \sum_{i=0}^{d-1} \alpha_i |i\rangle, \quad (1.66)$$

donde los coeficientes α_i son amplitudes de probabilidad que satisfacen la relación de normalización $\sum_{i=0}^{d-1} |\alpha_i|^2 = 1$.

La descripción de sistemas compuestos requiere combinar los espacios de Hilbert asociados a cada subsistema. En particular, cuando se consideran dos qubits, el espacio de estados resultante es el producto tensorial $\mathcal{H}_2 \otimes \mathcal{H}_2$, de dimensión cuatro. La base computacional de este espacio está formada por los vectores

$$|00\rangle, \quad |01\rangle, \quad |10\rangle, \quad |11\rangle, \quad (1.67)$$

donde se emplea la notación compacta $|ij\rangle := |i\rangle_A \otimes |j\rangle_B$, con $i, j \in \{0, 1\}$.

Un estado general normalizado en $\mathcal{H}_2 \otimes \mathcal{H}_2$ es

$$|\Psi\rangle_{AB} = \sum_{i,j \in \{0,1\}} c_{ij} |ij\rangle, \quad \sum_{i,j} |c_{ij}|^2 = 1, \quad (1.68)$$

con $c_{ij} \in \mathbb{C}$. Cuando el estado admite la factorización $|\Psi\rangle_{AB} = |\psi\rangle_A \otimes |\phi\rangle_B$, se denomina *estado producto* y no presenta entrelazamiento entre los subsistemas. En caso contrario, el estado es *entrelazado*: las mediciones locales sobre A y B pueden exhibir correlaciones que no se explican únicamente a partir de una factorización del estado global [4, 34]. El estudio de estos estados y de sus propiedades de correlación será el objeto central del **Capítulo 2**.

1.6. Mediciones proyectivas de qubit

Los postulados de la mecánica cuántica establecen que medir un observable $\hat{A}$ sobre un estado $|\Psi\rangle$ produce uno de sus autovalores con una probabilidad determinada por la regla de Born. Para sistemas de dos niveles, los observables con autovalores ± 1 son de especial importancia, ya que describen mediciones binarias. Este tipo de medición aparece de manera natural en los escenarios de Bell, donde las correlaciones se construyen a partir de resultados locales asociados a distintas configuraciones de medición. En esta sección se introducen las matrices de Pauli y los observables proyectivos de qubit que serán utilizados posteriormente para formular dichos correladores.

1.6.1. Matrices de Pauli

Para un sistema cuántico de dos niveles, con espacio de estados $\mathcal{H} = \mathbb{C}^2$, los observables se representan mediante operadores hermíticos de 2×2 . Todo operador hermítico sobre este espacio puede escribirse como una combinación lineal real de la identidad y de las tres matrices de Pauli:

$$M = c_0 \mathbb{1} + c_1 \sigma_1 + c_2 \sigma_2 + c_3 \sigma_3, \quad c_j \in \mathbb{R}. \quad (1.69)$$

cuya elección estándar está dada por

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (1.70)$$

donde $\sigma_1, \sigma_2, \sigma_3$ reciben el nombre de *matrices de Pauli* [39]. Se utilizan también las etiquetas $\sigma_x := \sigma_1$, $\sigma_y := \sigma_2$, $\sigma_z := \sigma_3$, asociadas a las mediciones de espín sobre los ejes x, y y z . Todas las matrices de Pauli son hermíticas ($\sigma_j^\dagger = \sigma_j$), tienen traza nula $\text{Tr}[\sigma_j] = 0$ y autovalores $+1$ y -1 . Por el teorema de descomposición espectral,

$$\sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0|, \quad \sigma_y = -i|0\rangle\langle 1| + i|1\rangle\langle 0|, \quad \sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|. \quad (1.71)$$

Sus acciones sobre la base computacional $\{|0\rangle, |1\rangle\}$ y la base de Hadamard $\{|+\rangle, |-\rangle\}$ son:

Teorema 1.5 (Propiedades algebraicas). *Las matrices de Pauli satisfacen las siguientes relaciones:*

- **Cuadrado:** $\sigma_j^2 = \mathbb{1}$ para $j = x, y, z$.
- **Anticonmutación:** $\{\sigma_i, \sigma_j\} := \sigma_i \sigma_j + \sigma_j \sigma_i = 2\delta_{ij} \mathbb{1}$.

Estado	σ_x	σ_y	σ_z
$ 0\rangle$	$ 1\rangle$	$i 1\rangle$	$ 0\rangle$
$ 1\rangle$	$ 0\rangle$	$-i 0\rangle$	$- 1\rangle$
$ +\rangle$	$ +\rangle$	$-i -\rangle$	$ -\rangle$
$ -\rangle$	$- -\rangle$	$i +\rangle$	$ +\rangle$

Tabla 1.1: Acción de las matrices de Pauli sobre la base computacional $\{|0\rangle, |1\rangle\}$ y la base de Hadamard $\{|+\rangle, |-\rangle\}$. **Fuente:** Elaboración propia.

- **Conmutación:** $[\sigma_i, \sigma_j] := \sigma_i \sigma_j - \sigma_j \sigma_i = 2i\varepsilon_{ijk}\sigma_k$, donde ε_{ijk} es el símbolo de Levi-Civita.
- **Traza del producto:** $\text{Tr}[\sigma_i \sigma_j] = 2\delta_{ij}$.

Todo observable de qubit con autovalores ± 1 puede escribirse como proyección de las matrices de Pauli sobre una dirección arbitraria. Dado un vector unitario $\hat{n} = (n_x, n_y, n_z) \in \mathbb{R}^3$, el observable

$$\hat{A}(\hat{n}) = \hat{n} \cdot \vec{\sigma} := n_x \sigma_x + n_y \sigma_y + n_z \sigma_z \quad (1.72)$$

es hermítico y tiene autovalores ± 1 . Sus autovectores corresponden a los puntos $\pm \hat{n}$ de la esfera de Bloch, de modo que elegir una configuración de medición equivale a elegir una dirección $\hat{n}$. El valor de expectación es

$$\langle \hat{A}(\hat{n}) \rangle = \langle \psi | \hat{n} \cdot \vec{\sigma} | \psi \rangle = \hat{n} \cdot \hat{n}_\psi, \quad (1.73)$$

donde $\hat{n}_\psi = (\langle \sigma_x \rangle, \langle \sigma_y \rangle, \langle \sigma_z \rangle)$ es el vector de Bloch del estado $|\psi\rangle$. En los escenarios de Bell, los correladores adoptan precisamente esta forma: $E_{ij} = \langle \Psi | \hat{A}_i \otimes \hat{B}_j | \Psi \rangle$, con $\hat{A}_i = \hat{n}_i \cdot \vec{\sigma}$ y $\hat{B}_j = \hat{m}_j \cdot \vec{\sigma}$ para distintas elecciones de $\hat{n}_i$ y $\hat{m}_j$.

Definición 1.24 (Proyectores de qubit). Dado el observable $\hat{A}(\hat{n}) = \hat{n} \cdot \vec{\sigma}$ con autovalores ± 1 , los operadores

$$\Pi_{\pm}(\hat{n}) = \frac{1 \pm \hat{n} \cdot \vec{\sigma}}{2} \quad (1.74)$$

son los *proyectores* sobre los subespacios propios de autovalor ± 1 . Satisfacen las relaciones de completitud, idempotencia y ortogonalidad:

$$\Pi_+(\hat{n}) + \Pi_-(\hat{n}) = 1, \quad \Pi_{\pm}^2(\hat{n}) = \Pi_{\pm}(\hat{n}), \quad \Pi_+(\hat{n}) \Pi_-(\hat{n}) = 0, \quad (1.75)$$

y permiten recuperar el observable como $\hat{A}(\hat{n}) = \Pi_+(\hat{n}) - \Pi_-(\hat{n})$. La probabilidad de obtener el resultado ± 1 al medir $\hat{A}(\hat{n})$ sobre el estado $|\psi\rangle$ es

$$P(\pm 1) = \langle \psi | \Pi_{\pm}(\hat{n}) | \psi \rangle = \frac{1 \pm \hat{n} \cdot \hat{n}_\psi}{2}, \quad (1.76)$$

donde $\hat{n}_\psi$ es el vector de Bloch de $|\psi\rangle$. Inmediatamente después de la medición, el estado colapsa al autoestado correspondiente, en acuerdo con el Quinto postulado.

⊗

Definición 1.25 (Compatibilidad de observables). Dos observables $\hat{A}$ y $\hat{B}$ son *compatibles* si pueden medirse simultáneamente sin perturbarse mutuamente, lo que ocurre si y solo si conmutan:

$$[\hat{A}, \hat{B}] = 0. \quad (1.77)$$

En caso contrario se denominan *incompatibles*.

⊗

Las matrices de Pauli son incompatibles por pares: de las relaciones de conmutación $[\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k$ se sigue que ningún par (σ_i, σ_j) con $i \neq j$ conmuta. Para observables generales de qubit, la compatibilidad entre $\hat{A}(\hat{n}) = \hat{n} \cdot \vec{\sigma}$ y $\hat{A}(\hat{m}) = \hat{m} \cdot \vec{\sigma}$ está determinada por el conmutador

$$[\hat{n} \cdot \vec{\sigma}, \hat{m} \cdot \vec{\sigma}] = 2i(\hat{n} \times \hat{m}) \cdot \vec{\sigma}. \quad (1.78)$$

Por tanto, estos observables son compatibles si y solo si $\hat{n} \times \hat{m} = 0$, es decir, cuando las direcciones $\hat{n}$ y $\hat{m}$ son paralelas o antiparalelas. En escenarios de Bell, la posibilidad de elegir mediciones locales incompatibles es un ingrediente esencial para obtener correlaciones que no admiten una descripción clásica en términos de variables ocultas locales.

1.7. La matriz densidad

El formalismo de estados puros introducido en los postulados resulta insuficiente en dos situaciones fundamentales. La primera ocurre cuando el estado del sistema no se conoce con certeza, sino que pertenece a un ensemble estadístico de posibles preparaciones. La segunda aparece al describir subsistemas de un sistema compuesto, especialmente cuando el estado global presenta entrelazamiento. En ambos casos, el objeto matemático adecuado es el *operador densidad* o *matriz densidad* ρ . Este formalismo permite calcular valores de expectación y correladores mediante expresiones de traza, lo que será esencial para describir mediciones locales en escenarios de Bell y redes cuánticas [34].

1.7.1. Ensemble de estados cuánticos

Suponga que un sistema cuántico puede estar en cualquiera de los estados $|\psi_k\rangle$ con probabilidades $p_k \geq 0$ y $\sum_k p_k = 1$. Al conjunto $\{p_k, |\psi_k\rangle\}$ se le denomina *ensemble de estados puros*.

Definición 1.26 (Operador densidad o matriz densidad). El *operador densidad* o *matriz densidad* asociado al ensemble $\{p_k, |\psi_k\rangle\}$ es el operador

$$\rho := \sum_k p_k |\psi_k\rangle\langle\psi_k|. \quad (1.79)$$

Un operador ρ es un operador densidad si y solo si satisface simultáneamente:

- **Hermítico:** $\rho^\dagger = \rho$.
- **Semidefinido positivo:** $\rho \geq 0$, es decir, todos sus autovalores son no negativos.
- **Traza unitaria:** $\text{Tr}[\rho] = 1$.

⊛

Una consecuencia directa de esta definición es la expresión general para el valor de expectación de un observable $\hat{A}$ en el estado ρ :

$$\langle\hat{A}\rangle = \text{Tr}[\rho \hat{A}]. \quad (1.80)$$

Para un estado puro $\rho = |\psi\rangle\langle\psi|$, esta expresión se reduce a $\langle\hat{A}\rangle = \langle\psi|\hat{A}|\psi\rangle$, recuperando la definición del Cuarto postulado. Para sistemas compuestos, los correladores de Bell adoptan la forma $E_{ij} = \text{Tr}[\rho(\hat{A}_i \otimes \hat{B}_j)]$, que será la expresión central del **Capítulo 2**.

Teorema 1.6 (Propiedades del operador densidad). Sean ρ_1 y ρ_2 dos operadores densidad y sea $\hat{U}$ un operador unitario. Entonces:

- i) $\rho_1 \otimes \rho_2$ es un operador densidad.
- ii) $\hat{U}\rho\hat{U}^\dagger$ es un operador densidad.

Definición 1.27 (Pureza). La *pureza* de un estado cuántico descrito por el operador densidad ρ se define como

$$\gamma(\rho) := \text{Tr}[\rho^2] \in \left[\frac{1}{d}, 1\right], \quad (1.81)$$

donde $d = \dim \mathcal{H}$. Un estado con $\gamma(\rho) = 1$ es *puro*; un estado con $\gamma(\rho) < 1$ es *mixto*. El estado maximalmente mixto $\rho = \mathbb{1}/d$ tiene la pureza mínima $\gamma = 1/d$. ⊛

La pureza cuantifica el grado de mezcla de un estado cuántico. Los estados puros satisfacen $\text{Tr}[\rho^2] = 1$, mientras que los estados mixtos satisfacen $\text{Tr}[\rho^2] < 1$. En particular, el estado maximalmente mixto $\rho = \mathbb{1}/d$ alcanza el valor mínimo $\gamma = 1/d$. Sus valores extremos caracterizan dos clases de estados cualitativamente distintas:

Teorema 1.7 (Propiedades de un estado puro). *Sea $\rho = |\psi\rangle\langle\psi|$ un estado puro. Entonces:*

- i) $\rho^2 = \rho$
- ii) $\text{Tr}[\rho] = 1$
- iii) $\text{Tr}[\rho^2] = 1$
- iv) $\rho \geq 0$

Teorema 1.8 (Propiedades de un estado mixto). *Sea ρ un estado mixto, es decir, una mezcla estadística no trivial de estados puros. Entonces:*

- i) $\rho^2 \neq \rho$
- ii) $\text{Tr}[\rho] = 1$
- iii) $\text{Tr}[\rho^2] < 1$
- iv) $\rho \geq 0$

Para el caso particular de un qubit ($d = 2$), el operador densidad admite una parametrización geométrica directa. Todo operador densidad sobre $\mathcal{H} = \mathbb{C}^2$ puede escribirse de manera única como

$$\rho = \frac{\mathbb{1} + \vec{r} \cdot \vec{\sigma}}{2}, \quad \vec{r} = (r_x, r_y, r_z) \in \mathbb{R}^3, \quad |\vec{r}| \leq 1, \quad (1.82)$$

donde $r_k = \text{Tr}[\rho \sigma_k]$ son los valores de expectación de las matrices de Pauli. El vector $\vec{r}$ es el *vector de Bloch* de ρ [4]:

- Si $|\vec{r}| = 1$, el estado es puro y corresponde a un punto sobre la superficie de la esfera de Bloch.
- Si $|\vec{r}| < 1$, el estado es mixto y corresponde a un punto en el interior de la esfera.
- El estado maximalmente mixto $\rho = \mathbb{1}/2$ corresponde al origen $\vec{r} = \vec{0}$.

Esta representación extiende la descripción geométrica de la esfera de Bloch al formalismo del operador densidad. En este caso, los estados puros se ubican sobre la superficie de la esfera, mientras que los estados mixtos se representan por puntos en su interior. La condición $|\vec{r}| \leq 1$ garantiza que ρ sea semidefinido positivo, y la normalización $\text{Tr}[\rho] = 1$ queda incorporada en la estructura de la parametrización.

1.7.2. Matriz densidad reducida

Cuando un sistema compuesto se encuentra en un estado ρ^{AB} , puede ser necesario describir solo uno de sus subsistemas. Por ejemplo, las mediciones realizadas exclusivamente sobre A quedan completamente determinadas por un operador densidad definido sobre $\mathcal{H}_A$, llamado *matriz densidad reducida* de A . Este operador se obtiene a partir del estado global mediante la *traza parcial*, operación que permite eliminar los grados de libertad del subsistema no observado [34].

Definición 1.28 (Matriz densidad reducida). Sean A y B dos subsistemas con espacio de estados conjunto $\mathcal{H}_A \otimes \mathcal{H}_B$, y sea ρ^{AB} el operador densidad del sistema compuesto. La *matriz densidad reducida* del subsistema A se define como

$$\rho^A := \text{Tr}_B(\rho^{AB}), \quad (1.83)$$

y análogamente, $\rho^B := \text{Tr}_A(\rho^{AB})$ es la matriz densidad reducida del subsistema B .

⊗

Definición 1.29 (Traza parcial). Sean $\{|j\rangle_B\}$ una base ortonormal de $\mathcal{H}_B$ y $\{|i\rangle_A\}$ una base ortonormal de $\mathcal{H}_A$. La *traza parcial* sobre el subsistema B está definida por

$$\text{Tr}_B(\rho^{AB}) := \sum_j (\mathbb{1}_A \otimes \langle j|_B) \rho^{AB} (\mathbb{1}_A \otimes |j\rangle_B), \quad (1.84)$$

y la traza parcial sobre el subsistema A está definida por

$$\text{Tr}_A(\rho^{AB}) := \sum_i (\langle i|_A \otimes \mathbb{1}_B) \rho^{AB} (|i\rangle_A \otimes \mathbb{1}_B). \quad (1.85)$$

⊗

La matriz densidad no solo permite describir estados mixtos desde un punto de vista formal, sino que también proporciona el lenguaje natural para estudiar sistemas cuánticos abiertos. En una implementación física real, un sistema cuántico rara vez evoluciona de manera completamente aislada. La interacción no controlada con grados de libertad externos, errores de compuerta, pérdidas, fluctuaciones del entorno y procesos de medición imperfectos producen transformaciones que, en general, no pueden describirse únicamente mediante evoluciones unitarias sobre el sistema de interés.

Desde esta perspectiva, la traza parcial tiene una interpretación operacional relevante. Si un sistema compuesto AB se encuentra descrito por un estado ρ_{AB} y solo se tiene acceso al subsistema A , entonces el estado efectivo observado es

$$\rho_A = \text{Tr}_B(\rho_{AB}). \quad (1.86)$$

Esta operación representa la pérdida de información asociada a ignorar los grados de libertad del subsistema B . En particular, si B se interpreta como un entorno, entonces la traza parcial permite describir cómo un estado puro global puede inducir un estado mixto local. Esta es una de las formas más simples de entender el origen formal de la decoherencia.

De manera más general, la evolución efectiva de un sistema cuántico abierto se describe mediante canales cuánticos. Un canal cuántico es una aplicación lineal

$$\mathcal{E} : \rho \longmapsto \mathcal{E}(\rho), \quad (1.87)$$

que transforma matrices densidad en matrices densidad. Para representar una evolución físicamente admisible, dicha aplicación debe ser completamente positiva y preservar la traza. Estas aplicaciones se conocen como mapas CPTP, por sus siglas en inglés: *completely positive and trace preserving maps* [34, 51].

Una representación usual de un canal cuántico está dada por operadores de Kraus $\{K_\mu\}_\mu$, de modo que

$$\mathcal{E}(\rho) = \sum_\mu K_\mu \rho K_\mu^\dagger, \quad \sum_\mu K_\mu^\dagger K_\mu = \mathbb{1}. \quad (1.88)$$

La primera expresión describe cómo el estado cambia bajo la acción del canal, mientras que la segunda condición garantiza la conservación de la traza, es decir, $\text{Tr}(\mathcal{E}(\rho)) = 1$ siempre que $\text{Tr}(\rho) = 1$ [34]. En este lenguaje, procesos como la despolarización, la pérdida de fase, la relajación de amplitud y los errores de medición pueden incorporarse directamente en la descripción del estado.

Este punto es especialmente importante en la era NISQ, *Noisy Intermediate-Scale Quantum*, donde los dispositivos cuánticos disponibles poseen un número intermedio de qubits, pero todavía están afectados por ruido, decoherencia y fidelidades imperfectas de operación [37]. En este régimen, los algoritmos cuánticos no pueden analizarse únicamente mediante estados puros y circuitos ideales. Es necesario describir cómo el ruido modifica las correlaciones observadas y cómo las mediciones reales se apartan del caso ideal.

Los algoritmos cuánticos variacionales, o VQA, se sitúan precisamente en este contexto [11]. En ellos se prepara una familia parametrizada de estados cuánticos, se realizan mediciones sobre el hardware y se utiliza un optimizador clásico para ajustar los parámetros del circuito. Por tanto, la cantidad relevante no es solo el estado ideal $|\psi(\theta)\rangle$, sino también el estado efectivo

$$\rho(\theta) = \mathcal{E}(|\psi(\theta)\rangle\langle\psi(\theta)|), \quad (1.89)$$

donde $\mathcal{E}$ representa, de forma compacta, la acción acumulada del ruido y de las imperfecciones experimentales.

En consecuencia, el formalismo de matriz densidad será esencial para conectar la descripción matemática de estados cuánticos con escenarios físicamente realizables. En particular, permitirá estudiar correlaciones cuánticas en presencia de subsistemas no observados, fuentes independientes, ruido efectivo y mediciones locales imperfectas. Esta conexión será relevante más adelante al analizar redes de Bell, donde las correlaciones observadas dependen no solo del estado preparado, sino también de los observables elegidos y de las limitaciones físicas del dispositivo utilizado para simular o implementar el escenario.

Capítulo 2

Correlaciones Cuánticas

Las correlaciones son el lenguaje natural con el que describimos sistemas físicos compuestos porque capturan directamente la información accesible cuando se realizan mediciones sobre partes distintas de un mismo sistema. En la práctica, rara vez accedemos a las variables intrínsecas de cada subsistema. Lo que se registra experimentalmente son resultados de medición y, sobre todo, las relaciones estadísticas entre ellos. En este sentido, una parte sustantiva de la física moderna puede leerse como una clasificación de qué patrones de correlación son compatibles con ciertas hipótesis de descripción, como independencia, causalidad, simetrías o restricciones impuestas por la dinámica.

En información cuántica, las correlaciones no se consideran solo como un rasgo estadístico, sino como una ventana directa a la estructura del estado físico. En particular, cuando el sistema se encuentra en un estado entrelazado, las correlaciones entre mediciones locales pueden ser incompatibles con una descripción basada en mezclas clásicas de estados producto. Por ello, además de cuantificar su magnitud, resulta esencial caracterizar su origen y determinar si admiten un modelo local con incertidumbre clásica, o si requieren coherencia cuántica y entrelazamiento. En el primer caso, se asume que los resultados están determinados por variables compartidas antes de la medición y que la correlación surge solo por desconocimiento sobre dichas variables, sin necesidad de recursos cuánticos no locales. En el segundo, la correlación refleja que el estado compuesto no puede reconstruirse a partir de descripciones independientes de A y B , y que las estadísticas dependen de la estructura global del estado. Esta distinción enlaza con la paradoja de EPR, donde el entrelazamiento permite predecir resultados remotos sin interacción directa, tensionando la compatibilidad entre localidad y una noción de propiedades físicas preexistentes [21].

2.1. Estados cuánticos y correlaciones clásicas

Antes de introducir correlaciones genuinamente cuánticas, conviene precisar que el formalismo de la teoría cuántica también describe situaciones donde las correlaciones observadas tienen un origen estrictamente clásico.

Consideremos un sistema bipartito formado por A y B . La fuente genera un estado compuesto y distribuye sus partes hacia ambos subsistemas. En cada realización experimental, la fuente no prepara necesariamente el mismo estado, sino que selecciona al azar, según una variable clásica λ , una de varias preparaciones posibles, con $p(\lambda) \geq 0$ y $\sum_{\lambda} p(\lambda) = 1$. En este caso, el estado del sistema puede escribirse como

$$\rho_{AB} = \sum_{\lambda} p(\lambda) \rho_A^{(\lambda)} \otimes \rho_B^{(\lambda)}, \quad (2.1)$$

el cual corresponde a un estado separable [50]. La correlación aparece porque A y B reciben estados que dependen de la misma etiqueta clásica λ .

Desde una perspectiva operativa, esta descomposición se refleja directamente en las estadísticas de medición. Si A y B eligen configuraciones de medición x e y y obtienen resultados a y b , la distribución conjunta puede escribirse como

$$p(a, b \mid x, y) = \sum_{\lambda} p(\lambda) p(a \mid x, \lambda) p(b \mid y, \lambda). \quad (2.2)$$

La ecuación (2.1) describe la estructura del estado como una mezcla de preparaciones producto, mientras que la ecuación (2.2) describe las estadísticas observables asociadas a mediciones locales. En esta última, λ funciona como una causa común clásica que correlaciona las respuestas locales, de modo que la correlación observada surge del promedio sobre λ y no requiere coherencia cuántica compartida entre A y B . Esta factorización constituirá el marco de referencia que permitirá formular modelos locales de variables ocultas en escenarios de Bell [3].

Un caso particular de estado separable ocurre cuando los estados locales $\{\rho_A^{(\lambda)}\}$ conmutan entre sí, y lo mismo sucede con $\{\rho_B^{(\lambda)}\}$. En tal situación existe una base producto local común que diagonaliza simultáneamente todas las preparaciones posibles, de modo que la matriz densidad no presenta términos fuera de la diagonal en dicha base,

$$\langle ij \mid \rho_{AB} \mid k\ell \rangle = 0 \quad \text{si} \quad (i, j) \neq (k, \ell). \quad (2.3)$$

Este caso particular se ilustra con el siguiente ejemplo mínimo de estado mixto clásicamente correlacionado,

$$\rho_{cc} = \frac{1}{2} (|00\rangle\langle 00| + |11\rangle\langle 11|). \quad (2.4)$$

Este estado puede interpretarse como una mezcla estadística en la que la fuente prepara $|00\rangle$ con probabilidad $1/2$ y $|11\rangle$ con probabilidad $1/2$. En consecuencia, si ambos observadores miden en la base computacional, los resultados quedan perfectamente correlacionados, mientras que la preparación no requiere entrelazamiento.

En la base computacional $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, el estado ρ_{cc} adopta una forma diagonal. En efecto, su representación matricial es

$$\rho_{cc} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}, \quad (2.5)$$

lo que muestra explícitamente la ausencia de términos fuera de la diagonal en esta base. Esta forma diagonal refleja que no existe coherencia cuántica entre las ramas $|00\rangle$ y $|11\rangle$, sino únicamente una incertidumbre clásica sobre cuál de ellas fue preparada. El contraste con los estados entrelazados surge precisamente cuando el estado compuesto no puede entenderse como una mezcla estadística de estados producto.

Esta forma diagonal no es incidental: las correlaciones de ρ_{cc} están confinadas a la base de preparación, y se degradan frente a elecciones de medición en bases distintas, como se ilustrará en la sección siguiente. Esta limitación de las correlaciones clásicas es, precisamente, lo que motiva la introducción de los estados entrelazados.

2.2. Estados entrelazados

Un estado bipartito ρ_{AB} se define como *entrelazado* cuando sus correlaciones no pueden explicarse como el resultado de una mezcla estadística de preparaciones locales independientes. Formalmente, esto significa que ρ_{AB} no admite una descomposición convexa en estados producto, es decir, no existen números $p_k \geq 0$ con $\sum_k p_k = 1$ y estados locales $\rho_A^{(k)}, \rho_B^{(k)}$ tales que

$$\rho_{AB} = \sum_k p_k \rho_A^{(k)} \otimes \rho_B^{(k)}. \quad (2.6)$$

Si tal representación existe, el estado se denomina separable; si no existe, se clasifica como entrelazado [30, 50].

Esta definición es operativa porque separa dos escenarios: en un estado separable, puede imaginarse un procedimiento de preparación donde, en cada corrida, una variable clásica k selecciona un par de estados locales $\rho_A^{(k)}$ y $\rho_B^{(k)}$, de modo que cualquier correlación observada proviene del promedio sobre k . En cambio, en un estado entrelazado no es posible reconstruir las estadísticas como un promedio de preparaciones producto. En ese sentido, el estado conjunto contiene una estructura global que no se reduce a información local más azar clásico compartido.

Para estados puros $|\psi\rangle_{AB}$, la condición se simplifica: $|\psi\rangle_{AB}$ es entrelazado si no puede escribirse como producto $|\psi\rangle_A \otimes |\phi\rangle_B$. Equivalentemente, al calcular el estado reducido $\rho_A = \text{Tr}_B(|\psi\rangle\langle\psi|)$, este resulta mixto, lo cual indica que el sistema A no tiene una descripción pura por sí solo [34]. Un ejemplo estándar es el denominado *estado de Bell*,

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}} (|0_A 0_B\rangle + |1_A 1_B\rangle), \quad (2.7)$$

el cual no factoriza. Este tipo de estados exhibe correlaciones que no pueden entenderse como el resultado de una selección aleatoria de estados producto, y por ello pueden manifestarse de manera consistente bajo distintas elecciones de medición local.

Un rasgo distintivo del entrelazamiento es que ciertas correlaciones pueden mantenerse bajo distintas elecciones de medición local. Por ejemplo, para el estado de Bell (2.7), si ambos miden en la base computacional (medición Z), los resultados quedan perfectamente correlacionados, con valor de expectación $\langle Z \otimes Z \rangle = 1$; y si ambos miden en la base X , la correlación se mantiene, con $\langle X \otimes X \rangle = 1$. A diferencia de esto, el estado mixto clásicamente correlacionado ρ_{cc} también da $\langle Z \otimes Z \rangle = 1$, pero al medir en la base X la correlación desaparece, $\langle X \otimes X \rangle = 0$.

Esta diferencia muestra que el entrelazamiento no solo modifica la intensidad de las correlaciones, sino también su comportamiento frente a distintas elecciones de medición local. Precisamente esta característica hizo que los estados entrelazados adquirieran un papel central en la discusión sobre la completitud de la teoría cuántica. Si las mediciones realizadas sobre un subsistema permiten predecir con certeza resultados asociados a otro subsistema espacialmente separado, surge la pregunta de si tales propiedades estaban previamente determinadas o si la descripción cuántica debe entenderse como una descripción completa del sistema compuesto. Esta tensión conceptual constituye el punto de partida de la paradoja EPR [21].

Estados en descomposición de Schmidt

Un caso particularmente relevante de estado bipartito puro es aquel que se define directamente a través de sus coeficientes de Schmidt, en lugar de mediante una preparación física específica. Retomando el teorema de descomposición de Schmidt establecido en el Capítulo 1, todo estado bipartito puro $|\psi\rangle_{AB}$ admite la forma

$$|\psi\rangle_{AB} = \sum_{i=1}^d \sqrt{\alpha_i} |i\rangle_A \otimes |i\rangle_B, \quad \sum_{i=1}^d \alpha_i = 1, \quad (2.8)$$

con $d = \min(d_A, d_B)$ [34]. Fijar directamente los coeficientes $\{\alpha_i\}$ permite parametrizar familias de estados con un grado de entrelazamiento controlado, sin necesidad de especificar las bases de Schmidt $\{|i\rangle_A\}$, $\{|i\rangle_B\}$ asociadas.

La principal característica de estos estados es que su entrelazamiento queda completamente determinado por los coeficientes $\{\alpha_i\}$. La entropía de entrelazamiento, definida como la entropía de von Neumann de cualquiera de los estados reducidos ρ_A , ρ_B ,

$$S(|\psi\rangle_{AB}) = -\text{Tr}(\rho_A \log_2 \rho_A) = -\sum_{i=1}^d \alpha_i \log_2 \alpha_i, \quad (2.9)$$

es, en consecuencia, una función exclusiva de estos coeficientes [34]. Esta cantidad se anula si y solo si algún $\alpha_i = 1$ (estado producto). Alcanza su valor máximo $\log_2 d$ cuando todos los coeficientes son iguales, $\alpha_i = 1/d$ para todo i , caso que se denomina estado maximalmente entrelazado.

