



UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

Estimación de la desigualdad CHSH para estados cuánticos desconocidos vía optimización estocástica

Profesor Guía: **Dr. Aldo 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 Magíster en Ciencias con mención en Física
de la Universidad de Concepción

Código Programa: 4102

por

JEAN FRANCO CORTÉS VEGA
CONCEPCIÓN - CHILE
Abril 2018



Comisión examinadora : Dr. Aldo Delgado Hidalgo

Dr. Carlos Saavedra Rubilar

Dr. Esteban Sepúlveda Gómez

Agradecimientos

En primer lugar, quiero agradecer a mis padres Ricardo y Marisol, a mis hermanos Andrea y Eduardo, a mi cuñado Cristian (el cumpita) y a mis sobrinos por estar siempre presentes y pendientes durante toda esta larga travesía. Su apoyo y cariño han sido fundamentales en todo este proceso. Quiero agradecer especialmente a mi polola Alexandra, quién me ha acompañado durante todos mis estudios, gracias a su amor y comprensión he podido lograr muchas de las metas que me he propuesto. Agradezco también todos los momentos inolvidables que he vivido a su lado, todas las risas, conversaciones, películas, series y siestas que hemos compartido. Por su puesto que le agradezco también a mi hijo no-humano Chappie, por siempre estar presente, por mover su colita y saludarme cada vez que llegaba agotado a casa. Aunque no pueda leer esto su compañía y cariño fue importante en estos últimos años. A mi amigo Alexis que a pesar de estar viviendo en el extranjero ha estado siempre preocupado por mi bienestar.

También quiero agradecer, en segundo lugar, al profesor Dr. Aldo Delgado y a mi compañero de carrera/magíster Luciano Pereira, pues sin la colaboración, paciencia y preocupación de ambos este trabajo no hubiese sido posible. Estoy muy agradecido que los investigadores Dr. Carlos Saavedra y al Dr. Estéban Sepúlveda hayan querido ser parte de nuestro trabajo conformando la comisión examinadora de mi tesis. Agradezco también al grupo de *exiliados* del sexto piso: Nayda, Natalia, Nicole, Sebastián, Luciano y Gerald por su constante apoyo, simpatía, risas, almuerzos y discusiones sobre la *importancia de lo que no importa*, sin ellos todo el proceso de la tesis hubiese sido una pesadilla. Debo agradecer también al *grupo de investigación*: profesor Aldo, Sebastián, Mario, Marco, Aníbal, Luciano y Leonardo por tan amenas y productivas reuniones, realmente es un agrado poder compartir con ustedes ~~y hablar de películas, libros o política $\approx 1/3$ de cada reunión~~. Finalmente, pero no menos importante, quiero agradecer a Calu por sus palabras de aliento y por aquellas extensas conversaciones. Se agradece el apañe de LEGION al momento de realizar simulaciones. Agradezco también el financiamiento otorgado por el *Núcleo Milenio de Óptica Avanzada* y al *Centro de Óptica y Fotónica* (CEFOP) durante todo el desarrollo de esta tesis.

Tabla de contenido

Agradecimientos	iii
Índice de tablas	vi
Índice de figuras	vii
Resumen	xi
1. Introducción	1
2. Fundamentos matemáticos	4
2.1. Álgebra Lineal	4
2.1.1. Espacios vectoriales	4
2.1.2. Producto interior	6
2.1.3. Espacios de Hilbert	7
2.1.4. Álgebra de operadores	9
2.1.5. Autovalores y autovectores de un operador	14
2.1.6. Descomposición espectral de un operador	15
3. Introducción a la Mecánica Cuántica	17
3.1. Postulados de la Mecánica Cuántica	17
3.2. Qubit: El bit cuántico	20
3.3. Operador densidad	21
3.3.1. Propiedades generales del operador densidad	23
3.3.2. Operador densidad reducido	24
3.4. La descomposición de Schmidt	24
3.5. Estados Isotrópicos y estados de Werner	25
4. Paradoja Einstein-Podolsky-Rosen y Desigualdades de Bell	27
4.1. Paradoja Einstein-Podolsky-Rosen	27
4.2. La versión de Bohm-Aharonov (1957)	34
4.3. Desigualdad de Bell (1965)	35

4.4. Desigualdad Clauser-Horne-Shimony-Holt (CHSH) (1969)	38
4.5. Máxima violación de la desigualdad CHSH para estados cuánticos de dimensión arbitraria	40
5. Algoritmos de optimización	42
5.1. Introducción: algoritmo del gradiente descendente	42
5.2. Simultaneous Perturbation Stochastic Approximation (SPSA)	43
5.2.1. Aplicaciones de SPSA en Información Cuántica: SGQT	46
5.3. Complex Simultaneous Perturbation Stochastic Approximation (CSPSA)	50
5.3.1. Aplicaciones de CSPSA en Información Cuántica: Tomografía	52
6. Desigualdad CHSH vía optimización estocástica	55
6.1. Introducción	55
6.2. Desigualdad CHSH vía SPSA	58
6.3. Desigualdad CHSH vía CSPSA	69
6.4. Estimación de la desigualdad CHSH para estados cuánticos desconocidos	85
6.4.1. Simulaciones SPSA	85
6.4.2. Simulaciones CSPSA y CSPSAN	89
6.5. Simulaciones con ensemble finito	93
6.5.1. Ensemble con $N = 10^2$	96
6.5.2. Ensemble con $N = 10^3$	103
6.5.3. Ensemble con $N = 10^4$	111
6.6. Resumen de resultados	120
7. Conclusión	129
A. Tomografía de estados cuánticos	131
B. Códigos de Matlab	133
B.0.1. Códigos SPSA	133
B.0.2. Código CSPSA	135
Bibliografía	137

Índice de tablas

2.1. Tabla de notaciones usadas en Información Cuántica.	9
3.1. Resumen de propiedades de los estados de Werner	26
5.1. Elección usual de ganancias para el algoritmo de SPSA.	45
5.2. Elección de ganancias para CSPSA	51
6.1. Tabla comparativa con parámetros de ajuste de la Figura 6.61 . . .	126
6.2. Tabla comparativa con parámetros de ajuste de la Figura 6.62 . . .	127



Índice de figuras

3.1. Representación de un qubit en la esfera de Bloch	21
5.1. Diagrama del método del gradiente descendente.	43
5.2. Representación de SGQT sobre la esfera de Bloch.	48
5.3. Gráfico infidelidad vs número de iteraciones para SGQT	49
5.4. Gráfico de infidelidad vs número de iteraciones para tomografía vía CSPA.	54
6.1. Representación de estados generados al azar sobre la esfera de Bloch.	56
6.2. Representación de estados generados uniformemente sobre la esfera de Bloch.	57
6.3. Esquema estados cuánticos desconocidos y algoritmos de optimización.	58
6.4. Trayectorias de SPSA para diferentes ganancias	62
6.5. Simulaciones de SPSA para un estado maximalmente entrelazado para 5×10^3 realizaciones.	63
6.6. Gráfica de los estados de Werner mediante SPSA para las ganancias usuales.	65
6.7. Gráfica de los estados de Werner mediante SPSA para las ganancias modificadas.	66
6.8. Estado en descomposición de Schmidt simulado con SPSA con las ganancias usuales	67
6.9. Estado en descomposición de Schmidt simulado con SPSA con las ganancias modificadas.	68
6.10. Trayectorias de CSPA para las ganancias estándar.	73
6.11. Trayectorias de CSPA para las ganancias asintóticas.	74
6.12. Simulación de CSPAy CSPSAN usando ganancias estandar para un estado maximalmente entrelazado con 5×10^3 realizaciones.	75
6.13. Simulación de CSPAy CSPSAN usando ganancias asintóticas para un estado maximalmente entrelazado con 5×10^3 realizaciones.	76
6.14. Promedio de las simulaciones realizadas conCSPA y CSPSAN aplicada estados de Werner para parámetros asintóticos.	77
6.15. Mediana de las simulaciones realizadas con CSPA y CSPSAN aplicada estados de Werner para parámetros asintóticos.	78

6.16. Promedio de las simulaciones realizadas con CSPA y CPSAN aplicada estados de Werner para parámetros estándar.	79
6.17. Mediana de las simulaciones realizadas con CSPA y CPSAN aplicada estados de Werner para parámetros estándar.	80
6.18. Promedio de las simulaciones realizadas con CSPA y CPSAN aplicada a estados descritos en descomposición de Schmidt para ganancias asintóticas.	81
6.19. Mediana de las simulaciones realizadas con CSPA y CPSAN aplicada a estados descritos en descomposición de Schmidt para ganancias asintóticas.	82
6.20. Promedio de las simulaciones realizadas con CSPA y CPSAN aplicada a estados descritos en descomposición de Schmidt para ganancias estándar.	83
6.21. Mediana de las simulaciones realizadas con CSPA y CPSAN aplicada a estados descritos en descomposición de Schmidt para ganancias estándar.	84
6.22. Simulaciones de SPSA para las ganancias asintóticas sobre 10^3 estados generados uniformemente.	86
6.23. Error relativo de SPSA para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente.	87
6.24. Simulaciones de SPSA para las ganancias estándar sobre 10^3 estados puros generados uniformemente.	88
6.25. Error relativo de SPSA para las ganancias estándar sobre 10^3 estados puros generados uniformemente.	89
6.26. Simulaciones de CSPA y CPSAN para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente.	90
6.27. Error relativo de CSPA y CPSAN para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente.	91
6.28. Simulaciones de CSPA y CPSAN para las ganancias estándar sobre 10^3 estados generados uniformemente.	92
6.29. Error relativo de CSPA y CPSAN para las ganancias estándar sobre 10^3 estados puros generados uniformemente.	93
6.30. Diagrama de las zonas para emular detecciones experimentales.	95
6.31. Simulaciones de SPSA para las ganancias asintóticas sobre 10^3 estados generados uniformemente con un tamaño de muestra $N = 10^2$	96
6.32. Error relativo de SPSA para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^2$	97
6.33. Simulaciones de SPSA para las ganancias estándar sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^2$	98
6.34. Error relativo de SPSA para las ganancias estándar sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^2$	99

6.35. Simulaciones de CSPA y CPSAN para las ganancias asintóticas sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^2$.	100
6.36. Error relativo de CSPA y CPSAN para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^2$.	101
6.37. Simulaciones de CSPA y CPSAN para las ganancias estándar sobre 10^3 estados generados uniformemente y $N = 10^2$ experimentos.	102
6.38. Error relativo de CSPA y CPSAN para las ganancias estándar sobre 10^3 estados puros generados uniformemente y $N = 10^2$ experimentos.	103
6.39. Simulaciones de SPSA para las ganancias asintóticas sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^3$.	104
6.40. Error relativo de SPSA para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^3$.	105
6.41. Simulaciones de SPSA para las ganancias estándar sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^3$.	106
6.42. Error relativo de SPSA para las ganancias estándar sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^3$.	107
6.43. Simulaciones de CSPA y CPSAN para las ganancias asintóticas sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^3$.	108
6.44. Error relativo de CSPA y CPSAN para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^3$.	109
6.45. Simulaciones de CSPA y CPSAN para las ganancias estándar sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^3$.	110
6.46. Error relativo de CSPA y CPSAN para las ganancias estándar sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^3$.	111
6.47. Simulaciones de SPSA para las ganancias asintóticas sobre 10^3 estados generados uniformemente y $N = 10^3$ experimentos.	112
6.48. Error relativo de SPSA para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^4$.	113
6.49. Simulaciones de SPSA para las ganancias estándar sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^4$.	114
6.50. Error relativo de SPSA para las ganancias estándar sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^4$.	115
6.51. Simulaciones de CSPA y CPSAN para las ganancias asintóticas sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^4$.	116
6.52. Error relativo de CSPA y CPSAN para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^4$.	117
6.53. Simulaciones de CSPA y CPSAN para las ganancias estándar sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^4$.	118

6.54. Error relativo de CSPA y CSPAN para las ganancias estándar sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^4$.	119
6.55. Rapidez de convergencia para un ensemble de $N = 10^2$.	120
6.56. Error relativo para un ensemble de $N = 10^2$.	121
6.57. Rapidez de convergencia para un ensemble de $N = 10^3$.	122
6.58. Error relativo para un ensemble de $N = 10^3$.	123
6.59. Rapidez de convergencia para un ensemble de $N = 10^4$.	124
6.60. Error relativo para un ensemble de $N = 10^4$.	125
6.61. Gráfico error relativo vs ensemble N	126
6.62. Gráfico error relativo vs ensemble N	127



Resumen

Las desigualdades de Bell nos proporcionan un esquema de trabajo en el cual podemos clasificar la naturaleza de diferentes estados cuánticos. Si se le aplica un test de Bell a un estado cuántico conocido y éste viola la desigualdad, podemos afirmar que dicho estado es entrelazado que no puede ser descrito por teorías clásicas. De esta forma podemos agrupar fácilmente a los estados cuánticos de acuerdo a su naturaleza. El ingrediente principal para evaluar las desigualdades de Bell es *conocer el estado de antemano*. Se deben encontrar las bases exactas donde las correlaciones de las mediciones exhiban una violación en las desigualdades de Bell. Si el estado es conocido entonces se puede abordar el problema mediante simulaciones numéricas. En particular se puede emplear algoritmos de programación semidefinida para optimizar sobre las bases de medición que permitan encontrar la máxima violación de Bell para el estado. Si el estado es desconocido este método no puede ser utilizado.

En el siguiente trabajo nosotros abordaremos el problema de la máxima violación de Bell para estados cuánticos desconocidos acudiendo a algoritmos iterativos de optimización llamados *Simultaneous Perturbation Stochastic Approximation* (SPSA) y *Complex Simultaneous Perturbation Stochastic Approximation* (CSPSA) que permiten encontrar la máxima violación de Bell aunque el estado sea completamente desconocido. Para mostrar su funcionamiento se ha probado inicialmente su desempeño sobre estados cuyas propiedades son ampliamente conocidas, logrando reproducir las máximas violaciones permitidas que predice la teoría. Por otra parte se ha implementado el algoritmo para estados cuánticos desconocidos generados uniformemente para una gran cantidad de simulaciones, consiguiendo que en mediana los algoritmos alcancen en pocas iteraciones las violaciones de Bell y errores relativos del orden de 10^{-6} . Por otra parte se ha implementado una situación hipotética de laboratorio donde se utilizó un tamaño de ensemble finito para estimar los valores esperados de los observables involucrados en la desigualdad CHSH, donde CSPSA en su versión normalizada obtiene un mejor desempeño en rapidez y precisión que SPSA.

Capítulo 1

Introducción

Desde su publicación en 1964 [2], las desigualdades de Bell han jugado un rol importante en el desarrollo de la Información Cuántica. Estas han permitido demostrar que la naturaleza exhibe un comportamiento *no-local* relacionado con la propiedad de *entrelazamiento cuántico*. En particular, existen estados cuánticos entrelazados cuyas predicciones no pueden ser empleadas por una teoría clásica. El concepto de *no-local* hace alusión a una propiedad que permite mantener un misterioso enlace entre dos sistemas que están alejados entre si lo suficiente como para garantizar que no exista una conexión causal entre ellos. Gracias al entrelazamiento cuántico, ha sido posible el desarrollo de la *computación cuántica* y sus mejoras [15] o de procesos exclusivamente cuánticos como la *criptografía cuántica* [12] o la *teleportación cuántica de estados* [3, 5]. Debido a la gran cantidad de aplicaciones en los que se pueden emplear los estados entrelazados, surge la necesidad de poder distinguirlos de estados que posean características locales. Así, las desigualdades de Bell nos proveen un marco de trabajo para clasificar a los estados cuánticos de acuerdo a su naturaleza. Así, si un estado viola la desigualdad de Bell podemos afirmar que el estado es entrelazado¹.

En la práctica el método usual para implementar un *test de Bell* es, dado un estado cuántico entrelazado de prueba, fijo y conocido se debe realizar una búsqueda de las bases de los observables donde las correlaciones de las detecciones exhiban una violación de las desigualdades de Bell. Es más, es posible encontrar un conjunto de bases donde se viole *maximalmente* la desigualdad de Bell, es decir, el valor máximo posible de la violación de Bell para el estado de prueba. Por ejemplo, si consideramos un estado maximalmente entrelazado entonces la máxima violación de la desigualdad

¹Debemos tener presente que no todo estado entrelazado viola las desigualdades de Bell. Existen estados, como por ejemplo los de Werner [24], que a pesar de ser entrelazados, no violan las desigualdades de Bell. De esta forma que un estado sea entrelazado no garantiza necesariamente una violación en las desigualdades de Bell.

CHSH para dicho estado es $2\sqrt{2}$, la cual es conocida como la cota de Cirel'son [8]. Naturalmente si el estado entrelazado de prueba se sustituye por otro, entonces se debe realizar otra búsqueda independiente donde las correlaciones muestren una violación de Bell. De esta forma es posible interpretar este procedimiento como un problema de optimización sobre las bases que caracterizan a los observables que permitan obtener una máxima violación de Bell. Para estos casos donde el estado cuántico de prueba es fijo y conocido, existen algoritmos que realizan la optimización mediante *programación semidefinida* [1, 23].

Nosotros hemos planteado un escenario complementario al descrito anteriormente, es decir, poder verificar si un estado cuántico desconocido puede violar la desigualdad Clauser-Horne-Shimony-Holt (CHSH). Debido a que no se conoce de antemano las propiedades del estado cuántico de prueba, no es posible utilizar los métodos basados en programación semidefinida. Hemos abordado este problema mediante el uso de dos algoritmos iterativos de optimización llamados *Simultaneous Perturbation Stochastic Approximation* (SPSA) [20] y *Complex Simultaneous Perturbation Stochastic Approximation* (CSPSA) [21, 22], los cuales han sido implementados en Información Cuántica para resolver el problema de la *tomografía de estados cuánticos* [13, 22]. Mediante estos métodos resulta posible encontrar las bases de medición sin tener la necesidad de conocer el estado cuántico con anterioridad.

En este trabajo mostraremos el detalle de la implementación y como ambos algoritmos pueden ser utilizados para la resolución del problema de la máxima violación de Bell para estados cuánticos desconocidos. Para el punto de partida de nuestra investigación se estudio la violación de Bell para estados cuánticos de prueba cuyas propiedades son ampliamente conocidas. Se consideró una base de Bell, los estados de Werner y estados en descomposición de Schmidt. Para estos casos los algoritmos mostraron tener la capacidad de poder reproducir las violaciones de Bell predichas por la teoría. El siguiente paso fue adaptar los algoritmos SPSA y CSPSA para estudiar las violaciones de Bell de estados cuánticos desconocidos generados uniformemente para un método exacto, donde se considera un experimento numérico donde la función objetivo (CHSH) es evaluada sin errores, y para un método aproximado, que utiliza un ensemble de tamaño finito para emular el proceso de detección. De acuerdo a los resultados obtenidos para el método exacto, el rendimiento de los algoritmos de SPSA y CSPSA son similares entre si, mientras que para el método aproximado, CSPSA en su versión normalizada posee un mayor rendimiento, tanto en rapidez de convergencia como un menor error relativo si es comparado con SPSA.

La estructura de esta tesis consta de 8 capítulos en total. En el **Capítulo 2** se describirán los fundamentos matemáticos y notaciones que se utilizarán a lo largo de toda la tesis. En el **Capítulo 3** se introducirán los conceptos fundamentales de la Mecánica Cuántica y se dará una descripción detallada de los Postulados que componen los cimientos de la teoría. En el **Capítulo 4** se revisarán exhaustivamente

de los conceptos que permitieron el desarrollo de las desigualdades de Bell original y la desigualdad CHSH. En el **Capítulo 5** se realizará una descripción detallada de los algoritmos de optimización utilizados. En el **Capítulo 6** se presentarán las implementaciones de los algoritmos al problema de las desigualdad de Bell para estados cuánticos desconocidos, además se mostrarán los resultados finales de nuestra investigación. Finalmente en el **Capítulo 7** se presentarán las conclusiones de nuestro trabajo.



Capítulo 2

Fundamentos matemáticos

La Mecánica Cuántica es una de las últimas grandes ramas de la física la cual fue desarrollada a principio del siglo XX. Esta teoría permite poder describir el átomo y su núcleo, también permite describir fenómenos que para la teoría clásica resultaría imposible de realizar (entrelazamiento, radiación de cuerpo negro) y repercutió en otras áreas como la teoría de la información, criptografía, entre otros. La formulación matemática más estructurada fue llevada a cabo por P. M. Dirac y J. von Neumann, los cuales utilizaron el álgebra lineal como marco de trabajo.

En el presente capítulo realizaremos una breve revisión de las propiedades matemáticas más usadas durante esta tesis. Comenzaremos por definir los vectores y sus propiedades, operadores lineales, las propiedades de espacios de Hilbert y la notación que adoptaremos a lo largo de todo nuestro trabajo.

2.1. Álgebra Lineal

2.1.1. Espacios vectoriales

Un **espacio vectorial** sobre un cuerpo \mathbb{K} es un conjunto no vacío V , definido por dos operaciones llamadas suma (operación interna) y producto por escalar (operación externa), las cuales están definidas, respectivamente, como:

$$+ : V \times V \longrightarrow V, \tag{2.1}$$

$$\bullet : \mathbb{K} \times V \longrightarrow V. \tag{2.2}$$

Sean $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V$ y $\alpha, \beta \in \mathbb{K}$. Un espacio vectorial con las operaciones definidas en (2.1) y (2.2) debe satisfacer las siguientes propiedades¹:

- (i).- $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$. [Commutatividad de la suma]
- (ii).- $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$. [Asociatividad de la suma]
- (iii).- $\mathbf{0} + \mathbf{a} = \mathbf{a}$. [Existencia del elemento neutro para la suma]
- (iv).- $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$ [Inverso aditivo]
- (v).- $\alpha \cdot (\mathbf{a} + \mathbf{b}) = \alpha \cdot \mathbf{a} + \alpha \cdot \mathbf{b}$ [Dist. del producto c/r a la suma de vectores]
- (vi).- $(\alpha + \beta) \cdot \mathbf{a} = \alpha \cdot \mathbf{a} + \beta \cdot \mathbf{a}$ [Dist. del producto c/r a la suma de escalares]
- (vii).- $(\alpha\beta) \cdot \mathbf{a} = \alpha \cdot (\beta \cdot \mathbf{a})$. [Asociatividad del producto por escalar]
- (viii).- $1 \cdot \mathbf{a} = \mathbf{a}$. [Existencia del elemento neutro multiplicativo]

Los elementos de V se conocen como **vectores**. De acuerdo a la notación precedente los vectores son descritos en negrita.

Una **combinación lineal** de vectores $\mathbf{a}, \mathbf{b}, \dots$ es una expresión de la forma

$$\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} + \dots \quad (2.3)$$

Un vector \mathbf{v} se dice **linealmente independiente** de $\mathbf{a}, \mathbf{b}, \dots$ si no puede ser escrito como combinación lineal de ellos. Un conjunto de vectores son **generadores** de V si *cualquier* vector del espacio puede ser escrito como una combinación lineal de los generadores². El conjunto de vectores linealmente independientes que genera el espacio se llaman **base** del espacio V . La **dimensión** del espacio es descrita por la cantidad de elementos que contiene la base. Por último, cualquier vector en V puede ser expandido como una combinación lineal de elementos de una base, es decir,

$$\mathbf{a} = \sum_{i=1}^n \alpha_i \mathbf{e}_i \quad (2.4)$$

donde $\alpha_i \in \mathbb{K}$, $n = \text{Dim}(V)$ y $\{\mathbf{e}_i\}$ es una base ortonormal de V .

¹Aquí $\mathbf{0}$ representa el vector nulo o $\vec{0}$. Dist.: Distributividad. c/r: con respecto.

²El ejemplo más típico es la base del espacio tridimensional euclídeo $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$.