Para el caso particular de dos qubits, con un único coeficiente de Schmidt independiente $\alpha \in [0, 1]$,

$$|\Psi(\alpha)\rangle = \sqrt{\alpha} |0_1 0_2\rangle + \sqrt{1-\alpha} |1_1 1_2\rangle, \quad (2.10)$$

el entrelazamiento puede caracterizarse también mediante la concurrencia,

$$C(\alpha) = 2\sqrt{\alpha(1-\alpha)} \in [0, 1], \quad (2.11)$$

que se anula en los extremos $\alpha = 0, 1$ (estados producto) y alcanza su valor máximo $C = 1$ en $\alpha = 1/2$, el caso maximalmente entrelazado [52]. A diferencia de la entropía de entrelazamiento, la concurrencia admite una extensión directa a estados mixtos de dos qubits mediante la construcción de Wootters. Esto la vuelve una herramienta particularmente utilizada en la caracterización de entrelazamiento bipartito.

Estados de Werner

Los estados de Werner [50] constituyen una familia paramétrica de estados mixtos de dos qubits, introducidos originalmente como el primer ejemplo de un estado mixto entrelazado que admite un modelo local de variables ocultas para mediciones proyectivas.

Definición 2.1 (Estado de Werner). Para dos qubits, el estado de Werner con parámetro $\lambda \in [0, 1]$ se define como la mezcla convexa entre el estado de Bell $|\Phi^+\rangle$ y el estado maximalmente mixto,

$$\rho_\lambda = \lambda |\Phi^+\rangle\langle\Phi^+| + \frac{1-\lambda}{4} \mathbb{1}, \quad (2.12)$$

donde $\mathbb{1}$ es el operador identidad en el espacio de dos qubits [50].

El parámetro λ cuantifica la proporción de ruido blanco mezclado con el estado maximalmente entrelazado: $\lambda = 1$ recupera el estado de Bell puro, mientras que $\lambda = 0$ corresponde al estado maximalmente mixto, que es separable de forma trivial. El estado de Werner ρ_λ es separable si y solo si $\lambda \leq 1/3$, y entrelazado para $\lambda > 1/3$ [50].

El entrelazamiento de ρ_λ puede cuantificarse mediante la extensión de la concurrencia a estados mixtos, definida por Wootters como

$$C(\rho) = \max(0, \sqrt{\mu_1} - \sqrt{\mu_2} - \sqrt{\mu_3} - \sqrt{\mu_4}), \quad (2.13)$$

donde $\mu_1 \geq \mu_2 \geq \mu_3 \geq \mu_4$ son los autovalores, en orden decreciente, del operador $\rho(\sigma_y \otimes \sigma_y) \rho^*(\sigma_y \otimes \sigma_y)$, con ρ^* el complejo conjugado de ρ en la base computacional [52]. Para el estado de Werner, esta expresión se reduce a la forma cerrada

$$C(\rho_\lambda) = \max\left(0, \frac{3\lambda - 1}{2}\right), \quad (2.14)$$

que se anula para $\lambda \leq 1/3$, consistentemente con el criterio de separabilidad enunciado arriba, y alcanza su valor máximo $C = 1$ en $\lambda = 1$, el caso del estado de Bell puro.

Los estados de Werner poseen además una propiedad estructural que motivó su introducción original: son, por construcción, los únicos estados bipartitos invariantes bajo la acción simétrica de cualquier unitaria local,

$$(U \otimes U) \rho_\lambda (U \otimes U)^\dagger = \rho_\lambda \quad \forall U, \quad (2.15)$$

de modo que la forma dada para ρ_λ es, en realidad, la consecuencia de exigir esta invariancia y no una elección arbitraria [50].

El resultado central del artículo de Werner es, sin embargo, de naturaleza distinta a la sola caracterización de separabilidad: para el rango $1/3 < \lambda \leq 1/2$, el estado ρ_λ ya es entrelazado, pero admite igualmente un modelo local de variables ocultas para cualquier medición proyectiva, de modo que jamás viola una desigualdad de Bell de este tipo pese a estar entrelazado [50]. La violación efectiva de la desigualdad CHSH, presentada más adelante en este capítulo, solo ocurre para $\lambda > 1/\sqrt{2}$, con un valor máximo de la función CHSH dado por $S(\rho_\lambda) = 2\sqrt{2}\lambda$. Estos tres regímenes, separable ($\lambda \leq 1/3$), entrelazado pero localmente explicable ($1/3 < \lambda \leq 1/2$), y no local ($\lambda > 1/\sqrt{2}$), muestran que el entrelazamiento y la no localidad son propiedades distintas, siendo la primera estrictamente necesaria pero no suficiente para la segunda.

2.3. La crítica EPR a la completitud de la mecánica cuántica

En 1935, Albert Einstein, Boris Podolsky y Nathan Rosen publicaron el artículo *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?*, en el cual cuestionaron si la función de onda proporciona una descripción completa de la realidad física [21]. Su crítica no estaba dirigida a la capacidad de la mecánica cuántica para predecir correctamente los resultados experimentales. De hecho, los autores distinguieron explícitamente dos cuestiones diferentes al evaluar una teoría física: si la teoría es correcta, en el sentido de que sus predicciones concuerdan con la experiencia, y si la descripción que proporciona es completa. Es esta segunda cuestión la que constituye el núcleo de su análisis.

Para precisar el significado de completitud, EPR establecieron como condición necesaria que todo elemento de la realidad física debe poseer una contraparte dentro de la teoría. Sin embargo, para aplicar esta condición era necesario determinar bajo qué circunstancias puede atribuirse realidad física a una determinada magnitud. Con este propósito, propusieron el siguiente criterio: si es posible predecir con certeza, es decir, con probabilidad igual a la unidad, el valor de una cantidad física sin perturbar de ninguna manera el sistema, entonces existe un elemento de realidad física correspondiente a dicha cantidad. Los propios autores señalaron que este criterio no pretende constituir una definición exhaustiva de la realidad, sino una condición suficiente para reconocer ciertos elementos de ella.

A partir de este criterio, EPR consideraron dos magnitudes físicas representadas por operadores no conmutativos. De acuerdo a los principios de la mecánica cuántica,

el conocimiento preciso de una de estas magnitudes impide atribuir simultáneamente un valor definido a la otra mediante una misma función de onda. De esta observación surge una alternativa fundamental, o bien la descripción de la realidad proporcionada por la función de onda no es completa, o bien dos cantidades físicas cuyos operadores no conmutan no pueden poseer realidad simultánea.

Para examinar esta alternativa, Einstein, Podolsky y Rosen consideraron dos sistemas, denominados I y II, que interactúan durante un cierto intervalo de tiempo y posteriormente dejan de hacerlo. Una vez finalizada la interacción, el sistema compuesto permanece descrito por una única función de onda Ψ , aunque ya no exista interacción entre sus partes. La medición de una determinada magnitud física sobre el sistema I permite entonces asignar al sistema II una función de onda condicionada por el resultado obtenido. Si sobre el sistema I se elige medir una magnitud diferente, el mismo sistema II puede quedar asociado a otra función de onda distinta.

El ejemplo central del artículo considera dos partículas correlacionadas en posición y momento. La función de onda conjunta es elegida de modo que, mediante una medición apropiada sobre la primera partícula, sea posible predecir con certeza el momento de la segunda o, alternativamente, su posición. En la notación utilizada por EPR, el estado conjunto puede escribirse como

$$\Psi(x_1, x_2) = \int_{-\infty}^{\infty} e^{\frac{2\pi i}{h}(x_1 - x_2 + x_0)p} dp. \quad (2.16)$$

Si se mide el momento de la primera partícula y se obtiene el valor p , puede predecirse con certeza que el momento de la segunda es $-p$. Por otro lado, si se mide la posición de la primera partícula y se obtiene el valor x , puede predecirse con certeza que la posición de la segunda es $x + x_0$. De acuerdo con el criterio de realidad formulado por EPR, tanto el momento como la posición de la segunda partícula deben entonces corresponder a elementos de realidad física.

El punto crucial de la crítica reside en que la elección de qué magnitud medir sobre el sistema I puede realizarse cuando ambos sistemas han dejado de interactuar. En consecuencia, EPR sostienen que ninguna acción realizada sobre el primero puede producir un cambio real en el segundo. Por tanto, la realidad física atribuida a la segunda partícula no debería depender de cuál de las dos mediciones se decida efectuar sobre la primera. Sin embargo, el formalismo cuántico asocia las dos predicciones posibles con funciones de onda diferentes y con observables que no conmutan.

EPR llegan así a su conclusión principal. Si se supone que la función de onda proporciona una descripción completa de la realidad física, entonces dos cantidades representadas por operadores no conmutativos no deberían poseer realidad simultánea. No obstante, el análisis de los dos sistemas permite, según su criterio de

realidad, atribuir simultáneamente realidad a ambas magnitudes. En consecuencia, la descripción cuántica proporcionada por la función de onda no puede considerarse completa.

La conclusión de EPR no establece cuál debería ser la teoría más completa ni demuestra la existencia de una formulación alternativa específica. Los propios autores dejaron abierta esta cuestión, aunque expresaron su convicción de que una descripción más completa era posible, en la que cada sistema poseyera propiedades adicionales, no contenidas en la función de onda, que determinaran de antemano los resultados de cualquier medición. Su crítica identificó así una tensión fundamental entre la completitud de la función de onda, la atribución de elementos de realidad física y la independencia de sistemas que ya no interactúan. Décadas más tarde, Bell retomaría precisamente este problema, formalizando esa idea bajo el nombre de variables ocultas, para determinar si una descripción más completa y compatible con una estructura local podía reproducir todas las predicciones de la mecánica cuántica [3].

Esta formulación original, sin embargo, se plantea en términos de variables continuas como la posición y el momento, cuyo tratamiento resulta menos manejable al momento de derivar predicciones cuantitativas contrastables experimentalmente. En 1951, David Bohm propuso una versión equivalente del argumento EPR utilizando un sistema de dos partículas de espín $1/2$ en el estado singlete, reemplazando las variables continuas originales por observables discretos [5]. Esta reformulación preservó la estructura conceptual de la paradoja, la posibilidad de predecir con certeza, sobre un sistema distante, el resultado de una entre varias mediciones incompatibles, pero la trasladó a un escenario de variables discretas. Fue precisamente esta versión, y no la original con variables continuas, la que permitió a Bell formular su teorema en términos de observables de espín y derivar una desigualdad cuantitativa contrastable experimentalmente.

2.4. El teorema de Bell y las correlaciones cuánticas

En 1964, John S. Bell publicó el artículo *On the Einstein Podolsky Rosen Paradox*, en el cual dio forma cuantitativa a la cuestión de completitud planteada por EPR y examinó si la mecánica cuántica podía complementarse mediante una descripción más completa del sistema físico [3]. Su contribución consistió en transformar una discusión principalmente conceptual en un problema cuantitativo, al establecer restricciones que permiten analizar la compatibilidad entre las predicciones cuánticas y

ciertas hipótesis sobre la naturaleza de la realidad física.

Bell consideró la versión del experimento EPR propuesta para partículas de espín $1/2$. Una fuente prepara dos partículas en el estado singlete

$$|\psi^-\rangle = \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle), \quad (2.17)$$

y las envía hacia dos regiones espacialmente separadas. En una de ellas se mide la componente del espín en una dirección representada por el vector unitario $\vec{a}$, mientras que en la otra se mide en una dirección $\vec{b}$. Cada medición produce uno de los resultados $+1$ o -1 .

Siguiendo la motivación de EPR, Bell consideró la posibilidad de que la función de onda no constituyera una descripción completa del sistema. En tal caso, podrían existir variables adicionales, representadas colectivamente por λ , que especificaran propiedades físicas no contenidas en el estado cuántico. Estas cantidades reciben el nombre de *variables ocultas*, no porque deban ser necesariamente inaccesibles, sino porque no aparecen explícitamente en la descripción proporcionada por la función de onda.

Una teoría de variables ocultas se denomina local cuando el resultado obtenido en cada región depende únicamente de la configuración de medición local y de la variable λ . Los resultados individuales pueden representarse entonces mediante las funciones

$$A(\vec{a}, \lambda) = \pm 1, \quad B(\vec{b}, \lambda) = \pm 1. \quad (2.18)$$

La condición de localidad implica que $A(\vec{a}, \lambda)$ no depende de la dirección distante $\vec{b}$ y que $B(\vec{b}, \lambda)$ no depende de la dirección $\vec{a}$, de modo que la elección de la orientación de un dispositivo no modifica el resultado producido en la región espacialmente separada. Esta condición corresponde a lo que Shimony denominó *independencia de parámetros*. Una condición distinta, la *independencia de resultados*, exigiría además que, fijado λ , el resultado obtenido en una región no dependiera del resultado obtenido en la otra; en el marco determinista considerado aquí esta segunda condición se satisface automáticamente, pues $A(\vec{a}, \lambda)$ y $B(\vec{b}, \lambda)$ son funciones fijas de λ sin dependencia residual entre sí. La distinción entre ambas adquiere relevancia en formulaciones estocásticas del teorema, como la utilizada más adelante para describir correlaciones clásicamente correlacionadas [42].

El conjunto de variables ocultas puede variar entre las distintas realizaciones del experimento. Se asume, además, que su distribución $\rho(\lambda)$ es independiente de la configuración de medición que finalmente se elija en cada región, es decir, $\rho(\lambda |$

$\vec{a}, \vec{b}) = \rho(\lambda)$. Esta condición, conocida como *libertad de elección* o *independencia de medición*, garantiza que λ pueda interpretarse como una propiedad del sistema fijada antes de que se determinen las configuraciones experimentales, y no como una variable correlacionada de antemano con dichas elecciones. Bajo estos supuestos, la correlación predicha por una teoría local de variables ocultas puede escribirse como

$$P(\vec{a}, \vec{b}) = \int d\lambda \rho(\lambda) A(\vec{a}, \lambda) B(\vec{b}, \lambda). \quad (2.19)$$

En el artículo original, la letra P representa el valor medio del producto de los resultados y no una probabilidad conjunta. Esta expresión contiene las características fundamentales del modelo: los resultados están determinados por una descripción más completa del sistema y cada resultado depende exclusivamente de la dirección local de medición y de la variable compartida λ .

Para el estado singlete, la mecánica cuántica predice la correlación

$$P_{\text{QM}}(\vec{a}, \vec{b}) = -\vec{a} \cdot \vec{b}. \quad (2.20)$$

En particular, cuando ambos dispositivos se orientan en la misma dirección, se obtiene una anticorrelación perfecta. Si una de las partículas produce el resultado $+1$, la otra debe producir necesariamente el resultado -1 . En el razonamiento de Bell, esta anticorrelación perfecta, combinada con la condición de localidad, implica que los resultados correspondientes a las diferentes direcciones de medición deben encontrarse fijados por la descripción completa representada por λ .

A partir de estas condiciones, Bell mostró que las correlaciones asociadas a tres direcciones arbitrarias $\vec{a}$, $\vec{b}$ y $\vec{c}$ deben satisfacer

$$1 + P(\vec{b}, \vec{c}) \geq |P(\vec{a}, \vec{b}) - P(\vec{a}, \vec{c})|. \quad (2.21)$$

Esta expresión corresponde a la desigualdad original de Bell. Su importancia reside en que no depende de la naturaleza específica de las variables ocultas ni de la forma particular de su distribución. Constituye una restricción general impuesta por la combinación de variables ocultas, resultados predeterminados y localidad.

Las correlaciones predichas por la mecánica cuántica no satisfacen esta desigualdad para todas las direcciones posibles. Por ejemplo, pueden elegirse tres direcciones coplanarias tales que $\vec{a}$ y $\vec{c}$ sean ortogonales y que $\vec{b}$ forme un ángulo de 45° con cada una. Para esta elección,

$$\vec{a} \cdot \vec{c} = 0, \quad \vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{c} = \frac{1}{\sqrt{2}}, \quad (2.22)$$

y la desigualdad de Bell exigiría

$$1 - \frac{1}{\sqrt{2}} \geq \frac{1}{\sqrt{2}}, \quad (2.23)$$

relación que no se cumple.

Esta violación muestra que las correlaciones cuánticas del estado singlete no pueden reproducirse mediante ninguna teoría de variables ocultas locales que posea la estructura considerada por Bell. El resultado no descarta únicamente una elección particular de variables ocultas, sino toda la clase de modelos en los que los resultados se encuentran determinados por λ y permanecen independientes de la configuración elegida en la región distante.

El teorema de Bell no demuestra que sea imposible formular una teoría más completa que la mecánica cuántica. De hecho, existen teorías de variables ocultas, como la teoría de Bohm, capaces de reproducir las predicciones cuánticas [6]. Sin embargo, estas teorías deben incorporar una estructura no local. Por consiguiente, no es posible completar la descripción cuántica mediante variables ocultas y conservar simultáneamente la localidad asumida en el razonamiento de EPR.

La importancia conceptual del teorema reside en que transforma sustancialmente los términos en que EPR habían planteado la cuestión de completitud. La cuestión ya no consiste únicamente en determinar si la función de onda es completa. Bell muestra que una descripción más completa basada en propiedades preexistentes no puede reproducir todas las correlaciones cuánticas si se exige, al mismo tiempo, que los resultados locales sean independientes de las elecciones de medición realizadas a distancia. De esta forma, las correlaciones cuánticas revelan una incompatibilidad entre las predicciones de la mecánica cuántica y la estructura de las teorías de variables ocultas locales.

2.5. La desigualdad CHSH y su alcance experimental

En 1969, John Clauser, Michael Horne, Abner Shimony y Richard Holt propusieron una formulación del teorema de Bell que permite analizar las correlaciones obtenidas en un escenario bipartito sin exigir la anticorrelación perfecta empleada en la desigualdad original [3, 14]. Esta formulación, conocida como desigualdad CHSH, conserva el conflicto fundamental entre las predicciones cuánticas y las teorías

de variables ocultas locales, pero lo expresa mediante cuatro correlaciones asociadas a distintos ajustes de medición.

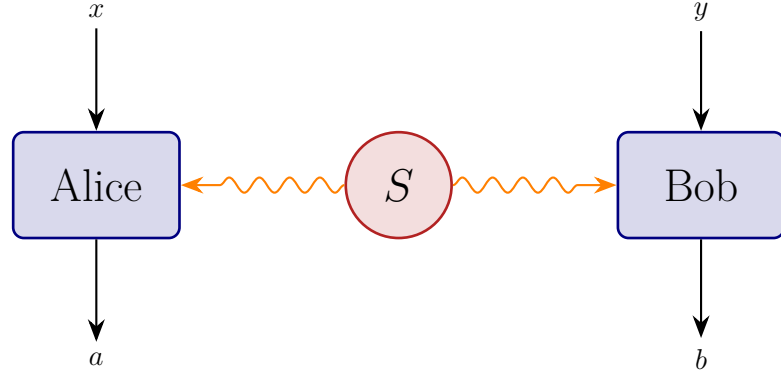


Figura 2.1: Escenario de Bell bipartito. Una fuente S distribuye sistemas cuánticos hacia dos observadores espacialmente separados, Alice y Bob, quienes seleccionan localmente un ajuste de medición (x e y , respectivamente) y obtienen un resultado (a y b). **Fuente:** Elaboración propia.

El escenario considera dos observadores espacialmente separados, Alice y Bob. En cada realización, Alice elige un ajuste de medición $x \in \{0, 1\}$ y Bob elige un ajuste $y \in \{0, 1\}$. Estas elecciones pueden asociarse, por ejemplo, a las direcciones $\vec{a}_0$, $\vec{a}_1$, $\vec{b}_0$ y $\vec{b}_1$. Los resultados obtenidos son dicotómicos, es decir, $a, b \in \{-1, +1\}$.

Las estadísticas del experimento quedan descritas por las distribuciones condicionales $p(a, b \mid x, y)$. Para cada combinación de ajustes, la correlación entre los resultados se define como

$$E_{xy} = \sum_{a, b = \pm 1} ab p(a, b \mid x, y). \quad (2.24)$$

En una teoría de variables ocultas locales, las correlaciones se explican mediante una variable compartida λ , distribuida según $\rho(\lambda)$, y mediante respuestas locales que dependen únicamente del ajuste elegido en cada región. La distribución conjunta debe adoptar la forma

$$p(a, b \mid x, y) = \int d\lambda \rho(\lambda) p(a \mid x, \lambda) p(b \mid y, \lambda). \quad (2.25)$$

Esta factorización expresa que, una vez especificada la variable λ , el resultado de Alice no depende de la elección de Bob y el resultado de Bob no depende de la elección de Alice. Las posibles correlaciones surgen exclusivamente de la información compartida en la preparación de los sistemas.

Clauser, Horne, Shimony y Holt consideraron la combinación

$$S = |E_{00} + E_{01} + E_{10} - E_{11}|. \quad (2.26)$$

Toda teoría de variables ocultas locales debe satisfacer

$$S \leq 2. \quad (2.27)$$

Esta cota no depende de la naturaleza particular de λ ni de los detalles del mecanismo que determina las respuestas locales. Por tanto, constituye una restricción sobre toda la clase de distribuciones que admiten una descomposición local de la forma anterior.

La mecánica cuántica describe las mediciones mediante observables locales A_0 , A_1 , B_0 y B_1 , cuyos valores propios son ± 1 . Para un estado bipartito ρ_{AB} , las correlaciones se expresan como

$$E_{xy} = \text{Tr} [\rho_{AB} (A_x \otimes B_y)]. \quad (2.28)$$

La combinación CHSH puede representarse mediante el operador

$$\mathcal{B}_{\text{CHSH}} = A_0 \otimes (B_0 + B_1) + A_1 \otimes (B_0 - B_1), \quad (2.29)$$

de modo que

$$S = |\text{Tr} (\rho_{AB} \mathcal{B}_{\text{CHSH}})|. \quad (2.30)$$

A diferencia de las teorías de variables ocultas locales, la mecánica cuántica permite alcanzar valores superiores a 2. Su valor máximo está limitado por la cota de Tsirelson,

$$S \leq 2\sqrt{2}. \quad (2.31)$$

Este máximo puede obtenerse utilizando un estado maximalmente entrelazado y configuraciones de medición apropiadas [13]. Para el estado singlete, una elección posible está dada por

$$A_0 = \sigma_z, \quad A_1 = \sigma_x, \quad B_0 = \frac{\sigma_z + \sigma_x}{\sqrt{2}}, \quad B_1 = \frac{\sigma_z - \sigma_x}{\sqrt{2}}, \quad (2.32)$$

con la cual se obtiene

$$S_{\text{QM}} = 2\sqrt{2}. \quad (2.33)$$

El contraste entre las cotas 2 y $2\sqrt{2}$ muestra que las correlaciones cuánticas ocupan una región más amplia que las correlaciones compatibles con variables ocultas

locales. Una violación de CHSH no indica simplemente que los resultados estén fuertemente correlacionados. Existen correlaciones clásicas perfectas que satisfacen la desigualdad. Lo que se descarta es la posibilidad de reproducir simultáneamente las cuatro correlaciones mediante una única variable compartida y respuestas locales para todas las configuraciones posibles.

La formulación CHSH también permite distinguir entre entrelazamiento y no localidad de Bell. Todo estado que viola la desigualdad CHSH es entrelazado, pero no todo estado entrelazado produce una violación para cualquier elección de mediciones. La violación depende tanto del estado como de los observables seleccionados. En consecuencia, CHSH no constituye una medida general de entrelazamiento, sino un criterio para identificar correlaciones incompatibles con una descripción local de variables ocultas.

Desde el punto de vista experimental, cada correlación E_{xy} se estima utilizando un subconjunto de realizaciones en el que se empleó la combinación correspondiente de configuraciones. Si N_{ab}^{xy} representa el número de eventos en los que se obtuvieron los resultados a y b , puede utilizarse el estimador

$$E_{xy} = \frac{N_{++}^{xy} + N_{--}^{xy} - N_{+-}^{xy} - N_{-+}^{xy}}{N_{++}^{xy} + N_{--}^{xy} + N_{+-}^{xy} + N_{-+}^{xy}}. \quad (2.34)$$

A partir de las cuatro correlaciones se calcula el parámetro S y se compara con la cota local. La conclusión experimental requiere que el valor observado sea superior a 2 por una cantidad mayor que las incertidumbres estadísticas y sistemáticas.

Las primeras pruebas experimentales se realizaron principalmente con pares de fotones correlacionados en polarización. En estos experimentos, las orientaciones de los analizadores representan las configuraciones locales, mientras que las coincidencias entre detectores permiten estimar las correlaciones [1, 24]. La formulación CHSH resultó especialmente relevante porque permitió efectuar esta comparación sin exigir correlaciones perfectas, las cuales se ven inevitablemente degradadas por el ruido, las pérdidas y las limitaciones instrumentales. De este modo, la desigualdad podía evaluarse a partir de cantidades accesibles experimentalmente, lo que permitió transformar el conflicto entre la mecánica cuántica y las teorías de variables ocultas locales en una prueba cuantitativa.

La interpretación de una violación requiere, sin embargo, controlar posibles explicaciones alternativas. Una eficiencia de detección insuficiente puede provocar que la muestra registrada no represente al conjunto total de pares emitidos, originando la laguna de detección. Asimismo, si las elecciones de medición y los resultados de ambos

observadores no se encuentran suficientemente separados, podría existir comunicación causal entre las regiones, dando lugar a la laguna de localidad. También debe garantizarse que las configuraciones sean elegidas independientemente de las variables asociadas a la fuente. El desarrollo experimental posterior permitió mejorar las fuentes, los detectores, la elección rápida de las configuraciones y la separación espacial entre los dispositivos, conduciendo a experimentos capaces de abordar simultáneamente las principales lagunas [27, 29, 40]. Como resultado, la desigualdad CHSH se consolidó como una herramienta operacional para excluir teorías de variables ocultas locales mediante observaciones controladas y no solo como una consecuencia formal del teorema de Bell.

Además de su relevancia fundacional, el desarrollo teórico y experimental de la desigualdad CHSH permitió establecer un marco preciso para identificar correlaciones incompatibles con descripciones locales y para certificar propiedades de estados y dispositivos a partir de los datos observados. Este avance amplió progresivamente el estudio de la no localidad más allá de los escenarios bipartitos formados por dos observadores y una única fuente, abriendo paso a configuraciones con varios nodos y fuentes independientes, en las cuales la estructura causal de la red impone nuevas restricciones sobre las correlaciones posibles.

2.6. El intercambio de entrelazamiento entre sistemas independientes

El intercambio de entrelazamiento, denominado *entanglement swapping* en inglés, es un esquema de preparación condicionada de estados entrelazados. Una medición conjunta sobre dos subsistemas intermedios proyecta a los subsistemas extremos sobre uno de varios estados maximalmente entrelazados. El estado obtenido depende del resultado de esta medición y puede identificarse una vez que dicho resultado es comunicado clásicamente a los nodos extremos. Estos subsistemas no han interactuado entre sí ni provienen de una fuente común. El entrelazamiento entre ellos se establece mediante la medición conjunta, a partir del entrelazamiento distribuido inicialmente por dos fuentes independientes. Este esquema fue propuesto por Żukowski, Zeilinger, Horne y Ekert [54] y realizado experimentalmente por Pan et al. [35]. Además, constituye un recurso fundamental para el estudio de redes cuánticas con fuentes independientes, como el escenario bilocal y la red triangular [7, 25].

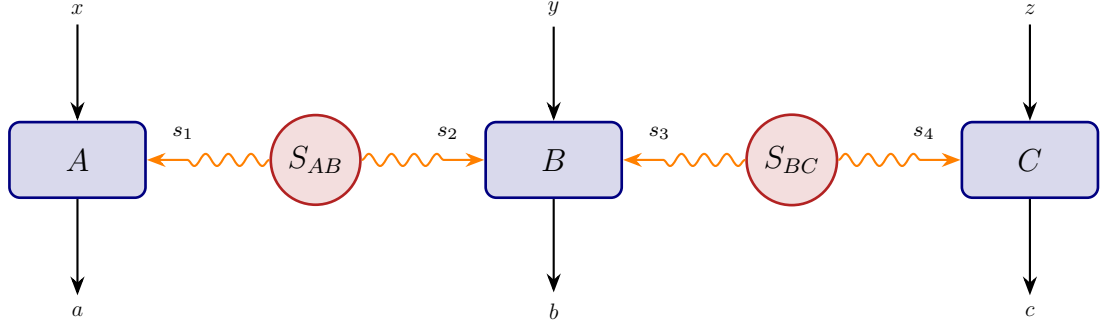


Figura 2.2: Configuración lineal de intercambio de entrelazamiento. La fuente S_{AB} distribuye los subsistemas s_1 y s_2 hacia los nodos A y B , respectivamente. La fuente S_{BC} distribuye los subsistemas s_3 y s_4 hacia los nodos B y C , respectivamente. Los símbolos x , y y z representan las configuraciones de medición seleccionadas en cada nodo, mientras que a , b y c corresponden a los resultados obtenidos. El nodo central B recibe los subsistemas s_2 y s_3 , provenientes de fuentes diferentes. **Fuente:** Elaboración propia.

Consideramos una configuración lineal de tres nodos A , B y C , junto con dos fuentes independientes denotadas por S_{AB} y S_{BC} . Denotaremos por s_1, s_2, s_3, s_4 los cuatro subsistemas cuánticos distribuidos por estas fuentes. La fuente S_{AB} prepara los subsistemas s_1 y s_2 , enviando el primero hacia A y el segundo hacia B . De manera análoga, la fuente S_{BC} prepara los subsistemas s_3 y s_4 , enviando el primero hacia B y el segundo hacia C . Por tanto, el nodo central B recibe los subsistemas s_2 y s_3 , mientras que los nodos extremos A y C reciben s_1 y s_4 , respectivamente. Cada fuente prepara sus dos subsistemas en un estado singlete. Para simplificar la notación, en las expresiones matemáticas se conservarán únicamente los subíndices numéricos. De este modo, $|\psi_{12}^-\rangle$ y $|\psi_{34}^-\rangle$ representan los estados singlete de los pares s_1, s_2 y s_3, s_4 , respectivamente.

$$|\psi_{12}^-\rangle = \frac{1}{\sqrt{2}} (|0_1\rangle |1_2\rangle - |1_1\rangle |0_2\rangle), \quad |\psi_{34}^-\rangle = \frac{1}{\sqrt{2}} (|0_3\rangle |1_4\rangle - |1_3\rangle |0_4\rangle), \quad (2.35)$$

de modo que, al operar las fuentes independientemente, el estado global es

$$|\Psi_{1234}^{\text{in}}\rangle = |\psi_{12}^-\rangle \otimes |\psi_{34}^-\rangle. \quad (2.36)$$

El paso decisivo del protocolo es posible porque los subsistemas s_2 y s_3 llegan al nodo B . Al disponer de ambos en un mismo lugar, este nodo puede realizar una medición conjunta en la base de Bell. Esta base está formada por los estados

$$|\phi_{23}^{\pm}\rangle = \frac{1}{\sqrt{2}} (|0_2\rangle |0_3\rangle \pm |1_2\rangle |1_3\rangle), \quad (2.37)$$

$$|\psi_{23}^{\pm}\rangle = \frac{1}{\sqrt{2}} (|0_2\rangle |1_3\rangle \pm |1_2\rangle |0_3\rangle). \quad (2.38)$$

El estado global puede expresarse en esta base como

$$|\Psi_{1234}^{\text{in}}\rangle = \frac{1}{2} \left(|\phi_{23}^+\rangle |\phi_{14}^+\rangle - |\phi_{23}^-\rangle |\phi_{14}^-\rangle - |\psi_{23}^+\rangle |\psi_{14}^+\rangle + |\psi_{23}^-\rangle |\psi_{14}^-\rangle \right). \quad (2.39)$$

Aquí, $|\phi_{14}^{\pm}\rangle$ y $|\psi_{14}^{\pm}\rangle$ son los estados de Bell de los subsistemas s_1 y s_4 . Esta descomposición muestra que cada resultado $r \in \{\phi^+, \phi^-, \psi^+, \psi^-\}$ obtenido en B proyecta, con probabilidad $1/4$, a los subsistemas s_1 y s_4 sobre el estado puro maximalmente entrelazado correspondiente

$$\rho_{14}^r = |\Psi_{14}^r\rangle\langle\Psi_{14}^r|. \quad (2.40)$$

Los subsistemas s_1 y s_4 quedan así entrelazados, aunque nunca hayan interactuado ni provengan de una fuente común. Este entrelazamiento se establece mediante la medición conjunta realizada sobre los subsistemas intermedios s_2 y s_3 . El estado específico obtenido depende del resultado r y puede identificarse cuando este es comunicado clásicamente a los nodos extremos A y C . Esta dependencia distingue la preparación condicionada de una preparación determinista. Mientras el resultado r sea desconocido, los observadores no pueden determinar cuál de los cuatro estados ρ_{14}^r comparten. Cada uno de estos estados es puro y maximalmente entrelazado, por lo que permite, en principio, violar una desigualdad de Bell entre A y C , aun cuando estos nodos nunca hayan compartido una fuente común. Este mecanismo constituye una base operacional para el estudio de la no localidad en redes con fuentes independientes.

2.6.1. Ignorancia del resultado y pérdida aparente de entrelazamiento

Es importante distinguir entre la ausencia *física* de intercambio de entrelazamiento y su ausencia *por desconocimiento*. La primera situación ocurre cuando la operación realizada en B no proyecta a los subsistemas extremos sobre un estado entrelazado. En ese caso, el intercambio no tiene lugar. La segunda situación es diferente. El intercambio sí ocurre, pero los observadores de A y C todavía no conocen el estado resultante.

Cuando el nodo B realiza una medición conjunta sobre los subsistemas s_2 y s_3 en la base de Bell, obtiene uno de los cuatro resultados posibles, denotado por r . Este resultado proyecta a los subsistemas s_1 y s_4 sobre el estado puro maximalmente entrelazado correspondiente ρ_{14}^r . Por tanto, el intercambio ocurre en cada repetición del experimento, aunque el estado obtenido depende del resultado registrado en B . Si este resultado no se comunica, los observadores de A y C no pueden identificar cuál de los cuatro estados fue obtenido. La incertidumbre no recae sobre la ocurrencia del intercambio, sino sobre el estado entrelazado que comparten los subsistemas extremos.

En efecto, si se promedia sobre los cuatro resultados posibles sin condicionar respecto de ellos y sin que B comunique el valor de r , la descripción conjunta de los subsistemas s_1 y s_4 está dada por

$$\rho_{14}^{\text{out}} = \frac{1}{4} \sum_r \rho_{14}^r = \frac{\mathbb{1}_1 \otimes \mathbb{1}_4}{4}. \quad (2.41)$$

Este es un estado producto de dos subsistemas maximalmente mixtos, sin entrelazamiento ni correlaciones estadísticas. El resultado no significa que el intercambio haya fallado. Surge al promediar sobre el resultado r , que en cada repetición está correlacionado con el estado obtenido por los subsistemas s_1 y s_4 . Al ignorar esta información, la correlación desaparece de la descripción promedio, aunque cada repetición haya producido un estado puro entrelazado.

El estado promedio ρ_{14}^{out} coincide con el estado reducido de los subsistemas s_1 y s_4 antes de la medición en B . Como $\rho_{1234}^{\text{in}} = \rho_{12} \otimes \rho_{34}$, al trazar sobre los subsistemas s_2 y s_3 se obtiene

$$\rho_{14}^{\text{in}} = \text{Tr}_{23}(\rho_{1234}^{\text{in}}) = \text{Tr}_2(\rho_{12}) \otimes \text{Tr}_3(\rho_{34}) = \frac{\mathbb{1}_1}{2} \otimes \frac{\mathbb{1}_4}{2} = \rho_{14}^{\text{out}}. \quad (2.42)$$

La igualdad $\rho_{14}^{\text{out}} = \rho_{14}^{\text{in}}$ es una manifestación del *principio de no-signaling* [26]. Cuando no se condiciona respecto del resultado obtenido, ninguna operación local realizada en B puede alterar las estadísticas observables de los subsistemas distantes. Por ello, el nodo B no puede transmitir información hacia A o C sin comunicar clásicamente su resultado. Si el estado promedio de los subsistemas s_1 y s_4 cambiara sin necesidad de comunicar r , este principio se violaría. El entrelazamiento no aparece en el estado promedio, sino en la relación entre el resultado r y el estado condicionado correspondiente. Esta relación se vuelve accesible y utilizable cuando r es comunicado a los nodos extremos.

Una situación distinta ocurre cuando B realiza mediciones locales e independientes sobre los subsistemas s_2 y s_3 en la base $\{|0\rangle, |1\rangle\}$, en lugar de realizar una medición

conjunta en la base de Bell. En este caso, los subsistemas s_1 y s_4 no quedan entrelazados en ninguna repetición, incluso cuando se conoce el resultado obtenido. A lo sumo, pueden generarse correlaciones clásicas entre ellos. Esta situación corresponde a una ausencia física del intercambio de entrelazamiento. Su ocurrencia depende de la medición realizada en B y no del resultado particular obtenido.

El esquema descrito, formado por dos fuentes independientes y un nodo central, corresponde a la estructura de una red bilocal. La descripción clásica que incorpora la independencia de estas fuentes se presenta en la sección siguiente. Posteriormente, la incorporación de una tercera fuente que conecte los nodos extremos permitirá construir la red triangular.

2.7. Bilocalidad

El escenario de intercambio de entrelazamiento descrito en la sección anterior está formado por dos fuentes independientes denotadas por S_{AB} y S_{BC} . La fuente S_{AB} distribuye los subsistemas s_1 y s_2 hacia los nodos A y B , respectivamente. Por su parte, S_{BC} distribuye los subsistemas s_3 y s_4 hacia los nodos B y C . Esta conectividad determina qué fuentes pueden influir sobre cada nodo y delimita la información disponible al realizar las mediciones.

Una vez establecida la descripción cuántica del protocolo, resulta natural preguntarse si las correlaciones observadas pueden reproducirse mediante un modelo clásico que respete la misma estructura de la red, sin recurrir a una causa común compartida por los tres nodos. Para representar esta exigencia, se asocia una variable oculta distinta a cada fuente. La variable λ_{AB} corresponde a S_{AB} y puede influir únicamente en los nodos A y B . De manera análoga, λ_{BC} corresponde a S_{BC} y puede influir únicamente en los nodos B y C .

En lo que sigue, la letra q denota las distribuciones de probabilidad de las variables ocultas generadas por las fuentes. La letra p se reserva para las probabilidades de respuesta de los nodos y para la distribución de los resultados observados. Así, $q(\lambda_{AB}, \lambda_{BC})$ representa la distribución conjunta de las variables ocultas, mientras que $q_{AB}(\lambda_{AB})$ y $q_{BC}(\lambda_{BC})$ describen individualmente las variables asociadas a cada fuente. La independencia física entre las fuentes exige que la distribución conjunta se factorice como

$$q(\lambda_{AB}, \lambda_{BC}) = q_{AB}(\lambda_{AB}) q_{BC}(\lambda_{BC}). \quad (2.43)$$

Además, se supone que las variables ocultas generadas por las fuentes son estadísticamente independientes de los ajustes de medición elegidos en los nodos. Esta

condición se expresa como

$$q(\lambda_{AB}, \lambda_{BC} \mid x, y, z) = q(\lambda_{AB}, \lambda_{BC}) = q_{AB}(\lambda_{AB}) q_{BC}(\lambda_{BC}). \quad (2.44)$$

La ecuación (2.43) representa la libertad de elección y establece que los ajustes x , y y z no proporcionan información sobre las variables ocultas. La ecuación (2.44) expresa la independencia entre las fuentes. En conjunto, ambas condiciones permiten interpretar λ_{AB} y λ_{BC} como variables generadas independientemente antes de seleccionar los ajustes de medición. Estas condiciones forman parte del modelo bilocal, cuya descomposición completa se introduce a continuación [7].

Sean x , y y z los ajustes de medición seleccionados en los nodos A , B y C , respectivamente. Los símbolos a , b y c representan los resultados correspondientes. En el escenario presentado en la sección anterior, el nodo B realiza una medición de Bell completa y fija. Por tanto, en ese caso no se introduce un ajuste y en B . La definición general de bilocalidad también permite que este nodo seleccione entre distintas mediciones mediante el ajuste y . Una distribución observable $p(a, b, c \mid x, y, z)$ se denomina bilocal si admite la descomposición [7]

$$\begin{aligned} p(a, b, c \mid x, y, z) = & \int d\lambda_{AB} d\lambda_{BC} q_{AB}(\lambda_{AB}) q_{BC}(\lambda_{BC}) \\ & \times p_A(a \mid x, \lambda_{AB}) p_B(b \mid y, \lambda_{AB}, \lambda_{BC}) p_C(c \mid z, \lambda_{BC}). \end{aligned} \quad (2.45)$$

El símbolo p del lado izquierdo representa la distribución de los resultados observados. Las funciones p_A , p_B y p_C describen las probabilidades de respuesta de cada nodo. Su dependencia está determinada por las conexiones de la red. El nodo A recibe información únicamente de S_{AB} , por lo que su respuesta depende de x y λ_{AB} . El nodo C recibe información únicamente de S_{BC} , por lo que su respuesta depende de z y λ_{BC} . En cambio, el nodo B recibe un subsistema de cada fuente y su respuesta puede depender de y , λ_{AB} y λ_{BC} . Finalmente, el producto $q_{AB}(\lambda_{AB}) q_{BC}(\lambda_{BC})$ representa las distribuciones independientes de las variables ocultas generadas por las fuentes.

La independencia entre las fuentes distingue la bilocalidad de la localidad convencional. En un modelo local usual, las correlaciones pueden atribuirse a una única variable oculta λ compartida por los tres nodos. Su distribución se representa mediante $q(\lambda)$ y el modelo adopta la forma

$$p(a, b, c \mid x, y, z) = \int d\lambda q(\lambda) p_A(a \mid x, \lambda) p_B(b \mid y, \lambda) p_C(c \mid z, \lambda). \quad (2.46)$$

Este modelo permite que las respuestas de los tres nodos estén coordinadas mediante una causa común representada por λ . El modelo bilocal excluye esta causa

común global y exige que las correlaciones se expliquen mediante las variables independientes λ_{AB} y λ_{BC} asociadas a las dos fuentes. Toda distribución bilocal es local en el sentido convencional. En efecto, ambas variables pueden agruparse en una sola variable $\lambda = (\lambda_{AB}, \lambda_{BC})$ cuya distribución satisface $q(\lambda) = q_{AB}(\lambda_{AB})q_{BC}(\lambda_{BC})$. Sin embargo, no toda distribución local es bilocal. Existen distribuciones locales que solo pueden reproducirse mediante variables correlacionadas entre las fuentes. Por ello, el conjunto bilocal constituye un subconjunto estricto del conjunto local [7].

Esta restricción tiene además una consecuencia estructural relevante. El conjunto de correlaciones locales convencionales forma un politopo convexo, mientras que el conjunto de correlaciones bilocales no es convexo [7]. En efecto, una combinación convexa de distribuciones bilocales no conserva necesariamente la factorización de las distribuciones asociadas a las fuentes. Por tanto, la distribución resultante puede quedar fuera del conjunto bilocal. Como consecuencia, las desigualdades lineales no permiten por sí solas caracterizar este conjunto. La independencia entre las fuentes conduce de manera natural a restricciones no lineales.

Branciard, Gisin y Pironio derivaron una desigualdad no lineal que toda distribución bilocal debe satisfacer cuando los tres nodos poseen ajustes y resultados binarios dados por $x, y, z, a, b, c \in \{0, 1\}$ [7, 8]. Para una distribución $p(a, b, c \mid x, y, z)$ se define la correlación tripartita

$$\langle A_x B_y C_z \rangle = \sum_{a,b,c} (-1)^{a+b+c} p(a, b, c \mid x, y, z). \quad (2.47)$$

A partir de estas correlaciones se construyen las combinaciones

$$I = \frac{1}{4} \sum_{x,z \in \{0,1\}} \langle A_x B_0 C_z \rangle, \quad J = \frac{1}{4} \sum_{x,z \in \{0,1\}} (-1)^{x+z} \langle A_x B_1 C_z \rangle. \quad (2.48)$$

La cantidad I reúne las cuatro combinaciones de ajustes de A y C evaluadas para $y = 0$. De manera análoga, J reúne las combinaciones evaluadas para $y = 1$ e incorpora el factor $(-1)^{x+z}$. Toda distribución bilocal satisface la desigualdad

$$\sqrt{|I|} + \sqrt{|J|} \leq 1. \quad (2.49)$$

La dependencia mediante raíces cuadradas convierte esta expresión en una desigualdad no lineal. Esta estructura refleja la independencia entre las fuentes y la geometría no convexa del conjunto bilocal.

Esta desigualdad es además ajustada en el plano de valores (I, J) . Todo par que satisface $\sqrt{|I|} + \sqrt{|J|} \leq 1$ puede obtenerse mediante una distribución bilocal [8]. Por

tanto, la cota 1 no puede reducirse sin excluir distribuciones bilocales. Cualquier valor de $\sqrt{|I|} + \sqrt{|J|}$ superior a 1 certifica que la distribución observada no admite un modelo bilocal.

Para ilustrar una violación cuántica, consideremos dos fuentes que preparan estados singlete. Los observables de los nodos A y C se eligen como

$$A_0 = C_0 = \frac{\sigma_z + \sigma_x}{\sqrt{2}}, \quad A_1 = C_1 = \frac{\sigma_z - \sigma_x}{\sqrt{2}}. \quad (2.50)$$

El nodo B selecciona entre dos observables separables sobre los subsistemas que recibe

$$B_0 = \sigma_z \otimes \sigma_z, \quad B_1 = \sigma_x \otimes \sigma_x. \quad (2.51)$$

Los valores propios de estos observables son ± 1 y se relacionan con los resultados binarios mediante la codificación $(-1)^a$, $(-1)^b$ y $(-1)^c$. Para esta configuración se obtiene $I = J = \frac{1}{2}$. En consecuencia,

$$\sqrt{|I|} + \sqrt{|J|} = \sqrt{2} > 1. \quad (2.52)$$

Este resultado viola la desigualdad bilocal. Aunque el nodo B opera sobre los dos subsistemas que recibe, cada observable B_y es un producto tensorial de observables locales y no implementa una medición entrelazante. La construcción puede relacionarse con dos pruebas de tipo CHSH, una asociada a cada fuente, cuyos resultados se combinan en B [8]. Por ello, este ejemplo ilustra la violación de una desigualdad bilocal no lineal, pero no implementa un protocolo de intercambio de entrelazamiento basado en una medición de Bell completa.

El modelo bilocal permite determinar si una distribución observada admite una explicación clásica compatible con la independencia de las fuentes y con las conexiones de la red. Si la distribución posee una descomposición bilocal, sus correlaciones son compatibles con este modelo clásico. Si no admite dicha descomposición y posee una realización cuántica, presenta correlaciones cuánticas no bilocales.

2.8. Red triangular

El escenario de *entanglement swapping* constituye un primer paso en el estudio de redes cuánticas con fuentes independientes. Su estructura lineal, formada por dos fuentes y tres nodos, presenta un nodo central distinguido que recibe sistemas

de ambas fuentes. Este escenario conduce de manera natural a explorar redes con estructuras más complejas, en las cuales las fuentes y los nodos se organizan formando ciclos.

La extensión cíclica más sencilla se obtiene al incorporar una tercera fuente independiente que conecta los nodos extremos de la red lineal. Se obtiene así la red triangular, llamada comúnmente C_3 , considerada por Fritz [25], en la que cada par de nodos está conectado por una fuente que distribuye un estado bipartito. Esta disposición elimina la presencia de un nodo central distinguido y da lugar a una estructura causal diferente de la asociada al *entanglement swapping*.

El interés de esta red no reside únicamente en su geometría, sino también en las restricciones que la independencia de las tres fuentes impone sobre las correlaciones observadas. El teorema de Fritz establece que existen correlaciones cuánticas en C_3 que no admiten una descripción clásica compatible con la estructura de la red [25]. Este resultado fundamenta el estudio de la no-localidad triangular y motiva la exploración de los fenómenos que surgen de esta disposición de fuentes, estados y nodos de medición. En particular, interesa determinar cómo los estados distribuidos y las mediciones realizadas condicionan las correlaciones obtenidas, así como identificar las circunstancias en las que estas correlaciones no pueden reproducirse mediante un modelo clásico triangular.

2.8.1. Estructura de la red triangular

La red triangular está constituida por tres nodos A , B y C , cada uno de los cuales corresponde a un lugar donde se realiza una medición. Entre cada par de nodos se ubica una fuente independiente: S_{AB} conecta los nodos A y B , S_{BC} conecta B y C , mientras que S_{AC} conecta A y C . De esta manera, las tres fuentes forman una estructura cíclica, como se representa en la siguiente figura.

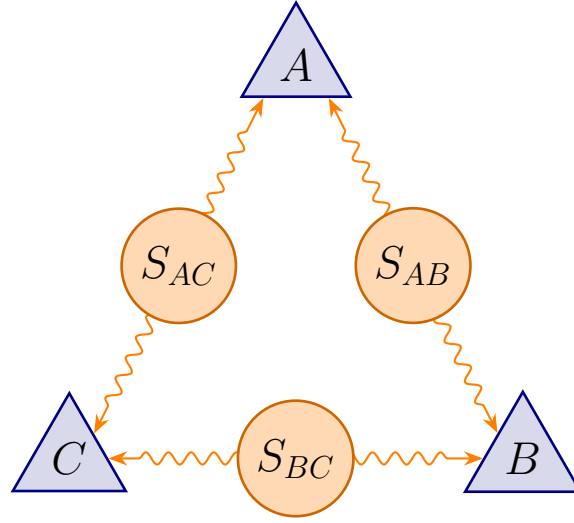


Figura 2.3: Estructura cíclica de la red triangular. Tres fuentes independientes, S_{AB} , S_{BC} y S_{AC} , distribuyen sistemas bipartitos entre los nodos A , B y C , de modo que cada nodo recibe un sistema de cada una de las dos fuentes a las que está conectado y realiza sobre ellos una única medición conjunta. **Fuente:** Elaboración propia.

Cada fuente distribuye un estado bipartito, enviando uno de sus sistemas a cada uno de los nodos que conecta. En particular, el nodo A recibe un sistema de S_{AB} y otro de S_{AC} , el nodo B recibe sistemas de S_{AB} y S_{BC} , mientras que el nodo C recibe sistemas de S_{AC} y S_{BC} . Por consiguiente, cada nodo dispone de dos sistemas provenientes de fuentes diferentes, sobre los cuales realiza una única medición conjunta y fija, registrando los resultados a , b y c , respectivamente.

A diferencia del escenario de Bell convencional, los dispositivos de medición no reciben *inputs* externos para ajustar las mediciones de los detectores: la medición realizada en cada nodo queda determinada únicamente por los dos sistemas que recibe. Ninguna fuente se encuentra conectada simultáneamente con los tres nodos, de modo que la estructura de la red restringe de antemano qué recursos pueden intervenir en la generación de cada resultado. A partir de esta estructura, en la subsección siguiente se establecerán las descripciones clásica y cuántica de las correlaciones producidas en la red.

2.8.2. Correlaciones clásicas y cuánticas en la red triangular

Los resultados registrados por los tres nodos se describen mediante la distribución conjunta $p(a, b, c)$. El carácter clásico o cuántico de esta correlación no está determinado únicamente por su forma, sino por los recursos empleados para generarla y por las restricciones que la red triangular impone sobre el acceso de cada nodo a dichos

recursos.

Este arreglo de fuentes y nodos define la estructura causal de la red triangular, es decir, la especificación de qué fuentes influyen causalmente sobre el resultado de cada nodo, junto con la ausencia de toda causa común entre las fuentes mismas. En consecuencia, cada nodo tiene acceso únicamente a la información proveniente de las dos fuentes que lo conectan, mientras que las tres fuentes actúan de manera independiente entre sí. Una correlación es clásica en la red triangular si puede reproducirse mediante tres variables ocultas λ_{AB} , λ_{BC} y λ_{AC} , generadas por las fuentes correspondientes. Debido a la independencia de las fuentes, su distribución conjunta debe satisfacer

$$q(\lambda_{AB}, \lambda_{BC}, \lambda_{AC}) = q_{AB}(\lambda_{AB})q_{BC}(\lambda_{BC})q_{AC}(\lambda_{AC}). \quad (2.53)$$

Cada nodo produce su resultado utilizando únicamente las variables ocultas proporcionadas por las dos fuentes adyacentes, sin acceso a la variable de la fuente con la que no está conectado. Por tanto, una correlación clásica compatible con la red triangular admite la descomposición

$$\begin{aligned} p(a, b, c) = & \int d\lambda_{AB} d\lambda_{BC} d\lambda_{AC} q_{AB}(\lambda_{AB})q_{BC}(\lambda_{BC})q_{AC}(\lambda_{AC}) \\ & \times p_A(a \mid \lambda_{AB}, \lambda_{AC})p_B(b \mid \lambda_{AB}, \lambda_{BC})p_C(c \mid \lambda_{AC}, \lambda_{BC}). \end{aligned} \quad (2.54)$$

Las distribuciones condicionales p_A , p_B y p_C representan las respuestas locales de los nodos. Cada una depende exactamente de las dos variables asociadas a las fuentes adyacentes a ese nodo, y no de la tercera. Esta expresión incorpora simultáneamente la independencia de las fuentes y la restricción estructural de la red.

En una descripción cuántica, las fuentes independientes distribuyen estados bipartitos ρ_{AB} , ρ_{BC} y ρ_{AC} . La independencia entre ellas se representa mediante el estado global

$$\rho_{\Delta} = \rho_{AB} \otimes \rho_{BC} \otimes \rho_{AC}. \quad (2.55)$$

En los vértices del arreglo triangular se realiza una única medición conjunta, dada por los POVM $\{M_a^A\}_a$, $\{M_b^B\}_b$ y $\{M_c^C\}_c$, de modo que la correlación resultante está dada por la regla de Born,

$$p(a, b, c) = \text{Tr} \left[\left(M_a^A \otimes M_b^B \otimes M_c^C \right) (\rho_{AB} \otimes \rho_{BC} \otimes \rho_{AC}) \right]. \quad (2.56)$$

La independencia de las fuentes se refiere únicamente a la ausencia de correlaciones entre ellas, y no impide que cada estado bipartito distribuido esté entrelazado.

Si se denotan por $\mathcal{C}_\Delta$ y $\mathcal{Q}_\Delta$ los conjuntos de correlaciones clásicas y cuánticas, respectivamente, toda correlación clásica puede reproducirse mediante una realización cuántica apropiada. En consecuencia,

$$\mathcal{C}_\Delta \subseteq \mathcal{Q}_\Delta. \quad (2.57)$$

Sin embargo, que una correlación posea una realización cuántica no significa necesariamente que carezca de una descripción clásica. Una correlación cuántica presenta no-localidad triangular cuando no admite ninguna descomposición clásica compatible con la independencia de las fuentes, es decir, cuando pertenece a $\mathcal{Q}_\Delta$ pero no a $\mathcal{C}_\Delta$.

2.8.3. La distribución de Fritz

La existencia de distribuciones que pertenecen a $\mathcal{Q}_\Delta$, pero no a $\mathcal{C}_\Delta$, fue demostrada por primera vez por Fritz [25], quien abordó este problema dentro de un marco más amplio denominado escenarios de correlación, desarrollado a partir de una idea propuesta previamente por Branciard, Rosset, Gisin y Pironio [8]. En estos escenarios no se introducen configuraciones de medición como variables de entrada externas, de modo que cada observador realiza una única medición fija. La independencia entre las fuentes sustituye entonces a la hipótesis de independencia entre las configuraciones de medición y la fuente compartida por los observadores, mientras se mantiene la estructura causal local exigida por el modelo clásico.

Desde esta perspectiva, Fritz mostró que todo escenario de Bell puede incorporarse en un escenario de correlación asociado, de manera que una violación de Bell puede traducirse en una distribución no clásica cuando se exige que las fuentes sean independientes. Entre los escenarios de correlación que no proceden directamente de un escenario de Bell bipartito, la red triangular ocupa un lugar destacado tanto por su simetría como por constituir el escenario más pequeño en el que pueden existir correlaciones no clásicas. Esto llevó a Fritz a investigar si esta red podía admitir correlaciones cuánticas incompatibles con cualquier descripción clásica que respetara su estructura causal.

Para responder esta pregunta, Fritz partió de la descripción clásica establecida en la subsección anterior, según la cual toda distribución $p(a, b, c)$ admite una descomposición en las variables ocultas mutuamente independientes λ_{AB} , λ_{BC} y λ_{AC} . Luego consideró distribuciones en las que ciertas funciones de los resultados de dos nodos están perfectamente correlacionadas. Esta correlación no constituye una nueva hipótesis de clasicidad, sino una propiedad particular de la distribución analizada

que permite deducir la independencia estadística entre una de esas funciones y una variable oculta a la que el otro nodo no tiene acceso.

Considérese que una función $f(a)$ del resultado de A coincide con certeza con una función $g(c)$ del resultado de C . Como el resultado de C depende únicamente de λ_{AC} y λ_{BC} , ambas independientes de λ_{AB} , la función $g(c)$ es estadísticamente independiente de λ_{AB} . La correlación perfecta obliga entonces a que lo mismo ocurra con $f(a)$, de modo que

$$p(f(a) = g(c)) = 1 \implies f(a) \perp \lambda_{AB}.$$

Este resultado corresponde al Corolario 2.15 de Fritz [25]. Cuando $f(a)$ representa el valor que determina la medición realizada en A , esta independencia cumple el papel de la independencia de medición requerida en un test de Bell convencional.

En la realización considerada, la fuente AB distribuye un estado maximalmente entrelazado de dos qubits. La fuente AC genera los bits uniformes y perfectamente correlacionados x y x' , registrados por A y C , mientras que la fuente BC genera los bits y y y' , registrados por B y C . Estas correlaciones pueden obtenerse mediante estados maximalmente entrelazados o estados separables clásicamente correlacionados medidos en la base computacional. El nodo A utiliza x para determinar la medición aplicada sobre el qubit procedente de la fuente AB , mientras que B utiliza y con el mismo propósito. Las mediciones generan la distribución

$$p_{\text{CHSH}}(\alpha, \beta | x, y) = \frac{1}{4} \left[1 + \frac{(-1)^{\alpha+\beta+xy}}{\sqrt{2}} \right].$$

Para esta distribución, la expresión CHSH alcanza el valor óptimo $S_{\text{CHSH}} = 2\sqrt{2}$. Finalmente, los nodos anuncian los resultados $a = (x, \alpha)$, $b = (y, \beta)$ y $c = (x', y')$. En conjunto, el procedimiento realizado en cada nodo corresponde a una única medición fija con cuatro resultados posibles.

En cualquier modelo clásico compatible con la red triangular, el resultado $c = (x', y')$ depende únicamente de λ_{AC} y λ_{BC} , por lo que es independiente de λ_{AB} . Como las correlaciones perfectas implican que $(x, y) = (x', y')$ con certeza, el par (x, y) también debe ser independiente de λ_{AB} . La distribución resultante puede escribirse como

$$p_{\text{F}}((x, \alpha), (y, \beta), (x', y')) = \frac{1}{4} \delta_{x, x'} \delta_{y, y'} p_{\text{CHSH}}(\alpha, \beta | x, y).$$

El factor $1/4$ proviene de la distribución uniforme e independiente de x e y , mientras que las deltas de Kronecker representan sus correlaciones perfectas con los bits registrados en C . Esta distribución es conocida como la distribución de Fritz [25, Teorema 2.16].

Supóngase que la distribución de Fritz admite una descomposición clásica compatible con la red triangular. Al condicionar respecto de x e y e integrar las variables auxiliares λ_{AC} y λ_{BC} , la independencia entre (x, y) y λ_{AB} permite obtener

$$p_{\text{CHSH}}(\alpha, \beta | x, y) = \int d\lambda_{AB} q_{AB}(\lambda_{AB}) p_A(\alpha | x, \lambda_{AB}) p_B(\beta | y, \lambda_{AB}).$$

Las funciones p_A y p_B representan las respuestas locales efectivas obtenidas después de integrar las variables auxiliares. Toda distribución que admite esta descomposición satisface la cota local $S_{\text{CHSH}} \leq 2$, lo que contradice el valor $S_{\text{CHSH}} = 2\sqrt{2}$ alcanzado por la distribución considerada. Por consiguiente, la distribución de Fritz no admite una descripción clásica triangular. Como la construcción anterior proporciona una realización cuántica explícita, se concluye que

$$p_F \in \mathcal{Q}_\Delta \setminus \mathcal{C}_\Delta.$$

Esto demuestra la existencia de no localidad en la red triangular.

2.8.4. Algunas estrategias de optimización para el estudio de la red triangular

Determinar si una distribución es compatible con la red triangular constituye un problema matemático complejo. En los escenarios de Bell con un número finito de entradas y resultados, el conjunto de correlaciones locales forma un politopo convexo, lo que permite estudiarlo mediante desigualdades de Bell y técnicas de programación lineal [9]. En la red triangular, la independencia de las tres fuentes impide introducir una variable aleatoria común que coordine mezclas entre diferentes realizaciones. Por esta razón, una combinación convexa de dos modelos triangulares no necesariamente conserva la independencia de las fuentes. En consecuencia, los conjuntos $\mathcal{C}_\Delta$ y $\mathcal{Q}_\Delta$ no son, en general, convexos [23, 25, 45]. Esta diferencia impide trasladar directamente las herramientas utilizadas en los escenarios de Bell convencionales y ha motivado el desarrollo de diversas estrategias numéricas y analíticas para estudiar la compatibilidad causal y la no localidad en la red triangular.

Uno de los enfoques numéricos desarrollados para estudiar la red triangular recurre a redes neuronales que actúan como oráculos computacionales¹. Esta estrategia fue desarrollada en el trabajo de Kriváchy et al. [32], en el que cada función de respuesta depende únicamente de las variables asociadas a las dos fuentes conectadas con el nodo correspondiente. El procedimiento comienza fijando una distribución objetivo $p_t(a, b, c)$, que contiene las probabilidades cuya compatibilidad con un modelo

¹En este contexto, un oráculo es una herramienta numérica que intenta determinar si una distribución puede reproducirse mediante un modelo causal dado. No corresponde a un procedimiento de decisión exacto.

clásico triangular se desea determinar. Por su parte, las redes neuronales producen una distribución $p_M(a, b, c)$ al combinar las funciones de respuesta de los tres nodos y promediarlas sobre muestras de las variables ocultas independientes. El entrenamiento modifica los pesos de las redes para que las probabilidades de p_M se aproximen a las correspondientes probabilidades de p_t . Como la arquitectura neuronal respeta la conectividad y la independencia de las fuentes, toda distribución p_M obtenida mediante este procedimiento pertenece a $\mathcal{C}_\Delta$. Si p_M coincidiera exactamente con p_t , se habría construido un modelo clásico triangular para la distribución objetivo. En la práctica, un error pequeño solo muestra que el método encontró una distribución clásica cercana a p_t dentro de la precisión numérica alcanzada. Por el contrario, si el error permanece por encima del valor de tolerancia adoptado, únicamente puede concluirse que ese proceso de entrenamiento no logró reproducir la distribución objetivo. Esto no demuestra que p_t sea no local, pues el resultado también puede deberse a las limitaciones de las redes o a que la optimización no alcanzó la mejor solución posible. La certificación rigurosa de no localidad requiere entonces un criterio adicional que permita descartar todos los modelos clásicos compatibles con la red triangular.

Otra aplicación de las redes neuronales al estudio de la red triangular fue desarrollada en la investigación de Bäumer et al. [10]. Los autores analizaron distribuciones invariantes al intercambiar los observadores o al cambiar simultáneamente las etiquetas de sus resultados. El estudio se centró en $p_{\text{EJM}}(a, b, c)$, que reúne las probabilidades obtenidas al distribuir tres estados singlete y aplicar una *Elegant Joint Measurement* en cada nodo². Mediante la combinación de modelos analíticos y búsquedas neuronales, los autores no encontraron una realización clásica de p_{EJM} . Este resultado reforzó la conjetura de su no localidad, pero no demostró que $p_{\text{EJM}} \notin \mathcal{C}_\Delta$, pues los procedimientos empleados no exploran exhaustivamente todos los modelos clásicos posibles. Además, cada distribución objetivo requiere entrenamientos independientes y repeticiones desde distintas condiciones iniciales. Por ello, estudiar una región amplia exige numerosas optimizaciones y evaluaciones, lo que limita la escalabilidad del método.