2.1.2. Producto interior

Definición 2.1.1. El **producto escalar**, **producto punto** o **producto interior** es una aplicación lineal, tal que³

$$\begin{aligned} (\cdot|\cdot) : V \times V &\longrightarrow \mathbb{K}, \\ (\mathbf{a}, \mathbf{b}) &\longmapsto \alpha = (\mathbf{a}|\mathbf{b}) \end{aligned} \quad (2.5)$$

donde $\mathbf{a}, \mathbf{b} \in V$ y $\alpha \in \mathbb{K}$, el cual debe satisfacer las siguientes condiciones:

1. Linealidad por la izquierda: $(\alpha\mathbf{a} + \mathbf{b}|\mathbf{c}) = \alpha(\mathbf{a}|\mathbf{c}) + (\mathbf{b}|\mathbf{c})$. $\alpha \in \mathbb{K}, \mathbf{a}, \mathbf{b}, \mathbf{c} \in V$
2. Linealidad por la derecha: $(\mathbf{a}|\alpha\mathbf{b} + \mathbf{c}) = \bar{\alpha}(\mathbf{a}|\mathbf{b}) + (\mathbf{a}|\mathbf{c})$. $\alpha \in \mathbb{K}, \mathbf{a}, \mathbf{b}, \mathbf{c} \in V$
3. Hermiticidad: $(\mathbf{a}|\mathbf{b}) = \overline{(\mathbf{b}|\mathbf{a})}$. $\mathbf{a}, \mathbf{b} \in V$
4. Positividad: $\mathbf{a} \in V, \mathbf{a} \neq 0 \implies (\mathbf{a}|\mathbf{a}) > 0$.

La barra ($\bar{}$) representa el complejo conjugado de un número complejo, esto es, si $\alpha = \beta + i\gamma \in \mathbb{C}$, con $\beta, \gamma \in \mathbb{R}$, entonces $\bar{\alpha} = \beta - i\gamma \in \mathbb{C}$.

Definición 2.1.2. Si V es un espacio vectorial y $\mathbf{a} \in V$, se llama **norma** o, más apropiadamente, **norma cuadrática** de \mathbf{a} a la operación

$$\|\mathbf{a}\| := \sqrt{(\mathbf{a}|\mathbf{a})}, \quad (2.6)$$

con $\|\mathbf{a}\| \in \mathbb{R}$. La norma generaliza la noción de longitud en los espacios vectoriales. Notemos que todo producto interior induce una norma sobre el espacio en el que está definido y sigue la regla de la eq. (2.6).

A partir de la definición 2.1.2 se enuncian las siguientes propiedades:

1. **Ortogonalidad de vectores:** Dos vectores \mathbf{a} y \mathbf{b} son ortogonales si y solo si

$$(\mathbf{a}|\mathbf{b}) = 0. \quad (2.7)$$

2. **Sistema ortonormal:** El conjunto de vectores $\{\mathbf{a}_n\}$ normalizados a 1 es un sistema ortonormal si y solo si

$$(\mathbf{a}_n|\mathbf{a}_m) = \delta_{nm}, \quad (2.8)$$

dónde δ_{nm} es la **delta de Kronecker**, definida de acuerdo a

$$\delta_{nm} = \begin{cases} 1 & n = m, \\ 0 & n \neq m. \end{cases} \quad (2.9)$$

³Notaciones alternativas: $\langle \mathbf{a}, \mathbf{b} \rangle, \mathbf{a} \cdot \mathbf{b}$.

3. **Sistema ortonormal completo:** El conjunto de vectores base $\{e_n\}$ forma un sistema ortonormal completo en V si un vector cualquiera \mathbf{a} de V puede expandido como

$$\mathbf{a} = \sum_n \alpha_n e_n, \quad (2.10)$$

con $\alpha_n \in \mathbb{K}$ y $\alpha_n = (e_n | \mathbf{a})$, por lo que

$$\mathbf{a} = \sum_n (e_n | \mathbf{a}) e_n. \quad (2.11)$$

2.1.3. Espacios de Hilbert

Definición 2.1.3. Un **espacio pre-hilbertiano** es un espacio vectorial \mathcal{H} dotado de un producto escalar definido de acuerdo a la eq. (2.5)

$$(\cdot | \cdot) : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{K}, \quad (2.12)$$

Un espacio pre-hilbertiano con $\mathbf{a}, \mathbf{b} \in \mathcal{H}$ y $\alpha \in \mathbb{K}$ debe satisfacer las siguientes propiedades:

1. $\|\mathbf{a}\| = 0$ si y solo si $\mathbf{a} = \mathbf{0}$,
2. $\|\alpha \mathbf{a}\| = |\alpha| \|\mathbf{a}\|$,
3. $|(a|b)| \leq \|a\| \|b\|$, [Desigualdad de Cauchy-Schwarz]
4. $\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|$, [Desigualdad triangular]
5. $|\|\mathbf{a}\| - \|\mathbf{b}\|| \leq \|\mathbf{a} - \mathbf{b}\|$, [Desigualdad triangular inversa]
6. $\|\mathbf{a} + \mathbf{b}\|^2 + \|\mathbf{a} - \mathbf{b}\|^2 = 2(\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2)$. [Identidad del paralelogramo]

Propiedades 2.1.1. Las propiedades enunciadas a continuación son fundamentales en la norma y el producto escalar.

1. El producto escalar es una aplicación continua $\mathcal{H} \times \mathcal{H}$ en \mathbb{C} .
2. La norma es una aplicación continua en \mathcal{H} en \mathbb{C} .

Ejemplo 2.1.1. Los siguientes espacios son pre-hilbertianos con los productos escalares definidos en cada caso:

1. $\mathcal{H} = \mathbb{R}^n$ con el producto interno definido como

$$(\mathbf{x} | \mathbf{y}) := \sum_{i=1}^n x_i y_i \quad (2.13)$$

con $\mathbf{x} := (x_1, \dots, x_n)$ y $\mathbf{y} := (y_1, \dots, y_n)$ y $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

2. $\mathcal{H} = \mathbb{C}^n$ con el producto interno definido como

$$(\mathbf{v}|\mathbf{w}) := \sum_{i=1}^n v_i^* w_i \quad (2.14)$$

con $\mathbf{v} := (v_1, \dots, v_n)$ y $\mathbf{w} := (w_1, \dots, w_n)$ y $\mathbf{v}, \mathbf{w} \in \mathbb{C}^n$.

3. $\mathcal{H} = C([a, b])$, el espacio vectorial de las funciones reales continuas en un intervalo cerrado $[a, b] \subset \mathbb{R}$, con

$$(f|g) := \int_a^b f(x)g(x)dx \quad (2.15)$$

4. $\mathcal{H} = l_2$, el espacio de las sucesiones reales $x = (x_i)$ de cuadrado sumable, es decir, tales que $\sum_{i=1}^{\infty} x_i^2 < \infty$, con

$$(x|y) = \sum_{i=1}^{\infty} x_i y_i \quad (2.16)$$

con $x, y \in l_2$.

5. $\mathcal{H} = L_2([a, b])$, el espacio de las funciones medibles Lebesgue $f : [a, b] \rightarrow \mathbb{R}$ de cuadrado integrable, es decir, tales que $\int_a^b f^2 < \infty$, donde la integral se entiende en el sentido de Lebesgue, con

$$(f|g) := \int_a^b f(x)g(x)dx \quad (2.17)$$

con $f, g \in L_2([a, b])$.

6. Todo subespacio vectorial de un espacio pre-hilbertiano es también un espacio vectorial pre-hilbertiano.

Definición 2.1.4. Un **espacio de Hilbert** es un espacio pre-hilbertiano que es completo para la norma cuadrática. Es decir, dado un elemento $\mathbf{a} \in \mathcal{H}$, este puede aproximarse por una serie $\{\mathbf{a}_n\}_{n \in \mathbb{N}}$, tal que

$$\lim_{n \rightarrow \infty} \|\mathbf{a} - \mathbf{a}_n\| = 0. \quad (2.18)$$

En resumen, todo espacio pre-hilbertiano es un espacio normado, con la norma asociada a su producto escalar, definida en la eq. (2.6). Si dicha norma es completa,

Notación	Descripción
z^*	Complejo conjugado de z . $(1 + i)^* = 1 - i$
$ \psi\rangle$	Vector o <i>ket</i> .
$\langle\psi $	Vector dual de $ \psi\rangle$ o <i>bra</i> .
$\langle\varphi \psi\rangle$	Producto punto entre $ \psi\rangle$ y $ \varphi\rangle$.
$ \varphi\rangle \otimes \psi\rangle$	Producto tensorial entre $ \psi\rangle$ y $ \varphi\rangle$.
$ \varphi\rangle \psi\rangle$ o $ \varphi\psi\rangle$	Abreviación del producto tensorial.
\mathbf{A}^*	Complejo conjugado de la matriz \mathbf{A} .
\mathbf{A}^T	Transpuesta de la matriz \mathbf{A} .
\mathbf{A}^\dagger	Hermítica conjugada o adjunta de la matriz \mathbf{A} , $\mathbf{A}^\dagger := (\mathbf{A}^T)^*$.
	$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^\dagger = \begin{bmatrix} a^* & c^* \\ b^* & d^* \end{bmatrix}$.
$\langle\varphi \mathbf{A} \psi\rangle$	Producto punto entre $ \varphi\rangle$ y $\mathbf{A} \psi\rangle$, o Producto punto entre $\mathbf{A}^\dagger \varphi\rangle$ y $ \psi\rangle$.

Tabla 2.1: Tabla de notaciones usadas en Información Cuántica. **Fuente:** Quantum information and Quantum Computation [17]

se dice que \mathcal{H} es un **espacio de Hilbert**. Este tipo de espacios es un caso particular de espacios de Banach, en los que la función $\mathbf{a} \mapsto \|\mathbf{a}\|^2$ es una forma cuadrática.

En el contexto de la mecánica cuántica, los espacios de Hilbert \mathcal{H} están definidos sobre el cuerpo de los números complejos \mathbb{C} . Los vectores y operaciones sobre ellos son descritas de acuerdo a la **notación de Dirac** [10]. En la table 2.1 se muestra un resumen de la notación estándar en mecánica cuántica y que será utilizada en los siguientes capítulos.

2.1.4. Álgebra de operadores

Espacio vectorial dual

Antes de comenzar a describir las propiedades fundamentales del álgebra de operadores, es necesario recordar el concepto de **espacio dual** de un espacio de Hilbert \mathcal{H} . Para cada vector $|\psi\rangle \in \mathcal{H}$ se puede asociar un **vector dual** descrito como $\{|\psi\rangle\}^\dagger := \langle\psi|$. De esta forma podemos definir el **espacio dual** \mathcal{H}^\dagger conformado por todos los vectores duales obtenidos a partir de los elementos de \mathcal{H} . Así, tenemos la siguiente asociación

$$|\psi\rangle = a|\alpha\rangle + b|\beta\rangle \longrightarrow \langle\psi| = \{a|\alpha\rangle + b|\beta\rangle\}^\dagger = a^*\langle\alpha| + \beta^*\langle\beta|. \quad (2.19)$$

Definimos, ahora, la aplicación producto punto entre los elementos de \mathcal{H}^\dagger y \mathcal{H} como

$$\begin{aligned} \langle \cdot | \cdot \rangle : \mathcal{H}^\dagger \times \mathcal{H} &\longrightarrow \mathbb{C} \\ \langle \phi | \psi \rangle &\longrightarrow \alpha \in \mathbb{C} \end{aligned} \quad (2.20)$$

dónde $\langle \phi | \in \mathcal{H}^\dagger$ y $|\psi\rangle \in \mathcal{H}$. Este producto punto debe satisfacer las mismas condiciones presentadas en la definición 2.1.2.

Observación: En general, el vector $|\psi\rangle$ puede ser representado como una matriz columna de $n \times 1$ componentes y un vector dual $\langle \phi |$ como una matriz fila de $n \times 1$ componentes, es decir

$$|\psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix}, \quad \langle \phi | = (\phi_1^* \quad \phi_2^* \quad \cdots \quad \phi_n^*). \quad (2.21)$$

A partir de eq. (2.21) podemos ver que el producto punto no es más que la multiplicación de una matriz fila $1 \times n$ por una matriz columna $n \times 1$,

$$\langle \phi | \psi \rangle = (\phi_1^* \quad \phi_2^* \quad \cdots \quad \phi_n^*) \cdot \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix} = \sum_{i=1}^n \psi_i \phi_i^* \quad (2.22)$$

cuya expresión es idéntica a la mostrada en la eq. (2.14).

Álgebra de operadores

Ahora que comprendemos la idea principal de los espacios duales, podemos comenzar con la descripción del álgebra de operadores, los cuales juegan un rol fundamental en la mecánica cuántica.

Un **operador lineal** es una transformación que toma un vector del espacio y lo convierte en otro $|\alpha\rangle \longrightarrow |\alpha'\rangle = \mathbf{A}|\alpha\rangle$. Es lineal debido puesto que preserva la linealidad en las combinaciones lineales de vectores, es decir,

$$\mathbf{A}(a|\alpha\rangle + b|\beta\rangle) = a\mathbf{A}|\alpha\rangle + b\mathbf{A}|\beta\rangle, \quad (2.23)$$

para cualquier vector $|\alpha\rangle, |\beta\rangle \in \mathcal{H}$ y cualquier escalar $a, b \in \mathbb{C}$.

Si el conjunto $\{|e_i\rangle\}_{i=1}^n$ es una base de \mathcal{H} y \mathbf{A} es una transformación lineal, entonces se tiene que

$$\mathbf{A}|e_j\rangle = \sum_{i=1}^n A_{ij}|e_i\rangle, \quad j = 1, 2, \dots, n. \quad (2.24)$$

Además, dado un vector $|\alpha\rangle = \sum_{j=1}^n a_j |e_j\rangle$,

$$\mathbf{A}|\alpha\rangle = \sum_{i=1}^n a_j [\mathbf{A}|e_j\rangle] = \sum_{j=1}^n \sum_{i=1}^n a_j A_{ij} |e_i\rangle = \sum_{j=1}^n \left[\sum_{i=1}^n A_{ij} a_j \right] |e_i\rangle. \quad (2.25)$$

De esta forma, de la eq. (2.25), podemos ver fácilmente que el operador \mathbf{A} transforma un vector con componentes a_1, a_2, \dots, a_n en otro con componentes dadas por

$$a'_i = \sum_{j=1}^n A_{ij} a_j. \quad (2.26)$$

Podemos ver que los n^2 elementos A_{ij} caracterizan de manera única al operador \mathbf{A} , con respecto a una base dada. Por otro lado, sabemos que, las n componentes a_i caracterizan de manera única al vector $|\alpha\rangle$, con respecto a esa misma base.

Si la base $\{|e_i\rangle\}_{i=1}^n$ es ortonormal, de la eq. (2.24) encontramos una expresión para calcular los **elementos de matriz** de \mathbf{A}

$$A_{ij} = \langle e_i | \mathbf{A} | e_j \rangle, \quad (2.27)$$

o en su forma matricial,

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}. \quad (2.28)$$

A continuación se mostrarán las aplicaciones que se pueden realizar sobre los operadores lineales.

1. **Operador transpuesto:** Es la aplicación de intercambiar filas por columnas.

$$\mathbf{A}^T = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}. \quad (2.29)$$

Una matriz es **simétrica** si $\mathbf{A}^T = \mathbf{A}$. Por otro lado es **antisimétrica** si $\mathbf{A}^T = -\mathbf{A}$.

2. **Operador conjugado:** Es la aplicación de conjugar cada elemento de la matriz.

$$\mathbf{A}^* = \begin{pmatrix} A_{11}^* & A_{12}^* & \cdots & A_{1n}^* \\ A_{21}^* & A_{22}^* & \cdots & A_{2n}^* \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1}^* & A_{n2}^* & \cdots & A_{nn}^* \end{pmatrix}. \quad (2.30)$$

Si todos los elementos de \mathbf{A} son reales, entonces $\mathbf{A}^* = \mathbf{A}$. Si todos los elementos son complejos, entonces $\mathbf{A}^* = -\mathbf{A}$.

3. **Operador adjunto:** Es la aplicación de transponer y conjugar una matriz al mismo tiempo.

$$\mathbf{A}^\dagger = (\mathbf{A}^T)^* = \begin{pmatrix} A_{11}^* & A_{21}^* & \cdots & A_{n1}^* \\ A_{12}^* & A_{22}^* & \cdots & A_{n2}^* \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n}^* & A_{2n}^* & \cdots & A_{nn}^* \end{pmatrix}. \quad (2.31)$$

Una matriz es **hermítica** si $\mathbf{A}^\dagger = \mathbf{A}$, es **antihermítica** si $\mathbf{A}^\dagger = -\mathbf{A}$.

4. La multiplicación de matrices, en general, no es conmutativa ($\mathbf{AB} \neq \mathbf{BA}$). Se define el **conmutador** entre \mathbf{A} y \mathbf{B} de acuerdo a

$$[\mathbf{A}, \mathbf{B}] := \mathbf{AB} - \mathbf{BA}. \quad (2.32)$$

Si dos operadores conmutan, entonces $[\mathbf{A}, \mathbf{B}] = 0$. Si esto ocurre entonces existe una base en \mathcal{H} en que ambos operadores son diagonales. En otras palabras, existe una base de \mathcal{H} cuyos vectores son simultáneamente autovectores de \mathbf{A} y \mathbf{B} .

5. La transpuesta del producto de dos operadores es la transpuesta de cada uno pero con orden invertido,

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T \quad (2.33)$$

6. Análogo al punto 6, el hermítico conjugado de dos operadores es,

$$(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger. \quad (2.34)$$

El operador de proyección

Definición 2.1.5. Si \mathcal{H} es la suma directa de n subespacios

$$\mathcal{H} = \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_n. \quad (2.35)$$

Se denomina **proyector** al operador lineal \mathbf{P}_i que hace corresponder a cada vector $|\alpha\rangle \in \mathcal{H}$ su componente $|\alpha_i\rangle \in \mathcal{H}_i$:

$$\begin{aligned} \mathbf{P}_i : \mathcal{H} &\longrightarrow \mathcal{H}_i \\ |\alpha\rangle &\longrightarrow |\alpha_i\rangle = \mathbf{P}_i |\alpha\rangle \end{aligned} \quad (2.36)$$

Proposición 2.1.1. Si \mathcal{H} es una suma directa de $\{\mathcal{H}_i\}_{i=1}^n$ y \mathbf{P}_i definido en la eq. (2.36). Entonces $\{\mathbf{P}_i\}_{i=1}^n$ es un conjunto completo de proyectores, es decir

$$\mathbf{P}_i \mathbf{P}_j = \delta_{ij} \mathbf{P}_i \quad \text{y} \quad \sum_{i=1}^n \mathbf{P}_i = \mathbf{1}. \quad (2.37)$$

Proposición 2.1.2. Todo operador es de proyección si y solo si es hermítico e idempotente, es decir, $\mathbf{P}^2 = \mathbf{P}$.

Proposición 2.1.3. Sea $\{|e_i\rangle\}_{i=1}^n$ una base ortonormal para \mathcal{H} . Entonces el conjunto $\{\mathbf{P}_i = |e_i\rangle\langle e_i|\}_{i=1}^n$ son proyectores mutuamente ortogonales, y

$$\sum_{i=1}^n \mathbf{P}_i = \sum_{i=1}^n |e_i\rangle\langle e_i| = \mathbf{1}. \quad (2.38)$$

Esta relación es conocida como **relación de completitud**.

El operador identidad

Consideremos la descomposición del estado $|\psi\rangle \in \mathcal{H}$ en la base ortonormal $\{|e_j\rangle_{j=1}^n\}$:

$$|\psi\rangle = \sum_{j=1}^n \psi_j |e_j\rangle, \quad (2.39)$$

tenemos que

$$\langle e_i | \psi \rangle = \sum_{j=1}^n \psi_j \langle e_i | e_j \rangle = \sum_{j=1}^n \psi_j \delta_{ij} = \psi_i. \quad (2.40)$$

Si reemplazamos la eq. (2.40) en la eq. (2.39), obtenemos

$$|\psi\rangle = \sum_{j=1}^n \langle e_j | \psi \rangle |e_j\rangle = \sum_{j=1}^n |e_j\rangle \langle e_j | \psi \rangle \quad (2.41)$$

$$= \left[\sum_{j=1}^n |e_j\rangle \langle e_j| \right] |\psi\rangle = \mathbb{1} |\psi\rangle. \quad (2.42)$$

A partir de la eq. (2.41) podemos identificar una expresión del **operador identidad**,

$$\mathbb{1} := \sum_{j=1}^n |e_j\rangle \langle e_j| \quad (2.43)$$

Operadores unitarios

Un operador es **unitario** si su inversa es igual a su hermítico conjugado

$$U^\dagger = U^{-1} \implies U^\dagger U = U U^\dagger = \mathbb{1} \quad (2.44)$$

En efecto, consideremos dos vectores $|\psi_1\rangle$ y $|\psi_2\rangle$ en \mathcal{H} , y sus transformaciones $|\Psi_1\rangle$ y $|\Psi_2\rangle$ bajo la acción de U :

$$\begin{aligned} |\Psi_1\rangle &= U |\psi_1\rangle \\ |\Psi_2\rangle &= U |\psi_2\rangle \end{aligned} \quad (2.45)$$

Así, tenemos que

$$\langle \Psi_1 | \Psi_2 \rangle = \langle \psi_1 | U^\dagger U | \psi_2 \rangle = \langle \psi_1 | \psi_2 \rangle. \quad (2.46)$$

De esta forma, una transformación unitaria **preserva el producto por escalar** en \mathcal{H} , transformando una base ortonormal en otra que también lo es.

2.1.5. Autovalores y autovectores de un operador

Consideremos un operador \mathbf{A} . Decimos que $|\alpha\rangle$ es un **autovector** del operador \mathbf{A} si satisface la ecuación,

$$\mathbf{A} |\alpha\rangle = \alpha |\alpha\rangle \quad (2.47)$$

donde $\alpha \in \mathbb{C}$ y es conocido como **autovalor** de \mathbf{A} . La eq. (2.47) es conocida como **ecuación de autovalores**. La eq. (2.47) puede ser escrita como (para $|\alpha\rangle \neq |0\rangle$ o vector nulo)

$$(\mathbf{A} - \alpha \mathbf{1}) |\alpha\rangle = 0. \quad (2.48)$$

De la eq. (2.48) se obtiene una expresión para encontrar α ,

$$\det(\mathbf{A} - \alpha \mathbf{1}) = \det \begin{pmatrix} (A_{11} - \alpha) & A_{12} & \cdots & A_{1n} \\ A_{21} & (A_{22} - \alpha) & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & (A_{nn} - \alpha) \end{pmatrix} = 0 \quad (2.49)$$

conocida como **ecuación secular**. A partir de ella obtenemos la **ecuación característica** que es un polinomio de grado n , cuyas soluciones permiten determinar los autovalores α ,

$$c_n \alpha^n + c_{n-1} \alpha^{n-1} + \cdots + c_1 \alpha + c_0 = 0, \quad (2.50)$$

donde c_n son parámetros que dependen de los elementos de \mathbf{A} . De las soluciones de la eq. (2.50) podemos ver que si todos los autovalores son distintos, entonces a cada uno de ellos le corresponde un único autovector. Si hay autovalores que coinciden (autovalor degenerado), puede ocurrir que la cantidad de autovectores sea menor que la dimensión del espacio. A pesar de esto siempre existirá al menos un autovector por cada autovalor.

2.1.6. Descomposición espectral de un operador

Definición 2.1.6. Se denomina operador **normal** al operador lineal que satisface

$$\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger\mathbf{A}. \quad (2.51)$$

Teorema 1. Sea \mathbf{A} un operador normal definido en un espacio de Hilbert de dimensión finita \mathcal{H} y sean $\{\alpha_i\}_{i=1}^n$ sus distintos autovalores, entonces \mathbf{A} puede ser descompuesto como

$$\mathbf{A} = \sum_{i=1}^n \alpha_i \mathbf{P}_i \quad (2.52)$$

dónde \mathbf{P}_i debe satisfacer $\sum_{i=1}^n \mathbf{P}_i = \mathbf{1}$ y $\mathbf{P}_i \mathbf{P}_j = 0$ para $i \neq j$.

Definición 2.1.7. Un operador positivo \mathbf{A} es aquel que $\langle \psi | \mathbf{A} | \psi \rangle \geq 0$ para todo $|\psi\rangle$.

Definición 2.1.8. Un **operador definido positivo** \mathbf{A} es aquel que $\langle \psi | \mathbf{A} | \psi \rangle > 0$ para todo $|\psi\rangle \neq 0$.

Observación: Todos los operadores positivos son hermíticos y tienen autovalores no negativos.

Definición 2.1.9. El **valor esperado** o **valor de expectación** de un operador \mathbf{A} en un estado normalizado $|\psi\rangle$ esta dado por

$$\langle \mathbf{A} \rangle_\psi := \langle \psi | \mathbf{A} | \psi \rangle. \quad (2.53)$$

De esta definición podemos mostrar que el valor esperado de un operador hermítico es real. Sea \mathbf{A} un operador hermítico, $\{|a_i\rangle\}$ una base ortonormal completa de \mathcal{H} y $|\psi\rangle \in \mathcal{H}$. Entonces

$$\langle \mathbf{A} \rangle_\psi = \langle \psi | \mathbf{A} | \psi \rangle = \sum_i \langle \psi | \mathbf{A} | a_i \rangle \langle a_i | \psi \rangle \quad (2.54)$$

$$= \sum_{i=1} \langle \psi | a_i \rangle a_i \langle a_i | \psi \rangle \quad (2.55)$$

$$(2.56)$$

$$\langle \mathbf{A} \rangle_\psi = \sum_i a_i |\langle a_i | \psi \rangle|^2 \in \mathbb{R}. \quad (2.57)$$

La eq. (2.57) se puede interpretar como el promedio ponderado de todos los valores posibles a_i , donde a cada a_i tiene un peso $|\langle a_i | \psi \rangle|^2$.

Definición 2.1.10. La **varianza** $\langle (\Delta \mathbf{A})^2 \rangle_\psi$ de un operador \mathbf{A} para un estado $|\psi\rangle$ está definida como

$$\langle (\Delta \mathbf{A})^2 \rangle_\psi := \langle \mathbf{A}^2 \rangle_\psi - \langle \mathbf{A} \rangle_\psi^2 \quad (2.58)$$

Capítulo 3

Introducción a la Mecánica Cuántica

Desde su fundación, la mecánica cuántica ha mostrado que las leyes de la naturaleza que rigen el mundo subatómico son completamente diferentes a las que observamos en el mundo macroscópico¹. En el presente capítulo mostraremos los *postulados de la mecánica cuántica* que son un marco matemático de referencia que permite el desarrollo de la teoría cuántica y, en particular, en la teoría cuántica de la información.

3.1. Postulados de la Mecánica Cuántica

Primer Postulado: *Cada sistema cuántico aislado tiene asociado un espacio vectorial complejo con producto interno conocido como Espacio de Hilbert \mathcal{H} o Espacio de Estados del sistema. El sistema físico está completamente definido por un vector estado que es unitario y que está contenido en \mathcal{H} .*

De acuerdo a la notación de Dirac, el estado cuántico es representado por un vector $|\psi\rangle \in \mathcal{H}$ llamado **estado** o **ket**. Cada vector en \mathcal{H} describe un posible estado del sistema. Además todos los estados obedecen el **principio de superposición** heredado del álgebra lineal. Para un espacio de Hilbert \mathcal{H}_2 de dimensión 2, un estado arbitrario $|\psi\rangle$ se puede escribir como una combinación lineal de dos estados

$$|\psi\rangle = a|\alpha\rangle + b|\beta\rangle \quad (3.1)$$

con $a, b \in \mathbb{C}$ y $|\alpha\rangle, |\beta\rangle \in \mathcal{H}$ son una base ortonormal. La condición que $|\psi\rangle$ es un vector unitario $\langle\psi|\psi\rangle = 1$ es completamente equivalente a $|a|^2 + |b|^2 = 1$.

¹Usualmente este concepto se menciona como *mundo clásico*.

Esta condición es conocida como **condición de normalización**.

Segundo Postulado: *La evolución de un sistema cuántico cerrado es descrito por una transformación unitaria U .* Esto es, el estado $|\psi\rangle$ de un sistema en un tiempo inicial t_0 está relacionado con el estado $|\psi'\rangle$ del mismo sistema pero en un tiempo posterior fijo t por una transformación unitaria $U := U(t_0, t)$,

$$|\psi'\rangle = U(t_0, t) |\psi\rangle. \quad (3.2)$$

Otra versión de este postulado puede ser formulada para un sistema cuántico para un tiempo continuo. Así, la evolución temporal del estado de un sistema cuántico cerrado es descrito por la ecuación de Schrödinger,

$$i\hbar \frac{d|\psi\rangle}{dt} = H |\psi\rangle, \quad (3.3)$$

dónde en la eq. (3.3) $\hbar = 1,054571628(53) \times 10^{-34} J \cdot s$ es la constante de Planck reducida [19], [16] y H es el *Hamiltoniano* del sistema cerrado.

Tercer Postulado: *Las mediciones cuánticas son representadas por un conjunto de operadores de medición $\{M_m\}$.* Estos operadores actúan sobre el espacio de Hilbert del sistema sobre el cual deseamos medir. La etiqueta m hace alusión a los resultados de las mediciones que se pueden obtener en un experimento. Para un estado $|\psi\rangle$, la probabilidad, inmediatamente antes de realizar la medición, de obtener el resultado con la etiqueta m es dado por

$$P_m = \langle \psi | M_m^\dagger M_m | \psi \rangle. \quad (3.4)$$

Luego de efectuar la medición, el estado del sistema $|\psi'\rangle$ es dado por

$$|\psi'\rangle = \frac{M_m}{\sqrt{\langle \psi | M_m^\dagger M_m | \psi \rangle}} |\psi\rangle. \quad (3.5)$$

Además los operadores de medición satisfacen la *relación de completitud*,

$$\sum_m M_m^\dagger M_m = \mathbb{1} \quad (3.6)$$

Cuarto Postulado: *El espacio de Hilbert de un sistema físico compuesto es el producto tensorial de los espacios de Hilbert de cada subsistema físico.* Es

decir, $\forall |\psi_j\rangle \in \mathcal{H}_j$, con $j = 1, \dots, n$ y $\mathcal{H}_j \subset \mathcal{H}$, se tiene que el espacio de Hilbert del sistema compuesto está dado por

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \dots \otimes |\psi_n\rangle \in \mathcal{H}. \quad (3.7)$$

A partir de la definición del cuarto postulado, podemos clasificar dos tipos de estados que son ampliamente usados en mecánica cuántica: los estados *separables* y los estados *entrelazados*. Los estados separables son estados que se pueden descomponer siguiendo la eq. (3.7). Por ejemplo, consideremos los estados $|\Phi\rangle_1 \in \mathcal{H}_1$ y $|\Phi\rangle_2 \in \mathcal{H}_2$, entonces podemos escribir el estado del sistema compuesto $|\psi\rangle_{12} \in \mathcal{H}$ como

$$|\psi\rangle_{12} = |\Phi\rangle_1 \otimes |\Phi\rangle_2. \quad (3.8)$$

Por otro lado, los estados entrelazados son aquellos que no se pueden descomponer en la forma de la eq. (3.7). Por ejemplo, si tenemos un estado

$$|\psi\rangle_{12} = \frac{1}{\sqrt{2}}(|0\rangle_1 \otimes |0\rangle_2 + |1\rangle_1 \otimes |1\rangle_2) \quad (3.9)$$

podemos observar que la primera base del estado $|\psi\rangle$ Así, vemos que resulta imposible escribir la eq. (3.9) de acuerdo al cuarto postulado. Por esta razón se dice que estos estados están entrelazados, puesto que las bases de cada espacio de Hilbert están mezcladas.

Los estados entrelazados de sistemas de dos dimensiones son los *estados de Bell*

$$|\psi_+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \quad (3.10)$$

$$|\psi_-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \quad (3.11)$$

$$|\phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \quad (3.12)$$

$$|\phi_-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \quad (3.13)$$

3.2. Qubit: El bit cuántico

En la Teoría Clásica de la Información, John Wilder Tukey² acuñó el término de *bit*³ para referirse a un sistema binario de numeración que está compuesto solo por dos valores: 0 y 1, los cuales representan acciones opuestas, por ejemplo abrir-cerrar o encender-apagar. El bit es la *unidad mínima de información* empleada en informática, en cualquier dispositivo digital o de almacenamiento, o en la Teoría de la Información. Por su parte la Teoría Cuántica de la Información tiene su propia versión del bit: el *qubit*, el cual está compuesto por dos bases ortonormales $|0\rangle$ y $|1\rangle$ de un espacio 2-dimensional. La diferencia entre los modelos es notoria: para el caso clásico el bit se comporta como una moneda que posee solo dos alternativas excluyentes entre sí, mientras que para el caso cuántico el qubit puede estar en otro estado además de $|0\rangle$ o $|1\rangle$. Esto es posible gracias a que en mecánica cuántica podemos formar combinaciones lineales de estados, conocidas como *superposición de estados*:

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle. \quad (3.14)$$

Los números α y β son números complejos, por lo que un qubit es un vector en un espacio vectorial complejo de dimensión dos. Las bases $|0\rangle$ y $|1\rangle$ son conocidas como *base computacional de estados* por la analogía del uso de los bits en las ciencias de la computación. Si realizamos una medición sobre la eq. (3.14), obtendremos el valor 0 con probabilidad $|\alpha|^2$ o 1 con probabilidad $|\beta|^2$. Resulta evidente que $|\alpha|^2 + |\beta|^2 = 1$, la cual es conocida como la *condición de normalización*. Debido a esta propiedad, podemos reescribir la eq. (3.14)

$$|\psi\rangle = e^{i\gamma} \left(\cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle \right) \quad (3.15)$$

dónde $\theta, \varphi, \gamma \in \mathbb{R}$. Notemos que $e^{i\gamma}$ es una *fase global* que podemos ignorar, debido a que no proporciona ningún efecto observable en las mediciones, por lo que la eq. (3.15) se reduce a

$$|\psi\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle. \quad (3.16)$$

²John Wilder Tukey - (16 de junio de 1915 - 26 de julio de 2000) estadista estadounidense.

³Voz tomada del inglés bit –acrónimo de bi[nary digi]t–, que significa, en informática, ‘*unidad de medida de información equivalente a la elección entre dos posibilidades igualmente probables*’. Su plural es *bits*.

Los parámetros $\theta \in [0, \pi]$ y $\varphi \in [0, 2\pi]$ definen a los qubits como punto sobre la superficie de una esfera unitaria tridimensional, conocida como Esfera de Bloch⁴. En esta representación geométrica la base $|0\rangle$ se ubica en el polo norte, mientras que la base $|1\rangle$ le corresponde el polo sur, tal y como se muestra en la figura

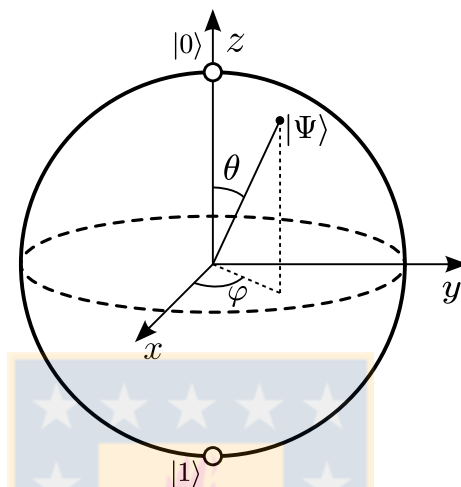


Figura 3.1: Representación de un qubit en la esfera de Bloch. Representación geométrica de un qubit en una esfera de Bloch. **Fuente:** Imagen generada por el autor.

Gracias a esta representación resulta más fácil observar cuales son los efectos de ciertas operaciones sobre los qubits. Es por esta razón que, más adelante, retomaremos este apartado para visualizar la generación de estados cuánticos utilizados para las simulaciones de nuestro trabajo.

3.3. Operador densidad

En las secciones anteriores se hizo una descripción de la mecánica cuántica en término del lenguaje de los vectores de estado. Otra formulación posible es mediante el uso del **operador densidad** o **matriz densidad**, la cual es utilizada cuando el estado del sistema no es completamente conocido. Supongamos que un sistema cuántico está, con probabilidad p_i , en el estado $|\psi_i\rangle$. Se llamará al conjunto $\{p_i, |\psi_i\rangle\}_{i=1}^n$ un **ensemble de estados puros**. El operador densidad para este sistema está definido de acuerdo a

$$\rho := \sum_{i=1}^n p_i |\psi_i\rangle\langle\psi_i|. \quad (3.17)$$

⁴Felix Bloch - (Zúrich, Suiza, 23 de octubre de 1905 - 10 de septiembre de 1983) fue un físico suizo que trabajó en energía nuclear en el Laboratorio Nacional de Los Álamos y que obtuvo el Premio Nobel de Física en 1952.

Si un sistema cuántico está con probabilidad igual a 1 en el estado $|\psi\rangle$ decimos que ρ representa a un **estado puro** descrito por,

$$\rho_{\text{puro}} = |\psi\rangle\langle\psi|. \quad (3.18)$$

Por otro lado, si un sistema es representado como en (3.17) decimos que es un **estado mezcla**, debido a que ρ puede ser descrito como una mezcla de diferentes estados puros del ensemble $\{p_i, |\psi_i\rangle\}_{i=1}^n$.

Supongamos que la evolución de un sistema cuántico es descrito por medio de un operador unitario U . Si el sistema está inicialmente en un estado $|\psi_i\rangle$ con probabilidad p_i entonces el estado final del sistema se encontrará en $U|\psi_i\rangle$ con probabilidad p_i . Así, la evolución del operador densidad es

$$\rho = \sum_{i=1}^N p_i |\psi_i\rangle\langle\psi_i| \xrightarrow{U} \sum_{i=1}^N p_i U |\psi_i\rangle\langle\psi_i| U^\dagger = U\rho U^\dagger. \quad (3.19)$$

Si realizamos una medición descrita por el operador M_m y el estado inicial es $|\psi_i\rangle$, la probabilidad de obtener el resultado m es

$$p(m|i) = \langle\psi_i|M_m^\dagger M_m|\psi_i\rangle = \text{Tr} \{M_m^\dagger M_m |\psi_i\rangle\langle\psi_i|\}, \quad (3.20)$$

dónde se ha utilizado la relación $\text{Tr} \{\mathbf{A} |\psi\rangle\langle\psi|\} = \langle\psi|\mathbf{A}|\psi\rangle$. Por otro lado, la probabilidad de obtener el resultado m es

$$p(m) = \sum_i p(m|i)p_i \quad (3.21)$$

$$= \sum_i p_i \text{Tr} \{M_m^\dagger M_m |\psi_i\rangle\langle\psi_i|\} \quad (3.22)$$

$$= \text{Tr} \{M_m^\dagger M_m \rho\}. \quad (3.23)$$

Si el estado inicial es $|\psi_i\rangle$, entonces el estado luego de obtener el resultado m es

$$|\psi_i^m\rangle = \frac{M_m |\psi_i\rangle}{\sqrt{\langle\psi_i|M_m^\dagger M_m|\psi_i\rangle}}. \quad (3.24)$$

Luego de las mediciones, se genera un ensemble de estados $\{p(i|m), |\psi_i^m\rangle\}_{i=1}^n$. El operador densidad ρ_m es, por lo tanto

$$\rho_m = \sum_i p(i|m) |\psi_i^m\rangle\langle\psi_j^m| = \sum_i p(i|m) \frac{M_m |\psi_i\rangle\langle\psi_i| M_m^\dagger}{\langle\psi_i|M_m^\dagger M_m|\psi_i\rangle}. \quad (3.25)$$

Utilizando la propiedad $p(i|m) = p(m, i)/p(m) = p(m|i)p_i/p(m)$ y los resultados obtenidos en (3.20) y (3.23), se tiene

$$\rho_m = \sum_i p_i \frac{M_m |\psi_i\rangle\langle\psi_i| M_m^\dagger}{\text{Tr} \{ M_m^\dagger M_m \rho \}} \quad (3.26)$$

$$= \frac{M_m \rho M_m^\dagger}{\text{Tr} \{ M_m^\dagger M_m \rho \}} \quad (3.27)$$

De esta forma, es posible reformular los postulados de la mecánica cuántica en el lenguaje de los operadores densidad.

3.3.1. Propiedades generales del operador densidad

Los tipos de operadores densidad son caracterizados de acuerdo al siguiente teorema

Teorema 2. *Un operador ρ es un operador densidad asociado a un ensemble $\{p_i, |\psi_i\rangle\}$ si y solo si satisface las siguientes condiciones:*

1. *La traza del operador ρ es igual a uno. $\text{Tr}\{\rho\} = 1$. [condición de traza]*
2. *ρ es un operador positivo. $\langle\psi|\rho|\psi\rangle \geq 0, \forall\psi$ [Condición de positividad]*

Si ρ es un operador que satisface las condiciones del teorema 2, entonces se puede aplicar el teorema de descomposición espectral

$$\rho = \sum_j \lambda_j |j\rangle\langle j| \quad (3.28)$$

donde los vectores $|j\rangle$ son ortogonales y λ_j son autovalores reales y positivos de ρ . De la condición de la traza, vemos que $\sum_i \lambda_i = 1$. Luego, un sistema, que se encuentra en el estado $|j\rangle$ con probabilidad λ_j , tendrá asociado un operador ρ . Dicho de otra forma, el ensemble $\{\lambda_j, |j\rangle\}$ es un ensemble de estados que da lugar a un operador densidad ρ .

Teorema 3. *El operador densidad ρ es un estado puro si y solo si $\text{Tr}\{\rho^2\} = 1$.*

Una forma geométrica de representar el operador densidad es en una esfera de Bloch, para ello un estado se puede parametrizar como:

$$\rho = \frac{1}{2} (\mathbb{1} + \vec{r} \cdot \vec{\sigma}) \quad (3.29)$$

dónde $\vec{\sigma} := (\sigma_x, \sigma_y, \sigma_z)$ con

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (3.30)$$

En esta representación los estados puros deben satisfacer $\|\vec{r}\| = 1$ y los estados mixtos $\|\vec{r}\| < 1$. Esto indica que los estados puros son representados como puntos sobre la esfera de Bloch, mientras que los estados mixtos son puntos en el interior.

3.3.2. Operador densidad reducido

Consideremos un sistema físico compuesto por A y B . El estado total del sistema lo podemos representar mediante un operador densidad etiquetado como ρ^{AB} . El operador **operador densidad reducida** para el sistema A es definido como

$$\rho^A := \text{Tr}_B \{ \rho^{AB} \}, \quad (3.31)$$

donde Tr_B se calcula como

$$\text{Tr} \{ |a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2| \} = |a_1\rangle\langle a_2| \text{Tr} \langle b_1| \langle b_2| \rangle \quad (3.32)$$

con $|a_1\rangle$ y $|a_2\rangle$ son dos vectores en el espacio de estados de A y $|b_1\rangle$ y $|b_2\rangle$ son dos vectores en el espacio de estados de B .

3.4. La descomposición de Schmidt

Es una descomposición que permite expresar cualquier estado de un sistema bipartito como una superposición de productos tensoriales de estados de cada subsistema.

Teorema 4. *Suponga que $|\psi\rangle$ es un estado puro de un sistema compuesto AB . Entonces existe una base ortonormal de estados $|i_A\rangle \in A$, y una base ortonormal de estados $|i_B\rangle \in B$ tal que*

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle \otimes |i_B\rangle, \quad (3.33)$$

dónde $\lambda_i \in \mathbb{R}_+$ que satisface $\sum_i \lambda_i^2 = 1$ los cuales son conocidos como coeficientes de Schmidt.

Este teorema tiene una consecuencia muy útil para abordar o resolver problemas en mecánica cuántica. Consideremos, por ejemplo, un estado puro $|\phi\rangle$ del sistema compuesto AB . Entonces, por la descomposición de Schmidt, tenemos que $\rho^A = \sum_i \lambda_i^2 |i_A\rangle\langle i_A|$ y $\rho^B = \sum_i \lambda_i^2 |i_B\rangle\langle i_B|$, entonces los autovalores para ρ^A y ρ^B son los mismos, es decir λ_i^2 para ambas matrices densidad en cada subsistema.

3.5. Estados Isotrópicos y estados de Werner

Definimos un *estado isotrópico* ρ_α^d sobre un espacio de Hilbert compuesto por dos subespacios de Hilbert de dimensión d tal que $\mathcal{H} = \mathcal{H}^d \otimes \mathcal{H}^d$, como

$$\rho_\alpha^d = \alpha |\phi_+^d\rangle\langle\phi_+^d| + \frac{1-\alpha}{d^2} \mathbb{1} \otimes \mathbb{1}, \quad -\frac{1}{d^2-1} \leq \alpha \leq 1, \quad (3.34)$$

donde el estado $|\phi_+^d\rangle$ es un estado maximalmente entrelazado definido por

$$|\phi_+^d\rangle := \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle \otimes |i\rangle, \quad (3.35)$$

y $\{|i\rangle\}$ es una base en \mathcal{H}^d .

Estos estados son llamados *isotrópicos* puesto que son invariantes bajo transformaciones $U \otimes U^*$ ⁵, es decir

$$(U \otimes U^*) \rho_\alpha^d (U \otimes U^*)^\dagger = \rho_\alpha^d. \quad (3.36)$$

Los estados mostrados en la eq. (3.34), de acuerdo al criterio de reducción, satisfacen

$$\begin{aligned} \rho_\alpha^d \text{ es separable si } & -\frac{1}{d^2-1} \leq \alpha \leq \frac{1}{d+1}, \\ \rho_\alpha^d \text{ es entrelazado si } & \frac{1}{d+1} < \alpha \leq 1. \end{aligned} \quad (3.37)$$

De la eq. (3.34) con $d = 2$ obtenemos los estados de Werner

⁵ U es un operador unitario y U^* es su complejo conjugado

$$\rho_W = \alpha |\phi_+\rangle\langle\phi_+| + \frac{1-\alpha}{4} \mathbb{1} \otimes \mathbb{1}, \quad -\frac{1}{3} \leq \alpha \leq 1, \quad (3.38)$$

los cuales satisfacen las propiedades mostradas en la siguiente tabla

Características de ρ_w	Valor de α
Separable	$-\frac{1}{3} \leq \alpha \leq \frac{1}{3}$
Entrelazado que no viola Bell	$\frac{1}{3} < \alpha < \frac{1}{\sqrt{2}}$
Entrelazado que viola Bell	$\frac{1}{\sqrt{2}} \leq \alpha \leq 1$

Tabla 3.1: Resumen de propiedades de los estados de Werner. Propiedades de los estados de Werner en función del parámetro α . Estas propiedades representan estados cuánticos separables, entrelazados que no violan las desigualdades de Bell y entrelazados que si las violan. **Fuente:** Elaboración propia.

Capítulo 4

Paradoja Einstein-Podolsky-Rosen y Desigualdades de Bell

En la primera mitad del siglo XX, no todos los científicos estaban de acuerdo con las consecuencias que dejaban las predicciones de la mecánica cuántica a las leyes de la física. Albert Einstein, uno de los científicos más importantes de nuestra historia, no estaba ajeno a ello y tenía profundas críticas respecto a esta interpretación, en particular a la descripción de fenómenos no locales (la llamada *acción fantasmagórica a distancia*) y al principio de incertidumbre de Heisenberg. Él estaba convencido de que «*la descripción de la realidad es dada por una función de onda que no es completa*», lo cual queda demostrado en su célebre artículo publicado en 1935. Posterior a esto, David Bohm realiza una reinterpretación de lo propuesto en EPR dejando planteada la posibilidad de la necesidad de incorporar variables ocultas auxiliares que permitían completar la función de onda y así ésta pueda dar una descripción completa de la realidad. Finalmente, con el trabajo publicado por John Bell en 1964, la discusión se termina: él demuestra, mediante el uso de variables ocultas, que *ninguna teoría determinista y local puede reproducir todos los resultados de la Mecánica Cuántica* demostrando así que la mecánica cuántica no es una teoría incompleta y que el estado cuántico sí da una descripción completa de los fenómenos.