Otro enfoque para estudiar la red triangular utiliza vectores entrópicos³. En lugar de analizar por separado todas las probabilidades de $p(a, b, c)$, estos métodos resumen la distribución mediante cantidades que miden la incertidumbre de los resultados individuales y conjuntos. En un modelo clásico triangular, la independencia de las fuentes y la dependencia local de cada resultado imponen determinadas relaciones entre estas entropías. Si una distribución con realización cuántica conocida

²La EJM es una medición conjunta de dos qubits con cuatro resultados que proyecta sobre una base de estados parcialmente entrelazados con simetría tetraédrica.

³Para tres variables observables, un vector entrópico reúne las cantidades $H(A)$, $H(B)$, $H(C)$, $H(A, B)$, $H(A, C)$, $H(B, C)$ y $H(A, B, C)$. Esta representación se utiliza en teoría de la información y en inferencia causal para expresar restricciones de independencia y compatibilidad causal mediante relaciones entre entropías.

violara una desigualdad entrópica válida para toda distribución de $\mathcal{C}_\Delta$, la violación descartaría cualquier modelo clásico triangular. Como la distribución posee además una realización cuántica, podría concluirse que pertenece a $\mathcal{Q}_\Delta \setminus \mathcal{C}_\Delta$. Por el contrario, satisfacer las desigualdades analizadas solo significa que este criterio no logró descartar su compatibilidad clásica. Esto no proporciona una descomposición en variables ocultas ni demuestra que la distribución pertenezca a $\mathcal{C}_\Delta$, pues distintas distribuciones pueden producir el mismo vector entrópico. Chaves et al. [12] demostraron que las relaciones basadas en la entropía de Shannon no distinguen las correlaciones clásicas de las cuánticas en la red triangular. Posteriormente, Weilenmann y Colbeck [49] y Vilasini y Colbeck [48] buscaron restricciones más fuertes mediante desigualdades no shannonianas y entropías de Tsallis⁴, aunque no obtuvieron una separación entre ambos conjuntos. Junto con esta limitación, su implementación exige combinar y eliminar un número de relaciones y cantidades no observables que aumenta rápidamente con la complejidad del problema, por lo que estos métodos resultan difíciles de aplicar.

Desde una metodología diferente, Fraser y Wolfe [23] desarrollaron una forma más directa de buscar certificaciones de no localidad en la red triangular. Para ello, aplicaron la técnica de inflación a su estructura causal y utilizaron la distribución de Fritz [25] como punto de partida. A grandes rasgos, esta técnica construye una red ampliada que contiene varias copias de las fuentes y de los nodos observables, organizadas de modo que conserven las relaciones de dependencia e independencia de la red original. Si una distribución admite un modelo clásico triangular, debe ser posible asignar una distribución conjunta a estas copias que reproduzca de forma consistente las probabilidades observadas. En el caso de la distribución de Fritz, estas condiciones no pueden satisfacerse simultáneamente. Esta incompatibilidad permitió derivar desigualdades polinómicas válidas para toda distribución clásica triangular, entre las cuales se encuentra la desigualdad Wagon-Wheel.

Una ventaja central de Wagon-Wheel es que está formulada únicamente en términos de las probabilidades observables de la red. Para una realización cuántica determinada, estas probabilidades se obtienen directamente a partir de los estados y los operadores de medición mediante la regla de Born y luego se sustituyen en el funcional. Por tanto, su evaluación no requiere reconstruir el conjunto $\mathcal{C}_\Delta$ ni construir un modelo clásico explícito. Como la desigualdad posee una cota clásica fija, cualquier violación proporciona una certificación rigurosa de no localidad y permite utilizar el funcional como objetivo para explorar diferentes estados y mediciones.

⁴Las entropías de Tsallis forman una familia de medidas de incertidumbre parametrizada por un número real q . Para $q > 0$ y $q \neq 1$, se definen mediante $S_q(X) = (1 - \sum_x p(x)^q)/(q - 1)$ y en el límite $q \rightarrow 1$ recuperan la entropía de Shannon. Al variar q cambia la sensibilidad de la entropía frente a las distintas probabilidades de la distribución, lo que permite buscar restricciones que no aparecen al considerar únicamente la entropía de Shannon.

Capítulo 3

Optimización variacional de la desigualdad Wagon-Wheel en la red triangular

Certificar que una distribución no admite una descripción clásica compatible con la red triangular puede abordarse mediante estrategias diferentes. La construcción propuesta por Fritz [25], presentada en el [Capítulo 2](#), proporciona un ejemplo de ello. En este caso, la incompatibilidad de la distribución se demuestra aprovechando ciertas correlaciones perfectas que permiten reducir el problema a un escenario bipartito de Bell. Sin embargo, este argumento depende de propiedades específicas de dicha distribución y no proporciona un criterio directo para estudiar otras realizaciones. Fraser y Wolfe [23] abordaron este problema proponiendo la desigualdad de Wagon-Wheel, la cual permite evaluar la compatibilidad causal de una distribución de probabilidades en la red triangular mediante su comparación con una cota clásica. La violación de esta cota certifica que la distribución evaluada es incompatible con el modelo clásico triangular.

En este capítulo se explicará cómo la desigualdad Wagon-Wheel puede aplicarse en conjunto con CSPA, un método estocástico de optimización formulado para variables complejas. Esta combinación permitirá explorar el espacio de mediciones de la red triangular con el propósito de encontrar configuraciones que produzcan las mayores violaciones de la cota clásica para los estados considerados.

3.1. El funcional Wagon-Wheel

Antes de explicar el origen y la forma del funcional Wagon-Wheel, conviene precisar la configuración general de la red triangular. Los tres nodos corresponden a sitios de medición en los que se encuentran los detectores A , B y C . Las fuentes X , Y y Z conectan por pares estos nodos. Cada fuente distribuye un estado cuántico bipartito de dos qubits y envía uno de ellos a cada detector conectado. En consecuencia, cada detector recibe dos qubits, uno procedente de cada una de las fuentes a las que está conectado, como se muestra en la [Figura 3.1](#).

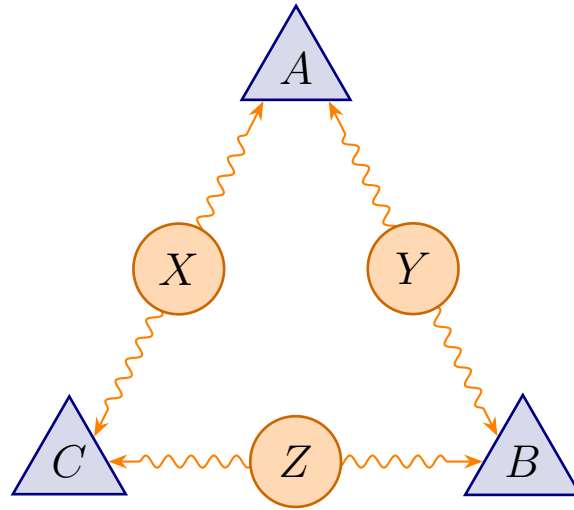


Figura 3.1: Estructura cíclica de la red triangular. Tres fuentes independientes, X , Y y Z , distribuyen sistemas bipartitos entre los nodos A , B y C , de modo que cada nodo recibe un sistema de cada una de las dos fuentes a las que está conectado y realiza sobre ellos una única medición conjunta. **Fuente:** Elaboración propia.

Sobre los dos qubits que recibe, cada detector realiza una única medición conjunta y fija. En este trabajo, dicha medición se define a partir de una base ortonormal del espacio de Hilbert conjunto de ambos qubits. Como este espacio tiene dimensión $2 \times 2 = 4$, la base contiene cuatro vectores y la medición presenta cuatro resultados posibles. Para el detector A , estos resultados se denotan por $\mathcal{A}_0$, $\mathcal{A}_1$, $\mathcal{A}_2$ y $\mathcal{A}_3$. De manera análoga, los detectores B y C presentan cuatro resultados cada uno.

Cada uno de los cuatro resultados de A admite una representación mediante dos bits. Los resultados $\mathcal{A}_0$, $\mathcal{A}_1$, $\mathcal{A}_2$ y $\mathcal{A}_3$ corresponden, respectivamente, a las combinaciones $(A_l, A_r) = (0, 0), (0, 1), (1, 0), (1, 1)$. Sus probabilidades se escriben como

$$\begin{aligned} P(\mathcal{A}_0) &= P_{A_l A_r}(00), & P(\mathcal{A}_1) &= P_{A_l A_r}(01), \\ P(\mathcal{A}_2) &= P_{A_l A_r}(10), & P(\mathcal{A}_3) &= P_{A_l A_r}(11). \end{aligned}$$

El subíndice $A_l A_r$ indica que la probabilidad está definida sobre este par de bits. Los dígitos del argumento se leen en el mismo orden que los subíndices. El primero corresponde a A_l y el segundo corresponde a A_r . Estas probabilidades satisfacen la condición de normalización

$$P_{A_l A_r}(00) + P_{A_l A_r}(01) + P_{A_l A_r}(10) + P_{A_l A_r}(11) = 1.$$

La misma representación se utiliza para los resultados de los detectores B y C .

Para abreviar la notación se introducen las contracciones

$$A_{lr} \equiv A_l A_r, \quad B_{lr} \equiv B_l B_r, \quad C_{lr} \equiv C_l C_r.$$

En estas expresiones, la yuxtaposición representa la concatenación ordenada de los dos bits y no su multiplicación. Por ejemplo, $A_{lr} = 10$ indica que $A_l = 1$ y $A_r = 0$.

La misma lógica se extiende a las probabilidades conjuntas entre distintos detectores. Por ejemplo, $P_{A_l B_l}(10)$ denota la probabilidad de obtener simultáneamente $A_l = 1$ y $B_l = 0$. En las expresiones que contienen más bits, la cadena se lee agrupando los dígitos de acuerdo con el orden de las variables en el subíndice. En $P_{A_{lr} B_{lr} C_{lr}}(001001)$, los dos primeros bits establecen $A_{lr} = 00$ y corresponden al resultado $\mathcal{A}_0$. El par siguiente establece $B_{lr} = 10$ y corresponde a $\mathcal{B}_2$. El último par establece $C_{lr} = 01$ y corresponde a $\mathcal{C}_1$. Por lo tanto, esta expresión representa la probabilidad conjunta de obtener $\mathcal{A}_0$, $\mathcal{B}_2$ y $\mathcal{C}_1$ en los detectores A , B y C , respectivamente. De forma equivalente, corresponde a $P_{ABC}(a, b, c)$ evaluada en $(a, b, c) = (0, 2, 1)$.

La distribución conjunta completa $P_{ABC}(a, b, c)$ especifica la probabilidad asociada a cada combinación de resultados de los tres detectores. Una distribución marginal se obtiene al sumar esta distribución sobre las variables de resultado que no se desean conservar. Por ejemplo, $P_{A_l B_l}(10)$ representa la probabilidad conjunta de obtener $A_l = 1$ y $B_l = 0$. De acuerdo con la codificación binaria adoptada, esta marginal puede escribirse como

$$P_{A_l B_l}(10) = \sum_{a=2}^3 \sum_{b=0}^1 \sum_{c=0}^3 P_{ABC}(a, b, c).$$

Para ilustrar cómo se deduce una restricción clásica en este escenario, es necesario analizar el caso particular que motivó el funcional Wagon-Wheel. La distribución de Fritz [25], presentada en el **Capítulo 2**, constituyó el punto de partida utilizado por Fraser y Wolfe. En esta construcción, la fuente Y distribuye entre A y B un estado entrelazado de dos qubits. Las fuentes X y Z proporcionan dos bits aleatorios independientes compartidos, respectivamente, por los pares de detectores A, C y B, C . Estos bits determinan las mediciones que A y B realizan sobre el estado recibido desde Y , las cuales producen la máxima violación cuántica de CHSH. Los bits procedentes de X y Z se registran como A_l y B_l , mientras que los resultados de las mediciones sobre Y se registran como A_r y B_r . El detector C reporta los dos primeros bits como C_{lr} . De este modo se establecen las correlaciones perfectas $C_l = A_l$ y $C_r = B_l$.

Como consecuencia de estas correlaciones, solo 16 de las 64 combinaciones posibles de (a, b, c) presentan probabilidad distinta de cero. Ocho de ellas poseen el valor $(2 + \sqrt{2})/32$ y las ocho restantes poseen el valor $(2 - \sqrt{2})/32$. La distribución puede escribirse como

$$\begin{aligned} P_F(00, 00, 00) &= P_F(01, 01, 00) = P_F(00, 10, 01) = P_F(01, 11, 01) \\ &= P_F(10, 00, 10) = P_F(11, 01, 10) = P_F(10, 11, 11) = P_F(11, 10, 11) \\ &= \frac{1}{32} (2 + \sqrt{2}), \end{aligned}$$

$$\begin{aligned} P_F(00, 01, 00) &= P_F(01, 00, 00) = P_F(00, 11, 01) = P_F(01, 10, 01) \\ &= P_F(10, 01, 10) = P_F(11, 00, 10) = P_F(10, 10, 11) = P_F(11, 11, 11) \\ &= \frac{1}{32} (2 - \sqrt{2}). \end{aligned}$$

Todas las combinaciones que no aparecen en estas expresiones tienen probabilidad igual a cero.

Fraser y Wolfe [23] utilizaron estos valores como entrada de un problema de marginales construido mediante la técnica de inflación. Para ello consideraron una ampliación de la red formada por dos copias de A , dos copias de B y cuatro copias de C . Esta ampliación, representada en la **Figura 3.2** del trabajo citado, recibe el nombre de inflación Wagon-Wheel debido a la forma de su representación gráfica. Si P_F admitiera un modelo clásico triangular, debería existir una distribución conjunta sobre las variables observadas de la red inflada cuyas marginales reprodujeran las distribuciones exigidas por la estructura original.

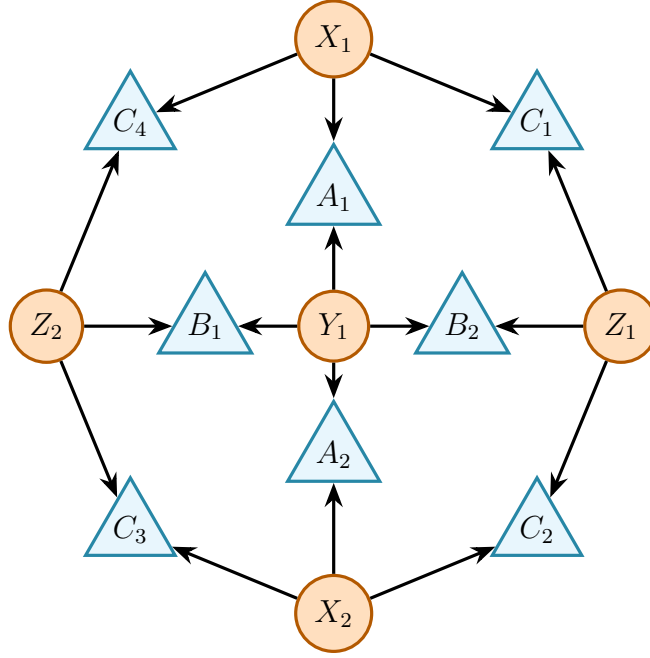


Figura 3.2: Estructura de la inflación Wagon-Wheel utilizada para derivar la desigualdad homónima. La red ampliada contiene dos copias de A , dos copias de B y cuatro copias de C , conectadas mediante copias de las fuentes independientes X , Y y Z . **Fuente:** Imagen elaborada por el autor y basada en la Figura 8(b) del artículo de Fraser y Wolfe [23].

En la red ampliada, las condiciones de positividad, normalización y consistencia entre las marginales son lineales. Esto permite formular un problema convexo de factibilidad que impone condiciones necesarias para la compatibilidad con el modelo clásico triangular. Al incorporar las marginales determinadas por la distribución de Fritz, Fraser y Wolfe comprobaron que ninguna distribución conjunta podía satisfacer simultáneamente todas las restricciones. El certificado de inviabilidad obtenido puede reescribirse mediante las factorizaciones derivadas de la independencia de las fuentes y la eliminación de los índices asociados a las copias. De este procedimiento se obtiene una desigualdad polinómica definida únicamente sobre las probabilidades observables de la red original, correspondiente a la desigualdad de Wagon-Wheel presentada a continuación.

Con la notación establecida, el funcional asociado a este certificado se escribe como

$$\begin{aligned}
I_{w-w} = & -P_{A_l B_l}(10) + P_{A_l B_l C_{lr}}(1010) \\
& -P_{A_l B_l}(01)P_{C_{lr}}(10) - P_{C_{lr}}(00)P_{C_{lr}}(11) \\
& -P_{C_{lr}}(01)P_{A_{lr} B_{lr} C_{lr}}(100110) - P_{C_{lr}}(01)P_{A_{lr} B_{lr} C_{lr}}(110010) \\
& + P_{C_{lr}}(00)P_{A_{lr} B_{lr} C_{lr}}(101111) + P_{C_{lr}}(00)P_{A_{lr} B_{lr} C_{lr}}(111011) \\
& + P_{C_{lr}}(10)P_{A_{lr} B_{lr} C_{lr}}(001001) + P_{C_{lr}}(10)P_{A_{lr} B_{lr} C_{lr}}(011101) \\
& + P_{C_{lr}}(11)P_{A_{lr} B_{lr} C_{lr}}(000000) + P_{C_{lr}}(11)P_{A_{lr} B_{lr} C_{lr}}(010100).
\end{aligned} \tag{3.1}$$

Los cuatro primeros términos de I_{w-w} contienen probabilidades marginales o productos de ellas. Los ocho términos restantes corresponden a productos entre una probabilidad marginal de C_{lr} y un valor de la distribución conjunta completa $P_{A_{lr} B_{lr} C_{lr}}$. En consecuencia, I_{w-w} no es un funcional lineal de las probabilidades, sino una expresión polinómica de grado dos.

Toda distribución P_{ABC} compatible con el modelo clásico triangular, formado por tres fuentes mutuamente independientes, debe satisfacer la cota clásica

$$I_{w-w} \leq 0. \tag{3.2}$$

En lo que sigue, la condición expresada en la (3.2) se denominará desigualdad de Wagon-Wheel.

Si para una distribución se obtiene $I_{w-w} > 0$, queda certificada su incompatibilidad con el modelo clásico triangular. Por el contrario, el cumplimiento de la cota $I_{w-w} \leq 0$ no basta para concluir que la distribución admite una descripción clásica.

Al evaluar la distribución de Fritz en el funcional se obtiene

$$I_{w-w}^{\text{Fritz}} = \frac{1}{16} (\sqrt{2} - 1) \approx 0,02589.$$

Este valor es positivo y certifica la incompatibilidad de P_F con el modelo clásico triangular. Fraser y Wolfe también exploraron numéricamente diferentes elecciones de estados y mediciones mediante varios métodos de optimización, pero no encontraron una realización cuántica que produjera una violación mayor. Por esta razón, $I_{w-w} \approx 0,02589$ debe interpretarse como el mayor valor obtenido durante su búsqueda numérica. Este resultado no constituye una cota superior para todas las realizaciones cuánticas de la red, ya que los autores no demostraron que ninguna otra

configuración pudiera superarlo. La distribución de Fritz representa, por tanto, la mejor realización encontrada y una candidata a alcanzar el máximo cuántico global.

La dificultad de explorar el espacio de mediciones y la ausencia de una demostración del máximo cuántico motivan el estudio de una estrategia de optimización alternativa. A continuación se presenta CSPA y posteriormente se desarrolla su aplicación a la maximización de I_{w-w} .

3.1.1. CSPA como método de optimización estocástica

CSPA (*Complex Simultaneous Perturbation Stochastic Approximation*) [47] es un algoritmo de optimización estocástica diseñado para maximizar o minimizar una función real que depende de un vector de variables complejas, en situaciones donde no se dispone de una expresión analítica cerrada para su gradiente. Es la extensión, al dominio complejo, del algoritmo SPSA (*Simultaneous Perturbation Stochastic Approximation*), originalmente formulado por Spall [43] para variables reales.

La dificultad que motiva este tipo de métodos aparece al estimar el gradiente de una función objetivo f respecto de un vector de parámetros $z \in \mathbb{C}^d$. Cada componente compleja contiene una parte real y una parte imaginaria que deben perturbarse por separado. Una estimación mediante diferencias finitas centrales requiere cuatro evaluaciones de f por cada componente compleja, por lo que el número total de evaluaciones crece linealmente con d . CSPA evita este costo al perturbar simultáneamente todas las componentes de z en una dirección aleatoria y construir una estimación del gradiente completo a partir de solo dos evaluaciones de la función objetivo.

Formalmente, sea $f : \mathbb{C}^d \rightarrow \mathbb{R}$ la función a optimizar. Como f es real y $z \in \mathbb{C}^d$, f no es, en general, una función holomorfa de z ; por esta razón, CSPA recurre al cálculo de Wirtinger, que trata a f como una función de dos argumentos formalmente independientes, z y su conjugado z^* , escrita $f(z, z^*)$, y que permite definir una derivada parcial bien definida $\partial f / \partial z^*$ incluso cuando f no es holomorfa. Es precisamente este gradiente de Wirtinger el que CSPA estima y utiliza para actualizar los parámetros.

Sea $z_k \in \mathbb{C}^d$ la estimación del parámetro óptimo en la iteración k . En cada iteración, se genera aleatoriamente un vector $\Delta_k \in \mathbb{C}^d$, cuyas d componentes se obtienen de forma independiente mediante un sorteo con distribución uniforme discreta sobre el conjunto $\{1, i, -1, -i\}$, y se perturba z_k en ambos sentidos a lo largo de Δ_k :

$$z_k^+ = z_k + c_k \Delta_k, \quad z_k^- = z_k - c_k \Delta_k,$$

donde $c_k > 0$ es un coeficiente de perturbación que decrece con el número de iteración k . A partir de las dos evaluaciones de la función objetivo en estos puntos, posiblemente afectadas por ruido de estimación, $f(z_k^+, z_k^{+*}) + \epsilon_{k,+}$ y $f(z_k^-, z_k^{-*}) + \epsilon_{k,-}$, se construye, componente a componente, el estimador del gradiente de Wirtinger $\partial f / \partial z^*$,

$$\hat{g}_{k,i} = \frac{f(z_k^+, z_k^{+*}) + \epsilon_{k,+} - f(z_k^-, z_k^{-*}) - \epsilon_{k,-}}{2c_k \Delta_{k,i}^*},$$

donde $\Delta_{k,i}^*$ es el conjugado complejo de la componente $\Delta_{k,i}$ y $\epsilon_{k,+}$ y $\epsilon_{k,-}$ representan posibles errores en las evaluaciones de la función objetivo. Aunque $\hat{g}_k$ se construye a partir de una única dirección aleatoria, su valor esperado se aproxima al gradiente de Wirtinger $\partial f / \partial z^*$ cuando c_k tiende a cero. En este sentido, el estimador es asintóticamente insesgado. La convergencia de la secuencia z_k requiere además condiciones de regularidad sobre la función objetivo, las perturbaciones aleatorias, las secuencias de ganancia y el ruido de evaluación. Bajo estas condiciones, CSPSA puede converger en media hacia un punto estacionario asociado a un óptimo local [43, 47].

Con $\hat{g}_k$ ya calculado, el parámetro se actualiza siguiendo la dirección de dicho gradiente,

$$z_{k+1} = z_k + a_k \hat{g}_k,$$

donde $a_k > 0$ es el coeficiente de paso del gradiente en la iteración k (el signo se invierte si el objetivo es minimizar en vez de maximizar f). Las secuencias a_k y c_k se conocen, en la literatura de aproximación estocástica, como secuencias de ganancia, y toman la forma

$$a_k = \frac{a}{(k+1+A)^s}, \quad c_k = \frac{b}{(k+1)^r},$$

con a, b, A, s, r constantes positivas fijadas de antemano. Para que la convergencia asintótica mencionada arriba esté garantizada, estas secuencias deben decrecer a una tasa particular, ni demasiado rápida ni demasiado lenta: los exponentes s, r deben satisfacer las condiciones de Robbins-Monro [38], que exigen que $\sum_k a_k = \infty$ (para que el algoritmo pueda seguir avanzando indefinidamente) junto con $\sum_k a_k^2 < \infty$ y $\sum_k (a_k/c_k)^2 < \infty$ (para que el ruido acumulado no impida la convergencia). Los valores $s = 0,602$ y $r = 0,101$, usados en este trabajo, son los recomendados por Spall [43] por satisfacer estas condiciones manteniendo, en la práctica, una velocidad de convergencia adecuada.

Una de las principales ventajas de CSPSA es que la estimación del gradiente requiere únicamente dos evaluaciones de la función objetivo por iteración, independientemente del número de parámetros d . Por tanto, el número de evaluaciones necesarias

para construir esta estimación no aumenta con la dimensión del espacio de parámetros. El costo de cada evaluación y de las operaciones realizadas sobre el vector z todavía puede depender de d . Esta reducción en el número de evaluaciones hace que CSPSA resulte especialmente conveniente cuando la función objetivo es costosa de calcular.

3.1.2. Aplicación a la maximización de I_{w-w}

La función objetivo utilizada por CSPSA es el funcional Wagon-Wheel, definido como $f(z) = I_{w-w}(z)$. Durante cada ejecución se mantiene fijo el estado conjunto ρ_{tot} distribuido por las tres fuentes. El vector z reúne los parámetros complejos de las mediciones conjuntas realizadas por los detectores A , B y C . La optimización busca entonces la configuración de mediciones que produzca el mayor valor del funcional para el estado considerado. Este problema se expresa como

$$z_{\text{opt}} = \arg \max_z I_{w-w}(z).$$

En esta primera etapa, cada evaluación de $I_{w-w}(z)$ se realiza directamente mediante la regla de Born, sin recurrir a una estimación basada en un número finito de mediciones. Para un valor dado de z , las mediciones de los tres detectores quedan determinadas por las unitarias $U_A(z_A)$, $U_B(z_B)$ y $U_C(z_C)$. La transformación conjunta aplicada antes de medir en la base computacional factoriza según

$$U(z) = U_A(z_A) \otimes U_B(z_B) \otimes U_C(z_C).$$

Si ρ_{tot} denota el estado conjunto de los seis qubits distribuidos por las fuentes, la probabilidad de obtener los resultados a , b y c está dada por

$$P_{ABC}(a, b, c|z) = \text{Tr} \left[(|a\rangle\langle a| \otimes |b\rangle\langle b| \otimes |c\rangle\langle c|) U(z) \rho_{\text{tot}} U(z)^\dagger \right].$$

Cada ket de esta expresión representa un vector de la base computacional de los dos qubits recibidos por el detector correspondiente. Esta formulación es válida tanto para estados puros como para estados mixtos.

Las probabilidades obtenidas mediante la regla de Born se sustituyen directamente en la expresión de I_{w-w} . Las distribuciones marginales se calculan sumando sobre los resultados que no se desean conservar. Por ejemplo,

$$P_{A_l B_l}(10|z) = \sum_{a=2}^3 \sum_{b=0}^1 \sum_{c=0}^3 P_{ABC}(a, b, c|z).$$

Los valores $a \in \{2, 3\}$ corresponden a $A_l = 1$, mientras que los valores $b \in \{0, 1\}$ corresponden a $B_l = 0$. El resultado de C puede tomar cualquiera de sus cuatro valores posibles.

Los términos cuadráticos se obtienen multiplicando las probabilidades correspondientes. Por ejemplo, uno de los términos del funcional es

$$P_{C_{lr}}(00|z)P_{A_{lr}B_{lr}C_{lr}}(101111|z) = P_{C_{lr}}(00|z)P_{ABC}(2, 3, 3|z).$$

La cadena 101111 establece $A_{lr} = 10$, $B_{lr} = 11$ y $C_{lr} = 11$. La marginal $P_{C_{lr}}(00|z)$ se obtiene sumando $P_{ABC}(a, b, 0|z)$ sobre todos los valores de a y b . Al repetir este procedimiento para los doce términos y combinarlos con sus signos correspondientes se obtiene el valor exacto de $I_{w-w}(z)$. En esta etapa no existe ruido de muestreo, por lo que $\epsilon_{k,+} = \epsilon_{k,-} = 0$.

Como $I_{w-w}(z)$ no es, en general, una función cóncava de sus parámetros, CSPSA puede converger hacia un máximo local. Para explorar diferentes regiones del espacio de mediciones, el algoritmo se ejecuta N_{rep} veces a partir de inicializaciones aleatorias e independientes. Al finalizar las ejecuciones se conserva la configuración que alcanza el mayor valor de I_{w-w} . Cada inicialización contiene tres matrices complejas de dimensión 4×4 , denotadas por $Z_{A,0}$, $Z_{B,0}$ y $Z_{C,0}$. Las partes reales e imaginarias de sus entradas se generan de forma independiente mediante una distribución normal estándar. Como estas matrices no son necesariamente unitarias, cada una se proyecta al conjunto de matrices unitarias antes de utilizarla como medición inicial.

La actualización aditiva de CSPSA puede hacer que las matrices asociadas a z_{k+1} dejen de ser unitarias. Para obtener mediciones válidas, cada matriz se proyecta nuevamente al conjunto de matrices unitarias mediante una descomposición polar. Si $\tilde{Z}_{x,k+1}$ representa la matriz obtenida después de la actualización para el detector $x \in \{A, B, C\}$, la matriz proyectada se define como

$$Z_{x,k+1} = \tilde{Z}_{x,k+1} \left(\tilde{Z}_{x,k+1}^\dagger \tilde{Z}_{x,k+1} \right)^{-1/2}.$$

La matriz resultante es unitaria y corresponde a la matriz unitaria más cercana a $\tilde{Z}_{x,k+1}$ en norma de Frobenius. La misma proyección se aplica a las matrices obtenidas en los puntos perturbados z_k^+ y z_k^- antes de evaluar la regla de Born. De este modo, todas las evaluaciones de I_{w-w} se realizan utilizando mediciones físicamente válidas.

El criterio de detención empleado es, a su vez, simplemente ejecutar un número fijo K de iteraciones por cada una de las N_{rep} corridas, en lugar de un criterio de convergencia adaptativo¹.

Reuniendo todos estos elementos, el procedimiento completo aplicado a la maximización de $I_{w-w}(z)$ queda como se muestra en el Algoritmo 1.

¹Por ejemplo, detener el algoritmo cuando el cambio entre iteraciones sucesivas cae bajo cierto umbral.