En las siguientes secciones se realizará una discusión detallada de todo lo planteado en éste preámbulo.

4.1. Paradoja Einstein-Podolsky-Rosen

Albert Einstein publica en 1935, junto a Boris Podolsky y Nathan Rosen, su crítica más seria a la mecánica cuántica y sus fundamentos, el artículo llamado “*Can*

Quantum-Mechanical Description of Reality be Considered Complete?" [11], donde cuestionan y ponen de manifiesto un problema aparente contenido en la teoría. Los autores estaban convencidos de que la mecánica cuántica, de ser local, daría una descripción incompleta de la realidad física, concluyendo que «*la descripción de la realidad es dada por una función de onda que no es completa*». Para comprender esta conclusión, debemos definir algunos conceptos fundamentales descritos en el artículo EPR.

Para intentar juzgar el éxito de una teoría física, debemos hacernos dos preguntas

- (1) ¿Es una teoría *correcta*?
- (2) ¿Es *completa* la descripción dada por esta teoría?

Este contexto, una teoría *correcta* se define como el *grado de acuerdo* entre la conclusión de una teoría y la experiencia humana. En física la experiencia que nos permite inferir la realidad objetiva son los *experimentos* y las *mediciones*. Es por esto que una teoría correcta debe ser capaz de calcular y predecir hechos comprobables experimentalmente. Por otro lado, una teoría completa debe cumplir necesariamente lo siguiente: «*todo elemento de realidad física debe tener una contraparte en la teoría física*». Esta afirmación es conocida como **condición de completitud**.

Los *elementos de realidad física* no pueden ser determinados *a priori* por consideraciones filosóficas, deben encontrarse apelando a experimentos y mediciones. Así, utilizaremos el siguiente criterio que consideraremos razonable «*Si, sin perturbar de forma alguna un sistema, podemos predecir con certeza (es decir, con probabilidad igual a la unidad) el valor de una cantidad física, entonces un elemento de realidad física correspondiente a esta cantidad física*». Esta criterio es una condición de suficiencia, no de necesidad.

Para ilustrar estas ideas consideremos el movimiento de una partícula en una dimensión, donde el *estado*, que suponemos que está completamente caracterizado por la función de onda Ψ , es una función que describe el comportamiento de la partícula. A cada cantidad físicamente observable \mathbf{A} le corresponde un operador, el cual será etiquetado con la misma letra.

Si Ψ es una autofunción del operador \mathbf{A} , entonces

$$\mathbf{A}\Psi = a\Psi, \tag{4.1}$$

donde a es el autovalor asociado. Así, la cantidad física \mathbf{A} tiene con certeza el valor a cuando la partícula está en el estado dado por Ψ . De acuerdo con el criterio de realidad, para una partícula en el estado Ψ que satisface la eq. (4.1), hay un elemento de realidad física correspondiente a la cantidad física \mathbf{A} . Si, por ejemplo la función de onda Ψ es

$$\Psi(x) = e^{(2\pi i/h)p_0 x} \quad (4.2)$$

donde h es la constante de Planck, p_0 es algún número constante y x es la variable independiente. Debido a que el operador correspondiente al momentum de la partícula es

$$\mathbf{P} = \left(\frac{h}{2\pi i} \right) \frac{\partial}{\partial x}, \quad (4.3)$$

obtenemos

$$\mathbf{P}\Psi = \left(\frac{h}{2\pi i} \right) \frac{\partial \Psi}{\partial x} = p_0 \Psi. \quad (4.4)$$

Así, en la eq. (4.4) vemos que el momentum tiene indudablemente el valor p_0 . Con este resultado tiene sentido decir que el momentum de la partícula con función de onda eq. (4.2) es un elemento de realidad.

Por otro lado, si la eq. (4.1) no se satisface, entonces no podemos asegurar que \mathbf{A} tiene un valor en particular. Para el ejemplo anterior, no podemos decir nada sobre la posición de la partícula. El operador correspondiente a la posición es representado por \mathbf{Q} . Luego,

$$\mathbf{Q}\Psi = x\Psi \neq a\Psi. \quad (4.5)$$

donde a es el autovalor presentado en la eq. (4.1). De acuerdo a la mecánica cuántica, solo podemos decir que la probabilidad relativa de que una medida de la posición nos de un resultado entre a y b es

$$P(a, b) = \int_a^b \bar{\Psi}\Psi dx = \int_a^b dx = b - a. \quad (4.6)$$

Debido a que la probabilidad es independiente de a pero que dependen solo de la diferencia $b - a$, podemos ver que todos los valores de las posiciones son igualmente probables.

Un valor definido de la posición, para una partícula dada por la función de onda definida en la eq. (4.2), no es predecible, pero podría ser obtenido solo por una medición directa. Dicha medición perturba a la partícula y altera su estado. Así la posición queda determinada pero la partícula ya no estará en el estado dado por la

eq. (4.2). A partir de este resultado concluimos que **cuando el momentum de la partícula es conocido, su posición no posee realidad física.**

Como es sabido, si dos operadores corresponden a dos cantidades físicas, llamadas \mathbf{A} y \mathbf{B} y no conmutan $\mathbf{AB} \neq \mathbf{BA}$, entonces solo podemos conocer con precisión una de ellas a la vez. Dicho de otra forma, el conocimiento de una implica el desconocimiento de la otra. Esta afirmación es conocida en mecánica cuántica como el *principio de incertidumbre de Heisenberg* y puede ser representada para los operadores posición y momentum como

$$[\mathbf{Q}, \mathbf{P}] = i\hbar. \quad (4.7)$$

A partir de lo expuesto anteriormente, se observa que los criterios de realidad física y completitud implican a una disyuntiva excluyente entre dos posibilidades:

- (1) la descripción de realidad dada por la función de onda en mecánica cuántica no es completa
- (2) cuando los operadores correspondientes a dos cantidades físicas no conmutan dichas cantidades no poseen realidad física simultánea.

Para ilustrar esto, consideremos un sistema bipartito o de dos partes, I y II , las cuales permitiremos que interactúen entre los tiempos $t = 0$ y $t = T$. Posterior a este intervalo de tiempo los subsistemas no interactúan entre si. Supondremos, además, que el estado de los dos sistemas antes de $t = 0$ eran conocidos. Sea Ψ la función de onda correspondiente al sistema completo, a_1, a_2, \dots los autovalores de alguna cantidad A del sistema I y $u_1(x_1), u_2(x_1), u_3(x_1), \dots$ las autofunciones correspondientes, donde x_1 indica las variables definidas para el primer sistema. Entonces la función de onda Ψ es una función de x_1 y puede ser expandida como

$$\Psi(x_1, x_2) = \sum_{n=1}^{\infty} \psi_n(x_2)u_n(x_1), \quad (4.8)$$

donde x_2 señala las variables que describen el segundo sistema. Aquí, $\psi_n(x_2)$ debe ser considerado como el coeficiente de la expansión de Ψ en una serie de funciones ortogonales $u_n(x_1)$. Supongamos ahora que la cantidad A es medida y el resultado de aquella medición es a_k . Después de la medición el estado del primer sistema es dado por la función de onda $u_k(x_1)$ y el estado del segundo sistema es dado por $\psi_k(x_2)$. De esta forma el estado total queda reducido a

$$\Psi(x_1, x_2) \propto \psi_k(x_2)u_k(x_1). \quad (4.9)$$

Notemos que el valor de $u_n(x_1)$ queda determinado por la elección de la cantidad física A . Por otro lado, consideremos la cantidad B que es medida sobre el primer sistema, con autovalores b_1, b_2, \dots y autofunciones $v_1(x_1), v_2(x_1), \dots$, tendremos otra expansión para Ψ en lugar de la mostrada en la eq. (4.8), la cual es dada por

$$\Psi(x_1, x_2) = \sum_{s=1}^{\infty} \varphi_s(x_2) v_s(x_1) \quad (4.10)$$

donde los φ_s son nuevos coeficientes. Supongamos que medimos la cantidad B con resultado b_r . De esta forma, el estado del primer sistema queda en $v_r(x_1)$ mientras que el segundo sistema es dejado en $\varphi_s(x_2)$.

De lo expuesto se desprende que al realizar dos mediciones diferentes sobre el primer sistema, el segundo sistema puede quedar en estados con dos funciones de onda distintas. Por otro lado, en ausencia de interacción entre los sistemas, ningún cambio real debiese ocurrir en el segundo sistema como consecuencias de realizar alguna medida sobre el primero. Así, *es posible asignar dos funciones de onda diferentes* (en lo anterior ψ y φ) *a la misma realidad* (el segundo sistema luego de interactuar con el primero).

Ahora consideremos que ψ y φ son autofunciones de dos operadores que no conmutan, como el operador posición (\mathbf{Q}) y el operador momentum (\mathbf{P}), respectivamente. Consideremos que los dos sistemas son dos partículas y que

$$\Psi(x_1, x_2) = \int_{-\infty}^{\infty} e^{(2\pi i/h)(x_1 - x_2 + x_0)p} dp, \quad (4.11)$$

donde x_0 es alguna constante. Sea A el momentum de la primera partícula, entonces, como vimos en la ecuación eq. (4.4), sus autofunciones serán

$$u_p(x_1) = e^{(2\pi i/h)px_1} \quad (4.12)$$

correspondiente al autovalor p . Debido a que se está estudiando el caso sobre un espectro continuo, la eq. (4.8) puede ser escrita como

$$\Psi(x_1, x_2) = \int_{-\infty}^{\infty} \psi_p(x_2) u_p(x_1) dp \quad (4.13)$$

donde ahora tenemos que

$$\psi_p(x_2) = e^{-(2\pi i/h)(x_2 - x_0)p}. \quad (4.14)$$

Luego, ψ_p es la autofunción del operador

$$\mathbf{P} = \frac{\hbar}{2\pi i} \frac{\partial}{\partial x_2} \quad (4.15)$$

correspondiente al autovalor $-p$ del operador momentum de la segunda partícula. Por otro lado, si B es la posición de la primera partícula, esta tiene autofunciones

$$v_x(x_1) = \delta(x_1 - x) \quad (4.16)$$

con autovalor x , donde δ es la delta de Dirac. Así, la versión continua de la eq. (4.8) es

$$\Psi(x_1, x_2) = \int_{-\infty}^{\infty} \varphi_x(x_2) v_x(x_1) dx \quad (4.17)$$

donde

$$\varphi_x(x_2) = \int_{-\infty}^{\infty} e^{(2\pi i/\hbar)(x-x_2+x_0)p} dp = \hbar \delta(x - x_2 + x_0). \quad (4.18)$$

Sin embargo, φ_x es autofunción del operador

$$\mathbf{Q} = x_2 \quad (4.19)$$

correspondiente al autovalor $x + x_0$ de la posición de la segunda partícula. Ya que

$$\mathbf{PQ} - \mathbf{QP} = i\hbar, \quad (4.20)$$

se ha mostrado que, en general, es posible que ψ_k y φ_r sean autofunciones de dos operadores que no conmutan, correspondientes a cantidades físicas.

De acuerdo los casos planteados en las ecuaciones (4.8) y (4.10), asumimos que ψ_k y φ_r son en realidad autofunciones de operadores que no conmutan \mathbf{P} y \mathbf{Q} , correspondientes a los autovalores p_k y q_r , respectivamente. Así, si medimos A o B , podemos predecir con certeza y sin perturbar de forma alguna el segundo sistema, ya sea el valor de la cantidad \mathbf{P} (que es p_k) o el valor de la cantidad \mathbf{Q} (que es q_r). Así, de acuerdo con el criterio de realidad planteado anteriormente, en el primer caso la cantidad \mathbf{P} tiene un elemento de realidad, en el segundo caso la cantidad \mathbf{Q} es un elemento de realidad, pero, como vimos, ambas funciones ψ_k y φ_r pertenecen a la misma realidad.

Como partimos suponiendo que la función de onda da una completa descripción de la realidad física, se llega a la conclusión que las dos cantidades, con dos operadores que no conmutan, pueden tener la misma realidad física. Así, la negación de la afirmación (1) conduce a la negación de la afirmación (2), concluyendo finalmente que, entonces, la realidad física dada por la función de onda no es completa.



4.2. La versión de Bohm-Aharonov (1957)

David Bohm y Yakir Aharonov en 1957 [4] proponen una versión alternativa y simplificada de la crítica realizada por EPR, donde utiliza para su estudio las componentes de espín de dos partículas con espín $1/2$.

El planteamiento de Bohm-Aharonov considera dos partículas que se alejan entre sí desde la fuente que las produce, cuyos espines son opuestos de tal modo que el espín total del sistema completo sea cero. A una cierta distancia, se ubican dos observables llamados Alice y Bob compuestos cada uno por aparatos de Stern-Gerlach. En Alice el aparato puede rotar en 90° permitiendo medir las componentes de espín en z o x , mientras que para Bob el aparato solo puede medir en la dirección z .

El estado que describe esta configuración es dado por el estado singlete, el cual es un estado maximalmente entrelazado,

$$|\Psi_s\rangle = \frac{1}{\sqrt{2}} (|+\rangle_z |-\rangle_z - |-\rangle_z |+\rangle_z). \quad (4.21)$$

Para ver el efecto de este experimento es necesario realizar medidas en los dos sistemas. Si, por ejemplo, Alice mide con $S_{1,z}$ en su sistema, obtendrá el resultado $+1$ y el estado singlete queda reducido a $|+\rangle_z |-\rangle_z$ perdiendo el entrelazamiento y las quedando las componentes de espín perfectamente definidas. Si, ahora, Bob mide con $S_{2,z}$ en su sitio obtendrá, sin lugar a dudas, -1 sin afectar en ningún caso al sistema. De esta forma $S_{2,z}$ tiene realidad física para Bob.

Si Alice decide medir la componente $S_{1,x}$, para obtener los resultados $+1$ o -1 el tratamiento es ligeramente distinto. Vemos que $|\pm\rangle_z$ se puede re escribir como

$$|\pm\rangle_z = \frac{1}{\sqrt{2}} (|+\rangle_x \pm |-\rangle_x). \quad (4.22)$$

Así, $|\Psi_s\rangle$ queda

$$|\Psi_s\rangle = \frac{1}{2} [|+\rangle_x \otimes (|-\rangle_z + |+\rangle_z) - |-\rangle_x \otimes (|-\rangle_z + |+\rangle_z)] \quad (4.23)$$

Si, por ejemplo, Alice mide y obtiene -1 , entonces el estado es

$$|\Psi_s\rangle \longrightarrow |-\rangle_x \otimes \frac{|-\rangle_z + |+\rangle_z}{\sqrt{2}} \quad (4.24)$$

Usando la eq. (4.22), la expresión anterior puede ser representada como

$$|\Psi_s\rangle \longrightarrow |-\rangle_x \otimes \frac{1}{\sqrt{2}} \left[\frac{1}{\sqrt{2}} (|+\rangle_x - |-\rangle_x) + \frac{1}{\sqrt{2}} (|+\rangle_x + |-\rangle_x) \right] = |-\rangle_x |+\rangle_x. \quad (4.25)$$

De la eq. (4.25) vemos que la partícula del sistema 2 tiene un valor bien definido para la componente $S_{2,x}$. Por otro lado, si Bob decide medir $S_{2,z}$ puede obtener los valores $+1$ o -1 con igual probabilidad. Por lo tanto, podemos asegurar que $S_{2,x}$ posee realidad física (puesto que se conoce su resultado de antemano), mientras que $S_{2,z}$ no.

De acuerdo a lo expuesto por EPR, al momento de realizar una medida sobre Alice, la realidad física de Bob está siendo afectada al instante a pesar de que se encuentran en sitios diferentes. Debido a esto se debe abandonar el concepto de localidad para esta situación. Por otro lado, una vez que Alice realiza sus mediciones sobre su sistema, las cantidades $S_{2,x}$ y $S_{2,z}$ estarían bien definidas en la realidad de Bob. Finalmente la única posibilidad es que el estado $|\Psi\rangle$ no de una descripción completa de la realidad física del sistema propuesto. Esta conclusión implica que la mecánica cuántica no es capaz de dar una descripción completa de la realidad física de los sistemas, requiriendo *variables ocultas* adicionales que permitan complementar la descripción dada por el estado $|\Psi\rangle$.

4.3. Desigualdad de Bell (1965)

A partir de lo expuesto por David Bohm, John Bell en 1965 publica un artículo donde demuestra que *toda teoría de variables ocultas que sea determinista y local tiene necesariamente algunas predicciones incompatibles con la Mecánica Cuántica*, garantizando que existe una incompatibilidad entre ésta y cualquier teoría determinista local de variables ocultas. Esto queda expuesto mediante una demostración matemática, que conduce a una inequación, la cual que permite enunciar posteriormente el **teorema de Bell**.

Supongamos el mismo escenario planteado por Bohm. A la misma distancia de la fuente pero en direcciones opuestas, se ubican Alice y Bob provistos con dos aparatos que miden de forma simultánea las componentes de espín de cada partícula en dos direcciones \hat{a} y \hat{b} . Llamaremos $A(\hat{a})$ al resultado de la medición realizada por Alice a la partícula 1 en la dirección \hat{a} y $B(\hat{b})$ el resultado la medición sobre la partícula 2 de Bob en la dirección \hat{b} . Si se mide sobre el estado singlete en la locación de Alice en la dirección \hat{a} [$A(\hat{a})$] y el resultado es $+1$, entonces, según la mecánica cuántica, la medida en Bob en la dirección \hat{a} [$B(\hat{a})$] debe ser -1 . Debido a que es posible predecir el resultado de $B(\hat{a})$ a partir del resultado obtenido en $A(\hat{a})$ y como además las partículas están separadas, podemos afirmar que $B(\hat{a})$ es un elemento de realidad

de acuerdo a lo expuesto por EPR. Notemos que también esto ocurre en el sentido inverso, es decir, $A(\hat{a}) = -1$ y $B(\hat{a}) = +1$. Además, debido a que \hat{a} es arbitrario, para cada elección del vector unitario obtendremos elementos de realidad para la segunda partícula.

De la conclusión del trabajo de Bohn, utilizaremos una variable auxiliar desconocida λ llamada *variable oculta*, la cual nos permitirá describir de manera *completa* el estado del sistema. Al considerar λ estamos asumiendo una hipótesis **determinista** adicional al planteado en los trabajos anteriores. De esta forma, si consideramos el escenario expuesto anteriormente, y añadiendo esta variable oculta, la medida sobre la primera partícula queda determinada por la dirección de medición y la variable oculta λ . Así, $A(\hat{a}, \lambda)$ indica el resultado (que puede ser $+1$ o -1) de medir a la primera partícula en la dirección \hat{a} , análogamente $B(\hat{b}, \lambda)$ representa el resultado de medir la segunda partícula en la dirección \hat{b} . De esta forma,

$$A(\hat{a}, \lambda) = \pm 1, \quad B(\hat{b}, \lambda) = \pm 1. \quad (4.26)$$

Supondremos que cada medida esta determinada por el parámetro λ . De acuerdo al principio de localidad, el resultado de la medición sobre la primera partícula $A(\hat{a}, \lambda)$ no depende de lo que se mida en la segunda, ni que $B(\hat{b}, \lambda)$ dependa de lo que se mida en la primera.

Asumiremos, además, que las variables ocultas tienen distribución de probabilidad $\rho(\lambda)$ que satisface

$$\int \rho(\lambda) d\lambda = 1. \quad (4.27)$$

De lo anterior, definimos la *función de correlación* entre los resultados de las mediciones realizadas sobre cada partícula, como

$$P(\hat{a}, \hat{b}) = \int \rho(\lambda) A(\hat{a}, \lambda) B(\hat{b}, \lambda) d\lambda. \quad (4.28)$$

Consideremos, una medición sobre la segunda partícula orientada sobre la dirección \hat{b}' . Notemos que solo se puede realizar una medición por vez tanto en Alice como en Bob, pero esta medición adicional se puede realizar sobre otro estado idénticamente preparado. De esta manera, podemos calcular

$$P(\hat{a}, \hat{b}) - P(\hat{a}, \hat{b}') = \int \left[A(\hat{a}, \lambda) B(\hat{b}, \lambda) - A(\hat{a}, \lambda) B(\hat{b}', \lambda) \right] \rho(\lambda) d\lambda. \quad (4.29)$$

Debido a que el espín total del sistema es cero, tenemos

$$A(\hat{b}, \lambda) = -B(\hat{b}, \lambda), \quad A(\hat{b}', \lambda) = -B(\hat{b}', \lambda). \quad (4.30)$$

De acuerdo a las relaciones (4.30) podemos reescribir la eq. (4.29),

$$P(\hat{a}, \hat{b}) - P(\hat{a}, \hat{b}') = - \int \left[A(\hat{a}, \lambda) A(\hat{b}, \lambda) - A(\hat{a}, \lambda) A(\hat{b}', \lambda) \right] \rho(\lambda) d\lambda. \quad (4.31)$$

Considerando que $A(\hat{b}, \lambda)A(\hat{b}, \lambda) = 1$, la eq. (4.31) queda

$$P(\hat{a}, \hat{b}) - P(\hat{a}, \hat{b}') = - \int A(\hat{a}, \lambda) A(\hat{b}, \lambda) \left[1 - A(\hat{b}, \lambda) A(\hat{b}', \lambda) \right] \rho(\lambda) d\lambda. \quad (4.32)$$

Puesto que $A(\hat{a}, \lambda) A(\hat{b}, \lambda)$ puede ser $+1$ o -1 , tendrá valor absoluto 1 y $\left[1 - A(\hat{a}, \lambda) A(\hat{b}', \lambda) \right]$ es no negativo y, por lo tanto, es igual a su valor absoluto. Considerando esto y tomando el valor absoluto en ambos lados, obtenemos

$$\left| P(\hat{a}, \hat{b}) - P(\hat{a}, \hat{b}') \right| \leq \int \left[1 - A(\hat{b}, \lambda) A(\hat{b}', \lambda) \right] \rho(\lambda) d\lambda = 1 + P(\hat{b}, \hat{b}'). \quad (4.33)$$

La eq. (4.33) es la desigualdad de Bell original mostrada en su artículo de 1964. Esta expresión nos indica que la formulación mediante el uso de variables ocultas locales está acotada. A partir de ella se puede demostrar que las variables ocultas deterministas son incompatibles con la mecánica cuántica, para ello debemos mostrar que la predicción cuántica supera o **viola** la desigualdad presentada en la eq. (4.33).

Supongamos que la función de correlación (4.28) coincide con el valor medio cuántico asociado a las medidas de espín simultáneas de cada partícula en dos direcciones \hat{a} y \hat{b} ,

$$\langle \Psi | A(\hat{a}) \otimes B(\hat{b}) | \Psi \rangle = -\hat{a} \cdot \hat{b} = -\cos \theta_{ab}. \quad (4.34)$$

Esta correlación cuántica solo depende de la orientación de las direcciones en las que se aplican las mediciones. Si se eligen tres orientaciones coplanares para $\hat{a}, \hat{b}, \hat{b}'$, de modo tal que formen ángulos de 60° entre ellos, se obtienen las correlaciones

$$P_{\text{MC}}(\hat{a}, \hat{b}) = P_{\text{MC}}(\hat{b}, \hat{b}') = -\frac{1}{2}, \quad P_{\text{MC}}(\hat{a}, \hat{b}') = \frac{1}{2}. \quad (4.35)$$

Finalmente con la relación (4.35), tenemos

$$\left| P_{\text{MC}}(\hat{a}, \hat{b}) - P_{\text{MC}}(\hat{a}, \hat{b}') \right| - P_{\text{MC}}(\hat{b}, \hat{b}') = \frac{3}{2} > 1, \quad (4.36)$$

violando la expresión (4.33). Este resultado puso en manifiesto que las teorías locales deterministas de variables ocultas son incompatibles con las predicciones que realiza la mecánica cuántica. Lo asombroso de este resultado es que plantea la posibilidad de una verificación experimental. Sin embargo, se requiere de correlaciones perfectas que nunca fueron realizadas experimentalmente. Debido a esto, en 1969 J. Clauser, M. Horne, A. Shimony y R. Holt proponen una alternativa, una nueva desigualdad de Bell que si puede ser testada con mayor facilidad en los laboratorios.

4.4. Desigualdad Clauser-Horne-Shimony-Holt (CHSH) (1969)

De acuerdo al artículo publicado por J. Clauser, M. Horne, A. Shimony y R. Holt en 1969 [9], es posible construir un experimento que permita verificar si la mecánica cuántica es o no consistente con una familia de teorías de variables ocultas (HVT).

Consideremos un sistema bipartito compuesto por Alice y Bob quienes están lo suficientemente separados para evitar cualquier tipo de interacción entre ellos (causalidad). Cada uno posee dos observables A_i y B_i los cuales realizan mediciones en las direcciones \hat{a} y \hat{b} respectivamente, que están configurados para tener dos salidas posibles y opuestas representadas como ± 1 . A cada parte se le envían dos partículas idénticas que provienen de una fuente en común.

Supongamos que una correlación estadística entre las mediciones de Alice $A(\hat{a})$ y Bob $B(\hat{b})$ es debido a la información llevada y localizada por cada partícula, y que, en algún momento del pasado, las partículas conformaron un par que estuvieron en contacto e interactuaron. Esta información no es cuántica y es parte de un conjunto de variables ocultas, las cuales son denotadas generalmente como λ . De esta forma los resultados de cada medición están dados por $A(\hat{a}, \lambda)$ y $B(\hat{b}, \lambda)$. Adicionalmente, requerimos que $A(\hat{a}, \lambda)$ no dependa de la dirección \hat{b} en Bob y que $B(\hat{b}, \lambda)$ no dependa de la dirección \hat{a} en Alice, puesto que están separadas a grandes distancias. Puesto que las partículas son generadas independientes de los ajustes \hat{a} y \hat{b} , se asume que la distribución de probabilidad $\rho(\lambda)$ que caracteriza el sistema solo depende de λ .

Definimos como $\langle A(\hat{a}, \lambda) \rangle$ y $\langle B(\hat{b}, \lambda) \rangle$ los valores promedio de los resultados $A(\hat{a}, \lambda)$ y $B(\hat{b}, \lambda)$. Definimos la función de correlación como

$$P(\hat{a}, \hat{b}) = \int d\lambda \rho(\lambda) \langle A(\hat{a}, \lambda) \rangle \langle B(\hat{b}, \lambda) \rangle \quad (4.37)$$

$$\begin{aligned} P(\hat{a}, \hat{b}) - P(\hat{a}, \hat{d}) &= \int d\lambda \rho(\lambda) \left[\langle A(\hat{a}, \lambda) \rangle \langle B(\hat{b}, \lambda) \rangle \left(1 \pm \langle A(\hat{c}, \lambda) \rangle \langle B(\hat{d}, \lambda) \rangle \right) \right] \\ &\quad - \int d\lambda \rho(\lambda) \left[\langle A(\hat{a}, \lambda) \rangle \langle B(\hat{d}, \lambda) \rangle \left(1 \pm \langle A(\hat{c}, \lambda) \rangle \langle B(\hat{b}, \lambda) \rangle \right) \right], \end{aligned} \quad (4.38)$$

entonces,

$$\left| P(\hat{a}, \hat{b}) - P(\hat{a}, \hat{d}) \right| \leq \int d\lambda \rho(\lambda) \left(2 \pm \langle A(\hat{c}, \lambda) \rangle \langle B(\hat{d}, \lambda) \rangle \pm \langle A(\hat{c}, \lambda) \rangle \langle B(\hat{b}, \lambda) \rangle \right) \quad (4.39)$$

Luego,

$$\left| P(\hat{a}, \hat{b}) - P(\hat{a}, \hat{d}) \right| \leq 2 \pm \left(P(\hat{c}, \hat{d}) + P(\hat{c}, \hat{b}) \right) \quad (4.40)$$

$$\left| P(\hat{a}, \hat{b}) - P(\hat{a}, \hat{d}) \right| + \left| P(\hat{c}, \hat{d}) + P(\hat{c}, \hat{b}) \right| \leq 2. \quad (4.41)$$

Utilizando la desigualdad triangular, se tiene que

$$-2 \leq P(\hat{a}, \hat{b}) - P(\hat{a}, \hat{d}) + P(\hat{c}, \hat{d}) + P(\hat{c}, \hat{b}) \leq 2, \quad (4.42)$$

la cual es la popular desigualdad CHSH.

Con frecuencia se encuentra representada como los valores de expectación de los valores obtenidos en un experimento, esto es

$$\mathcal{S} = |E_{11} + E_{12} + E_{21} - E_{22}| \leq 2, \quad (4.43)$$

con $E_{ij} = \text{Tr}\{\rho(A_i \otimes B_j)\}$ $i, j = 1, 2$, donde A_i y B_j tienen valores propios ± 1 . De esta manera y en concordancia con el teorema de Bell, si la parte izquierda de la eq. (4.43) es menor que 2, entonces el sistema puede ser descrito mediante teorías de variables ocultas locales y los estados que cumplan con esta característica se les denomina *estados separables*. Por otro lado, si $\mathcal{S} > 2$, entonces decimos que el estado **viola la desigualdad CHSH** y, por lo tanto, no puede ser descrito en término de teorías de variables ocultas locales.

4.5. Máxima violación de la desigualdad CHSH para estados cuánticos de dimensión arbitraria

N. Gisin y A. Peres [14] demuestran que es posible encontrar pares de observables cuyas correlaciones violen las desigualdades de Bell.

Teorema 5. *Sea $\psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$. Si ψ no es factorizable, existen observables $A \otimes B$, con autovalores ± 1 , cuyas correlaciones violen la desigualdad CHSH*

$$|E_{11} + E_{12} + E_{21} - E_{22}| \leq 2 \quad (4.44)$$

dónde $E_{ij} = \langle \psi | A_i \otimes B_j | \psi \rangle$.

Demostración: Cualquier estado ψ puede ser escrito en su descomposición de Schmidt de acuerdo a

$$\psi = \sum_i c_i u_i \otimes v_i, \quad (4.45)$$

donde $\{u_i\}$ y $\{v_i\}$ son dos bases ortogonales en \mathcal{H}_1 y \mathcal{H}_2 , respectivamente. Es posible elegir sus fases para que todos los c_i sean reales y positivos, y etiquetarlos como $c_1 \geq c_2 \geq \dots \geq 0$. Ahora podemos restringir nuestra atención al caso de subespacios N -dimensionales de \mathcal{H}_1 y \mathcal{H}_2 los cuales corresponden a términos no nulos de c_j . Un estado separable ocurre para $N > 1$.

Con las bases ortonormales definidas anteriormente, sean Γ_x y Γ_z matrices diagonales por bloque, donde cada bloque es una matriz de Pauli, σ_x y σ_y , respectivamente:

$$\Gamma_x := \begin{pmatrix} \sigma_x & 0 & 0 & \cdots \\ 0 & \sigma_x & 0 & \cdots \\ 0 & 0 & \sigma_x & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad \Gamma_z := \begin{pmatrix} \sigma_z & 0 & 0 & \cdots \\ 0 & \sigma_z & 0 & \cdots \\ 0 & 0 & \sigma_z & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (4.46)$$

Es decir, los elementos distintos de cero de Γ_x y Γ_z son

$$(\Gamma_x)_{2n-1,2n} = (\Gamma_x)_{2n,2n-1} = 1, \quad (\Gamma_z)_{n,n} = (-1)^{n-1}. \quad (4.47)$$

Si N es impar (el cual complica ligeramente la demostración) tomaremos $(\Gamma_z)_{NN} = 0$ en lugar de 1, y definimos otra matriz Π cuyos elementos distintos de cero son

$\Pi_{NN} = 1$. Si N es par, Π es la matriz nula. Además, es conveniente definir un número γ tal que

$$\gamma = c_N^2 \text{ (para } N \text{ impar),} \quad \gamma = 0 \text{ (para } N \text{ par).} \quad (4.48)$$

Con la notación anterior, consideremos los observables

$$\begin{aligned} A(\alpha_1) &:= \Gamma_x \sin \alpha_1 + \Gamma_z \cos \alpha_1 + \Pi \\ B(\beta_1) &:= \Gamma_x \sin \beta_1 + \Gamma_z \cos \beta_1 + \Pi \end{aligned} \quad (4.49)$$

Los autovalores de $A(\alpha)$ y $B(\beta)$, denotados por a y b respectivamente, son ± 1 y las correlaciones de esos observables son

$$E_{11} = \langle \psi | A(\alpha_1) \otimes B(\alpha_1) | \psi \rangle = (1 - \gamma) \cos \alpha_1 \cos \beta_1 + K \sin \alpha_1 \sin \beta_1 + \gamma \quad (4.50)$$

donde $A_i := A(\alpha_i)$, $B_i := B(\beta_i)$ y

$$K := 2(c_1 c_2 + c_3 c_4 + \dots) \quad (4.51)$$

si siempre es positivo para estados no factorizables. En particular, si elegimos $\alpha_1 = 0$, $\alpha_2 = \pi/2$, y $\beta_1 = -\beta_2 = \tan^{-1}[K/(1 - \gamma)]$, obtenemos

$$\begin{aligned} E_{11} + E_{12} + E_{21} - E_{22} &= 2[(1 - \gamma)^2 + K^2]^{1/2} + 2\gamma \\ &> 2(1 - \gamma) + 2\gamma, \end{aligned} \quad (4.52)$$

la cual contradice la eq. (4.44). \square

En el caso especial de un par de partículas de spin j en un estado singlete, tenemos

$$c_i = (2j + 1)^{-1/2}, \quad \forall i, \quad (4.53)$$

y, por tanto, $K = 1$ si $2j + 1$ es par. El lado derecho de la eq. (4.52) es, entonces, $2\sqrt{2}$ el cual es la máxima violación de la desigualdad permitida de acuerdo al teorema de Cirel'son [8]. Si, por otro lado, $2j$ es par, uno de los términos de la suma en la eq. (4.51) no tiene pareja y K es solo $2j/(2j + 1)$.

Capítulo 5

Algoritmos de optimización

5.1. Introducción: algoritmo del gradiente descendente

El *algoritmo del gradiente descendente* (AGD) es un método iterativo que permite hallar el punto $\tilde{p} \in \Omega$ tal que la *función de prueba* $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ alcance un mínimo. De lo anterior, f debe satisfacer la condición,

$$\nabla f(\tilde{p}) = 0. \tag{5.1}$$

Para abordar este problema consideramos una sucesión de puntos $\{\hat{p}\}$ tal que \hat{p} converge a \tilde{p} . La secuencia iterativa tiene la forma,

$$\hat{p}_{k+1} = \hat{p}_k - \alpha_k \nabla f(\hat{p}_k), \tag{5.2}$$

donde el parámetro α_k se elige de manera tal que $\hat{p}_{k+1} \in \Omega$ y $f(\hat{p}_{k+1}) < f(\hat{p}_k)$, $\Lambda_k \in \mathbb{R}^n$ es la *dirección del paso* y $\alpha_k \in \mathbb{R}$ es el *tamaño del paso*. Requerimos que el algoritmo se detenga cuando $\nabla f(\hat{p}_k)$ sea muy cercano a cero, esto ocurre cuando $\|\nabla f(\hat{p}_k)\| < \epsilon$, con $\epsilon > 0$. Usualmente se utiliza el *tamaño del paso constante* $\alpha_k = \alpha$ y *tamaño del paso decreciente*, es decir $\lim_{k \rightarrow \infty} \alpha_k = 0$ cuando $\sum_{k=1}^{\infty} \alpha_k$ diverge.

Por otra parte, existe el *algoritmo del gradiente ascendente* el cual es un método iterativo análogo en estructura al AGD, donde se busca el punto donde función de prueba f alcanza un máximo. Para su implementación basta con cambiar de signo el gradiente en la ecuación (5.2), es decir,

$$\hat{p}_{k+1} = \hat{p}_k + \alpha_k \nabla f(\hat{p}_k), \quad (5.3)$$

Para este caso se mantienen las consideraciones expresadas para el método AGD.

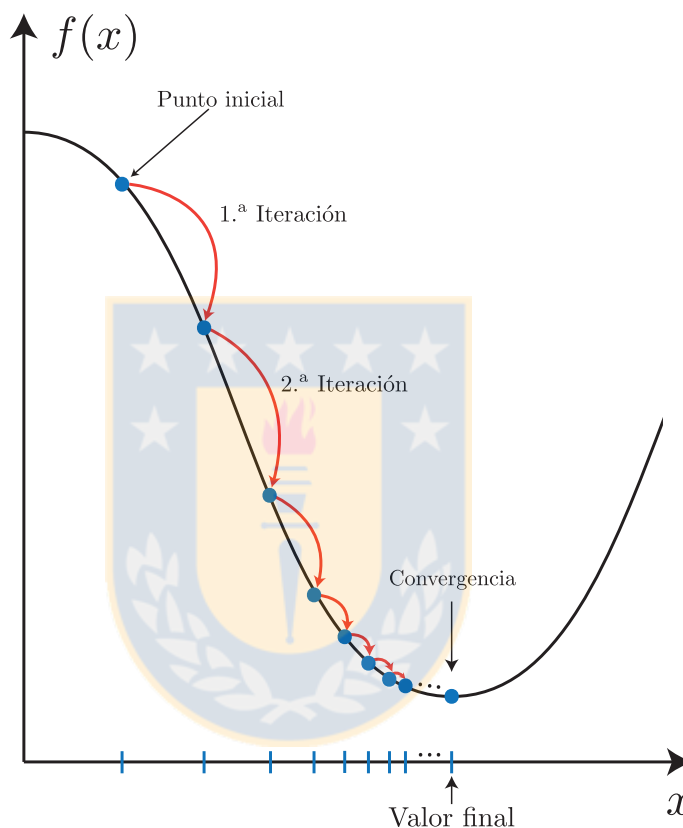


Figura 5.1: Diagrama del método del gradiente descendente. Representación gráfica del método del gradiente descendente. Dado un punto inicial de partida (guess), el algoritmo permite acercarse de manera asintótica y en una secuencia finita de iteraciones al punto donde la función alcanza un mínimo. **Fuente:** Imagen generada por el autor.

5.2. Simultaneous Perturbation Stochastic Approximation (SPSA)

El algoritmo *Simultaneous Perturbation Stochastic Approximation* (SPSA), presentado por J. C. Spall [20] en 1992, es una secuencia iterativa gradiente ascendente

que permite hallar el punto donde la función objetivo alcanza un mínimo. Las características principales de este método es que considera pasos que van disminuyendo en tamaño en cada iteración, el uso de múltiples parámetros desconocidos y la implementación de una aproximación del gradiente para la elección de la dirección del paso.

El problema se puede resumir en la búsqueda de un $\tilde{\mathbf{x}} \in \mathbb{R}^n$ tal que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ satisfaga

$$\tilde{\mathbf{g}}(\tilde{\mathbf{x}}) = \nabla f(\tilde{\mathbf{x}}) = 0. \quad (5.4)$$

El algoritmo de búsqueda es de la forma,

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}_k - \alpha_k \hat{\mathbf{g}}_k(\hat{\mathbf{x}}_k), \quad (5.5)$$

donde α_k es el tamaño del paso que disminuye en cada iteración y $\hat{\mathbf{g}}_k \in \mathbb{R}^n$ es la estimación del gradiente $\tilde{\mathbf{g}}$ en la k -ésima iteración. La i -ésima componente de $\hat{\mathbf{g}}_k$ se obtiene como

$$\hat{g}_{k,i}(\hat{\mathbf{x}}) = \frac{f(\hat{\mathbf{x}}_k + \beta_k \Delta_k) + \epsilon_{k,+} - f(\hat{\mathbf{x}}_k - \beta_k \Delta_k) - \epsilon_{k,-}}{2\beta_k \Delta_{k,i}}, \quad (5.6)$$

donde β_k es un parámetro positivo, $\Delta_k \in \mathbb{R}^n$ es un vector generado uniformemente cuyas componentes son representadas por $\Delta_{k,i}$ y $\epsilon_{k,\pm}$ son términos de ruido. Los parámetros α_k y β_k de las ecuaciones (5.5) y (5.6) se calculan de acuerdo a las siguientes expresiones,

$$\alpha_k = \frac{a}{(k+1+A)^s}, \quad \beta_k = \frac{b}{(k+1)^t}. \quad (5.7)$$

El conjunto de parámetros a , A , b , s y t de las ecuaciones (5.7) son llamados **ganancias** y son sumamente importantes para el desempeño de SPSA: una mala elección provoca resultados no deseados, como una lenta convergencia del algoritmo. Pese a esto, existen dos conjuntos particulares de ganancias recomendadas [20], [18] que han demostrado tener buen rendimiento aplicado a diferentes problemas.

	a	A	b	s	t
Estándar	3,0	0	0,1	0,602	0,101
Asintótico	3,0	0	0,1	1	1/6

Tabla 5.1: Elección usual de ganancias para el algoritmo de SPSA. **Fuente:** Elaboración propia.

Se debe tener en cuenta que estas ganancias son recomendadas, por lo que puede existir otro conjunto que permita un mejor rendimiento del algoritmo. Es por esto que la búsqueda de ganancias es otro problema de optimización más grande y complicado de resolver.

A continuación se mostrarán las definiciones y teoremas que garantizan la convergencia del algoritmo. El detalle de las demostraciones puede ser visto en [20]. El lema 1 muestra las condiciones bajo las cuales el sesgo en el estimador del gradiente $\hat{\mathbf{g}}_k(\cdot)$ de \mathbf{g} va a 0 cuando $k \rightarrow \infty$.

Teorema 6. El no sesgo en el estimador del gradiente Para α_0 , α_1 y α_2 constantes reales positivas y $\Omega \in \{\omega\}$ el espacio muestral que genera la secuencia $\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \dots$, considere todos los $k \geq K$ para algún $K < \infty$ y suponga que para cada k de $\{\Delta_{k,i}\}$ son independientes e idénticamente distribuidos (i.i.d.) y simétricamente distribuidos ($\mathbb{E}(\Delta_{k,i}) = 0$) cerca del 0 con $|\Delta_{k,i}| \leq \alpha_0$ casi seguramente (a.s.) y $E|\Delta_{k,i}^{-1}| \leq \alpha_1$. Para casi todos $\hat{\mathbf{x}}_k$ (en cada $k \geq K$) supongamos que para todo \mathbf{x} en una vecindad abierta de $\hat{\mathbf{x}}_k$ que no es una función de k o ω , $f \in \mathcal{C}^3(\mathbb{R}^n)$ con $|\partial_{x_i} \partial_{x_j} \partial_{x_l} f(\mathbf{x})| \leq \alpha_2$. Entonces, para casi todos $\omega \in \Omega$

$$\mathbf{b}_k(\hat{\mathbf{x}}_k) = \mathcal{O}(\beta_k^2). \quad (5.8)$$

Lema 1. Dada las condiciones del teorema 6 y las siguientes suposiciones:

A1: $\alpha_k, \beta_k > 0 \forall k$; $\alpha_k \rightarrow 0, \beta_k \rightarrow 0$ cuando $k \rightarrow \infty$;

$$\sum_{k=0}^{\infty} \alpha_k = \infty, \quad \sum_{k=0}^{\infty} \left(\frac{\alpha_k}{\beta_k} \right)^2 < \infty \quad (5.9)$$

A2: Para algún $\gamma_0, \gamma_1, \gamma_2 > 0$ y $\forall k$, $\mathbb{E}(\epsilon_{k,\pm}^2) < \gamma_0$, $\mathbb{E}([f(\hat{\mathbf{x}}_k \pm \beta_k \Delta_k)]^2) \leq \gamma_1$, $\mathbb{E}(\Delta_{k,i}^{-2}) \leq \gamma_2$,

A3: $\|\hat{\mathbf{x}}_k\| < \infty$ a.s. $\forall k$,

A4: $\mathbf{x}(t) = \tilde{\mathbf{x}}$ es una solución asintóticamente estable de la ecuación diferencial $d\mathbf{x}(t)/dt = -\mathbf{g}(\mathbf{x})$,

A5: Sea $D(\tilde{\mathbf{x}}) = \left\{ \mathbf{x}_0 : \lim_{t \rightarrow \infty} \mathbf{x}(t|\mathbf{x}_0) = \tilde{\mathbf{x}} \right\}$ donde denota la solución de la ecuación diferencial de A4 basado en la condición inicial $\mathbf{x}(0) = \mathbf{x}_0$ (esto es $D(\tilde{\mathbf{x}})$ es un dominio de atracción). Existe un compacto $S \subseteq D(\tilde{\mathbf{x}})$ tal que $\tilde{\mathbf{x}} \in S$

entonces

$$\lim_{k \rightarrow \infty} \hat{\mathbf{x}}_k = \tilde{\mathbf{x}}. \quad (5.10)$$

De esta forma, se garantiza la convergencia del algoritmo.

5.2.1. Aplicaciones de SPSA en Información Cuántica: SGQT

El algoritmo SPSA ha mostrado¹ tener una gran versatilidad debido a que puede ser aplicada en diferentes áreas, como: óptica adaptativa, modelamiento atmosférico y planetario, economía, imágenes médicas, entre otras. En esta sección nos centraremos en la aplicación de SPSA al problema de tomografía de estados cuánticos.

El método *self-guided quantum tomography* (SGQT) presentado por Christopher Ferrie [13] muestra un proceso iterativo autoguiado basado en SPSA que permite estimar el estado cuántico desconocido (ρ) de un sistema a partir de un estado de partida generado al azar (σ). Para poder realizar una comparación entre los estados debemos fijar una medida de distancia entre ellos $m(\sigma, \rho)$ y la única restricción es que aquella distancia pueda ser estimada a partir experimentos, es por eso que se tiene acceso a

$$f(\sigma) = \langle m(\sigma, \rho) \rangle. \quad (5.11)$$

De esta manera el algoritmo SGQT provee iterativamente nuevos estados σ que van convergiendo a ρ mediante estimaciones de la función $f(\sigma)$. En este contexto SGQT solo necesita dos estados iniciales σ_{\pm} y estimaciones del experimento $f(\sigma_{\pm})$ para proporcionar una estimación no sesgada del gradiente, el cual nos indicará la dirección hacia el estado que deseamos conocer.

¹Lista de artículos donde ha sido usado SPSA: http://www.jhuapl.edu/SPSA/Pages/References-List_Ref.htm

Suponiendo que ya nos encontramos en la k -ésima iteración, el procedimiento es el siguiente:

1. Generar una dirección aleatoria definida por Δ_k ,
2. Calcular la estimación del gradiente en esa dirección,

$$g_k = \frac{f(\sigma_k + \beta_k \Delta_k) - f(\sigma_k - \beta_k \Delta_k)}{2\beta_k} \Delta_k, \quad (5.12)$$

3. Calcular el punto de partida de la siguiente iteración,

$$\sigma_{k+1} = \sigma_k + \alpha_k g_k, \quad (5.13)$$

donde α_k y β_k están definidas en la eq. (5.7). Así, conforme $k \rightarrow \infty$ entonces $\sigma \rightarrow \rho$. Un ejemplo de medida de distancia es la *infidelidad* entre dos estados [17],

$$m(\sigma, \rho) = 1 - F(\sigma, \rho). \quad (5.14)$$

Para estados puros esta expresión es equivalente a

$$m(\psi, \phi) = 1 - |\langle \psi | \phi \rangle|^2, \quad (5.15)$$

y puede ser estimada midiendo en la base que contiene $|\phi\rangle$, es decir, mediante el conteo del número de resultados en la dirección de $|\phi\rangle$, etiquetada como $n(\phi_+)$, por lo que se puede estimar de acuerdo a

$$m(\psi, \phi) \approx 1 - \frac{n(\phi_+)}{n(\phi_+) + n(\phi_-)}, \quad (5.16)$$

la cual puede variar debido a fluctuaciones estadísticas.