Algoritmo 1 CSPSA aplicado a la maximización de $I_{w-w}(z)$ con evaluación exacta

```

1: procedure CSPSA  $I_{ww}(\rho_{\text{tot}}, N_{\text{rep}}, K, \{a_k, c_k\})$ 
2:   for  $j = 1$  to  $N_{\text{rep}}$  do
3:     for  $x \in \{A, B, C\}$  do
4:       Generar una matriz compleja aleatoria  $\tilde{Z}_{x,0}$ 
5:        $Z_{x,0} \leftarrow \tilde{Z}_{x,0} (\tilde{Z}_{x,0}^\dagger \tilde{Z}_{x,0})^{-1/2}$ 
6:     end for
7:     Establecer  $z_0 \leftarrow \{Z_{A,0}, Z_{B,0}, Z_{C,0}\}$ 
8:     for  $k = 0$  to  $K - 1$  do
9:       Sortear  $\Delta_k$  con componentes uniformes sobre  $\{1, i, -1, -i\}$ 
10:      Calcular  $\tilde{z}_k^+ \leftarrow z_k + c_k \Delta_k$  y  $\tilde{z}_k^- \leftarrow z_k - c_k \Delta_k$ 
11:      for  $x \in \{A, B, C\}$  do
12:         $Z_{x,k}^\pm \leftarrow \tilde{Z}_{x,k}^\pm (\tilde{Z}_{x,k}^{\pm\dagger} \tilde{Z}_{x,k}^\pm)^{-1/2}$ 
13:      end for
14:      Establecer  $z_k^\pm \leftarrow \{Z_{A,k}^\pm, Z_{B,k}^\pm, Z_{C,k}^\pm\}$ 
15:      Calcular  $P_{ABC}(a, b, c | z_k^+)$  y  $P_{ABC}(a, b, c | z_k^-)$ 
16:      Evaluar  $I_{w-w}(z_k^+)$  e  $I_{w-w}(z_k^-)$ 
17:       $\hat{g}_k \leftarrow \frac{I_{w-w}(z_k^+) - I_{w-w}(z_k^-)}{2c_k \Delta_k^*}$ 
18:      Calcular  $\tilde{z}_{k+1} \leftarrow z_k + a_k \hat{g}_k$ 
19:      for  $x \in \{A, B, C\}$  do
20:         $Z_{x,k+1} \leftarrow \tilde{Z}_{x,k+1} (\tilde{Z}_{x,k+1}^\dagger \tilde{Z}_{x,k+1})^{-1/2}$ 
21:      end for
22:      Establecer  $z_{k+1} \leftarrow \{Z_{A,k+1}, Z_{B,k+1}, Z_{C,k+1}\}$ 
23:    end for
24:    Guardar  $z_K^{(j)} \leftarrow z_K$  y  $v^{(j)} \leftarrow I_{w-w}(z_K^{(j)})$ 
25:  end for
26:   $j_{\text{best}} \leftarrow \arg \max_j v^{(j)}$ 
27:   $z_{\text{best}} \leftarrow z_K^{(j_{\text{best}})}$ 
28:  return  $z_{\text{best}}, v^{(j_{\text{best}})}$ 
29: end procedure

```

3.1.3. Protocolo de Classical Shadows

Sea ρ un estado cuántico fijo de un sistema de n qubits. Se supondrá que este estado es desconocido o que su reconstrucción completa resulta costosa. El protocolo de Classical Shadows, propuesto por Huang, Kueng y Preskill [31], permite estimar el valor de expectación $\text{Tr}(O\rho)$ de un observable O de interés a partir de mediciones repetidas sobre copias independientes de ρ , sin necesidad de reconstruir explícitamente su matriz densidad.

Una posibilidad consiste en reconstruir el estado completo mediante tomografía cuántica y calcular posteriormente las probabilidades requeridas. Sin embargo, la cantidad de configuraciones y repeticiones necesarias para esta reconstrucción crece exponencialmente con el número de qubits, aun cuando solo se desea estimar un conjunto particular de propiedades. Esta dificultad motiva la búsqueda de un procedimiento más viable. El método Classical Shadows [31] utiliza mediciones aleatorias para construir una representación clásica del estado, a partir de la cual pueden estimarse diferentes valores de expectación sin realizar una tomografía completa. Sobre esta base, se propone estimar I_{w-w} mediante Classical Shadows e incorporar esta estimación al proceso de optimización realizado con CSPSA. Las subsecciones siguientes presentan el protocolo y desarrollan su aplicación a la estimación del funcional.

3.2. Estimación mediante Classical Shadows

Sea ρ un estado cuántico fijo de un sistema de n qubits. Se supondrá que este estado es desconocido o que su reconstrucción completa resulta costosa. El protocolo de Classical Shadows, propuesto por Huang, Kueng y Preskill [31], permite estimar el valor de expectación $\text{Tr}(O\rho)$ de un observable O a partir de mediciones repetidas sobre copias independientes de ρ , sin necesidad de reconstruir explícitamente su matriz densidad.

La primera etapa del protocolo corresponde a un procedimiento de medición aleatoria. En cada ejecución se selecciona al azar una unitaria $\mathcal{U}$ ² perteneciente al ensemble $\mathcal{U}$. Este ensemble reúne las unitarias disponibles para el procedimiento de medición. La selección se realiza de forma independiente respecto de las ejecuciones anteriores y siempre sigue la misma distribución de probabilidad establecida previamente. Por tanto, la unitaria seleccionada puede cambiar entre ejecuciones, aunque la probabilidad asignada a cada unitaria se mantiene invariable.

²La notación $\mathcal{U}$ identifica la unitaria aleatoria utilizada para construir la representación clásica del estado y la distingue de $U(z)$, que representa las mediciones ajustadas mediante CSPSA.

Una vez seleccionada, $\mathcal{U}$ transforma el estado según $\rho \mapsto \mathcal{U} \rho \mathcal{U}^\dagger$. A continuación, el estado transformado se mide en la base computacional. Como resultado se obtiene una cadena aleatoria de bits $\hat{b} \in \{0, 1\}^n$, en la que cada componente registra el resultado correspondiente a uno de los qubits. Para una unitaria $\mathcal{U}$ determinada, la probabilidad de obtener una cadena particular b está dada por

$$\Pr[\hat{b} = b \mid \mathcal{U}] = \langle b | \mathcal{U} \rho \mathcal{U}^\dagger | b \rangle.$$

Esta expresión muestra que la unitaria seleccionada y el resultado observado determinan el proyector $\mathcal{U}^\dagger |\hat{b}\rangle \langle \hat{b}| \mathcal{U}$ en la representación del estado original. La información proporcionada por una sola ejecución no es suficiente para reconstruir ρ . Sin embargo, el promedio de estos proyectores sobre las unitarias y los resultados posibles define el canal de medición

$$\mathcal{M}(\rho) = \mathbb{E}_{\mathcal{U}, \hat{b}} [\mathcal{U}^\dagger |\hat{b}\rangle \langle \hat{b}| \mathcal{U}].$$

Este valor esperado considera tanto la selección aleatoria de $\mathcal{U}$ como la distribución de resultados determinada por la regla de Born. La aplicación $\mathcal{M}$ recibe el nombre de *canal de medición* porque asigna al estado ρ el operador promedio producido por el procedimiento de medición aleatoria. Las mediciones generadas por $\mathcal{U}$ son tomográficamente completas cuando permiten distinguir cualquier par de estados y reconstruir de forma única el estado desconocido. Cuando se cumple esta condición, el canal $\mathcal{M}$ es invertible como aplicación lineal en el espacio de operadores.

La inversa del canal permite transformar el proyector determinado por una ejecución en un estimador del estado original. El operador aleatorio obtenido de esta forma se denomina *snapshot* y se define como

$$\hat{\rho} = \mathcal{M}^{-1} [\mathcal{U}^\dagger |\hat{b}\rangle \langle \hat{b}| \mathcal{U}].$$

Cada snapshot queda determinado por la unitaria seleccionada y la cadena de bits observada en una sola ejecución. Por tanto, diferentes ejecuciones producen en general distintos snapshots. Su interpretación como estimador es estadística y no implica que $\hat{\rho}$ represente un estado físico. En particular, la aplicación $\mathcal{M}^{-1}$ no tiene por qué ser completamente positiva y puede generar un operador con autovalores negativos. Esto no representa un inconveniente, ya que el snapshot se obtiene mediante posprocesamiento clásico y se utiliza únicamente para estimar propiedades de ρ .

Por construcción, el valor esperado del snapshot reproduce el estado original

$$\mathbb{E}[\hat{\rho}] = \mathcal{M}^{-1} [\mathcal{M}(\rho)] = \rho.$$

Para cualquier observable O se puede definir entonces

$$\hat{o} = \text{Tr}(O\hat{\rho}), \quad \mathbb{E}[\hat{o}] = \text{Tr}(O\rho).$$

Aunque $\hat{o}$ es un estimador insesgado, una sola ejecución puede presentar fluctuaciones considerables y no proporciona una estimación precisa.

Al repetir el procedimiento N veces sobre copias independientes de ρ se obtiene la colección

$$S_N(\rho) = \{\hat{\rho}_1, \dots, \hat{\rho}_N\}.$$

Esta colección constituye la Classical Shadow de ρ y N corresponde a su tamaño. En lo que sigue, N se denominará presupuesto de shadows³. Como cada ejecución utiliza una copia independiente de ρ y genera un snapshot, este presupuesto coincide con el número de copias medidas y de snapshots obtenidos. El valor de expectación de O se estima mediante el promedio

$$\bar{o}_N = \frac{1}{N} \sum_{i=1}^N \text{Tr}(O\hat{\rho}_i).$$

La independencia entre los snapshots permite reducir las fluctuaciones estadísticas al aumentar el presupuesto N .

La propuesta original también considera la mediana de promedios, denominada *median of means*, para reducir la influencia de valores atípicos. Para ello, los N snapshots se dividen en B bloques de tamaño similar. En cada bloque se calcula el promedio de $\text{Tr}(O\hat{\rho}_i)$ y la estimación final se obtiene tomando la mediana de los B resultados. Este procedimiento proporciona un control más robusto de la probabilidad de error. A diferencia del promedio simple, la mediana de promedios no tiene por qué conservar exactamente el carácter insesgado.

Una vez construida la Classical Shadow $S_N(\rho)$, el mismo conjunto de datos puede utilizarse para estimar los valores de expectación de múltiples observables, incluso cuando estos se seleccionan después de realizar las mediciones. Esto no significa que todos puedan estimarse con la misma precisión. La cantidad de mediciones necesaria depende del ensemble $\mathcal{U}$, de los observables considerados y de la precisión requerida.

Esta dependencia puede describirse mediante la shadow norm, que cuantifica la dificultad de estimar un observable para el esquema de medición escogido. Para

³El término presupuesto corresponde a la traducción literal del inglés *budget*. En este contexto no representa un costo monetario, sino la cantidad total de recursos experimentales destinados a construir la Classical Shadow.

estimar simultáneamente los valores de expectación de M observables $O_1, \dots, O_M$ con un error aditivo menor que ϵ y una probabilidad de fallo acotada por δ , un número suficiente de shots satisface una relación del orden

$$N = \mathcal{O} \left(\frac{\log(M/\delta)}{\epsilon^2} \max_{1 \leq i \leq M} \|O_i\|_{\text{shadow}}^2 \right).$$

Esta expresión muestra que el número de shots depende de la shadow norm y del inverso del cuadrado de la precisión requerida. Su dependencia respecto del número de observables es logarítmica. Por esta razón, Classical Shadows resulta especialmente eficiente cuando los observables poseen una shadow norm acotada para el esquema de medición utilizado. La elección concreta de las unitarias aleatorias se estudia en la subsección siguiente.

3.2.1. Classical Shadows con Pauli local

El ensemble de unitarias $\mathcal{U}$ empleado para nuestro propósito es el llamado ensemble de Pauli local, en el que la unitaria aplicada al sistema completo de n qubits factoriza como un producto de unitarias de un solo qubit,

$$U = U_1 \otimes U_2 \otimes \dots \otimes U_n,$$

con cada U_j elegida de forma independiente y uniforme del grupo de Clifford de un solo qubit. Este grupo está formado por las unitarias para las cuales la transformación $U_j^\dagger P U_j$ lleva cada operador de Pauli P a otro operador de Pauli, posiblemente con un cambio de signo. Por ejemplo, $U_j^\dagger Z U_j$ siempre corresponde a uno de los operadores $\pm X$, $\pm Y$ o $\pm Z$. El grupo contiene 24 elementos si se consideran equivalentes las unitarias que difieren únicamente en una fase global. Esta propiedad explica su utilidad en el procedimiento de medición. Aplicar U_j y medir después en la base computacional equivale a medir el estado original en una de las bases propias de X , Y o Z . La selección uniforme de las unitarias hace que las tres bases se elijan con la misma probabilidad. El posible cambio de signo solo intercambia la asignación de los resultados 0 y 1. Cada base contiene dos estados propios, de modo que las tres bases reúnen seis estados distintos.

Para el ensemble de unitarias $U = \bigotimes_{j=1}^n U_j$, con cada U_j seleccionada de forma independiente y uniforme del grupo de Clifford de un qubit, el canal de medición factoriza según $\mathcal{M}_P = \bigotimes_{j=1}^n \mathcal{M}_1$. La inversa del canal local aplicada al proyector asociado al resultado del qubit j está dada por

$$\mathcal{M}_1^{-1} \left(U_j^\dagger |\hat{b}_j\rangle \langle \hat{b}_j| U_j \right) = 3 U_j^\dagger |\hat{b}_j\rangle \langle \hat{b}_j| U_j - I,$$

donde I representa la identidad en el espacio de un qubit.

El snapshot del estado completo se obtiene, entonces, mediante el producto tensorial de estas contribuciones locales,

$$\hat{\rho} = \bigotimes_{j=1}^n \left(3 U_j^\dagger |\hat{b}_j\rangle \langle \hat{b}_j| U_j - I \right).$$

Cada factor individual posee autovalores 2 y -1 , por lo que no representa un estado físico válido. El snapshot resultante también presenta autovalores negativos. Sin embargo, la construcción sigue satisfaciendo $\mathbb{E}[\hat{\rho}] = \rho$, tal como se estableció para el protocolo general.

Esta elección de ensemble resulta especialmente conveniente cuando el observable O actúa de forma no trivial sobre un máximo de k qubits. Para las mediciones locales de Pauli, la varianza del estimador puede acotarse mediante la shadow norm de O , que satisface

$$\|O\|_{\text{shadow}}^2 \leq 4^k \|O\|_{\infty}^2,$$

donde $\|O\|_{\infty}$ es la norma de operador. El subíndice ∞ indica que, para el observable hermítico O , esta norma corresponde al mayor valor absoluto de sus autovalores. Esta cota es válida para cualquier observable con soporte en un máximo de k qubits.

Para una precisión y una confianza dadas, esta relación permite acotar el número suficiente de shots en función de k y de la norma del observable. Aquí, n representa el número total de qubits del sistema, mientras que $k \leq n$ cuenta únicamente aquellos sobre los que el observable actúa de forma no trivial. Por tanto, aumentar n no exige aumentar ese presupuesto si k y la norma del observable permanecen fijos. Sin embargo, el factor 4^k crece exponencialmente con k , de modo que la ventaja es especialmente relevante para observables que involucran pocos qubits.

3.2.2. Estimación de $I_{\text{w-w}}$ mediante Classical Shadows

Una vez establecido el protocolo de Classical Shadows con mediciones locales de Pauli, queda por precisar cómo se conecta con $I_{\text{w-w}}$. Este funcional está escrito en términos de probabilidades, mientras que Classical Shadows permite estimar valores de expectación de observables. El vínculo es directo para las mediciones proyectivas consideradas, ya que la probabilidad de cada resultado corresponde al valor de expectación de su proyector asociado. Consideremos como ejemplo $P_{C_{lr}}(00|z)$, donde z reúne los parámetros de las mediciones que se ajustan mediante CSPSA. Esta probabilidad corresponde al resultado $C_{lr} = 00$ de la medición definida por la unitaria

$U_C(z_C)$ seguida de una medición en la base computacional. Por la regla de Born,

$$P_{C_{lr}}(00|z) = \langle 00|U_C(z_C)\rho_C U_C^\dagger(z_C)|00\rangle,$$

donde ρ_C es el estado reducido de los dos qubits que recibe el detector C antes de aplicar la unitaria.

El resultado $C_{lr} = 00$ está asociado al proyector $\Pi_{00}^C(z_C) = U_C^\dagger(z_C)|00\rangle\langle 00|U_C(z_C)$, que actúa sobre los dos qubits del detector C . Este proyector permite expresar la probabilidad anterior como un valor de expectación sobre el estado original. Utilizando la propiedad cíclica de la traza se obtiene

$$P_{C_{lr}}(00|z) = \text{Tr} \left[\Pi_{00}^C(z_C) \rho_C \right].$$

La probabilidad anterior puede estimarse a partir de los N snapshots que componen la Classical Shadow del estado completo. Para ello, se obtiene de cada snapshot su reducción a los dos qubits del detector C , denotada por $\hat{\rho}_{C,i}$, y se calcula el promedio

$$\hat{P}_{C_{lr}}(00|z) = \frac{1}{N} \sum_{i=1}^N \text{Tr} \left[\Pi_{00}^C(z_C) \hat{\rho}_{C,i} \right].$$

La unitaria $U_C(z_C)$ interviene únicamente a través del proyector utilizado durante este cálculo. Por tanto, las probabilidades asociadas a distintas elecciones de medición pueden estimarse mediante el posprocesamiento de los mismos datos.

Para las marginales que conservan solo algunos bits se suman los proyectores de los resultados compatibles. Por ejemplo, $A_l = 1$ comprende los resultados $A_{lr} = 10$ y $A_{lr} = 11$, mientras que $B_l = 0$ comprende $B_{lr} = 00$ y $B_{lr} = 01$. Por tanto,

$$P_{A_l B_l}(10|z) = \text{Tr} \left[\left(\left[\Pi_{10}^A(z_A) + \Pi_{11}^A(z_A) \right] \otimes \left[\Pi_{00}^B(z_B) + \Pi_{01}^B(z_B) \right] \otimes I_C \right) \rho_{\text{tot}} \right],$$

donde ρ_{tot} es el estado conjunto de los seis qubits e I_C es la identidad sobre los dos qubits del detector C . Esta marginal se estima promediando el valor de expectación del observable indicado sobre los snapshots disponibles.

El procedimiento se extiende a las probabilidades conjuntas de los tres detectores. Por ejemplo, $P_{A_{lr} B_{lr} C_{lr}}(101111|z)$ corresponde a obtener $A_{lr} = 10$, $B_{lr} = 11$ y $C_{lr} = 11$. Su estimador se construye mediante

$$\hat{P}_{A_{lr} B_{lr} C_{lr}}(101111|z) = \frac{1}{N} \sum_{i=1}^N \text{Tr} \left[\left(\Pi_{10}^A(z_A) \otimes \Pi_{11}^B(z_B) \otimes \Pi_{11}^C(z_C) \right) \hat{\rho}_i \right].$$

Las contribuciones de los tres detectores se combinan dentro de cada snapshot antes de promediar, conservando así la información sobre sus correlaciones. El observable conjunto actúa sobre seis qubits, aunque se escriba como un producto de tres proyectores de dos qubits.

El soporte del observable depende del tipo de probabilidad que se desea estimar. El proyector asociado a un resultado de un solo detector actúa sobre dos qubits, mientras que una marginal conjunta de dos detectores puede actuar sobre cuatro. Las probabilidades conjuntas completas requieren observables con soporte sobre los seis qubits de la red. Por esta razón, el costo de estimar I_{w-w} mediante Classical Shadows está controlado principalmente por los observables de seis qubits.

El funcional I_{w-w} contiene dos términos lineales y diez productos de probabilidades. Para estimar estos últimos es necesario considerar que multiplicar dos estimadores obtenidos de los mismos datos puede introducir un sesgo, aun cuando cada estimador sea insesgado por separado. Una forma de evitarlo consiste en dividir los N snapshots independientes en dos grupos disjuntos de tamaños aproximadamente iguales y estimar cada factor con un grupo diferente.

Por ejemplo, el término $P_{C_{lr}}(00|z)P_{A_{lr}B_{lr}C_{lr}}(101111|z)$ se estima mediante

$$\hat{T}(z) = \hat{P}_{C_{lr}}^{(1)}(00|z)\hat{P}_{A_{lr}B_{lr}C_{lr}}^{(2)}(101111|z),$$

donde los superíndices identifican el grupo utilizado y cada promedio se normaliza por el tamaño de ese grupo. Para un valor de z fijado independientemente de los datos, ambos factores son estimadores independientes y el valor esperado de su producto coincide con el producto de las probabilidades. La misma partición puede utilizarse para todos los productos del funcional.

Al sustituir los doce términos por sus estimadores correspondientes se obtiene $\hat{I}_{w-w}(z)$. Bajo el modelo ideal de medición, con snapshots independientes y para un valor fijo de z , este estimador satisface

$$\hat{I}_{w-w}(z) = I_{w-w}(z) + \epsilon_N(z), \quad \mathbb{E}[\epsilon_N(z)] = 0, \quad \text{Var}[\epsilon_N(z)] = \mathcal{O}(N^{-1}).$$

Así, aumentar el número de shots permite reducir las fluctuaciones de la estimación.

El procedimiento completo se resume en el Algoritmo 2. Antes de iniciar CSPSA se adquieren los N snapshots y se dividen en dos grupos independientes, $\mathcal{S}_1$ y $\mathcal{S}_2$, de tamaños aproximadamente iguales. Esta partición permanece fija durante la búsqueda. El algoritmo recibe ambos grupos y los parámetros z de las mediciones que se desean evaluar. Las quince probabilidades distintas que aparecen en el funcional se denotan por $p_j(z)$ y sus observables asociados por $O_j(z)$.

Algoritmo 2 Estimación de $I_{w-w}(z)$ mediante Classical Shadows

```

1: procedure ESTIMAR  $I_{ww}(z, \mathcal{S}_1, \mathcal{S}_2)$ 
2:   Establecer  $N_1 = |\mathcal{S}_1|$ ,  $N_2 = |\mathcal{S}_2|$  y  $N = N_1 + N_2$ 
3:   for cada una de las quince probabilidades  $p_j(z)$  do
4:     Construir el observable  $O_j(z)$  correspondiente
5:     for  $g \in \{1, 2\}$  do
6:       Calcular  $\hat{p}_j^{(g)}(z) = \frac{1}{N_g} \sum_{\hat{\rho} \in \mathcal{S}_g} \text{Tr}[O_j(z)\hat{\rho}]$ 
7:     end for
8:     Calcular  $\hat{p}_j(z) = \frac{N_1 \hat{p}_j^{(1)}(z) + N_2 \hat{p}_j^{(2)}(z)}{N}$ 
9:   end for
10:  for cada uno de los doce términos de  $I_{w-w}$  do
11:    if el término es lineal then
12:      Utilizar la estimación  $\hat{p}_j(z)$  correspondiente
13:    else
14:      Estimar el producto  $p_j(z)p_\ell(z)$  mediante  $\hat{p}_j^{(1)}(z)\hat{p}_\ell^{(2)}(z)$ 
15:    end if
16:  end for
17:  Sumar los doce términos estimados con sus signos correspondientes
18:  return  $\hat{I}_{w-w}(z)$ 
19: end procedure

```

Durante la optimización, $\hat{I}_{w-w}$ reemplaza al funcional exacto en las evaluaciones de z_k^+ y z_k^- utilizadas por CSPSA. Si los mismos datos se reutilizan durante todas las iteraciones, los parámetros seleccionados dependen de las fluctuaciones de esa Classical Shadow. Por ello, el insesgamiento establecido para parámetros fijos no garantiza que los errores de evaluación mantengan media cero durante la búsqueda ni permite trasladar automáticamente las garantías de convergencia de CSPSA al funcional exacto.

La estimación mediante Classical Shadows permite seleccionar las mediciones que serán sometidas a una verificación final. En una implementación experimental, esta etapa consiste en fijar las mediciones obtenidas por CSPSA y realizar N_{st} repeticiones independientes de los datos utilizados durante la búsqueda. Las probabilidades se estiman directamente mediante las frecuencias observadas y se evalúa nuevamente I_{w-w} . Una violación estadísticamente significativa certifica la incompatibilidad con el modelo clásico triangular, bajo los supuestos causales de la red. La evaluación exacta mediante la regla de Born utilizada en las simulaciones constituye una referencia ideal de esta etapa.

Capítulo 4

Resultados numéricos de la maximización de la desigualdad Wagon-Wheel

En este capítulo se presentan los resultados obtenidos al aplicar el protocolo descrito en el **Capítulo 3** a la búsqueda de violaciones de la desigualdad Wagon-Wheel en la red triangular. En una primera etapa, el funcional I_{w-w} se evalúa de manera exacta a partir de las probabilidades $P_{ABC}(a, b, c)$. Esta evaluación permite comprobar el funcionamiento del método y estudiar su comportamiento tanto en configuraciones con una fuente separable como en aquellas donde las tres fuentes distribuyen estados pertenecientes a una misma familia.

Posteriormente, el funcional se estima mediante Classical Shadows. Esta segunda etapa permite analizar el desempeño del protocolo cuando las probabilidades no se conocen de manera exacta y deben inferirse a partir de un número finito de mediciones. La comparación entre ambos procedimientos permite evaluar el efecto de las fluctuaciones estadísticas sobre la optimización y sobre la certificación de la violación.

4.1. Aplicación del protocolo al problema de optimización

Consideremos la red triangular estándar, compuesta por tres fuentes independientes que distribuyen sistemas bipartitos de dos qubits entre Alicia, Bob y Charlie, de modo que cada parte recibe un qubit de cada una de las dos fuentes vecinas. El estado conjunto de los seis qubits se representa mediante una matriz densidad ρ definida

sobre $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, con $\dim \mathcal{H}_P = 4$ para $P \in \{A, B, C\}$. Cada parte aplica sobre sus dos qubits una transformación unitaria $U_P \in U(4)$, seguida de una medición en la base computacional. Este procedimiento define la medición proyectiva conjunta de cuatro resultados descrita en el [Capítulo 3](#).

Con esta notación, el máximo de I_{w-w} sobre las mediciones consideradas, para un estado ρ dado, se define como

$$I_{w-w}^{\max}(\rho) = \max_{U_A, U_B, U_C \in U(4)} I_{w-w}(\rho; U_A, U_B, U_C). \quad (4.1)$$

Este problema se aborda numéricamente mediante el protocolo CSPSA presentado en el [Capítulo 3](#), que actualiza de forma autoguiada las matrices U_A , U_B y U_C en cada iteración, sin fijar previamente una distribución objetivo hacia la cual converger. Debido a la naturaleza estocástica del algoritmo y a la dependencia de sus resultados respecto de las condiciones iniciales, cada optimización se realiza N_{rep} veces utilizando inicializaciones independientes, y cada repetición comprende K iteraciones. A partir de las trayectorias obtenidas se reportan la media, la mediana y los cuartiles Q_1 y Q_3 en cada iteración, así como los máximos alcanzados a lo largo de las trayectorias cuando corresponde. Salvo que se indique lo contrario, todas las simulaciones de este capítulo emplean $N_{\text{rep}} = 100$ repeticiones independientes y $K = 1000$ iteraciones de CSPSA por repetición.

4.2. Evaluación exacta del funcional wagon-wheel

Se consideran tres escenarios para los estados distribuidos por las fuentes, estados maximalmente entrelazados, estados con descomposición de Schmidt y estados de Werner.

4.2.1. Configuraciones con una fuente separable

En todas las simulaciones presentadas en esta subsección se consideran $K = 1000$ iteraciones y $N_{\text{rep}} = 100$ repeticiones independientes para cada configuración de estados. Como valor de referencia se utiliza

$$I_{w-w}^{\text{Fritz}} = \frac{\sqrt{2} - 1}{16} \approx 0,02589, \quad (4.2)$$

correspondiente al máximo valor del funcional Wagon-Wheel alcanzado por la distribución de Fritz.

Dos fuentes maximalmente entrelazadas

En esta simulación, dos de las tres fuentes de la red triangular, X e Y , distribuyen estados de Bell maximalmente entrelazados dados por

$$|\psi^-\rangle_X = |\psi^-\rangle_Y = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle), \quad (4.3)$$

mientras que la tercera fuente, Z , distribuye el estado separable $|00\rangle_Z$. Por tanto, el estado global de los seis qubits que llegan a los tres detectores queda dado por

$$|\Psi\rangle = |\psi^-\rangle_X \otimes |\psi^-\rangle_Y \otimes |00\rangle_Z. \quad (4.4)$$

Durante la optimización, los estados distribuidos por las fuentes permanecen fijos y CSPSA modifica únicamente las transformaciones unitarias U_A , U_B y U_C que definen las mediciones conjuntas realizadas en los tres detectores. En cada una de las N_{rep} repeticiones independientes, el algoritmo busca una configuración de mediciones que maximice $I_{\text{w-w}}$. Esta simulación permite examinar si la presencia de dos fuentes maximalmente entrelazadas es suficiente para obtener una violación de la desigualdad Wagon-Wheel cuando la tercera fuente permanece separable.

La [Figura 4.1](#) muestra que, para esta configuración, tanto la mediana como la media evolucionan desde valores negativos y se aproximan progresivamente a la cota clásica $I_{\text{w-w}} = 0$, sin superarla durante las K iteraciones simuladas. El recuadro interior permite observar que ambas cantidades permanecen en valores negativos del orden de 10^{-4} durante el tramo final de la optimización y, por tanto, alejadas de $I_{\text{w-w}}^{\text{Fritz}}$. Este comportamiento muestra que el protocolo no produce una violación en la tendencia central de las N_{rep} trayectorias. Sin embargo, la media y la mediana no permiten descartar que alguna repetición individual haya superado la cota clásica en una iteración intermedia, por lo que resulta necesario examinar el máximo alcanzado a lo largo de cada trayectoria.

La [Figura 4.2](#) presenta la distribución acumulada de los máximos alcanzados a lo largo de cada una de las N_{rep} trayectorias. De esta forma, el análisis permite determinar si alguna repetición superó la cota clásica en cualquier momento de la optimización, aunque posteriormente haya regresado a valores negativos. Los resultados muestran que el máximo de cada trayectoria permanece por debajo de $I_{\text{w-w}} = 0$. Incluso la realización que más se aproxima a la cota alcanza un valor negativo cuya distancia respecto de cero es del orden de 10^{-5} . Por consiguiente, ninguna de las $N_{\text{rep}}K$ evaluaciones exactas del funcional efectuadas en esta simulación produce una violación de la desigualdad Wagon-Wheel. Esta ausencia de violación dentro del presupuesto numérico considerado no permite concluir que la distribución subyacente admita una descripción clásica.

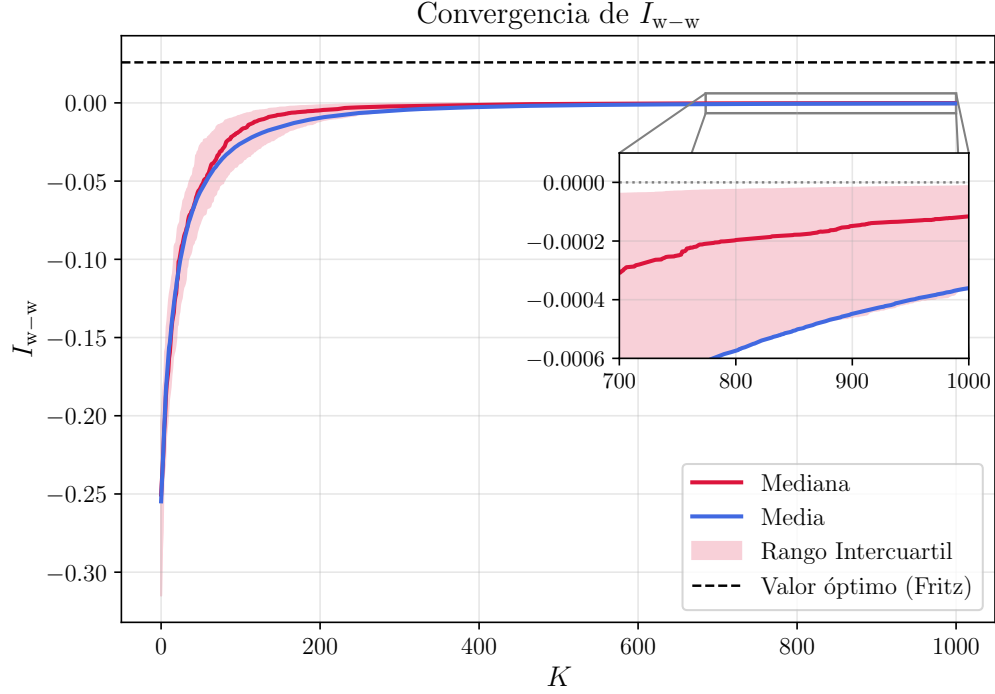


Figura 4.1: Evolución de I_{w-w} en función de K , considerando dos fuentes maximalmente entrelazadas y una fuente separable. La curva roja representa la mediana y la curva azul la media de las N_{rep} trayectorias independientes. Se muestran además el rango intercuartil, la cota clásica $I_{w-w} = 0$ e I_{w-w}^{Fritz} . El recuadro interior amplía el intervalo $K \in [700, 1000]$, en el que la mediana y la media permanecen por debajo de la cota clásica. **Fuente:** Elaboración propia.