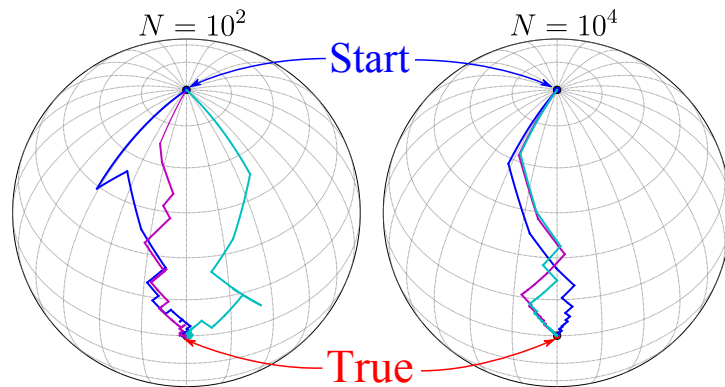


Figura 5.2: Representación de SGQT sobre la esfera de Bloch.. Diagrama de las trayectorias sobre la esfera de Bloch de la estimación de un qubit por medio de SGQT usando la fidelidad como función objetivo. En la figura de la izquierda, cada fidelidad estimada fue generada con $N = 10^2$ experimentos, mientras que a la derecha se utilizaron $N = 10^4$. Cada línea representa una corrida con $k = 10^3$ iteraciones. **Fuente:** C. Ferrie [13].

La **Figura 5.2** muestra las diferentes trayectorias tomadas para los casos en los que $N := n(\phi_+) + n(\phi_-) = 10^2$ y $N = 10^3$. Vemos que pese a lo errático de las curvas en todos los casos las trayectorias terminan en el mismo punto. Vemos, por otro lado, que la variabilidad de los caminos se va mitigando a medida que se aumenta el número de experimentos. Sin embargo, la merma en la variabilidad no disminuye considerablemente a medida que aumentamos la cantidad de experimentos.

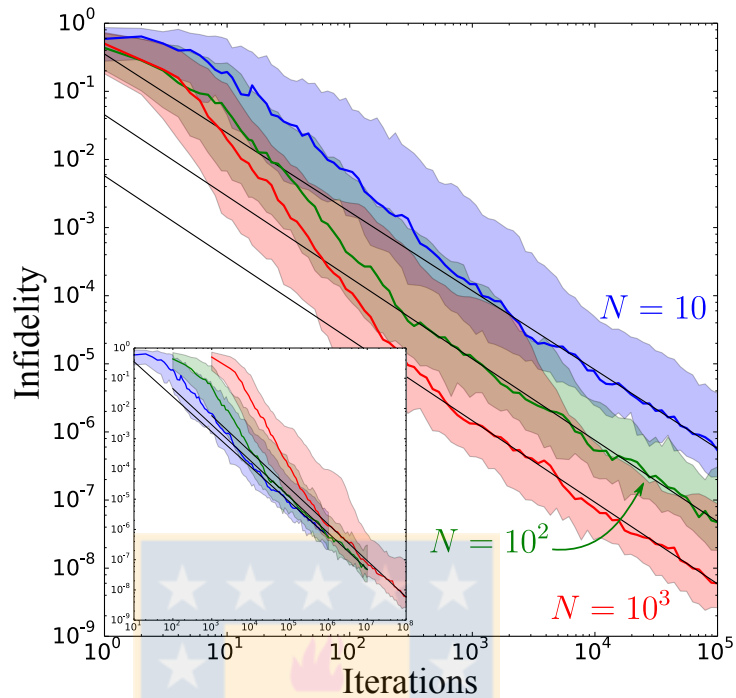


Figura 5.3: Gráfico infidelidad vs número de iteraciones para SGQT. La infidelidad versus número de iteraciones k realizadas por SGQT. Las zonas de colores representan el rango intercuartil de las infidelidades. Las líneas continuas interiores representan la mediana sobre 100 estados generados con distribución uniforme. El cuadro interior muestra el rendimiento como función de $N \cdot k$ (mejorar). Los parámetros de estas simulaciones son $a = 3$, $A = 0$ y $b = 0,1$.

Fuente: C. Ferrie [13].

La **Figura 5.3** muestra la comparación entre la infidelidad y el número de iteraciones conseguidos mediante SGQT. El gráfico es conseguido para 100 estados generados uniformemente. Las líneas interiores muestran la mediana y las sombras el rango intercuartil. Como se observa, a medida que aumenta el número de experimentos las curvas no se vuelven más suaves, pero si aumenta la precisión conforme aumenta N .

Por otro lado Chapman et al. [7] demostraron experimentalmente que el método SGQT es viable para una tomografía real, mediante qubits en polarización. Esto se consiguió optimizando sobre medidas de proyección sin ningún post procesamiento. Este resultado da un punto de partida para implementar SPSA a otros problemas en Información Cuántica, como el estudio de las desigualdades de Bell.

5.3. Complex Simultaneous Perturbation Stochastic Approximation (CSPSA)

El algoritmo CSPSA es la extensión de SPSA a variable compleja y fue desarrollado por A. Utreras-Alarcón et al. en su tesis [21]. El método está basado en el algoritmo de SPSA pero extendido a el espacio complejo \mathbb{C} y hereda todas las características de su versión real, esto es, el paso disminuye iteración a iteración, permite el uso de parámetros desconocidos y utiliza una aproximación del gradiente complejo para la dirección del paso siguiente.

En este nuevo método el problema consiste en buscar un vector $\tilde{\boldsymbol{\mu}} := (\tilde{\boldsymbol{z}}, \tilde{\boldsymbol{z}}^*)$ con $\tilde{\boldsymbol{z}} \in \mathbb{C}^n$, $f(\tilde{\boldsymbol{\mu}}): \text{Dom}(f) \subseteq \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{R}$ tal que satisfaga

$$\frac{\partial f}{\partial \boldsymbol{\mu}}(\tilde{\boldsymbol{\mu}}) = 0. \quad (5.17)$$

Esta expresión es equivalente a

$$\boldsymbol{g}(\tilde{\boldsymbol{\mu}}) = \frac{\partial}{\partial z} f(\tilde{\boldsymbol{\mu}}) = 0. \quad (5.18)$$

dónde ∂_{z^*} es la derivada de Wirtinger [25], la cuál está definida como

$$\partial_z f := \frac{1}{2} (\partial_x f - i \partial_y f), \quad \partial_{z^*} f := \frac{1}{2} (\partial_x f + i \partial_y f) \quad (5.19)$$

con $z = x + iy \in \mathbb{C}$ y $x, y \in \mathbb{R}$.

La secuencia iterativa es análoga a la eq. (5.5) para el caso de SPSA, esto es

$$\hat{\boldsymbol{z}}_{k+1} = \hat{\boldsymbol{z}}_k - \alpha_k \hat{\boldsymbol{g}}_k(\hat{\boldsymbol{\mu}}). \quad (5.20)$$

La estimación del gradiente en la eq. (5.20) es dada por

$$\hat{\boldsymbol{g}}_{k,i} = \frac{f(\hat{\boldsymbol{z}}_{k+}, \hat{\boldsymbol{z}}_{k+}^*) + \varepsilon_{k,+} - f(\hat{\boldsymbol{z}}_{k-}, \hat{\boldsymbol{z}}_{k-}^*) + \varepsilon_{k,-}}{2\beta_k \Delta_{k,i}^*}, \quad (5.21)$$

con $\hat{\boldsymbol{z}}_{k\pm} := \hat{\boldsymbol{z}}_k \pm \beta_k \boldsymbol{\Delta}_k$, $\boldsymbol{\Delta}_k \in \mathbb{C}^n$ es generado al azar, β_k un parámetro positivo y $\varepsilon_{k,\pm}$ son términos de ruido.

Los términos α_k y β_k al igual que en el caso de SPSA describen el tamaño del paso y la perturbación en la función, respectivamente. Estos términos se calculan de la misma forma que la presentada en la eq. (5.7)

$$\alpha_k = \frac{a}{(k+1+A)^s}, \quad \beta_k = \frac{b}{(k+1)^t}. \quad (5.22)$$

Los parámetros a , A , b , s y t también son llamados **ganancias** y su papel es vital para un correcto desempeño de CSPSA. A diferencia de SPSA, CSPSA es un algoritmo desarrollado recientemente y por lo tanto no posee, hasta ahora, un estudio tan riguroso para la búsqueda de ganancias apropiadas ni mucho menos se ha implementado a diferentes problemas de optimización. Debido a esto, no existen ganancias recomendadas para CSPSA por lo que es útil considerar las sugeridas para el algoritmo de SPSA,

	a	A	b	s	t
Estándar	3,0	0	0,1	0,602	0,101
Asintótico	3,0	0	0,1	1	1/6
Estándar Modificado	0,2	0	0,2	0,602	0,101
Asintótico Modificado	1,0	5	0,25	1	1/6

Tabla 5.2: Elección de ganancias para CSPSA. Se incluyen las versiones modificadas de las ganancias usuales. Estas mostraron, durante la investigación, tener un mejor rendimiento en las simulaciones. **Fuente:** Elaboración propia.

Es necesario destacar que, dependiendo del problema, las ganancias deben ser elegidas de acuerdo a la función objetivo que se desea optimizar por lo que, análogamente al algoritmo de SPSA, la búsqueda de ganancias para CSPAes otro problema de optimización mas complicado. Durante esta investigación se modificaron las ganancias anteriores con el objetivo de mejorar el algoritmo CSPA. En el capítulo siguiente se mostrarán simulaciones realizadas con estas ganancias.

A continuación se mostrarán las definiciones y teoremas que garantizan la convergencia del algoritmo, cuyas demostraciones puede ser encontradas en la tesis de Anibal Utreras-Alarcón «*An enhancement to self-guided quantum tomography*».

Teorema 7. El no-sesgo en el estimador del gradiente Sean α_0 , α_1 y α_2 constantes reales positivas y $\Omega \in \{\omega\}$ el espacio maestro que genera la secuencia $\hat{z}_1, \hat{z}_2, \dots$. Para algún $K < \infty$ considérese todos los enteros $k \geq K$ y supóngase que para para cada k de $\{\Delta_{k,i}\}$ son independientes e idénticamente distribuidos (i.i.d.), y simétricamente distribuidos ($\mathbb{E}(\Delta_{k,i}) = 0$) cerca del 0 con $|\Delta_{k,i}| \leq \alpha_0$ casi seguramente (a.s.) y $\mathbb{E}(|\Delta_{k,i}^{-1}|) \leq \alpha_1$ y $\mathbb{E}(e^{2i\phi_{k,i}})$, $\phi_{k,i}$ donde es la fase de $\Delta_{k,i}$ en su descomposición polar. Para todo $k \geq K$, para casi todo $\hat{\mu}_k$ supóngase que para todo

$\boldsymbol{\mu}$, $f \in \mathcal{C}^2(\mathbb{C}^n \times \mathbb{C}^n)$ con $|\partial_{\mu_i} \partial_{\mu_j} f(\boldsymbol{\mu})| \leq \alpha_2$. Entonces para casi todo $\omega \in \Omega$

$$\mathbf{b}_k(\hat{\boldsymbol{\mu}}_k) = \mathcal{O}(\beta_k^2). \quad (5.23)$$

Lema 2. Dada las condiciones del teorema 7 y las siguientes suposiciones:

B1: $\alpha_k, \beta_k > 0 \forall k$;

$$\lim_{k \rightarrow \infty} (|a_k| + |c_k|) = 0; \quad \sum_{k=0}^{\infty} \alpha_k = \infty; \quad \sum_{k=0}^{\infty} \left(\frac{\alpha_k}{\beta_k} \right)^2 < \infty, \quad (5.24)$$

B2: Para algún $\delta_0, \delta_1, \delta_2 > 0$ y $\forall k$;

$$\mathbb{E}(\epsilon_{k,\pm}^2) < \delta_0; \quad \mathbb{E}([f(\hat{\mathbf{z}}_{k+}, \mathbf{z}_{k+}^*)]^2) \leq \delta_1; \quad \mathbb{E}(\Delta_{k,i}^{-2}) \leq \delta_2, \quad (5.25)$$

B3: $\|\hat{\mathbf{z}}_k\| < \infty$ a.s. $\forall k$,

B4: $\mathbf{z}(t) = \tilde{\mathbf{z}}$ es una solución asintóticamente estable de la ecuación diferencial $d\mathbf{z}(t)/dt = -\mathbf{g}(\boldsymbol{\mu})$,

B5: Sea $D(\tilde{\mathbf{z}}) = \left\{ \mathbf{z}_0 : \lim_{t \rightarrow \infty} \mathbf{z}(t|\mathbf{z}_0) = \tilde{\mathbf{z}} \right\}$ donde denota la solución de la ecuación diferencial de B4 basado en la condición inicial $\mathbf{z}(0) = \mathbf{z}_0$ (esto es $D(\tilde{\mathbf{z}})$ es un dominio de atracción). Existe un compacto $S \subseteq D(\tilde{\mathbf{z}})$ tal que $\hat{\mathbf{z}}_k \in S$ a menudo infinitesimalmente para cada puntos de la muestra.

entonces las siguientes afirmaciones son ciertas

S1: $\|\mathbf{b}_k(\hat{\boldsymbol{\mu}}_k)\| \leq \infty$ y $\lim_{k \rightarrow \infty} \|\mathbf{b}_k(\hat{\boldsymbol{\mu}}_k)\| = 0$ casi seguramente,

S2: $\lim_{k \rightarrow \infty} P(\sup_{m \geq k} \|\sum_{i=k}^m a_i e_i(\hat{\boldsymbol{\mu}}_i)\| \geq \eta) = 0$ para todo $\eta > 0$.

Teorema 8. Si las hipótesis del lema 2 se cumplen, entonces

$$\lim_{k \rightarrow \infty} \hat{\mathbf{z}}_k = \tilde{\mathbf{z}}. \quad (5.26)$$

De esta forma, se garantiza la convergencia del algoritmo.

5.3.1. Aplicaciones de CSPA en Información Cuántica: Tomografía

La primera aplicación de CSPA ha sido para la tomografía de estados puros. En el artículo de A. Utreras-Alarcón et. al. [22] se realiza una comparación entre el

rendimiento de SPSA y CSPSA aplicado la tomografía de estados cuánticos de un qubit.

Para este caso se considera un estado cuántico puro desconocido $|\psi\rangle$. Éste estado puede ser completamente determinado mediante la minimización de la infidelidad, definida como

$$I(\mathbf{z}) = 1 - |\langle\psi|\phi(\mathbf{z})\rangle| \quad (5.27)$$

con respecto a la variable $z_i \in \mathbb{C}$ que define el estado cuántico puro conocido

$$|\phi(\mathbf{z})\rangle = \sum_i \frac{z_i}{\sqrt{\sum_i |z_i|^2}} |i\rangle. \quad (5.28)$$

Los coeficientes complejos del estado $|\psi\rangle$ que forman parte de $I(\mathbf{z})$ son considerados desconocidos pero fijos. Así, el mínimo global $I = 0$ se consigue cuando $|\psi\rangle = |\phi\rangle$, para cualquier $|\psi\rangle$.

La optimización de la infidelidad puede ser realizada también sobre los números reales. En dicho caso las componentes de \mathbf{z} son mapeadas sobre los números reales mediante ángulos escritos en coordenadas hiperesféricas y argumentos de fases complejas. De esta forma la infidelidad puede ser escrita en función de un vector real $I(\mathbf{x})$ y, con esto, es posible aplicar el algoritmo de optimización SPSA a la función objetivo.

La **Figura 5.4** muestra la infidelidad promedio versus el número de iteraciones de los algoritmos de CSPSA (líneas rojas continuas) y SPSA (líneas discontinuas azules). Desde arriba a abajo las líneas rojas (azules) representan el número de experimentos $N = 10, 10^2, 10^3$ y 10^4 . Como podemos ver el rendimiento de CSPSA es superior al de SPSA, debido a varios factores: en primer lugar, para la misma cantidad de iteraciones y experimentos en los dos métodos CSPSA consigue siempre una mayor precisión, eso se puede observar tanto en el gráfico exterior (promedio) como en el interior (mediana) las curvas rojas están por debajo de las azules, en segundo lugar las curvas del método complejo son más estrechas, dando lugar a una menor varianza (y mediana) con respecto al método real y en tercer lugar, la velocidad de convergencia CSPSA es mucho mayor a la de SPSA, esto se puede apreciar por simple inspección al observar las pendientes de las curvas en el gráfico exterior.

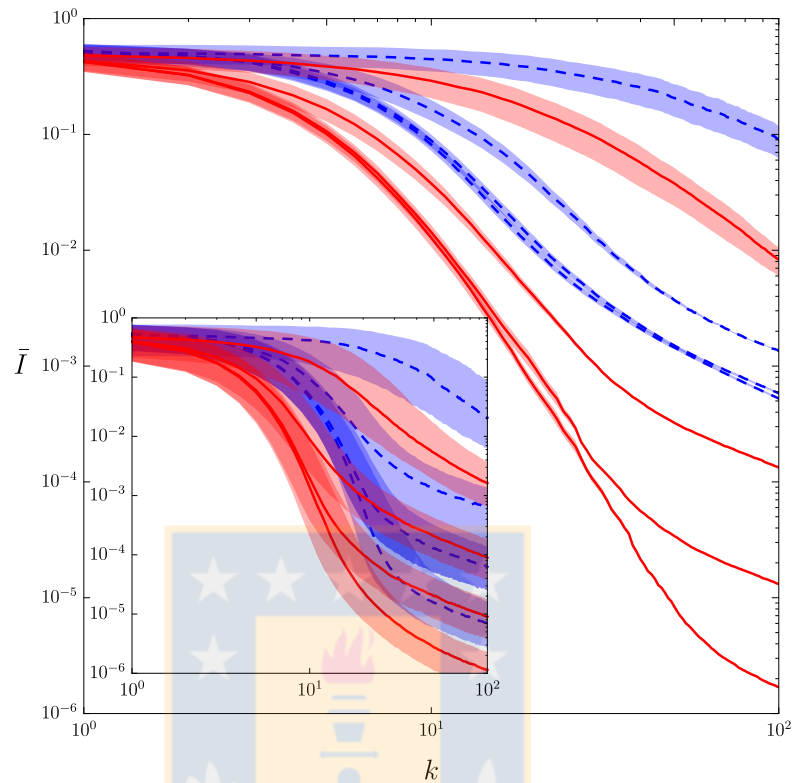
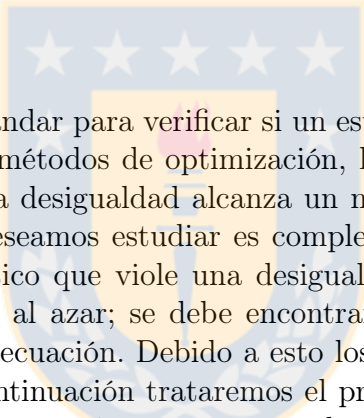


Figura 5.4: Gráfico de infidelidad vs número de iteraciones para tomografía vía CSPA. La figura muestra el valor promedio de infidelidad versus el número de iteraciones sobre 10^4 estados cuánticos desconocidos y guesses iniciales, para el proceso de tomografía cuántica de un qubit mediante CSPA (líneas rojas) y SPSA (líneas segmentadas azules). El sombreado indica la varianza en torno al promedio. En el gráfico interior se muestra la mediana y sus respectivos rangos intercuartiles. Desde arriba a abajo las líneas rojas (azules) representan $N = 10, 10^2, 10^3$ y 10^4 . Para CSPA se usaron las ganancias asíntóticas y para SPSA las estándar.

Fuente: A. Utreras-Alarcón et. al. [22].

Capítulo 6

Desigualdad CHSH vía optimización estocástica



El procedimiento estándar para verificar si un estado cuántico viola una desigualdad del Bell es utilizar métodos de optimización, los cuales realizan una búsqueda de las bases donde dicha desigualdad alcanza un máximo. Pero ¿cómo procedemos cuando el estado que deseamos estudiar es completamente desconocido? Como sabemos, un estado cuántico que viole una desigualdad de Bell no lo hace para un conjunto de observables al azar; se debe encontrar la base precisa para la cual el estado no satisfaga la inecuación. Debido a esto los procedimientos usuales no pueden ser utilizados. A continuación trataremos el problema de estimar el valor de la desigualdad de CHSH para un sistema cuyo estado cuántico es desconocido mediante el uso de los algoritmos presentados en el capítulo 5.

6.1. Introducción

Antes de comenzar con la descripción del procedimiento que permite obtener la máxima violación de Bell para estados desconocidos, es necesario aclarar algunos conceptos que se utilizarán con frecuencia en el presente capítulo.

En primer lugar, llamaremos al conjunto los parámetros definidos en la [Tabla 5.1](#) como **ganancias**. Dependiendo del caso, se usarán los términos *ganancias estándar*, *ganancias asintóticas*, *ganancias estándar modificadas* y *ganancias asintóticas modificadas* para referirnos al conjunto completo de parámetros definidos en la [Tabla 5.2](#). Por otro lado, llamaremos **guess** a los vectores que inician la secuencia iterativa, donde cada guess produce una *realización independiente* del cálculo numérico. De esta forma, el estudio de las simulaciones se realizará para una gran cantidad de

realizaciones independientes.

A lo largo de éste capítulo, se utilizarán vectores o estados cuánticos generados al azar en las simulaciones. Es necesario hacer la distinción entre vectores o estados cuánticos generados al azar o generados al azar uniformemente.

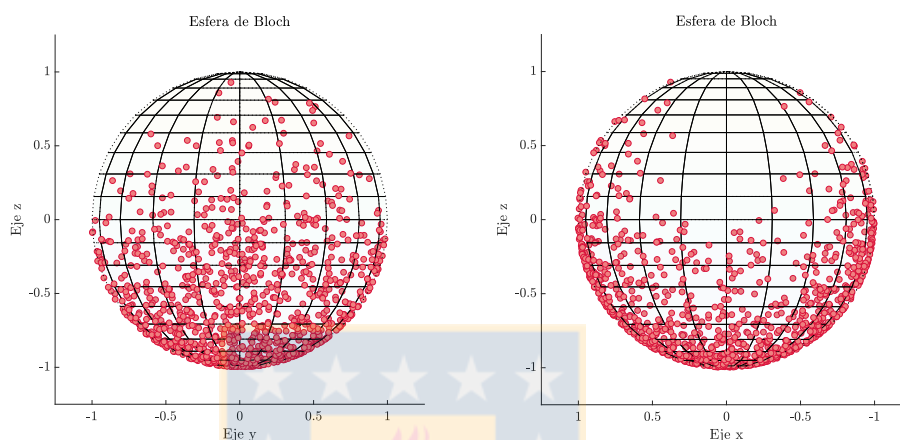


Figura 6.1: Representación de estados generados al azar sobre la esfera de Bloch. Los diagramas muestran cortes transversales de una esfera de Bloch, donde la figura de la izquierda tiene su eje x apuntando ortogonal saliendo del papel, mientras que en la esfera derecha el eje y apunta hacia adentro. Los puntos de color rojo representan 10^3 estados puros $|\psi\rangle = a|e\rangle + b|e^\perp\rangle$ generados al azar sobre la esfera. **Fuente:** Elaboración propia

La **Figura 6.1** muestra dos perspectivas de la misma esfera de Bloch donde fueron generados al azar 10^3 estados cuánticos puros $|\psi\rangle = a|e\rangle + b|e^\perp\rangle$. Como se puede observar los estados no están esparcidos de manera homogénea sobre la esfera, esto muestra gráficamente que los estados generados tienen un **sesgo muestral** hacia el polo sur de la esfera. De esta manera los datos de las simulaciones no serían completamente representativos.