Dos fuentes con estados en descomposición de Schmidt

Para extender el escenario anterior, se mantiene fija la fuente Z en el estado separable $|00\rangle_Z$, mientras que las fuentes X e Y distribuyen el mismo estado puro bipartito escrito en su descomposición de Schmidt,

$$|\psi(\alpha)\rangle_X = |\psi(\alpha)\rangle_Y = \sqrt{\alpha} |00\rangle + \sqrt{1-\alpha} |11\rangle, \quad \alpha \in [0, 1/2]. \quad (4.5)$$

El estado global de los seis qubits que llegan a los tres detectores queda dado por

$$|\Psi(\alpha)\rangle = |\psi(\alpha)\rangle_X \otimes |\psi(\alpha)\rangle_Y \otimes |00\rangle_Z. \quad (4.6)$$

El intervalo $\alpha \in [0, 1/2]$ permite recorrer todos los grados de entrelazamiento de esta familia sin pérdida de generalidad, puesto que los estados asociados a α y $1-\alpha$ son equivalentes mediante transformaciones unitarias locales. Para $\alpha = 0$, las fuentes

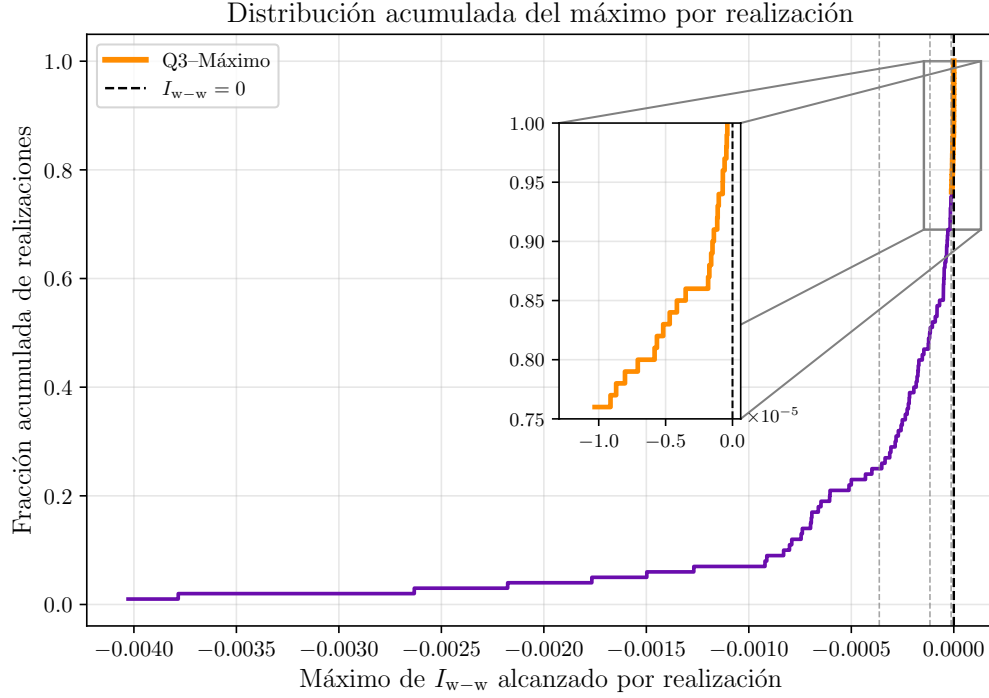


Figura 4.2: Función de distribución acumulada empírica de los máximos de I_{w-w} obtenidos para dos fuentes maximalmente entrelazadas y una fuente separable. Cada valor corresponde al máximo alcanzado durante las K iteraciones de una de las N_{rep} repeticiones independientes. El punto situado más a la derecha representa el mayor valor registrado entre las $N_{\text{rep}}K$ evaluaciones exactas del funcional. **Fuente:** Elaboración propia.

X e Y distribuyen estados producto y el estado global es completamente separable. Para $0 < \alpha < 1/2$, ambas fuentes distribuyen estados parcialmente entrelazados, mientras que $\alpha = 1/2$ corresponde al estado maximalmente entrelazado $|\phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, localmente equivalente al estado singlete considerado anteriormente. La optimización permite así estudiar cómo cambia el valor alcanzado por I_{w-w} al aumentar el entrelazamiento de las fuentes X e Y , manteniendo separable la fuente Z .

La **Figura 4.3** muestra que, para todos los valores de α considerados, la mediana de I_{w-w} evoluciona desde valores negativos y se aproxima a la cota clásica $I_{w-w} = 0$ sin superarla durante las K iteraciones. Este comportamiento extiende el resultado anterior a distintos grados de entrelazamiento de las fuentes X e Y . Incluso para $\alpha = 1/2$, correspondiente al caso maximalmente entrelazado, el protocolo no encuentra una violación mediana mientras la fuente Z permanece separable. Por tanto, dentro de la familia estudiada, aumentar el entrelazamiento de estas dos fuentes no basta

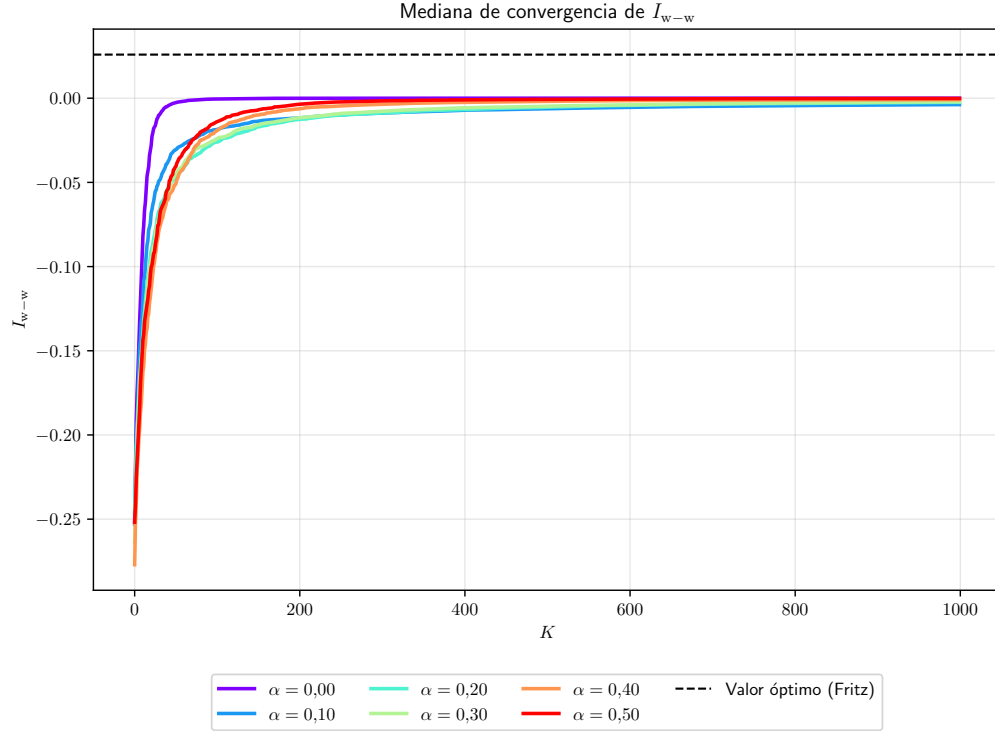


Figura 4.3: Mediana de I_{w-w} en función de K , considerando que las fuentes X e Y distribuyen estados con el mismo coeficiente de Schmidt α y que la fuente Z permanece fija en el estado separable $|00\rangle_Z$. Cada curva se obtiene a partir de N_{rep} repeticiones independientes y corresponde a uno de los valores $\alpha \in \{0,00, 0,10, 0,20, 0,30, 0,40, 0,50\}$. Se indican la cota clásica $I_{w-w} = 0$ e I_{w-w}^{Fritz} . **Fuente:** Elaboración propia.

para certificar la no localidad de la red mediante la desigualdad Wagon-Wheel.

Dos fuentes con estados de Werner

Para estudiar el efecto del ruido cuando una de las fuentes permanece separable, se considera que las fuentes X e Y distribuyen estados de Werner con un parámetro común λ ,

$$\rho_W(\lambda) = \lambda |\psi^-\rangle \langle \psi^-| + (1 - \lambda) \frac{\mathbb{I}_4}{4}, \quad 0 \leq \lambda \leq 1, \quad (4.7)$$

donde $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$. La tercera fuente se mantiene fija en el estado maximalmente mixto y separable $\rho_{\text{sep}} = \mathbb{I}_4/4$. Por tanto, los estados distribuidos por

las tres fuentes satisfacen

$$\rho_X(\lambda) = \rho_Y(\lambda) = \rho_W(\lambda), \quad \rho_Z = \rho_{\text{sep}} = \frac{\mathbb{I}_4}{4}, \quad (4.8)$$

y el estado global de los seis qubits queda dado por

$$\rho(\lambda) = \rho_W^X(\lambda) \otimes \rho_W^Y(\lambda) \otimes \rho_{\text{sep}}^Z. \quad (4.9)$$

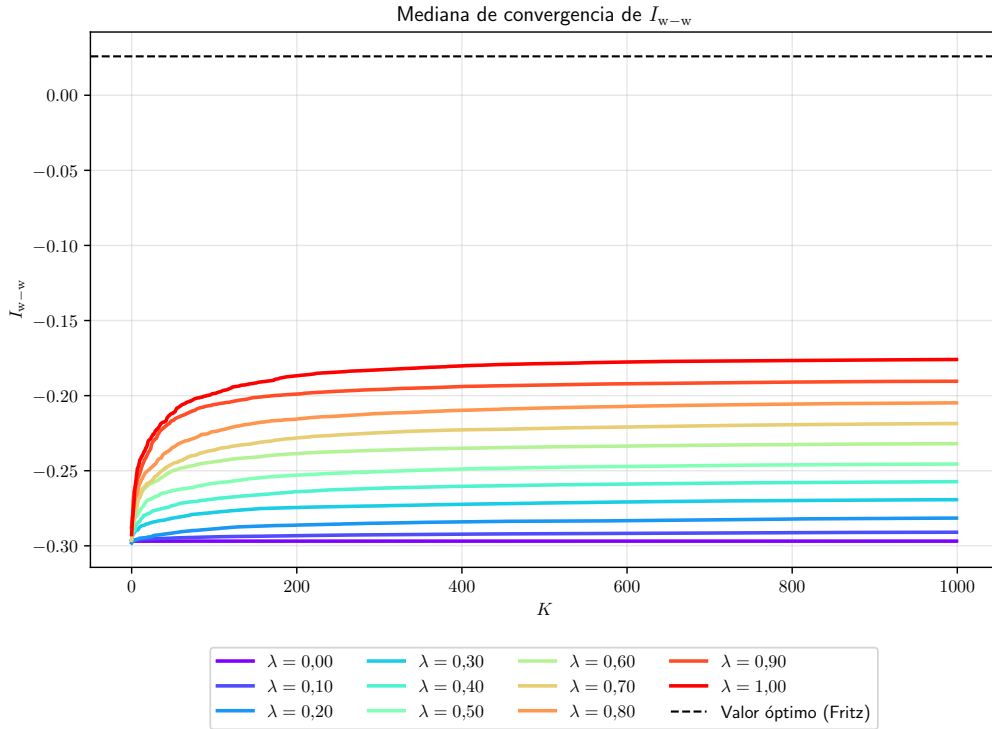


Figura 4.4: Mediana de I_{w-w} en función de K , obtenida a partir de N_{rep} repeticiones independientes, para dos fuentes con estados de Werner de parámetro común $\lambda \in [0, 1]$ y una fuente separable fija. El parámetro se modifica en pasos $\Delta\lambda = 0,10$ y la línea discontinua indica I_{w-w}^{Fritz} . **Fuente:** Elaboración propia.

El parámetro λ controla el peso del estado singlete en las fuentes X e Y . Para $\lambda = 0$, ambas distribuyen estados maximalmente mixtos, mientras que $\lambda = 1$ corresponde a dos estados singlete. Además, los estados de Werner son separables para $0 \leq \lambda \leq 1/3$ y entrelazados para $\lambda > 1/3$. Esta configuración permite estudiar si el aumento de la pureza y del entrelazamiento de las fuentes X e Y conduce a una violación de la desigualdad Wagon-Wheel cuando la fuente Z permanece separable.

La **Figura 4.4** muestra que la mediana alcanza valores progresivamente mayores a medida que aumenta λ , pero permanece por debajo de la cota clásica $I_{w-w} = 0$

durante las K iteraciones. Esto incluye el extremo $\lambda = 1$, en el que las fuentes X e Y distribuyen estados singlete y la fuente Z permanece en el estado separable maximalmente mixto. Por tanto, dentro de la familia estudiada, el aumento de la pureza y del entrelazamiento de estas dos fuentes no basta para que el protocolo certifique la no localidad de la red mediante la desigualdad Wagon-Wheel. Este resultado extiende la ausencia de violación observada anteriormente para una fuente separable pura al caso de una fuente separable maximalmente mixta.

4.2.2. Configuraciones con tres fuentes equivalentes

Después de estudiar las configuraciones en las que una fuente permanece separable, se consideran escenarios donde las tres fuentes distribuyen estados pertenecientes a una misma familia. Esto permite analizar cómo cambia el valor alcanzado por I_{w-w} cuando las tres conexiones de la red aportan recursos cuánticos equivalentes. Se estudian primero estados maximalmente entrelazados y luego estados en descomposición de Schmidt y estados de Werner.

En todas las simulaciones presentadas en esta subsección se consideran $K = 1000$ iteraciones y $N_{\text{rep}} = 100$ repeticiones independientes para cada configuración de estados. Los valores obtenidos se comparan con la cota clásica $I_{w-w} = 0$ y con I_{w-w}^{Fritz} .

Tres fuentes maximalmente entrelazadas

En esta simulación, las tres fuentes de la red triangular distribuyen el mismo estado de Bell maximalmente entrelazado,

$$|\psi^-\rangle_X = |\psi^-\rangle_Y = |\psi^-\rangle_Z = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \quad (4.10)$$

El estado global de los seis qubits que llegan a los tres detectores queda dado por

$$|\Psi\rangle = |\psi^-\rangle_X \otimes |\psi^-\rangle_Y \otimes |\psi^-\rangle_Z. \quad (4.11)$$

En este caso, las tres conexiones de la red se encuentran en igualdad de condiciones y distribuyen el mismo recurso cuántico. Esta configuración constituye el escenario de referencia para evaluar el funcionamiento del protocolo, pues permite determinar si CSPSA encuentra mediciones que violen la desigualdad Wagon-Wheel y recuperen el valor asociado a la distribución de Fritz.

La **Figura 4.5** muestra que la mediana y la media de las N_{rep} trayectorias evolucionan desde valores negativos, superan la cota clásica $I_{w-w} = 0$ durante las primeras etapas de la optimización y permanecen por encima de ella durante las iteraciones

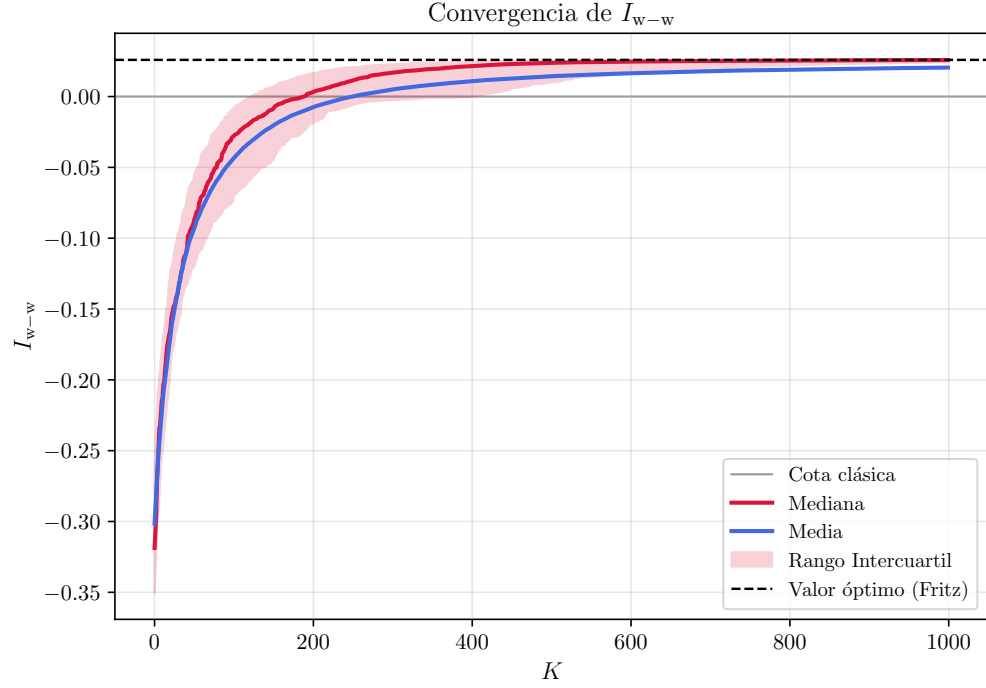


Figura 4.5: Evolución de I_{w-w} en función del número de iteraciones K , considerando que las tres fuentes de la red triangular distribuyen estados de Bell maximalmente entrelazados. La curva roja representa la mediana y la curva azul la media de las $N_{\text{rep}} = 100$ trayectorias independientes. Se muestran además el rango intercuartil, la cota clásica $I_{w-w} = 0$ y el valor de referencia de Fritz $(\sqrt{2} - 1)/16$. **Fuente:** Elaboración propia.

posteriores. Al finalizar las K iteraciones, la mediana se aproxima a I_{w-w}^{Fritz} , mientras que la media permanece ligeramente por debajo. Este comportamiento muestra que el protocolo encuentra de manera consistente mediciones capaces de violar la desigualdad Wagon-Wheel y recuperar numéricamente este valor.

Toda distribución compatible con un modelo causal clásico de la red triangular, en el que las tres fuentes operan de forma independiente, debe satisfacer $I_{w-w} \leq 0$. Por tanto, los valores positivos obtenidos certifican que las correlaciones generadas por los tres pares maximalmente entrelazados no admiten una descripción de esa forma. Esta certificación se obtiene en un escenario cerrado, donde cada detector realiza una medición fija y no dispone de diferentes ajustes seleccionados externamente. La proximidad de la mediana a I_{w-w}^{Fritz} debe interpretarse como la recuperación numérica de ese valor y no como una demostración de que corresponda al máximo cuántico global. La diferencia observada entre la mediana y la media indica, además, que las trayectorias no convergen de manera uniforme, aspecto que se analiza a

continuación.

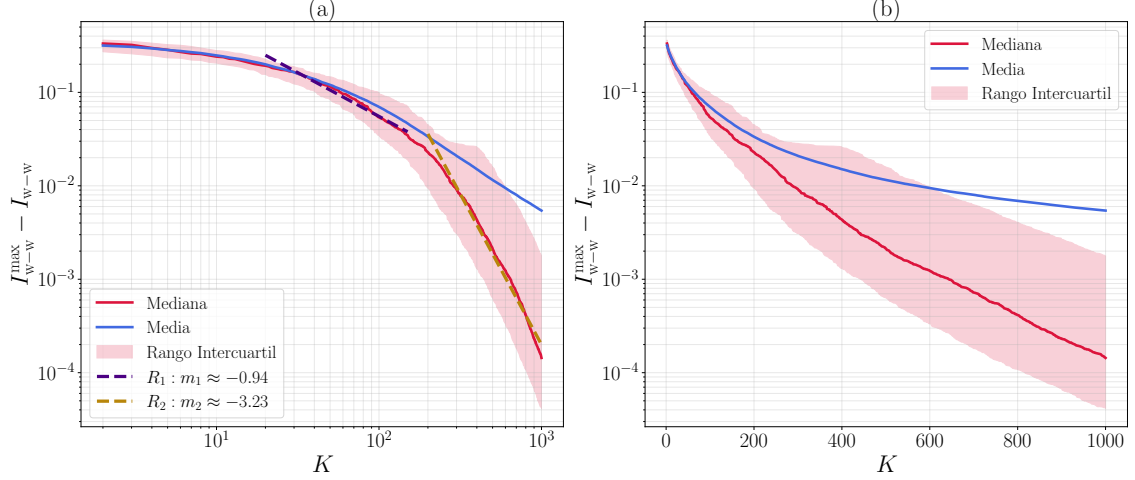


Figura 4.6: Diferencia entre I_{w-w}^{Fritz} y los valores representados por la mediana y la media de las N_{rep} trayectorias, junto con su rango intercuartil. El panel (a) presenta esta diferencia en escala log-log. Los intervalos R_1 y R_2 corresponden a los tramos situados antes y después del cambio de comportamiento observado durante la optimización, mientras que m_1 y m_2 representan las pendientes efectivas de los respectivos ajustes lineales. El panel (b) muestra la misma diferencia en escala semilogarítmica y permite examinar con mayor detalle su comportamiento durante el régimen tardío. **Fuente:** Elaboración propia.

La Figura 4.6 muestra que la aproximación a I_{w-w}^{Fritz} no ocurre con un ritmo uniforme. Como antecedente cuantitativo, los ajustes realizados en el panel (a) permiten distinguir dos intervalos con pendientes efectivas $m_1 \approx -0,94$ y $m_2 \approx -3,23$. Estos valores describen el comportamiento de la mediana dentro de los intervalos considerados y evidencian que su descenso se vuelve más pronunciado después de las primeras etapas de la optimización. No obstante, estas pendientes no se interpretan como exponentes que definan dos leyes de potencia asintóticas.

El aspecto central se observa en el panel (b). Para $K \gtrsim 400$, la diferencia entre la mediana e I_{w-w}^{Fritz} sigue una trayectoria aproximadamente lineal en escala semilogarítmica, lo que resulta compatible con un decaimiento exponencial durante el régimen tardío. La media, en cambio, se aproxima a I_{w-w}^{Fritz} con mayor lentitud y permanece a una distancia mayor al finalizar las K iteraciones. Esta separación muestra que la convergencia no se reproduce de manera uniforme en todas las trayectorias.

Tres fuentes con estados en descomposición de Schmidt

En esta simulación, las tres fuentes de la red triangular distribuyen el mismo estado en descomposición de Schmidt para dos qubits, parametrizado por un único coeficiente $\alpha \in [0, 1]$ común a las tres fuentes,

$$|\Psi(\alpha)\rangle_X = |\Psi(\alpha)\rangle_Y = |\Psi(\alpha)\rangle_Z = \sqrt{\alpha} |00\rangle + \sqrt{1-\alpha} |11\rangle. \quad (4.12)$$

El estado global de los seis qubits que llegan a los tres detectores es entonces

$$|\Psi\rangle = |\Psi(\alpha)\rangle_X \otimes |\Psi(\alpha)\rangle_Y \otimes |\Psi(\alpha)\rangle_Z. \quad (4.13)$$

El caso $\alpha = 1/2$ corresponde al estado maximalmente entrelazado considerado en la subsección anterior, mientras que $\alpha = 0$ corresponde a un estado producto sin entrelazamiento entre los dos qubits de cada fuente. Empleando el mismo protocolo, se evalúa la mediana de I_{w-w} a partir de las N_{rep} repeticiones para seis valores de α entre 0 y 0,5.

La [Figura 4.7](#) muestra que el valor final de la mediana no crece de forma monótona con α en todo el rango considerado. Para $\alpha = 0,30$, $\alpha = 0,40$ y $\alpha = 0,50$, la mediana supera la cota clásica $I_{w-w} = 0$ durante las K iteraciones y, para $\alpha = 0,50$, alcanza un valor final prácticamente indistinguible de I_{w-w}^{Fritz} . Para $\alpha = 0,10$ y $\alpha = 0,20$, en cambio, la mediana permanece por debajo de la cota clásica y su valor final resulta incluso inferior al alcanzado para $\alpha = 0,00$, pese a que este último corresponde al estado producto sin entrelazamiento. La curva asociada a $\alpha = 0,00$ converge rápidamente y se estabiliza cerca de $I_{w-w} = 0$ antes de $K \approx 50$, sin superar la cota de forma apreciable.

La superación de la cota clásica para $\alpha = 0,30$, $\alpha = 0,40$ y $\alpha = 0,50$ muestra que el protocolo encuentra mediciones capaces de generar correlaciones incompatibles con un modelo causal clásico de la red triangular basado en tres fuentes independientes. Además, para $\alpha = 0,50$, la mediana se aproxima a I_{w-w}^{Fritz} y reproduce el comportamiento observado en la subsección anterior. En cambio, para $\alpha = 0,00$, $\alpha = 0,10$ y $\alpha = 0,20$, la mediana no supera la cota clásica. Este resultado solo muestra que la tendencia central de las trayectorias no permite certificar la no localidad dentro del presupuesto numérico considerado.

El orden observado en el valor final de la mediana no coincide con el orden creciente de α para $\alpha \leq 0,20$. En particular, el resultado para $\alpha = 0,00$ queda por encima de los obtenidos para $\alpha = 0,10$ y $\alpha = 0,20$. Esta diferencia podría deberse a distintas velocidades de convergencia, más que a los valores asintóticos que alcanzaría cada configuración si se extendiera K , de forma análoga a los regímenes identificados en la subsección anterior.

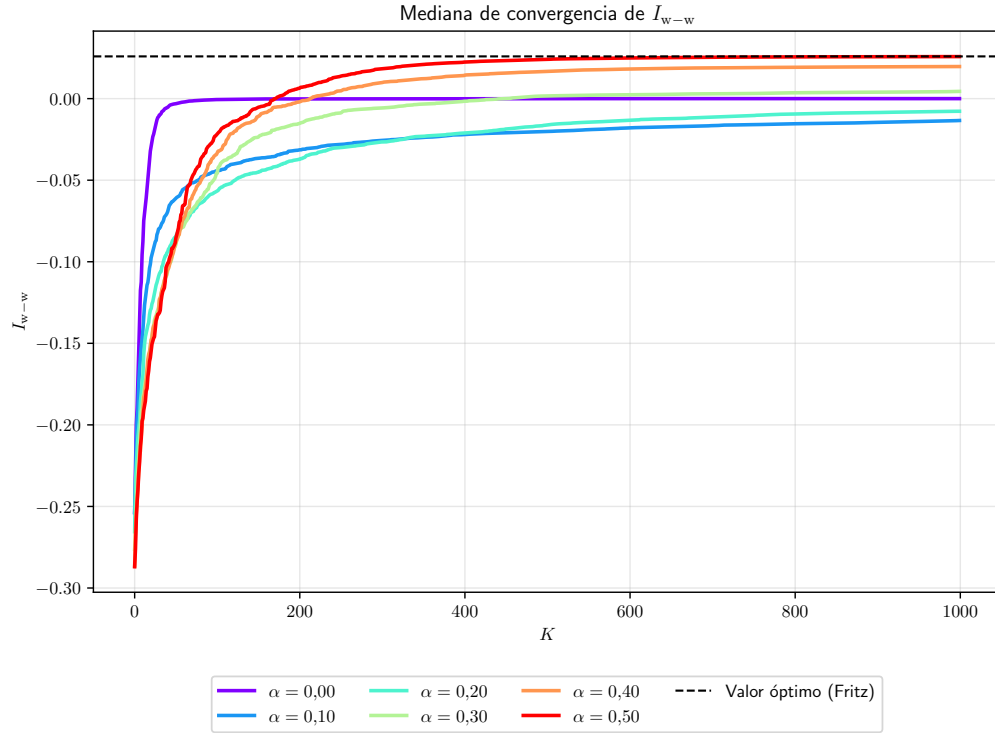


Figura 4.7: Mediana de I_{w-w} en función de K , obtenida a partir de N_{rep} repeticiones independientes, con las tres fuentes distribuyendo estados con el mismo coeficiente de Schmidt α . Se consideran los valores $\alpha \in \{0,00, 0,10, 0,20, 0,30, 0,40, 0,50\}$. La línea punteada indica I_{w-w}^{Fritz} . **Fuente:** Elaboración propia.

La Figura 4.8 muestra, con mayor resolución en α , la tendencia identificada en la Figura 4.7. A medida que α aumenta dentro del intervalo considerado, la mediana supera la cota clásica en una iteración más temprana y permanece por encima de ella durante una mayor fracción de las K iteraciones. Para $\alpha = 0,50$, la curva se estabiliza cerca de I_{w-w}^{Fritz} a partir de $K \approx 400$. Esta resolución adicional muestra una progresión ordenada hacia valores mayores de I_{w-w} conforme aumenta α , sin los cambios de orden observados para $\alpha \leq 0,20$ en la figura anterior.

Dentro de este subintervalo, a medida que aumenta el entrelazamiento de los estados distribuidos por las fuentes, el protocolo supera la cota clásica con mayor facilidad. Esto se manifiesta en un cruce más temprano de $I_{w-w} = 0$ y en una permanencia más prolongada por encima de esta cota.

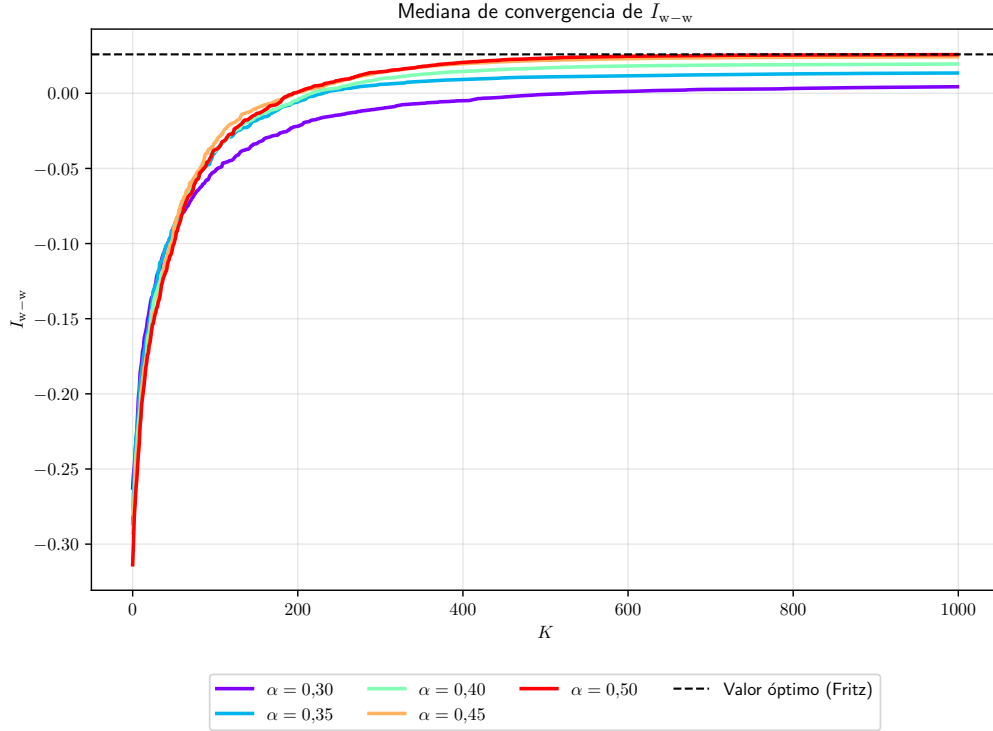


Figura 4.8: Ampliación de los resultados presentados en la [Figura 4.7](#), que incorpora la mediana de I_{w-w} en función de K para valores de $\alpha \in [0,30, 0,50]$ con paso $\Delta\alpha = 0,05$. Las curvas se obtienen mediante el mismo protocolo y a partir de N_{rep} repeticiones independientes. La línea punteada indica I_{w-w}^{Fritz} . **Fuente:** Elaboración propia.

Tres fuentes con estados de Werner

Para estudiar el efecto del ruido sobre la violación de la desigualdad Wagon-Wheel, se considera que las tres fuentes de la red triangular distribuyen el mismo estado de Werner,

$$\rho_W(\lambda) = \lambda |\psi^-\rangle \langle \psi^-| + (1 - \lambda) \frac{\mathbb{I}_4}{4}, \quad 0 \leq \lambda \leq 1, \quad (4.14)$$

donde $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ y λ determina el peso del estado singlete frente al estado maximalmente mixto. El estado global de los seis qubits queda dado por

$$\rho(\lambda) = \rho_W^X(\lambda) \otimes \rho_W^Y(\lambda) \otimes \rho_W^Z(\lambda). \quad (4.15)$$

Para $\lambda = 0$, cada fuente distribuye el estado maximalmente mixto $\mathbb{I}_4/4$, mientras

que $\lambda = 1$ recupera el caso de tres singletes analizado anteriormente. Además, los estados de Werner son separables para $0 \leq \lambda \leq 1/3$ y entrelazados para $\lambda > 1/3$. Para cada valor de λ , los estados permanecen fijos y CSPA optimiza las mediciones conjuntas U_A , U_B y U_C .

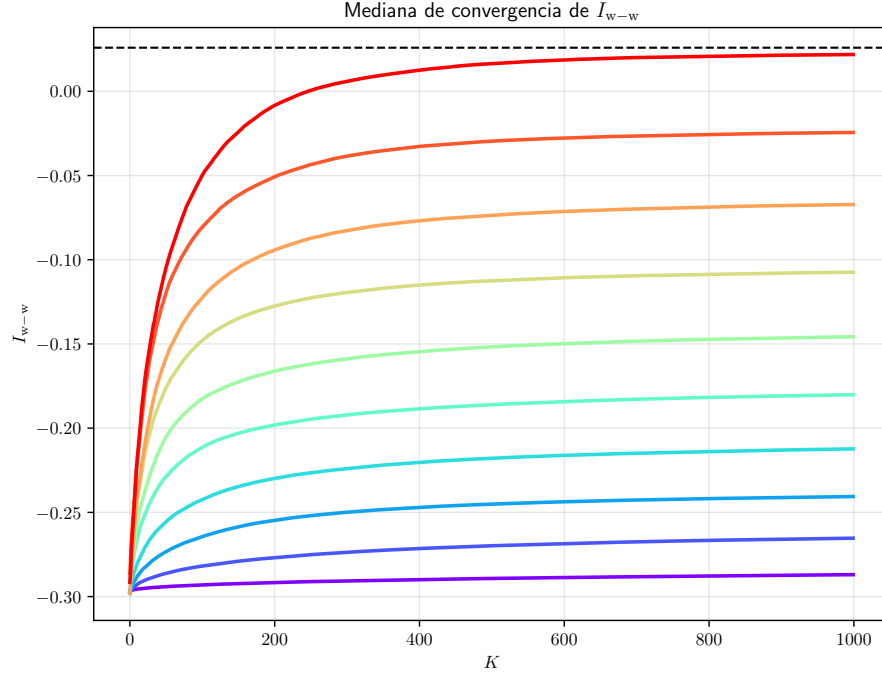


Figura 4.9: Mediana de I_{w-w} en función de K , obtenida a partir de N_{rep} repeticiones independientes, considerando que las tres fuentes distribuyen estados de Werner con un parámetro común $\lambda \in \{0,10, 0,20, \dots, 1,00\}$. La línea discontinua indica I_{w-w}^{Fritz} . **Fuente:** Elaboración propia.

La **Figura 4.9** muestra un incremento progresivo del valor final de la mediana a medida que aumenta λ . Para $\lambda = 0,10, \dots, 0,90$, las curvas permanecen por debajo de la cota clásica $I_{w-w} = 0$ durante las K iteraciones. En cambio, para $\lambda = 1,00$, la mediana supera esta cota y se aproxima a I_{w-w}^{Fritz} . Este último caso coincide con la configuración de tres fuentes maximalmente entrelazadas estudiada anteriormente.

El barrido inicial sitúa el cruce de la cota clásica por parte de la mediana entre $\lambda = 0,90$ y $\lambda = 1,00$. Para examinar esta región con mayor detalle, se repite la simulación utilizando valores separados por un paso $\Delta\lambda = 0,01$.

La **Figura 4.10** permite precisar el cambio de comportamiento observado en el primer barrido. Para $0,90 \leq \lambda \leq 0,94$, la mediana se estabiliza en valores negativos y no supera la cota clásica. A partir de $\lambda = 0,95$, las curvas superan $I_{w-w} = 0$ y

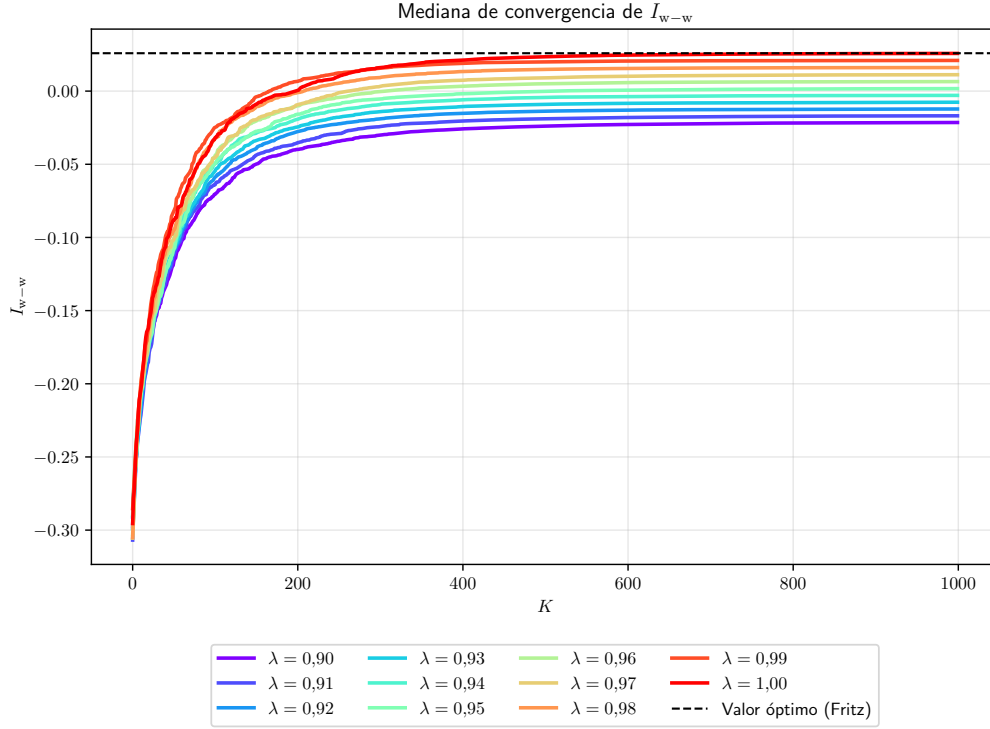


Figura 4.10: Mediana de I_{w-w} en función de K , obtenida a partir de N_{rep} repeticiones independientes, para $\lambda \in [0,90, 1,00]$ con paso $\Delta\lambda = 0,01$. La línea discontinua indica I_{w-w}^{Fritz} . **Fuente:** Elaboración propia.

terminan la optimización dentro de la región de violación. Asimismo, cuanto mayor es λ , más temprano ocurre este cruce y más próximo a I_{w-w}^{Fritz} se encuentra el resultado final.

Dentro de la resolución considerada, el umbral numérico de detección se sitúa entre $\lambda = 0,94$ y $\lambda = 0,95$. Para $\lambda \geq 0,95$, el valor positivo de la mediana indica que al menos la mitad de las repeticiones encuentra mediciones que generan distribuciones incompatibles con el modelo clásico de la red triangular. Para valores inferiores, el hecho de que la mediana no supere la cota clásica no establece la compatibilidad clásica de las correlaciones. Solo muestra que el protocolo no certifica su no localidad de red durante las K iteraciones.

4.3. Estimación de I_{w-w} mediante Classical Shadows

El análisis se extiende ahora al escenario en que el valor de I_{w-w} debe estimarse a partir de un número finito de mediciones. Para ello, se incorporan Classical Shadows como fuente de información durante la optimización. En los dos procedimientos considerados, CSPSA utiliza las estimaciones del funcional en los puntos perturbados z_k^+ y z_k^- para aproximar el gradiente y actualizar los parámetros de las mediciones. La diferencia entre ambos se encuentra en la evaluación de las mediciones actualizadas. En Classical Shadows con evaluación exacta, esta evaluación se realiza utilizando el estado exacto ρ y el valor obtenido se denota mediante I_{w-w}^{ex} . En Full Classical Shadows, la evaluación también se realiza mediante Classical Shadows, por lo que el valor correspondiente se denota como $\hat{I}_{w-w}^{\text{sh}}$.

El procedimiento Classical Shadows con evaluación exacta se desarrolla mediante las siguientes etapas generales.

1. Al comienzo de cada repetición se generan N_{sh} registros de Classical Shadows mediante mediciones locales aleatorias de Pauli sobre copias independientes del estado. Este conjunto se mantiene fijo durante las K iteraciones y permite estimar las probabilidades y los marginales que componen el funcional Wagon-Wheel para cada configuración de medición considerada.
2. Teniendo estas estimaciones, se calculan los valores aproximados del funcional. A partir de ellos, CSPSA construye el gradiente estimado y determina la actualización de los parámetros de medición z_{k+1} que se utilizarán en la siguiente iteración.
3. Una vez obtenida la medición actualizada z_{k+1} , esta se evalúa utilizando el estado exacto ρ . Mediante la regla de Born se calculan las probabilidades y los marginales exactos, con los cuales se obtiene $I_{w-w}^{\text{ex}}(z_{k+1})$. Este valor se almacena para analizar la medición encontrada y no interviene en la actualización posterior.

La secuencia de cálculo puede resumirse mediante

$$\begin{aligned} \hat{P}_\mu^{\text{sh}}(z_k^\pm) &\longrightarrow I_{w-w}[\hat{P}_\mu^{\text{sh}}(z_k^\pm)] \longrightarrow \hat{g}_k \longrightarrow z_{k+1}, \\ z_{k+1} &\longrightarrow P_\mu^{\text{ex}}(z_{k+1}; \rho) \longrightarrow I_{w-w}^{\text{ex}}(z_{k+1}; \rho). \end{aligned} \quad (4.16)$$

donde μ identifica las probabilidades y los marginales que componen el funcional Wagon-Wheel, mientras que ρ representa el estado cuántico global exacto utilizado en la simulación.

En Full Classical Shadows, tanto la búsqueda como el seguimiento de sus resultados se realizan a partir de la información estadística proporcionada por Classical Shadows. El procedimiento se organiza en las siguientes etapas.

1. Para cada repetición se recopilan N_{sh} registros mediante mediciones locales aleatorias de Pauli. Estos datos se conservan sin modificaciones y se consultan a lo largo de las K iteraciones.
2. En cada iteración, los registros permiten estimar las probabilidades y los marginales asociados a las configuraciones consideradas. Los valores aproximados de Wagon-Wheel contruidos con estas estimaciones permiten calcular el gradiente y obtener los nuevos parámetros de medición z_{k+1} .
3. A diferencia del procedimiento anterior, la configuración actualizada no se evalúa utilizando el estado exacto ρ . El mismo conjunto de Classical Shadows se emplea para estimar las probabilidades y los marginales correspondientes a z_{k+1} , con los cuales se calcula $\hat{I}_{\text{w-w}}^{\text{sh}}(z_{k+1})$.

La secuencia de cálculo queda representada por

$$\begin{aligned} \hat{P}_{\mu}^{\text{sh}}(z_k^{\pm}; \hat{\rho}_{\text{CS}}) &\longrightarrow I_{\text{w-w}} \left[\hat{P}_{\mu}^{\text{sh}}(z_k^{\pm}; \hat{\rho}_{\text{CS}}) \right] \longrightarrow \hat{g}_k \longrightarrow z_{k+1}, \\ z_{k+1} &\longrightarrow \hat{P}_{\mu}^{\text{sh}}(z_{k+1}; \hat{\rho}_{\text{CS}}) \longrightarrow \hat{I}_{\text{w-w}}^{\text{sh}}(z_{k+1}; \hat{\rho}_{\text{CS}}). \end{aligned} \quad (4.17)$$

donde μ identifica las probabilidades y los marginales que componen el funcional Wagon-Wheel, mientras que $\hat{\rho}_{\text{CS}}$ representa el estimador del estado cuántico global construido a partir de los N_{sh} registros de Classical Shadows.

Este procedimiento reproduce una situación en la que no se dispone del estado exacto para examinar las mediciones encontradas. En consecuencia, Classical Shadows determina tanto la dirección de la búsqueda como el valor utilizado para describir su desempeño. Esta característica permite estudiar una implementación basada completamente en datos de medición, aunque hace necesario verificar posteriormente la violación mediante un conjunto independiente de realizaciones.

Tres estados maximalmente entrelazados

La [Figura 4.11](#) presenta la mediana de $I_{\text{w-w}}^{\text{ex}}$ obtenida al evaluar con el estado exacto ρ las mediciones encontradas mediante la optimización guiada por Classical Shadows. Para $N_{\text{sh}} = 10^3$, la curva aumenta durante las primeras iteraciones, pero se estabiliza alrededor de $I_{\text{w-w}}^{\text{ex}} \approx -0,24$. Al aumentar el presupuesto a $N_{\text{sh}} = 10^4$, la mediana se aproxima más a la cota clásica, aunque permanece en valores negativos al finalizar la optimización. En ambos casos, las mediciones encontradas no permiten que la mediana de la evaluación exacta supere $I_{\text{w-w}} = 0$.

El comportamiento cambia para $N_{\text{sh}} = 10^5$ y $N_{\text{sh}} = 10^6$. La mediana supera la cota clásica aproximadamente en $K \approx 250$ y $K \approx 180$, respectivamente, y permanece positiva durante las iteraciones posteriores. Además, para $N_{\text{sh}} = 10^6$ se aproxima estrechamente a $I_{\text{w-w}}^{\text{Fritz}}$, mientras que para $N_{\text{sh}} = 10^5$ se estabiliza ligeramente por debajo. Estos resultados muestran que el aumento de N_{sh} mejora la capacidad de Classical Shadows para guiar la búsqueda hacia mediciones que producen una violación al ser evaluadas con el estado exacto. Por tanto, el cruce observado para los mayores presupuestos no corresponde a una sobreestimación del valor representado en la figura, pues cada punto de estas curvas se calcula mediante la regla de Born utilizando ρ .

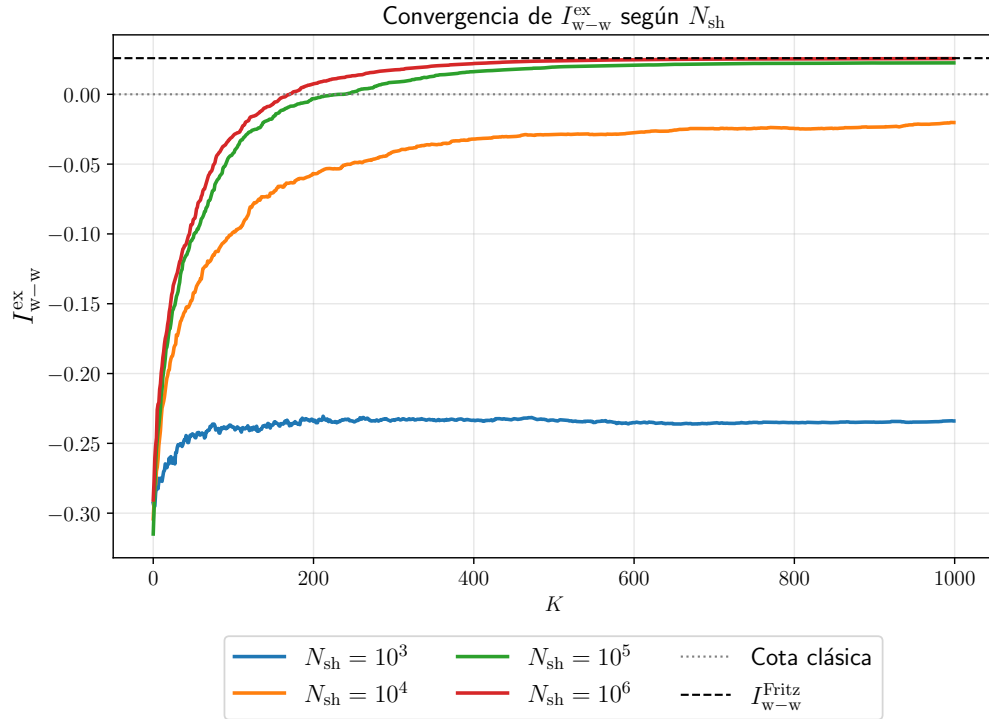


Figura 4.11: Mediana de $I_{\text{w-w}}^{\text{ex}}$ en función de K , obtenida al evaluar con el estado exacto ρ las mediciones encontradas mediante la optimización guiada por Classical Shadows. Las curvas corresponden a distintos valores de N_{sh} . Se indican la cota clásica $I_{\text{w-w}} = 0$ y el valor $I_{\text{w-w}}^{\text{Fritz}}$. **Fuente:** Elaboración propia.

La **Figura 4.12** examina con mayor detalle los presupuestos para los cuales la mediana de $I_{\text{w-w}}^{\text{ex}}$ supera la cota clásica. Para $N_{\text{sh}} = 10^6$, este cruce ocurre antes que para $N_{\text{sh}} = 10^5$ y la curva se aproxima más estrechamente a $I_{\text{w-w}}^{\text{Fritz}}$. Este comportamiento muestra que el aumento de N_{sh} permite que la optimización guiada por Classical Shadows encuentre con mayor rapidez mediciones que producen una violación al ser evaluadas con el estado exacto ρ .

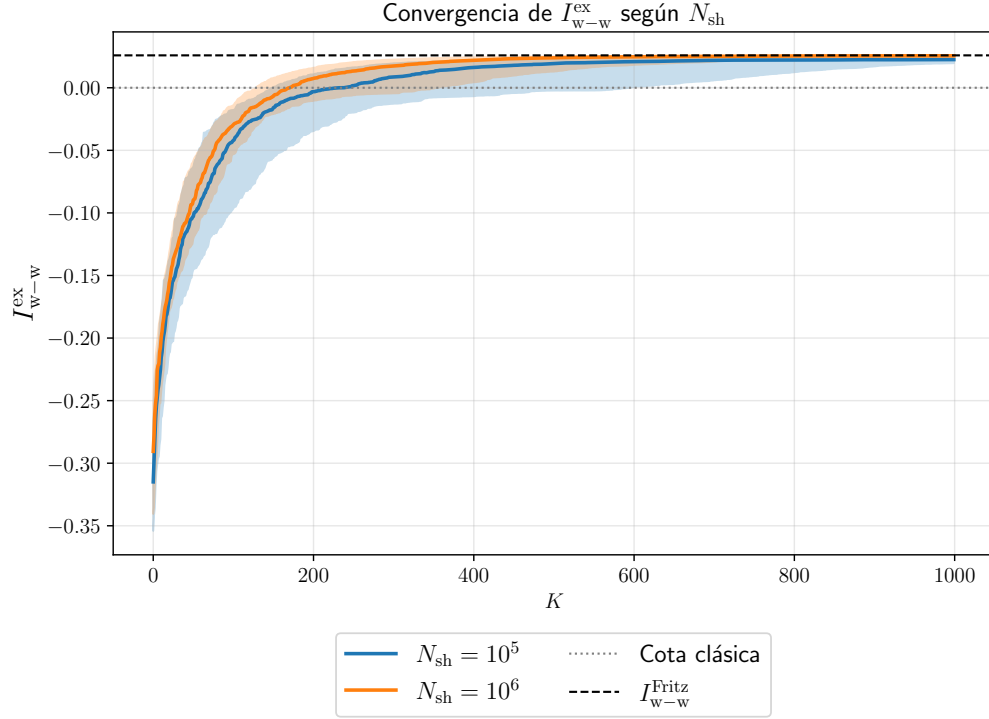


Figura 4.12: Mediana de I_{w-w}^{ex} en función de K , obtenida al evaluar con el estado exacto ρ las mediciones encontradas mediante la optimización guiada por Classical Shadows. Se consideran los presupuestos $N_{sh} = 10^5$ y $N_{sh} = 10^6$, mientras que las regiones sombreadas representan los respectivos rangos intercuartiles. Las líneas horizontales punteada y discontinua indican la cota clásica y I_{w-w}^{Fritz} , respectivamente. **Fuente:** Elaboración propia.

Los rangos intercuartiles permiten evaluar la consistencia de este resultado entre las distintas repeticiones. Para $N_{sh} = 10^6$, el rango intercuartil se desplaza completamente por encima de la cota clásica durante una etapa más temprana de la optimización. En cambio, para $N_{sh} = 10^5$, la mediana se vuelve positiva mientras el extremo inferior del rango intercuartil permanece bajo la cota durante un intervalo más prolongado. Cuando este extremo también supera $I_{w-w} = 0$, al menos el 75 % de las repeticiones presenta valores exactos positivos. Hacia el final de la optimización, esta condición se cumple para ambos presupuestos, aunque se alcanza antes y con una menor dispersión para $N_{sh} = 10^6$.

La **Figura 4.13** muestra que el aumento de N_{sh} reduce aproximadamente en un orden de magnitud la distancia final respecto de I_{w-w}^{Fritz} , desde una distancia del orden de 10^{-3} para $N_{sh} = 10^5$ hasta una del orden de 10^{-4} para $N_{sh} = 10^6$. Además, mientras la primera curva tiende a estabilizarse durante las últimas iteraciones, la

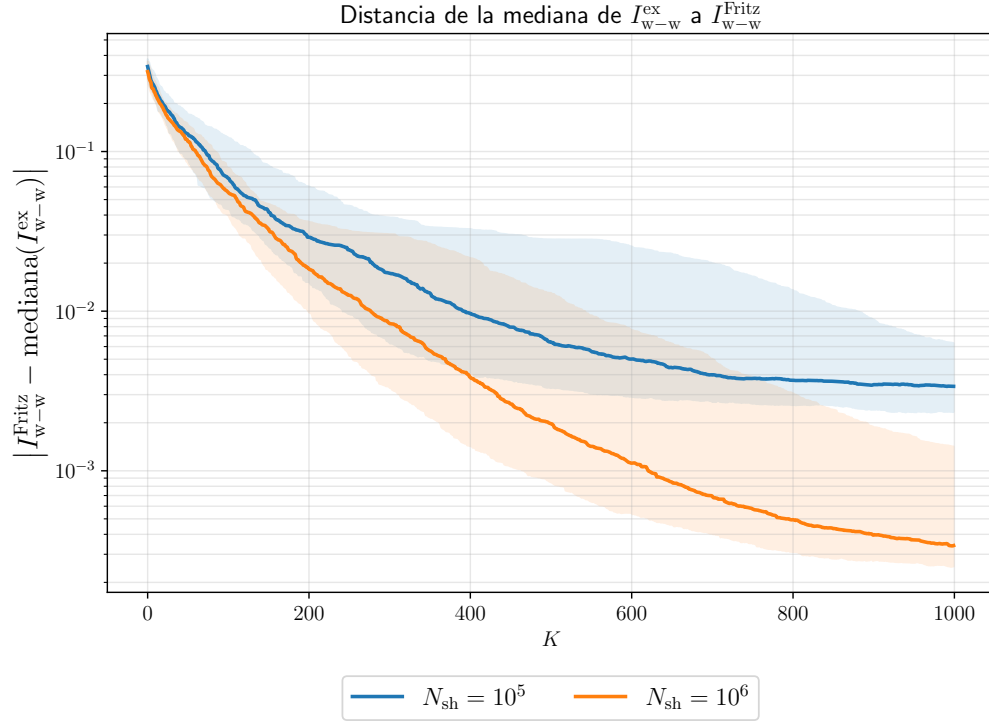


Figura 4.13: Diferencia absoluta entre I_{w-w}^{Fritz} y la mediana de I_{w-w}^{ex} en función de K , para los presupuestos $N_{sh} = 10^5$ y $N_{sh} = 10^6$. Las regiones sombreadas están delimitadas por las distancias del primer y tercer cuartil de I_{w-w}^{ex} respecto de I_{w-w}^{Fritz} . **Fuente:** Elaboración propia.

segunda mantiene su descenso durante un intervalo más prolongado. Por tanto, el mayor presupuesto de Classical Shadows permite que la mediana de I_{w-w}^{ex} alcance una aproximación más precisa a I_{w-w}^{Fritz} .

El análisis de la configuración en la que las tres fuentes distribuyen estados singlete continúa ahora mediante la comparación entre Classical Shadows con evaluación exacta y Full Classical Shadows. Para ello, se examinan los resultados obtenidos mediante ambos procedimientos para distintos valores de N_{sh} .

La **Figura 4.14** muestra que ambas evaluaciones difieren considerablemente para los presupuestos más bajos. Para $N_{sh} = 10^3$, la mediana de $\hat{I}_{w-w}^{sh}$ alcanza un valor cercano a 0,6, mientras que la evaluación exacta de las mismas mediciones se estabiliza alrededor de $-0,24$. Para $N_{sh} = 10^4$, Full Classical Shadows todavía supera I_{w-w}^{Fritz} , aunque la mediana exacta permanece cerca de $-0,02$. En cambio, para $N_{sh} = 10^5$ y $N_{sh} = 10^6$, la concordancia entre ambas evaluaciones mejora y el rango intercuartil de I_{w-w}^{ex} termina por superar la cota clásica. Esto indica que al menos el 75 % de las repeticiones encuentra mediciones que producen una violación al ser evaluadas con

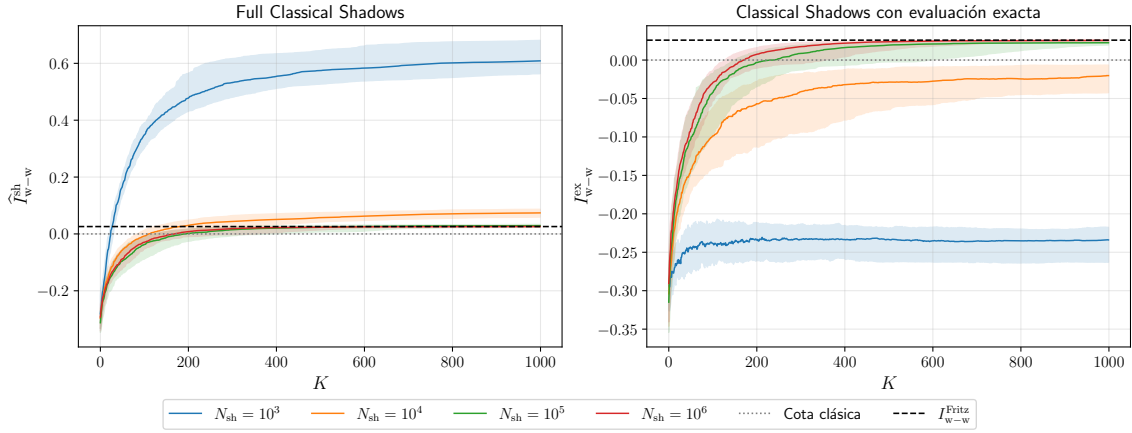


Figura 4.14: Mediana de I_{w-w} en función de K para distintos valores de N_{sh} , considerando que las tres fuentes distribuyen estados singlete. El panel izquierdo muestra $\hat{I}_{w-w}^{sh}$, utilizado para guiar y evaluar la optimización mediante Full Classical Shadows. El panel derecho presenta I_{w-w}^{ex} , obtenido al evaluar con el estado exacto ρ las mismas mediciones. Las regiones sombreadas representan los rangos intercuartiles, mientras que las líneas horizontales punteada y discontinua indican la cota clásica e I_{w-w}^{Fritz} , respectivamente. **Fuente:** Elaboración propia.

el estado exacto.

La discrepancia para los presupuestos menores se explica por la reutilización del mismo conjunto finito de Classical Shadows durante toda la optimización. CSPSA puede favorecer mediciones que explotan sus fluctuaciones estadísticas y elevar $\hat{I}_{w-w}^{sh}$ sin mejorar el valor exacto. Este comportamiento es consistente con la estimación de peor caso para los términos que actúan sobre los seis qubits de la red, según la cual resolver una violación del orden de I_{w-w}^{Fritz} requiere $N_{sh} \gtrsim 3^6 / (I_{w-w}^{Fritz})^2 \sim 10^6$. Por tanto, N_{sh} y K deben elegirse conjuntamente, pues aumentar las iteraciones con un presupuesto insuficiente puede intensificar el sesgo de optimización.

Tres estados en descomposición de Schmidt

Corresponde ahora analizar los estados escritos en descomposición de Schmidt, cuyo grado de entrelazamiento depende del parámetro común α . Para $N_{sh} = 10^5$, la Figura 4.15 muestra que la concordancia entre ambos procedimientos depende de este parámetro. Para $\alpha = 0,00$, Full Classical Shadows produce una mediana positiva cercana a I_{w-w}^{Fritz} , mientras que la evaluación exacta permanece junto a la cota clásica. Para $\alpha = 0,25$, ambas curvas permanecen bajo esta cota, aunque todavía presentan una diferencia. En cambio, para $\alpha = 0,50$, los dos procedimientos superan la cota clásica y se aproximan a I_{w-w}^{Fritz} . Por tanto, $N_{sh} = 10^5$ permite reproducir el

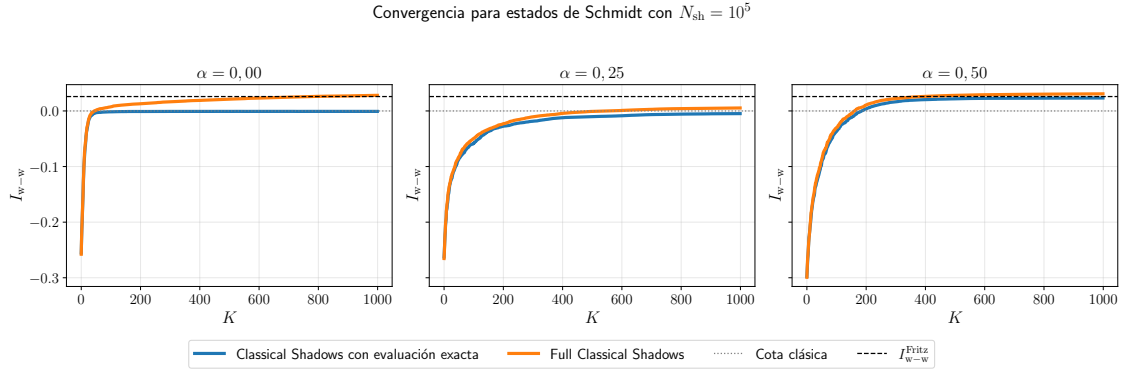


Figura 4.15: Mediana de I_{w-w} en función de K , considerando que las tres fuentes distribuyen estados con el mismo coeficiente de Schmidt α y que la optimización utiliza $N_{\text{sh}} = 10^5$. Los paneles corresponden a $\alpha \in \{0,00, , 0,25, , 0,50\}$. La curva azul representa Classical Shadows con evaluación exacta y la curva naranja Full Classical Shadows. Las líneas horizontales punteada y discontinua indican la cota clásica e I_{w-w}^{Fritz} , respectivamente. **Fuente:** Elaboración propia.