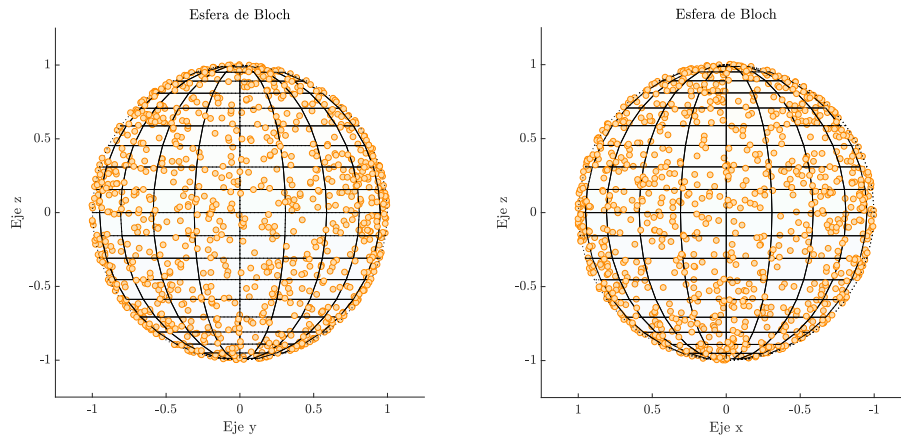


Figura 6.2: Representación de estados generados uniformemente sobre la esfera de Bloch. Los diagramas muestran cortes transversales de una esfera de Bloch, donde la figura de la izquierda tiene su eje x apuntando ortogonal saliendo del papel, mientras que en la esfera derecha el eje y apunta hacia adentro. Los puntos de color naranja representan 10^3 estados puros $|\psi\rangle = a|e\rangle + b|e^\perp\rangle$ generados con distribución uniforme sobre la esfera. **Fuente:** Imagen generada por el autor. **Fuente:** Elaboración propia

Por otro lado la [Figura 6.2](#) muestra dos perspectivas de la misma esfera de Bloch donde se generaron uniformemente 10^3 estados cuánticos puros $|\psi\rangle = a|e\rangle + b|e^\perp\rangle$. Como se puede observar los estados están distribuidos de manera completamente homogénea sobre la esfera, esto muestra gráficamente que los estados utilizados sobre las simulaciones son producidos uniformemente.

Es imprescindible definir, en este contexto, que entendemos por *estados cuánticos desconocidos*. En este escenario los estados cuánticos son generados en el ordenador, por lo que nosotros podríamos tener acceso total a sus propiedades en cualquier momento, así el estado cuántico es *conocido para el programador*. Por otra parte, los algoritmos son quienes realizan la estimación de la desigualdad de Bell *sin conocer el estado ni sus propiedades*, es más, los algoritmos no utilizan las propiedades del estado cuántico (autovalores y autovectores) para estimar \mathcal{S} . Así, el estado es *desconocido para los algoritmos de optimización*. Finalmente, ya que el programador tiene acceso al estado, se calcula de manera exacta el valor de la máxima violación de Bell \mathcal{S}_{teo} para el estado en cuestión y dicho valor se utilizará para cuantificar el rendimiento de los algoritmos. Estas consideraciones fueron mostradas por C. Caves et al. [6].

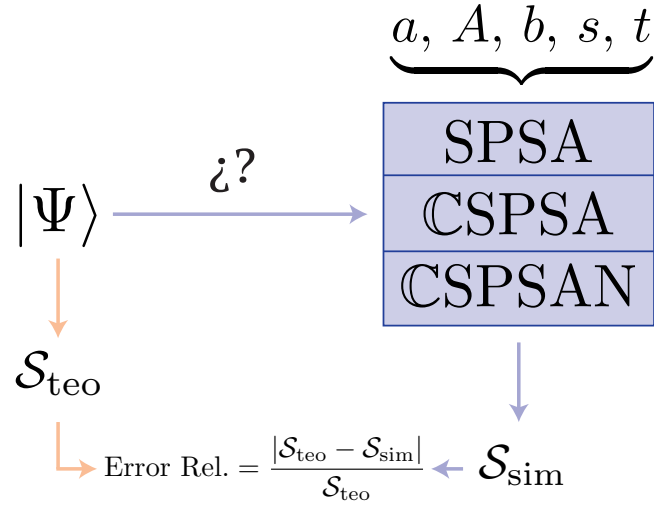


Figura 6.3: Esquema estados cuánticos desconocidos y algoritmos de optimización. La rama inferior representa la obtención del valor exacto de la desigualdad \mathcal{S}_{teo} . La rama derecha muestra el ingreso de un estado desconocido para los algoritmos SPSA, CSPSA y CSPSAN con ganancias a, A, b, s, t fijas para realizaciones independientes. \mathcal{S}_{sim} representa el valor de la desigualdad obtenido mediante simulaciones. El valor teórico es utilizado para cuantificar el rendimiento de los algoritmos. **Fuente:** Elaboración propia

6.2. Desigualdad CHSH vía SPSA

De acuerdo al esquema presentado por J. Clauser, M. Horne, A. Shimony y R. Holt, tenemos un sistema bipartito compuesto por Alice y Bob donde cada uno posee dos operadores locales de medición, etiquetados como A_1 y A_2 para Alice, y B_1 y B_2 para Bob. Debido a que cada operador tiene autovalores ± 1 , podemos utilizar la descomposición de observables en término de las matrices de Pauli, es decir

$$\mathcal{O} = \vec{r} \cdot \vec{\sigma}, \quad (6.1)$$

dónde $\vec{r} := \vec{r}(\theta, \phi) = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$ es un vector tridimensional parametrizado sobre la superficie de una esfera de radio 1 y $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ es el vector cuyas componentes son las matrices de Pauli. De esta forma, el operador $\mathcal{O} := (\theta, \phi)$ depende de dos parámetros reales definidos en $\theta \in [0, \pi[$ y $\phi \in [0, 2\pi[$.

De acuerdo a lo anterior, la forma explícita del operador \mathcal{O} es

$$\mathcal{O} := \mathcal{O}(\theta, \phi) = \begin{pmatrix} \cos \theta & \sin \theta \cos \phi - i \sin \theta \sin \phi \\ \sin \theta \cos \phi + i \sin \theta \sin \phi & -\cos \theta \end{pmatrix}. \quad (6.2)$$

Para la implementación del algoritmo SPSA es necesario generar, con distribución uniforme, el vector de partida que llamaremos **guess**

$$\hat{\Theta}_0 := (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4, \hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3, \hat{\phi}_4) \quad (6.3)$$

cuyas componentes están restringidas a las condiciones señaladas anteriormente para iniciar la secuencia iterativa. Así, podemos definir los operadores de medición que nos permitirán construir la función objetivo f de la eq. (5.5),

$$\begin{aligned} \hat{\mathbf{A}}_1 &:= \hat{\mathbf{A}}_1(\hat{\theta}_1, \hat{\phi}_1), & \hat{\mathbf{A}}_2 &:= \hat{\mathbf{A}}_2(\hat{\theta}_2, \hat{\phi}_2), \\ \hat{\mathbf{B}}_1 &:= \hat{\mathbf{B}}_1(\hat{\theta}_3, \hat{\phi}_3), & \hat{\mathbf{B}}_2 &:= \hat{\mathbf{B}}_2(\hat{\theta}_4, \hat{\phi}_4). \end{aligned} \quad (6.4)$$

donde $\hat{\mathbf{A}}_i, \hat{\mathbf{B}}_i$ tienen la forma presentada en la eq. (6.2). Definimos la función objetivo como,

$$S(\hat{\Theta}) := \left| \hat{\mathbf{E}}_{11} + \hat{\mathbf{E}}_{12} + \hat{\mathbf{E}}_{21} - \hat{\mathbf{E}}_{22} \right|, \quad (6.5)$$

dónde $\hat{\mathbf{E}}_{ij} = \text{Tr} \{ \rho(\hat{\mathbf{A}}_i \otimes \hat{\mathbf{B}}_j) \}$.

Entonces re definimos la secuencia iterativa mostrada en la eq. (5.5) para encontrar el máximo de la función S en la eq. (6.5)

$$\hat{\Theta}_{k+1} := \hat{\Theta}_k + \alpha_k \hat{\mathbf{g}}_k(\hat{\Theta}_k), \quad (6.6)$$

obteniendo la estimación del gradiente $\hat{\mathbf{g}}_k$ de acuerdo a,

$$\hat{g}_{k,i}(\hat{\Theta}) := \frac{S(\hat{\Theta}_k + \beta_k \Delta_k) + \epsilon_{k,+} - S(\hat{\Theta}_k - \beta_k \Delta_k) - \epsilon_{k,-}}{2\beta_k \Delta_{k,i}}, \quad (6.7)$$

donde Δ_k es un vector de ocho componentes que es generado uniformemente y cada una de ellas puede tomar valores $+1$ o -1 , y $\epsilon_{k,\pm}$ son términos de ruido. Las ganancias α_k y β_k se obtienen de acuerdo a la ecuación eq. (5.7). Los parámetros a, A, b, s y t usados son para el caso estándar y asintótico mostrados en la table 5.1. Así, para $k \rightarrow \infty$ se consigue el máximo de S para el estado ρ .

El pseudo código que permite realizar este proceso es mostrado en el cuadro de algoritmo 1

Algorithm 1 Algoritmo SPSA y función objetivo CHSH definida de acuerdo a la parametrización mostrada en la eq. (6.2).

Require: D: Dimensión del espacio. N: Número de iteraciones. Θ : Ángulos iniciales definidos en la eq. (6.3). ρ : Estado cuántico. a, b, s, t, A : parámetros de ajuste. En general se utilizan los asintóticos o estándar.

Ensure: Valor del parámetro S en cada iteración.

```

1: procedure SPSA(D,N, $\Theta,\rho,a,b,s,t,A$ )
2:   for  $k \leftarrow 1, K$  do
3:      $\alpha_k \leftarrow \frac{a}{(k+1+A)^s}$ 
4:      $\beta_k \leftarrow \frac{b}{(k+1)^t}$ 
5:      $\Delta_k \leftarrow$  vector de dimensión N generado uniformemente entre los valores
       $\pm 1$ 
6:      $\Theta_+ \leftarrow \Theta + \beta_k \Delta_k$ 
7:      $\Theta_- \leftarrow \Theta - \beta_k \Delta_k$ 
8:      $\text{CHSH}_+ \leftarrow \text{CHSH}(\Theta_+, \rho)$ 
9:      $\text{CHSH}_- \leftarrow \text{CHSH}(\Theta_-, \rho)$ 
10:     $g_k \leftarrow \frac{\text{CHSH}_+ - \text{CHSH}_-}{2\beta_k} \Delta_k$ 
11:     $\Theta \leftarrow \Theta + \alpha_k g_k$ 
12:    Out  $\leftarrow \text{CHSH}(\Theta, \rho)$ 
13:  end for
14:  return Out
15: end procedure

16: function CHSH( $\Theta, \rho$ )

17:   $\sigma_x \leftarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y \leftarrow \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z \leftarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ 
18:   $A_1 \leftarrow \sin(\Theta_1) \cos(\Theta_5) \sigma_x + \sin(\Theta_1) \sin(\Theta_5) \sigma_y + \cos(\Theta_1) \sigma_z$ 
19:   $A_2 \leftarrow \sin(\Theta_2) \cos(\Theta_6) \sigma_x + \sin(\Theta_2) \sin(\Theta_6) \sigma_y + \cos(\Theta_2) \sigma_z$ 
20:   $B_1 \leftarrow \sin(\Theta_3) \cos(\Theta_7) \sigma_x + \sin(\Theta_3) \sin(\Theta_7) \sigma_y + \cos(\Theta_3) \sigma_z$ 
21:   $B_2 \leftarrow \sin(\Theta_4) \cos(\Theta_8) \sigma_x + \sin(\Theta_4) \sin(\Theta_8) \sigma_y + \cos(\Theta_4) \sigma_z$ 

22:   $E_{11} \leftarrow \text{Tr}[(A_1 \otimes B_1)\rho], E_{12} \leftarrow \text{Tr}[(A_1 \otimes B_2)\rho]$ 
23:   $E_{21} \leftarrow \text{Tr}[(A_2 \otimes B_1)\rho], E_{22} \leftarrow \text{Tr}[(A_2 \otimes B_2)\rho]$ 
24:   $S \leftarrow |E_{11} + E_{21} + E_{12} - E_{22}|$ 
25:  return  $S$ 
26: end function

```

Como primer test de rendimiento del algoritmo de SPAS aplicado a la desigualdad CHSH, utilizaremos como estado de prueba el estado (3.10) mostrado en el capítulo 3,

$$|\phi_+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle). \quad (6.8)$$

Este estado es maximalmente entrelazado, es decir, posee el máximo grado de entrelazamiento entre dos qubits. Como sabemos, y de acuerdo al teorema de Cirel'son, la máxima violación de Bell que alcanzan estos estados es $2\sqrt{2}$ el cual es un límite que ningún estado puede superar.

Consideremos cuatro guesses $\hat{\Theta}_0$ generados de manera independiente y al azar, los cuales nos proporcionarán los puntos de partida para iniciar la secuencia iterativa. Veremos el comportamiento del algoritmo de SPSA para las cuatro ganancias mostradas en la [Tabla 5.2](#)



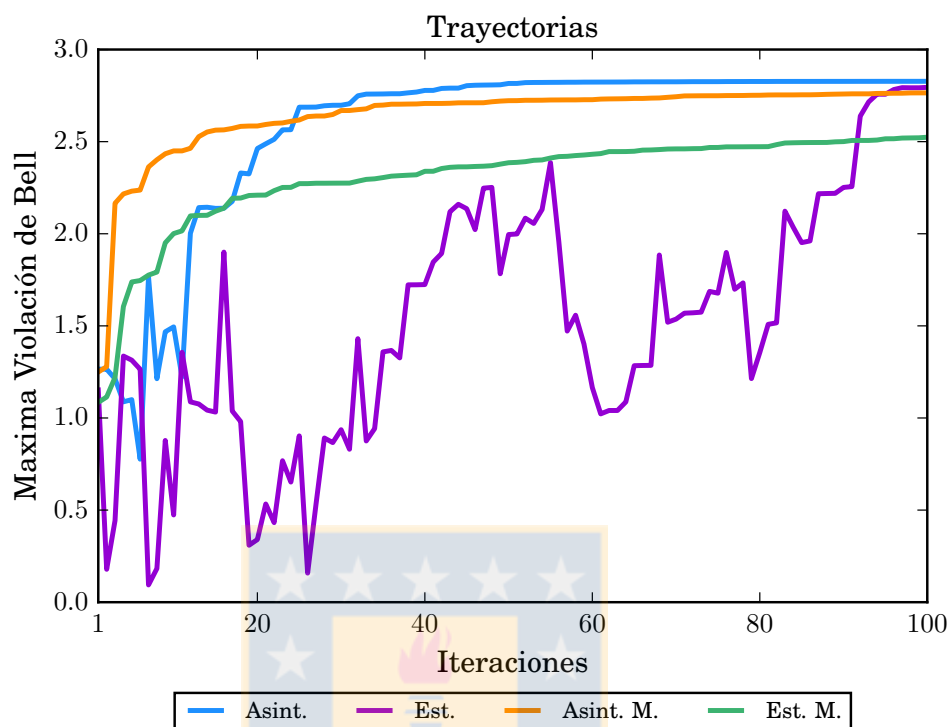


Figura 6.4: Trayectorias de SPSA para diferentes ganancias. El gráfico muestra la máxima violación de bell versus el número de iteraciones del algoritmo de SPSA. Cada trayectoria representa el camino que sigue cada simulación realizado con ganancias diferentes. La nomenclatura corresponde a Asint: Asintóticos, Est: Estandar, Asint. M.: Asintótico Modificado, Est. M.: Estandar Modificado. **Fuente:** Elaboración propia

En la **Figura 6.4** vemos las trayectorias producidas por SPSA para la elección de parámetros estándar, asintótico y sus modificaciones. En cada caso las curvas son caóticas al inicio, lo cual se debe a la naturaleza estocástica del algoritmo. A medida que aumenta el número de iteraciones la variabilidad comienza a disminuir haciendo más estables las curvas, las cuales convergen asintóticamente a la cota $2\sqrt{2}$. La característica más distintiva de este método es que, sin importar cual sea el punto de partida, el algoritmo se encarga de encontrar el máximo de la función objetivo, acercándose al punto donde alcanza el máximo de manera asintótica. Notamos que, de acuerdo a lo mencionado en el capítulo anterior, las ganancias juegan un rol importante en el rendimiento del algoritmo. Vemos que las ganancias estándar no son una buena elección si es comparada con las otras tres.

Ya que el método es estocástico no es posible utilizar solo una realización para describir completamente el rendimiento del algoritmo y lo más apropiado es realizar un conjunto grande de simulaciones con las cuales se obtendrá el rendimiento en

mediana de los algoritmos.

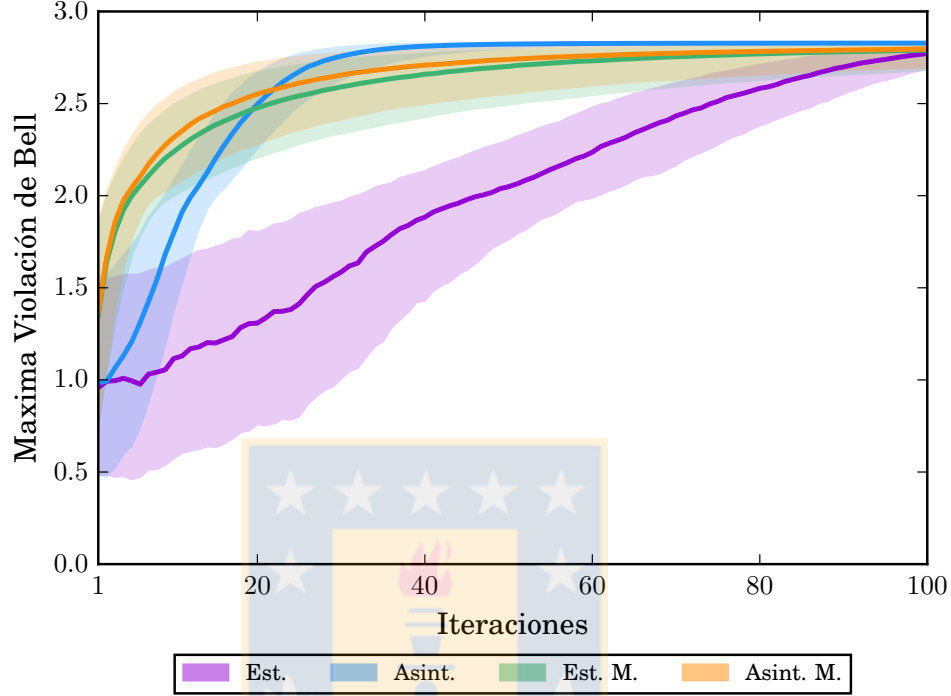


Figura 6.5: Simulaciones de SPSA para un estado maximalmente entrelazado para 5×10^3 realizaciones. La figura muestra la máxima violación de Bell versus el número de iteraciones para el estado (6.8) para 5×10^3 realizaciones. Las líneas continuas representan la mediana y las sombras representan el rango intercuartil. La nomenclatura corresponde a Asint: Asintótico, Est: Estandar, Asint. M.: Asintótico Modificado, Est. M.: Estandar Modificado. **Fuente:** Elaboración propia

La figura muestra 5×10^3 realizaciones producidas con gueses diferentes. Con cada color se muestra la mediana y el rango intercuartil correspondiente a cada conjunto de ganancias. Vemos que dependiendo de lo que deseemos estudiar debemos elegir el conjunto de ganancias apropiado. Por ejemplo si queremos determinar si un estado es entrelazado o no, es decir, si el estado supera el valor 2.0 en el eje de las ordenadas, entonces es recomendable elegir las ganancias estándar y asintóticas modificadas, mientras que si buscamos la mayor velocidad de convergencia hacia el máximo debemos elegir el conjunto de ganancias asintóticas.

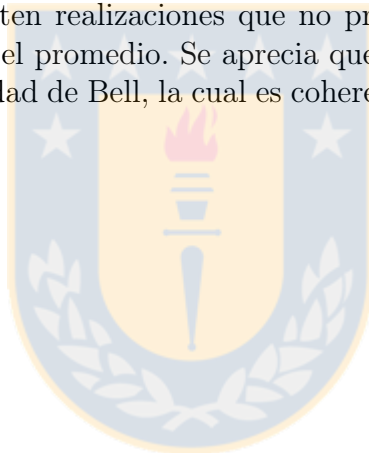
El segundo ejemplo que muestra el rendimiento de SPSA es considerar los estados de Werner presentados en la eq. (3.38) mostrado en el capítulo 3,

$$\rho_W = \alpha |\phi_+\rangle\langle\phi_+| + \frac{1-\alpha}{4} \mathbf{1} \otimes \mathbf{1}, \quad -\frac{1}{3} \leq \alpha \leq 1. \quad (6.9)$$

Como vimos, dependiendo del valor del parámetro α , los estados de Werner describen diferentes sistemas cuánticos. Esto queda resumido en la tabla 3.1 presentada en el capítulo 3. A partir de la eq. (3.38) podemos encontrar el valor de la máxima violación de Bell en función del parámetro α ,

$$S_{\text{Werner}}(\alpha) = 2\sqrt{2}\alpha. \quad (6.10)$$

La Figura 6.6 muestra la implementación de SPSA para evaluar si el conjunto de estados de Werner determinados por el parámetro α viola la desigualdad de Bell. Para esto se produjeron 10 estados variando el parámetro λ entre 0 y 1. Se puede ver que la mediana de simulaciones realizadas con las ganancias estándar y asintótica reproducen completamente la curva teórica. Por otra parte, vemos que con los promedios las mismas ganancias no se obtienen buenos resultados para $\alpha \leq 0,4$, esto se debe a que existen realizaciones que no proporcionan buenos resultados y, por lo tanto, empeoran el promedio. Se aprecia que a partir de $\alpha = 0,7$ se exhibe la violación en la desigualdad de Bell, la cual es coherente con lo presentado en la tabla 3.1.



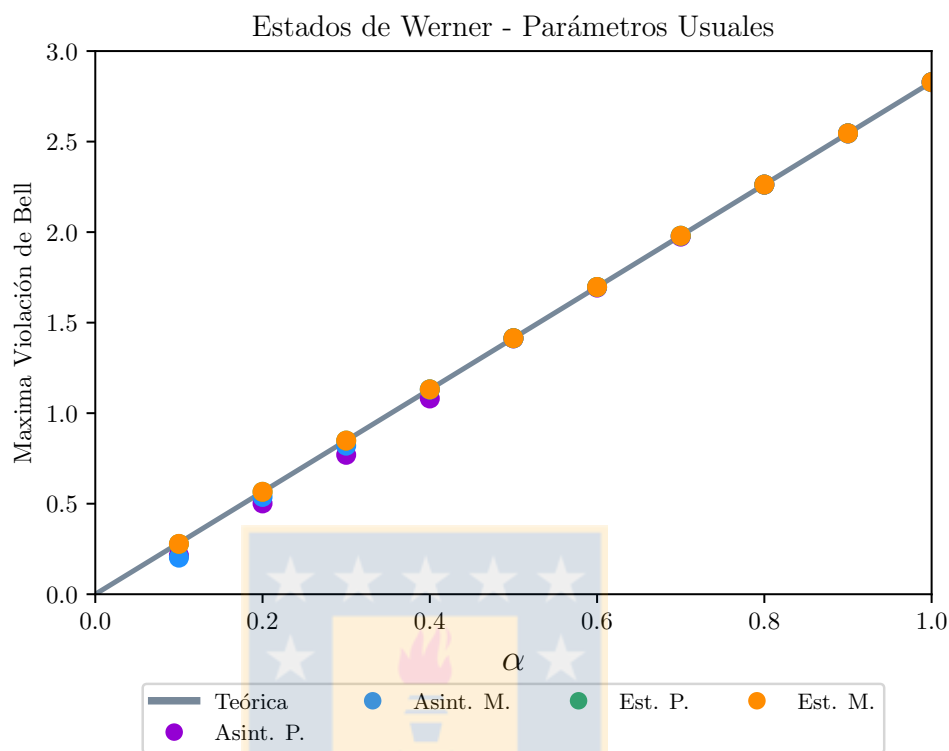


Figura 6.6: Gráfica de los estados de Werner mediante SPSA para las ganancias usuales. La figura muestra la máxima violación de Bell vs α , donde α define los distintos estados de Werner de acuerdo a la eq. (6.9). Cada punto representa el promedio o mediana del máximo valor conseguido a las 200 iteraciones, donde se realizaron 10 simulaciones con guesses distintos.