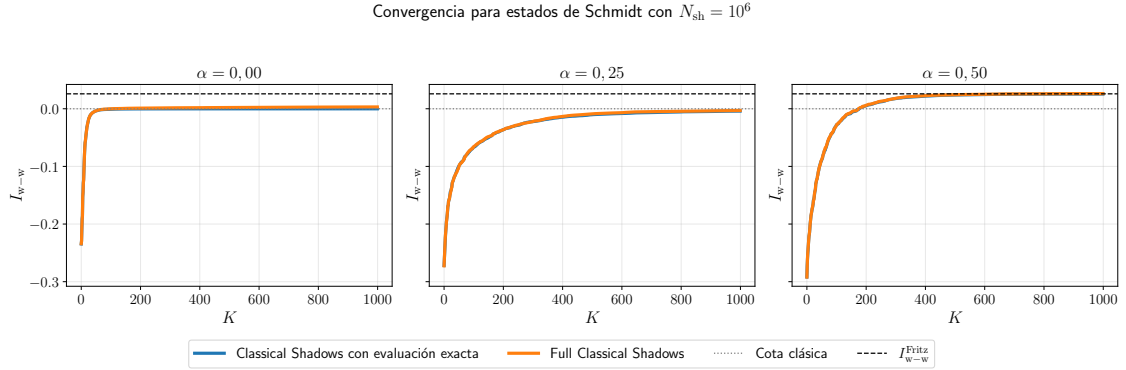


Figura 4.16: Mediana de I_{w-w} en función de K , considerando que las tres fuentes distribuyen estados con el mismo coeficiente de Schmidt α y que la optimización utiliza $N_{\text{sh}} = 10^6$. Los paneles corresponden a $\alpha \in \{0,00, , 0,25, , 0,50\}$. La curva azul representa Classical Shadows con evaluación exacta y la curva naranja Full Classical Shadows. Las líneas horizontales punteada y discontinua indican la cota clásica e I_{w-w}^{Fritz} , respectivamente. **Fuente:** Elaboración propia.

comportamiento del caso maximalmente entrelazado, pero no proporciona resultados consistentes para toda la familia.

Al aumentar el número de Classical Shadows a $N_{\text{sh}} = 10^6$, la **Figura 4.16** muestra una concordancia estrecha entre ambas curvas para los tres valores considerados. Los resultados se estabilizan junto a la cota clásica para $\alpha = 0,00$, permanecen bajo ella para $\alpha = 0,25$ y se aproximan a I_{w-w}^{Fritz} para $\alpha = 0,50$. En consecuencia, entre los

valores de N_{sh} estudiados, se requiere $N_{\text{sh}} = 10^6$ para que Full Classical Shadows reproduzca consistentemente la evaluación exacta a lo largo de esta familia.

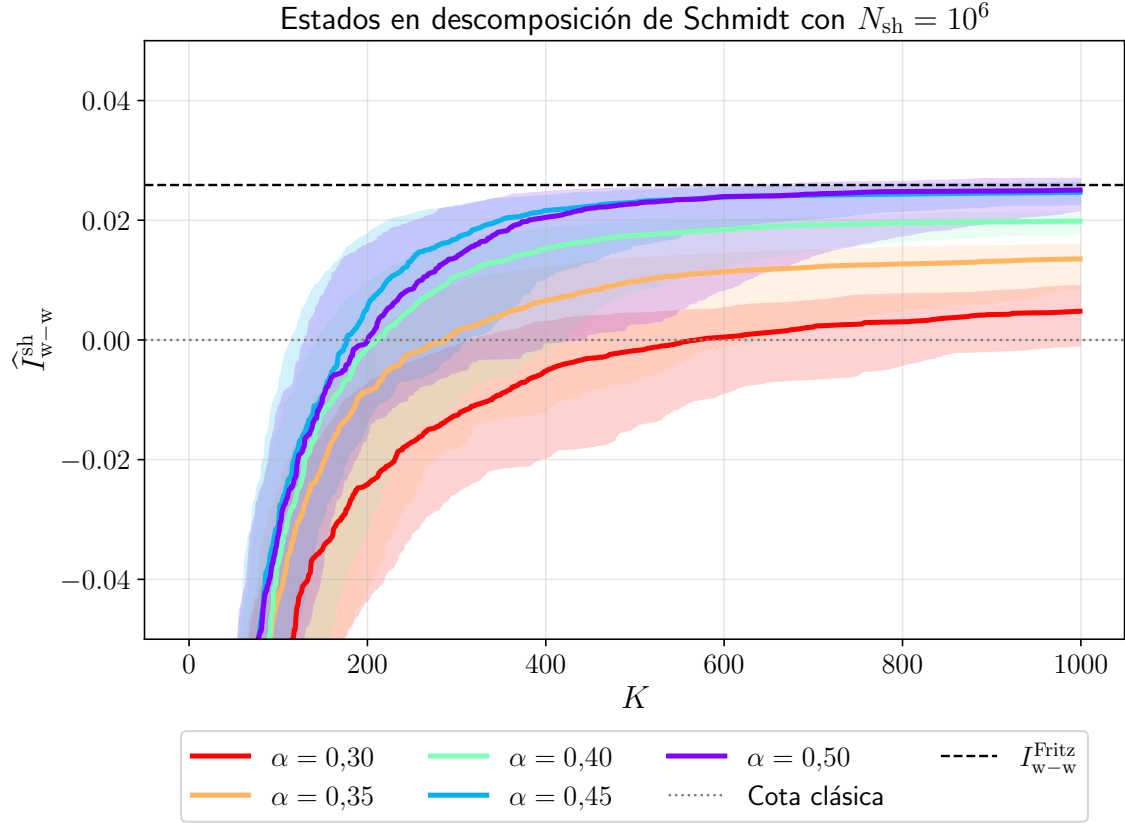


Figura 4.17: Mediana de $\hat{I}_{w-w}^{\text{sh}}$ en función de K , obtenida mediante Full Classical Shadows con $N_{\text{sh}} = 10^6$, para estados con coeficiente de Schmidt común $\alpha \in \{0,30, 0,35, 0,40, 0,45, 0,50\}$. Las regiones sombreadas representan los rangos intercuartiles de las N_{rep} trayectorias. La línea punteada indica la cota clásica y la línea discontinua corresponde a I_{w-w}^{Fritz} . **Fuente:** Elaboración propia.

La **Figura 4.17** muestra que la mediana de $\hat{I}_{w-w}^{\text{sh}}$ supera la cota clásica para todos los valores de α considerados, aunque la consistencia de este resultado varía entre las configuraciones. Para $\alpha \geq 0,35$, el rango intercuartil queda completamente por encima de la cota clásica antes de finalizar las K iteraciones. Este cruce ocurre cerca de $K \approx 250$ para $\alpha = 0,45$ y $\alpha = 0,50$, mientras que para $\alpha = 0,35$ se retrasa hasta aproximadamente $K \approx 600$. El caso $\alpha = 0,40$ presenta un comportamiento intermedio.

Para $\alpha = 0,30$, la mediana supera la cota clásica alrededor de $K \approx 560$, pero el límite inferior del rango intercuartil permanece bajo ella hasta el final de la simulación. Por tanto, esta configuración no satisface el criterio basado en que al menos

el 75 % de las trayectorias produzca valores positivos. En conjunto, los resultados muestran que, dentro del intervalo analizado y para este protocolo, el aumento de α permite superar la cota clásica en menos iteraciones y con mayor consistencia entre las repeticiones. En particular, las curvas correspondientes a los estados más entrelazados se aproximan con mayor rapidez a I_{w-w}^{Fritz} .

Tres estados de Werner

Corresponde ahora estudiar los estados de Werner mediante Full Classical Shadows. El análisis exacto presentado en la [Figura 4.9](#) mostró que la mediana de I_{w-w} aumenta con λ , mientras que la [Figura 4.10](#) situó el umbral numérico de detección entre $\lambda = 0,94$ y $\lambda = 0,95$. Por esta razón, no se repite el barrido completo y el análisis se concentra en valores a partir de $\lambda = 0,92$. Además, se utiliza $N_{\text{sh}} = 10^6$, ya que este número de Classical Shadows permitió obtener resultados consistentes en las configuraciones estudiadas anteriormente.

La [Figura 4.18](#) muestra que, para $\lambda = 0,92$ y $\lambda = 0,94$, la mediana de $\hat{I}_{w-w}^{\text{sh}}$ permanece bajo la cota clásica durante las K iteraciones. En cambio, para $\lambda = 0,96$, $\lambda = 0,98$ y $\lambda = 1,00$, la mediana supera esta cota y el rango intercuartil queda completamente por encima de ella. Este último comportamiento se observa alrededor de $K \approx 750$ para $\lambda = 0,96$, de $K \approx 600$ para $\lambda = 0,98$ y de $K \approx 200$ para $\lambda = 1,00$.

En consecuencia, cuanto mayor es λ , menos iteraciones requiere el protocolo para superar la cota clásica de forma consistente entre las repeticiones. El caso $\lambda = 1,00$ presenta la convergencia más rápida y se aproxima a I_{w-w}^{Fritz} , mientras que el aumento del ruido reduce la violación y retrasa su detección. Dentro de la resolución considerada en esta figura, el cambio de comportamiento de Full Classical Shadows se encuentra entre $\lambda = 0,94$ y $\lambda = 0,96$, en concordancia con la región identificada previamente mediante la evaluación exacta.

Este resultado presenta una proximidad numérica con el valor $\lambda \approx 0,94$ reportado por Sunilkumar, Shaji y Saha [\[44\]](#) para la detección de no localidad genuina de red en estados de Werner. Sin embargo, ambos valores corresponden a criterios diferentes. El resultado citado caracteriza la no localidad genuina de la red triangular, mientras que el análisis realizado aquí examina la violación de la desigualdad Wagon-Wheel mediante una estrategia asociada a la distribución de Fritz. Por tanto, la proximidad entre ambos valores debe interpretarse como una concordancia numérica sugestiva y no como la identificación de un mismo umbral.

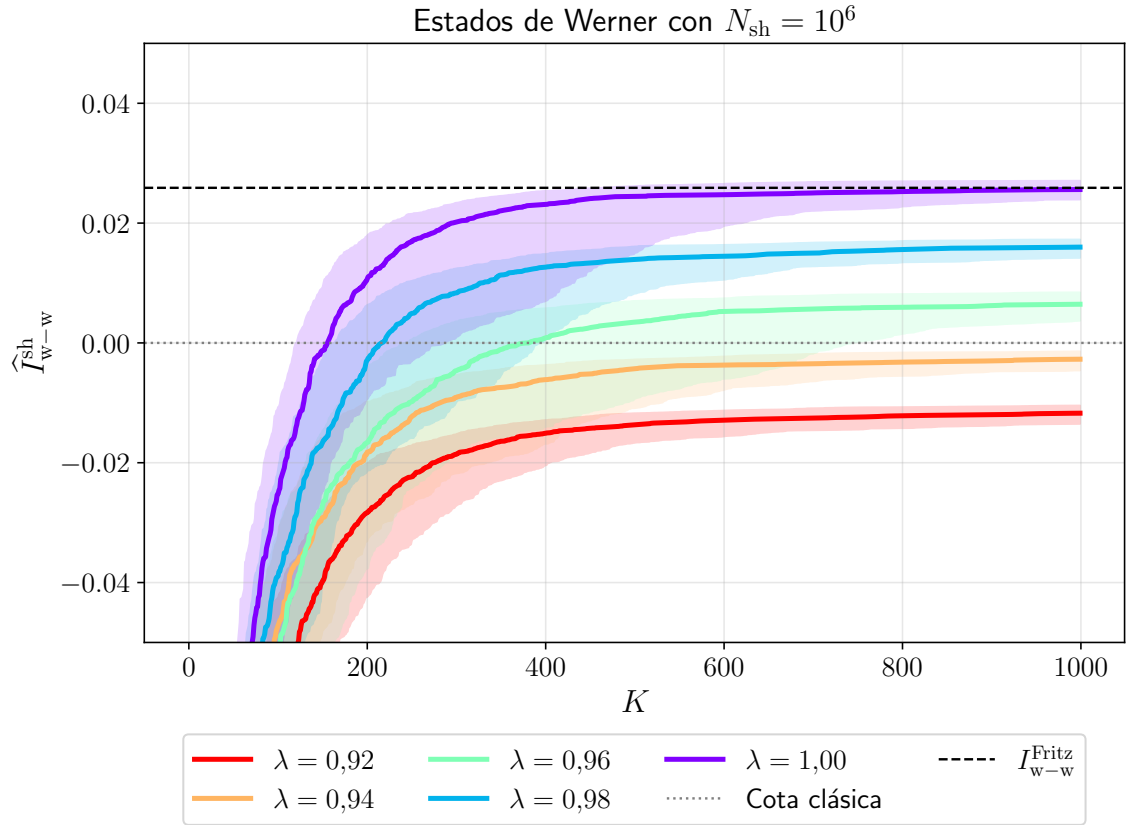


Figura 4.18: Mediana de $\hat{I}_{W-W}^{\text{sh}}$ en función de K , obtenida mediante Full Classical Shadows con $N_{\text{sh}} = 10^6$, para estados de Werner con parámetro común $\lambda \in \{0,92, 0,94, 0,96, 0,98, 1,00\}$. Las regiones sombreadas representan los rangos intercuartiles de las N_{rep} trayectorias. La línea punteada indica la cota clásica y la línea discontinua corresponde a I_{W-W}^{Fritz} . **Fuente:** Elaboración propia.

Capítulo 5

Conclusiones

El problema abordado en esta investigación extiende la lógica operacional de las pruebas de Bell hacia una red compuesta por varias fuentes independientes. En el escenario triangular, la descripción clásica debe respetar la estructura causal de la red y la independencia de las tres fuentes, lo que da lugar a un conjunto de correlaciones no convexo sujeto a restricciones no lineales. Por ello, el problema no consiste únicamente en determinar si los estados distribuidos están entrelazados, sino en encontrar mediciones conjuntas que produzcan una distribución $P_{ABC}(a, b, c)$ incompatible con el modelo clásico triangular. Los resultados del Capítulo 4 muestran que este problema puede abordarse mediante un protocolo autoguiado basado en CSPA. Para cada configuración fija de estados, el algoritmo explora las mediciones parametrizadas por $U_A, U_B, U_C \in U(4)$ y maximiza directamente el funcional Wagon-Wheel, sin imponer una distribución objetivo ni utilizar como punto de partida las mediciones asociadas a la distribución de Fritz. De este modo, se establece un procedimiento constructivo para encontrar mediciones que generen correlaciones incompatibles con la estructura causal clásica de la red.

La configuración formada por tres fuentes que distribuyen estados maximalmente entrelazados proporciona la principal validación del protocolo. Tanto la media como la mediana superan la cota clásica $I_{w-w} = 0$ durante las primeras iteraciones y permanecen en la región de violación. Al finalizar las $K = 1000$ iteraciones, la mediana se encuentra muy próxima al valor de referencia de Fritz, $(\sqrt{2} - 1)/16$. Los valores positivos obtenidos certifican que las mediciones encontradas generan correlaciones incompatibles con el modelo clásico triangular. El análisis de la distancia entre la mediana y el valor de Fritz muestra dos regímenes empíricos de convergencia. El primero se extiende aproximadamente hasta $100 \lesssim K \lesssim 150$ y presenta una pendiente $m_1 \approx -0,94$, mientras que el segundo tiene una pendiente $m_2 \approx -3,23$. Para $K \gtrsim 400$, esta distancia también resulta compatible con un decaimiento aproximadamente exponencial. Las diferencias entre las trayectorias explican la separación

observada entre la media y la mediana. Estos ajustes caracterizan el comportamiento numérico del protocolo, pero no establecen su tasa asintótica de convergencia.

Las familias parametrizadas permiten estudiar cómo cambia la detección frente al entrelazamiento parcial y al ruido. Cuando las tres fuentes distribuyen el mismo estado puro en descomposición de Schmidt, la mediana supera la cota clásica para $\alpha \geq 0,30$ dentro de la malla simulada y permanece negativa para $\alpha = 0,10$ y $\alpha = 0,20$. Por tanto, el cambio de signo se sitúa entre $\alpha = 0,20$ y $\alpha = 0,30$. Además, la violación aparece más tempranamente y alcanza valores finales mayores a medida que α se aproxima a $1/2$. Sin embargo, los valores negativos obtenidos para entrelazamientos pequeños no establecen una medida de este recurso. En particular, el caso $\alpha = 0$ solo refleja la saturación numérica de la cota clásica por un estado producto. Para tres estados de Werner, el barrido refinado sitúa el cambio de signo de la mediana entre $\lambda = 0,94$ y $\lambda = 0,95$, con valores negativos para $\lambda \leq 0,94$ y positivos para $\lambda \geq 0,95$. Puesto que estos estados son entrelazados para $\lambda > 1/3$, los resultados identifican una región extensa en la que el entrelazamiento de las fuentes no produce una violación mediana del funcional. Esto muestra que el entrelazamiento bipartito y la no localidad de red detectada mediante Wagon-Wheel son propiedades diferentes. La segunda depende conjuntamente de los estados, las mediciones y la estructura causal, además de la capacidad de detección del protocolo considerado.

Las configuraciones con una fuente separable presentan un patrón consistente entre las distintas familias estudiadas. Cuando dos fuentes distribuyen singletes y la tercera permanece fija en $|00\rangle$, la media y la mediana se aproximan a la cota clásica desde valores negativos sin superarla. Ninguna de las $N_{\text{rep}} = 100$ trayectorias produce una violación durante las $K = 1000$ iteraciones, lo que comprende un total de 10^5 evaluaciones. El mayor valor encontrado permanece aproximadamente 10^{-5} por debajo de la cota clásica. Tampoco se obtiene una violación mediana cuando dos fuentes distribuyen estados en descomposición de Schmidt y la tercera permanece en $|00\rangle$, incluso para $\alpha = 1/2$, ni cuando dos fuentes distribuyen estados de Werner y la tercera se mantiene en $\mathbb{I}_4/4$, aun para $\lambda = 1$. En conjunto, las violaciones medianas se detectan cuando las tres fuentes distribuyen los estados entrelazados estudiados con suficiente grado de entrelazamiento o resistencia al ruido, mientras que no se observan cuando una de ellas distribuye alguno de los estados separables considerados. Este comportamiento muestra la sensibilidad de las correlaciones de red al conjunto completo de enlaces. Sin embargo, no demuestra que el entrelazamiento de las tres fuentes sea una condición necesaria para toda posible violación ni permite extender el resultado a cualquier estado separable.

El alcance de esta primera fase queda delimitado por el dominio explorado, que comprende la evaluación exacta de las probabilidades, mediciones proyectivas conjuntas parametrizadas mediante elementos de $U(4)$, $N_{\text{rep}} = 100$ inicializaciones independientes, $K = 1000$ iteraciones y conjuntos discretos de valores para α y λ .

Mientras que $I_{w-w} > 0$ certifica la incompatibilidad con el modelo clásico triangular, un resultado no positivo solo indica que el protocolo no obtuvo dicha certificación bajo las condiciones consideradas. Por ello, los intervalos identificados para α y λ representan umbrales empíricos de detección y no fronteras exactas del conjunto de correlaciones cuánticas. Del mismo modo, la proximidad a $(\sqrt{2} - 1)/16$ muestra que el protocolo recupera el valor de referencia de Fritz, pero no establece que este sea el máximo cuántico global del funcional. Sobre esta base ideal, la segunda fase de la investigación incorporó Classical Shadows para estudiar el desempeño del protocolo cuando las probabilidades deben estimarse a partir de un número finito de mediciones.

La incorporación de Classical Shadows permitió reemplazar el acceso exacto a las probabilidades por estimaciones obtenidas a partir de un conjunto fijo de mediciones aleatorias locales de Pauli. Estos datos pueden reutilizarse durante las iteraciones de CSPSA para evaluar el funcional asociado a distintas mediciones conjuntas, sin reconstruir el estado completo ni generar un nuevo conjunto de datos para cada valor de los parámetros. En este esquema, Classical Shadows cumple una función de búsqueda, ya que sus estimaciones orientan al algoritmo hacia mediciones potencialmente adecuadas. La evaluación exacta de esas mismas mediciones, disponible en las simulaciones, permite determinar si la violación estimada corresponde a una propiedad de las correlaciones o al efecto de las fluctuaciones estadísticas. Esta distinción es esencial porque un valor positivo estimado mediante Classical Shadows no constituye por sí mismo una certificación. En una implementación experimental, las mediciones encontradas deben evaluarse posteriormente mediante frecuencias obtenidas de datos independientes. Solo un valor positivo del funcional calculado a partir de esta evaluación final certifica la incompatibilidad con el modelo clásico triangular.

La configuración con tres singletes muestra que el desempeño de Full Classical Shadows depende fuertemente del presupuesto N_{sh} . Para $N_{\text{sh}} = 10^3$, la estimación del funcional alcanza valores cercanos a 0,6, aunque la evaluación exacta de las mismas mediciones permanece alrededor de $-0,24$. Con $N_{\text{sh}} = 10^4$, esta sobreestimación disminuye, pero la evaluación exacta todavía se mantiene próxima a $-0,02$. En cambio, para $N_{\text{sh}} = 10^5$ y 10^6 , el rango intercuartil de la evaluación exacta queda completamente por encima de la cota clásica aproximadamente desde $K \approx 250$ y $K \approx 180$, respectivamente. Además, al aumentar el presupuesto de 10^5 a 10^6 , la distancia final entre la mediana y el valor de Fritz disminuye desde el orden de 10^{-3} hasta 10^{-4} . Estos resultados muestran que la reutilización de un mismo conjunto finito de shadows puede llevar a CSPSA a seleccionar mediciones favorecidas por sus fluctuaciones estadísticas. Por ello, aumentar el número de iteraciones sin disponer de un presupuesto suficiente puede intensificar la discrepancia entre el valor estimado y el exacto. La elección de N_{sh} y K debe realizarse, en consecuencia, de manera conjunta.

Con $N_{\text{sh}} = 10^6$, Full Classical Shadows también permite identificar cómo el entrelazamiento parcial y el ruido afectan la estabilidad de la búsqueda. Para los estados en descomposición de Schmidt, el rango intercuartil queda completamente por encima de la cota clásica para $\alpha \geq 0,35$ dentro de la malla simulada. En $\alpha = 0,30$, la mediana se vuelve positiva alrededor de $K \approx 560$, pero el extremo inferior del rango intercuartil permanece bajo la cota, por lo que no se satisface el criterio de consistencia adoptado. Para los estados de Werner, el rango intercuartil supera completamente la cota para $\lambda = 0,96$, $\lambda = 0,98$ y $\lambda = 1,00$ aproximadamente desde $K \approx 750$, $K \approx 600$ y $K \approx 200$, respectivamente. Esta condición no se alcanza para $\lambda = 0,92$ ni $\lambda = 0,94$ durante las 1000 iteraciones. Estos resultados no contradicen el cambio de signo situado entre $\lambda = 0,94$ y $\lambda = 0,95$ mediante evaluación exacta, ya que ambos análisis utilizan criterios diferentes. El primero exige que el rango intercuartil completo permanezca en la región de violación, mientras que el segundo considera únicamente el signo de la mediana. En consecuencia, la estimación con un número finito de mediciones introduce una exigencia estadística adicional para obtener una detección consistente.

Una primera línea de continuidad consiste en extender el protocolo a otras desigualdades no lineales de redes cuánticas que puedan expresarse mediante observables locales de Pauli. En sistemas de mayor tamaño será importante priorizar desigualdades formadas por términos de pocos cuerpos, ya que el número de shadows requerido aumenta rápidamente con el soporte de los observables. También resulta conveniente reemplazar la optimización sobre unitarias arbitrarias por ansätze físicamente motivados y menos expresivos. Estas modificaciones permitirían reducir el número de parámetros y mejorar la escalabilidad y la entrenabilidad del protocolo.

Los resultados obtenidos abren la posibilidad de extender el protocolo a desigualdades no lineales aplicables a redes cuánticas con un mayor número de nodos. Para mantener su viabilidad será importante priorizar desigualdades expresables mediante observables locales de Pauli y formadas por términos de pocos cuerpos, ya que el número de shadows requerido aumenta rápidamente con el soporte de los observables. Futuros trabajos también podrían sustituir la optimización sobre unitarias arbitrarias por ansätze físicamente motivados y menos expresivos. Estos desarrollos permitirían proyectar el protocolo desde el estudio de la red triangular hacia la certificación de recursos en arquitecturas cuánticas distribuidas más complejas, incluido un futuro internet cuántico y protocolos de comunicación cuántica segura cuyas garantías se sustenten en las correlaciones observadas.

Bibliografía

- [1] Alain Aspect, Jean Dalibard, and Gérard Roger. Experimental Test of Bell's Inequalities Using Time-Varying Analyzers. *Physical Review Letters*, 49(25):1804–1807, 1982.
- [2] Sheldon Axler. *Linear Algebra Done Right*. Undergraduate Texts in Mathematics. Springer, 3rd edition, 2015.
- [3] John S. Bell. On the Einstein Podolsky Rosen Paradox. *Physics Physique Fizika*, 1(3):195–200, 1964.
- [4] Ingemar Bengtsson and Karol Życzkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge University Press, 2006.
- [5] David Bohm. *Quantum Theory*. Prentice-Hall, Englewood Cliffs, NJ, 1951.
- [6] David Bohm. A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I. *Physical Review*, 85(2):166–179, 1952.
- [7] Cyril Branciard, Nicolas Gisin, and Stefano Pironio. Characterizing the Nonlocal Correlations Created via Entanglement Swapping. *Physical Review Letters*, 104(17):170401, 2010.
- [8] Cyril Branciard, Denis Rosset, Nicolas Gisin, and Stefano Pironio. Bilocal versus nonbilocal correlations in entanglement-swapping experiments. *Physical Review A*, 85(3):032119, 2012.
- [9] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Reviews of Modern Physics*, 86(2):419–478, 2014.
- [10] Elisa Bäumer, Victor Gitton, Tamás Kriváňchý, Nicolas Gisin, and Renato Renner. Exploring the Local Landscape in the Triangle Network. *Physical Review A*, 111:052453, 2025.
- [11] M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz

- Cincio, and Patrick J. Coles. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, 2021.
- [12] Rafael Chaves, Christian Majenz, and David Gross. Information-theoretic implications of quantum causal structures. *Nature Communications*, 6(1):5766, 2015.
 - [13] B. S. Cirel’son. Quantum generalizations of Bell’s inequality. *Letters in Mathematical Physics*, 4(2):93–100, 1980.
 - [14] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed Experiment to Test Local Hidden-Variable Theories. *Physical Review Letters*, 23(15):880–884, 1969.
 - [15] Claude Cohen-Tannoudji, Bernard Diu, and Franck Laloë. *Quantum Mechanics, Volume 1*. Wiley-VCH, 1991.
 - [16] D. Concha, L. Pereira, L. Zambrano, and A. Delgado. Training a quantum measurement device to discriminate unknown non-orthogonal quantum states. *Scientific Reports*, 13:7460, 2023.
 - [17] J. Cortés-Vega, J. F. Barra, L. Pereira, and A. Delgado. Detecting entanglement of unknown states by violating the Clauser-Horne-Shimony-Holt inequality. *arXiv preprint*, 2023.
 - [18] José Mário da Silva and Fernando Parisio. Numerically assisted determination of local models in network scenarios. *Physical Review A*, 108:052602, 2023.
 - [19] José Mário da Silva, Alejandro Pozas-Kerstjens, and Fernando Parisio. Local models and Bell inequalities for the minimal triangle network. 2025.
 - [20] Erwan Don, Jessica Bavaresco, Patryk Lipka-Bartosik, Nicolas Gisin, Nicolas Brunner, and Alejandro Pozas-Kerstjens. The minimal example of quantum network Bell nonlocality. 2026.
 - [21] A. Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review*, 47(10):777–780, 1935.
 - [22] Arthur Fine. Hidden variables, joint probability, and the bell inequalities. *Physical Review Letters*, 48(5):291–295, February 1982.
 - [23] Thomas C. Fraser and Elie Wolfe. Causal compatibility inequalities admitting quantum violations in the triangle structure. *Physical Review A*, 98:022113, 2018. Erratum: Phys. Rev. A 102, 029901 (2020).
 - [24] Stuart J. Freedman and John F. Clauser. Experimental Test of Local Hidden-Variable Theories. *Physical Review Letters*, 28(14):938–941, 1972.

- [25] Tobias Fritz. Beyond Bell’s theorem: correlation scenarios. *New Journal of Physics*, 14(10):103001, 2012.
- [26] GianCarlo Ghirardi, Alberto Rimini, and Tullio Weber. A general argument against superluminal transmission through the quantum mechanical measurement process. *Lettere al Nuovo Cimento*, 27(10):293–298, 1980.
- [27] Marissa Giustina et al. Significant-Loophole-Free Test of Bell’s Theorem with Entangled Photons. *Physical Review Letters*, 115(25):250401, 2015.
- [28] Geoffrey Grimmett and David Stirzaker. *Probability and Random Processes*. Oxford University Press, 3rd edition, 2001.
- [29] B. Hensen et al. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526(7575):682–686, 2015.
- [30] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Reviews of Modern Physics*, 81(2):865–942, 2009.
- [31] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16:1050–1057, 2020.
- [32] Tamás Kriváňchý, Yu Cai, Daniel Cavalcanti, Arash Tavakoli, Nicolas Gisin, and Nicolas Brunner. A neural network oracle for quantum nonlocality problems in networks. *npj Quantum Information*, 6:70, 2020.
- [33] A. D. Muñoz Moller, L. Pereira, L. Zambrano, J. Cortés-Vega, and A. Delgado. Variational determination of multielectron geometric entanglement. *Physical Review Applied*, 18:024048, 2022.
- [34] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.
- [35] Jian-Wei Pan, Dik Bouwmeester, Harald Weinfurter, and Anton Zeilinger. Experimental Entanglement Swapping: Entangling Photons That Never Interacted. *Physical Review Letters*, 80(18):3891–3894, 1998.
- [36] Asher Peres. *Quantum Theory: Concepts and Methods*. Kluwer Academic Publishers, 2002.
- [37] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, 2018.
- [38] Herbert Robbins and Sutton Monro. A stochastic approximation method. *The Annals of Mathematical Statistics*, 22(3):400–407, 1951.
- [39] J. J. Sakurai and Jim Napolitano. *Modern Quantum Mechanics*. Cambridge University Press, 3rd edition, 2020.

- [40] Lynden K. Shalm et al. Strong Loophole-Free Test of Local Realism. *Physical Review Letters*, 115(25):250402, 2015.
- [41] Ramamurti Shankar. *Principles of Quantum Mechanics*. Springer, 2nd edition, 1994.
- [42] Abner Shimony. Events and Processes in the Quantum World. In Roger Penrose and Christopher J. Isham, editors, *Quantum Concepts in Space and Time*, pages 182–203. Oxford University Press, 1986.
- [43] James C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, 37(3):332–341, 1992.
- [44] Anantha Krishnan Sunilkumar, Anil Shaji, and Debashis Saha. Quantum-informed learning of genuine network nonlocality beyond idealized resources, 2025.
- [45] Armin Tavakoli, Alejandro Pozas-Kerstjens, Ming-Xing Luo, and Marc-Olivier Renou. Bell nonlocality in networks. *Reports on Progress in Physics*, 85:056001, 2022.
- [46] The Nobel Foundation. The Nobel Prize in Physics 2022, 2022. Accessed: [fecha en que la consultes].
- [47] A. Utreras-Alarcón, M. Rivera-Tapia, S. Niklitschek, and A. Delgado. Stochastic optimization on complex variables and pure-state quantum tomography. *Scientific Reports*, 9:16143, 2019.
- [48] V. Vilasini and Roger Colbeck. Analyzing causal structures using Tsallis entropies. *Physical Review A*, 100(6):062108, 2019.
- [49] Mirjam Weilenmann and Roger Colbeck. Non-Shannon inequalities in the entropy vector approach to causal structures. *Quantum*, 2:57, 2018.
- [50] Reinhard F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Physical Review A*, 40(8):4277–4281, 1989.
- [51] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2nd edition, 2017.
- [52] William K. Wootters. Entanglement of formation of an arbitrary state of two qubits. *Physical Review Letters*, 80(10):2245–2248, 1998.
- [53] L. Zambrano, L. Pereira, S. Niklitschek, and A. Delgado. Estimation of pure quantum states in high dimension at the limit of quantum accuracy through complex optimization and statistical inference. *Scientific Reports*, 10:12781, 2020.

- [54] Marek Żukowski, Anton Zeilinger, Michael A. Horne, and Artur K. Ekert. “Event-ready-detectors” Bell experiment via entanglement swapping. *Physical Review Letters*, 71(26):4287–4290, 1993.