Nomenclatura: Asint. M.: Mediana Asintótico - Est. P.: Promedio Estándar - Est. Med. Mediana Estándar - Asint. P.: Promedio Asintótico. **Fuente:** Elaboración propia

La **Figura 6.7** muestra las simulaciones realizadas con SPSA pero usando las ganancias asintóticas y estándar modificadas, para 10 realizaciones y 200 iteraciones en cada punto. Como vemos, ninguno reproduce ni total ni parcialmente la curva teórica. De esta forma, para esta prueba, las ganancias modificadas no tienen un buen rendimiento.

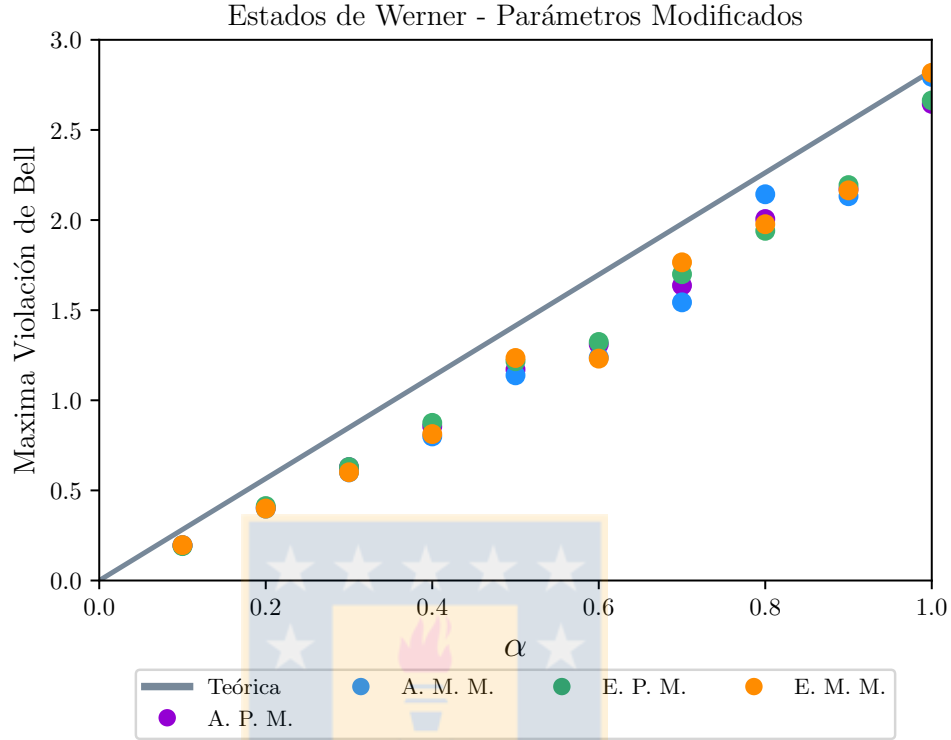


Figura 6.7: Gráfica de los estados de Werner mediante SPSA para las ganancias modificadas. La figura muestra la máxima violación de Bell vs α , donde α define los distintos estados de Werner de acuerdo a la eq. (6.9). Cada punto representa el promedio o mediana del máximo valor conseguido a las 200 iteraciones, donde se realizaron 10 simulaciones con guesses distintos.

Nomenclatura: A. M. M.: Mediana Asintótico - E. P. M.: Promedio Estándar - E. M. M.: Mediana Estándar - A. P. M.: Promedio Estándar. **Fuente:** Elaboración propia

Como último ejemplo, estudiaremos el caso de un estado cuántico descrito en su descomposición de Schmidt,

$$|\Psi\rangle = \sqrt{\lambda}|01\rangle + \sqrt{1-\lambda}|10\rangle. \quad (6.11)$$

A partir del resultado obtenido por N. Gisin y A. Peres [14] mostrado en la eq. (4.52) y considerando $\gamma = 0$, obtenemos que la máxima violación de Bell del estado eq. (6.11) en función de λ es dado por

$$S_{\text{Schmidt}}(\lambda) = 2\sqrt{1 + 4\lambda(1 - \lambda)} \quad (6.12)$$

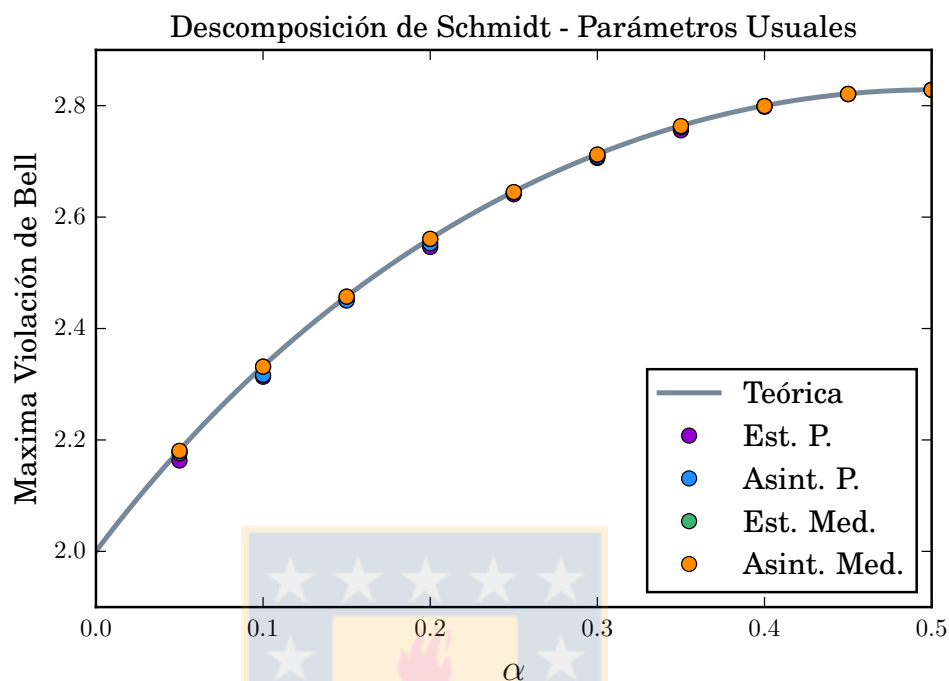


Figura 6.8: Estado en descomposición de Schmidt simulado con SPSA con las ganancias usuales. La figura muestra la máxima violación de Bell vs λ , donde λ describe diferentes estados en descomposición de Schmidt. Cada punto representa el promedio o mediana del máximo valor conseguido a las 200 iteraciones, donde se realizaron 10 simulaciones con guesses distintos. Nomenclatura: Est. P.: Estándar promediado, Asint. P. Asintótico promediado, Est. Med. Mediana Estándar, Asint. Med.: Mediana Asintótico. **Fuente:** Imagen generada por el autor.

La **Figura 6.8** muestra una comparación entre las predicciones hechas por la teoría sobre la máxima violación de Bell permitida en función de λ y los resultados de las simulaciones realizadas usando SPSA con las ganancias usuales. Para esto se produjeron 10 estados variando el parámetro α entre 0 y 1/2. El gráfico muestra los promedios y medianas calculados a partir del último valor conseguido para cada punto. Como vemos todas las ganancias usuales reproducen los resultados que predice la teoría, al igual que en el caso de los estados de Werner. Nuevamente vemos algunos puntos que se escapan de la curva, estos justamente fueron calculados con el promedio de las simulaciones.

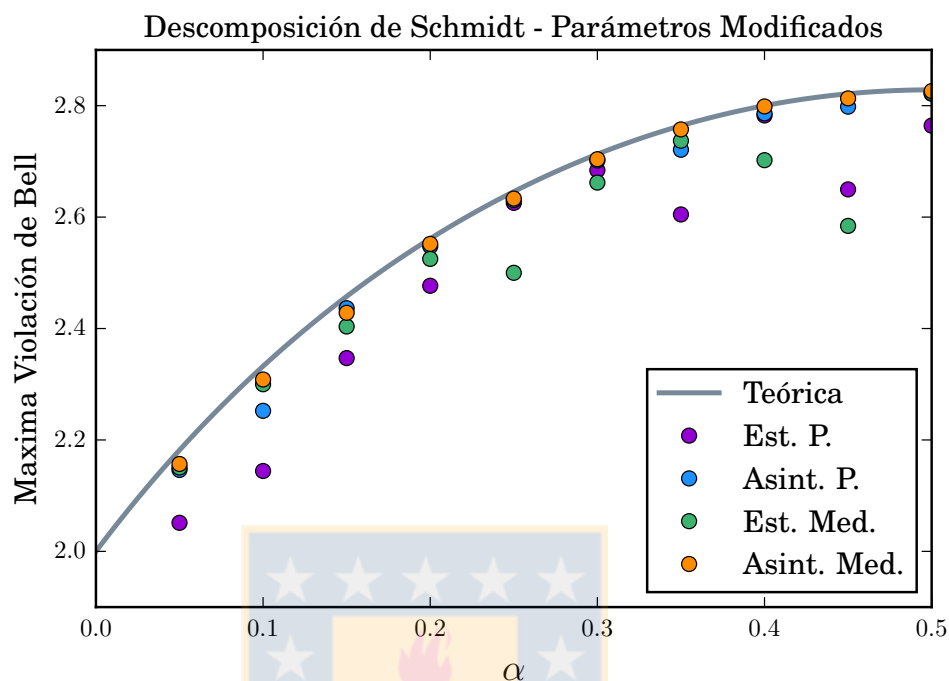


Figura 6.9: Estado en descomposición de Schmidt simulado con SPSA con las ganancias modificadas. La figura muestra la máxima violación de Bell vs λ , donde λ describe diferentes estados en descomposición de Schmidt. Cada punto representa el promedio o mediana del máximo valor conseguido a las 200 iteraciones, donde se realizaron 10 simulaciones con guesses distintos. Nomenclatura: Est. P.: Estándar promediado, Asint. P. Asintótico promediado, Est. Med. Mediana Estándar, Asint. Med.: Mediana Asintótico. **Fuente:** Imagen generada por el autor.

En la [Figura 6.9](#) se muestran los resultados de los cálculos realizados con las ganancias modificadas para los mismos estados, iteraciones y realizaciones de la figura anterior. Vemos que la única ganancia que muestra resultados más cercanos a la predicción teórica es la asintótica modificada, mientras que las otras no tienen un rendimiento apropiado.

De lo expuesto anteriormente vemos que no todas las ganancias son útiles para abordar estos problemas, por lo que se deben hacer pruebas rigurosas antes de elegir el grupo de parámetros que se utilizarán para realizar los cálculos.

6.3. Desigualdad CHSH vía CSPA

Como vimos en la section 5.3, el algoritmo de CSPA permite encontrar un vector $\hat{\boldsymbol{\mu}}_k$ en los complejos, tal que la función objetivo f alcance un mínimo (o máximo) para $k \rightarrow \infty$ en la secuencia iterativa. Siguiendo el procedimiento de de SPSA, debemos construir dos observables para cada sitio y que tengan autovalores ± 1 . Consideremos $\mathbf{z} = (z_1, z_2, \dots, z_8)$ con $z_i \in \mathbb{C}$ tal que con cada par de números z_i, z_j , definidos en el vector \mathbf{z} , podemos construir vectores unitarios de dimensión dos como los presentados a continuación

$$\begin{aligned} |\mathbf{e}_1\rangle &= \frac{1}{\sqrt{z_1 + z_2}}(z_1 |0\rangle + z_2 |1\rangle), & |\mathbf{e}_2\rangle &= \frac{1}{\sqrt{z_3 + z_4}}(z_3 |0\rangle + z_4 |1\rangle), \\ |\mathbf{e}_3\rangle &= \frac{1}{\sqrt{z_5 + z_6}}(z_5 |0\rangle + z_6 |1\rangle), & |\mathbf{e}_4\rangle &= \frac{1}{\sqrt{z_7 + z_8}}(z_7 |0\rangle + z_8 |1\rangle). \end{aligned} \quad (6.13)$$

Así, por cada $|\mathbf{e}_i\rangle$ calculamos los respectivos operadores de proyección como

$$\begin{aligned} P_1 &= |\mathbf{e}_1\rangle\langle\mathbf{e}_1|, & P_2 &= |\mathbf{e}_2\rangle\langle\mathbf{e}_2|, \\ P_3 &= |\mathbf{e}_3\rangle\langle\mathbf{e}_3|, & P_4 &= |\mathbf{e}_4\rangle\langle\mathbf{e}_4|. \end{aligned} \quad (6.14)$$

Se definen los operadores de medición locales como,

$$\mathcal{O} = (+1)P_i + (-1)P_i^\perp = P_i - (-P_i + \mathbb{1}_{2 \times 2}) = 2P_i - \mathbb{1}_{2 \times 2} \quad (6.15)$$

Luego, los operadores de medición para Alice y Bob están descritos por dos parámetros complejos, la misma cantidad de parámetros para la versión real de SPSA,

$$\begin{aligned} \mathcal{A}_1 &:= \mathcal{A}_1(z_1, z_2) = 2P_1 - \mathbb{1}_{2 \times 2}, & \mathcal{A}_2 &:= \mathcal{A}_2(z_3, z_4) = 2P_2 - \mathbb{1}_{2 \times 2} \\ \mathcal{B}_1 &:= \mathcal{B}_1(z_5, z_6) = 2P_3 - \mathbb{1}_{2 \times 2}, & \mathcal{B}_2 &:= \mathcal{B}_2(z_7, z_8) = 2P_4 - \mathbb{1}_{2 \times 2}. \end{aligned} \quad (6.16)$$

Ahora, para la adaptación del algoritmo de CSPA al problema del máximo valor de la función S , es necesario generar un vector inicial con distribución uniforme $\hat{\mathbf{z}}_0 := (\hat{z}_1, \dots, \hat{z}_8)$ con $z_i \in \mathbb{C}$ que nos permitirá escribir la función objetivo f en este caso,

$$\begin{aligned} \hat{\mathcal{A}}_1 &:= \hat{\mathcal{A}}_1(\hat{z}_1, \hat{z}_2), & \hat{\mathcal{A}}_2 &:= \hat{\mathcal{A}}_2(\hat{z}_3, \hat{z}_4), \\ \hat{\mathcal{B}}_1 &:= \hat{\mathcal{B}}_1(\hat{z}_5, \hat{z}_6), & \hat{\mathcal{B}}_2 &:= \hat{\mathcal{B}}_2(\hat{z}_7, \hat{z}_8). \end{aligned} \quad (6.17)$$

Definimos la función objetivo \mathcal{S} como

$$\mathcal{S}(\hat{\mathbf{z}}) := \left| \hat{\mathcal{E}}_{11} + \hat{\mathcal{E}}_{12} + \hat{\mathcal{E}}_{21} - \hat{\mathcal{E}}_{22} \right|, \quad (6.18)$$

con $\hat{\mathcal{E}}_{ij} = \text{Tr} \{ \rho(\hat{\mathcal{A}}_i \otimes \hat{\mathcal{B}}_j) \}$.

Re definimos la secuencia iterativa mostrada en la eq. (5.20) como

$$\hat{\mathbf{z}}_{k+1} = \hat{\mathbf{z}}_k + \alpha_k \hat{\mathbf{g}}_k(\hat{\mathbf{z}}_k). \quad (6.19)$$

La estimación del gradiente en la eq. (6.19) es dada por

$$\hat{\mathbf{g}}_{k,i} = \frac{\mathcal{S}(\hat{\mathbf{z}}_k + \beta_k \Delta_k) + \varepsilon_{k,+} - \mathcal{S}(\hat{\mathbf{z}}_k - \beta_k \Delta_k) + \varepsilon_{k,-}}{2\beta_k \Delta_{k,i}^*}, \quad (6.20)$$

en este caso Δ_k es un vector de ocho componentes generado uniformemente donde cada componente puede tomar los valores $\{+1, -1, +i, -i\}$ con $i := \sqrt{-1}$, y $\varepsilon_{k,\pm}$ es un termino de ruido. Las ganancias α_k y β_k se obtienen de acuerdo a la eq. (5.22). Los parámetros a , A , b , s y t usados son para el caso estándar y asintótico mostrados en la [Tabla 5.2](#). Así, para $k \rightarrow \infty$ se consigue el máximo de \mathcal{S} para el estado ρ .

En las siguientes tablas se muestra el pseudo código para los dos métodos que estudiaremos en el presente capítulo. En [Algoritmo 2](#) se muestra el procedimiento usado para encontrar el máximo de \mathcal{S} mediante CSPA. Cabe destacar que el guess \mathbf{z} no está normalizado en la evaluación de la función $\text{CHSH}(\mathbf{z}, \rho)$. Por otro lado, en [Algoritmo 3](#) muestra el código de CSPApero con la normalización del guess en la evaluación de la función $\text{CHSH}(\mathbf{z}, \rho)$ (línea 9). Llamaremos a esta versión CSPAN o CSPANormalizado para distinguirlo de la versión sin normalizar presentada en [Algoritmo 2](#). Como veremos más adelante, esta pequeña sutileza en el código se vuelve significativa al momento de comparar la precisión de los métodos.

Algorithm 2 Algoritmo CSPA y función objetivo CHSH definida de acuerdo a la parametrización mostrada en la eq. (6.16).

Require: D : Dimensión del espacio. N : Número de iteraciones. z : vector inicial complejo de ocho componentes. ρ : Estado cuántico. a, b, s, t, A : parámetros de ajuste.

Ensure: Valor del parámetro S en cada iteración.

```

1: procedure CSPA( $D, N, z, \rho, a, b, s, t, A$ )
2:   for  $k \leftarrow 1, K$  do
3:      $\alpha_k \leftarrow \frac{a}{(k+1+A)^s}, \quad \beta_k \leftarrow \frac{b}{(k+1)^t}$ 
4:      $\Delta_k \leftarrow$  vector de dimensión  $N$  generado uniformemente entre los valores
        $\{\pm 1, \pm i\}$ , con  $i = \sqrt{-1}$ .
5:      $z_+ \leftarrow z + \beta_k \Delta_k, \quad z_- \leftarrow z - \beta_k \Delta_k$ 
6:      $\text{CHSH}_+ \leftarrow \text{CHSH}(z_+, \rho) \quad \text{CHSH}_- \leftarrow \text{CHSH}(z_-, \rho)$ 
7:      $g_k \leftarrow \frac{\text{CHSH}_+ - \text{CHSH}_-}{2\beta_k \Delta_k^*}$ 
8:      $z \leftarrow z + \alpha_k g_k$ 
9:      $\text{Out} \leftarrow \text{CHSH}(z, \rho)$ 
10:  end for
11:  return  $\text{Out}$ 
12: end procedure

13: function CHSH( $z, \rho$ )
14:   $|0\rangle \leftarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad |1\rangle \leftarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ 
15:   $|e_1\rangle \leftarrow \frac{1}{\sqrt{z_1 + z_2}}(z_1 |0\rangle + z_2 |1\rangle), \quad |e_2\rangle \leftarrow \frac{1}{\sqrt{z_3 + z_4}}(z_3 |0\rangle + z_4 |1\rangle)$ 
16:   $|e_3\rangle \leftarrow \frac{1}{\sqrt{z_5 + z_6}}(z_5 |0\rangle + z_6 |1\rangle), \quad |e_4\rangle \leftarrow \frac{1}{\sqrt{z_7 + z_8}}(z_7 |0\rangle + z_8 |1\rangle)$ 
17:   $P_1 \leftarrow |e_1\rangle\langle e_1|, \quad P_2 \leftarrow |e_2\rangle\langle e_2|, \quad P_3 \leftarrow |e_3\rangle\langle e_3|, \quad P_4 \leftarrow |e_4\rangle\langle e_4|$ 
18:   $A_1 \leftarrow 2P_1 - \mathbb{1}_2, \quad A_2 \leftarrow 2P_2 - \mathbb{1}_2$ 
19:   $B_1 \leftarrow 2P_3 - \mathbb{1}_2, \quad B_2 \leftarrow 2P_4 - \mathbb{1}_2$ 
20:   $E_{11} \leftarrow \text{Tr}[(A_1 \otimes B_1)\rho], \quad E_{12} \leftarrow \text{Tr}[(A_1 \otimes B_2)\rho]$ 
21:   $E_{21} \leftarrow \text{Tr}[(A_2 \otimes B_1)\rho], \quad E_{22} \leftarrow \text{Tr}[(A_2 \otimes B_2)\rho]$ 
22:   $S \leftarrow |E_{11} + E_{21} + E_{12} - E_{22}|$ 
23:  return  $S$ 
24: end function

```

Algorithm 3 Algoritmo CSPA normalizado y función objetivo CHSH definida de acuerdo a la parametrización mostrada en la eq. (6.16). En este método se realiza una normalización al momento de actualizar el guess.

Require: D: Dimensión del espacio. N: Número de iteraciones. z : vector inicial complejo de ocho componentes. ρ : Estado cuántico. a, b, s, t, A : parámetros de ajuste.

Ensure: Valor del parámetro S en cada iteración.

```

1: procedure CSPA(D,N,z,ρ,a,b,s,t,A)
2:   for  $k \leftarrow 1, K$  do
3:      $\alpha_k \leftarrow \frac{a}{(k+1+A)^s}, \quad \beta_k \leftarrow \frac{b}{(k+1)^t}$ 
4:      $\Delta_k \leftarrow$  vector de dimensión N generado uniformemente entre los valores
        $\{\pm 1, \pm i\}$  con  $i = \sqrt{-1}$ .
5:      $z_+ \leftarrow z + \beta_k \Delta_k, \quad z_- \leftarrow z - \beta_k \Delta_k$ 
6:      $\text{CHSH}_+ \leftarrow \text{CHSH}(z_+, \rho), \quad \text{CHSH}_- \leftarrow \text{CHSH}(z_-, \rho)$ 
7:      $g_k \leftarrow \frac{\text{CHSH}_+ - \text{CHSH}_-}{2\beta_k \Delta_k^*}$ 
8:      $z \leftarrow z + \alpha_k g_k$ 
9:      $\zeta_1 \leftarrow \frac{(z_1, z_2)}{\sqrt{z_1 + z_2}}, \quad \zeta_2 \leftarrow \frac{(z_3, z_4)}{\sqrt{z_3 + z_4}}, \quad \zeta_3 \leftarrow \frac{(z_5, z_6)}{\sqrt{z_5 + z_6}}, \quad \zeta_4 \leftarrow \frac{(z_7, z_8)}{\sqrt{z_7 + z_8}}$ 
10:     $z \leftarrow [\zeta_1; \zeta_2; \zeta_3; \zeta_4]$ 
11:    Out  $\leftarrow \text{CHSH}(z, \rho)$ 
12:  end for
13:  return Out
14: end procedure

15: function CHSH( $z, \rho$ )
16:   $|0\rangle \leftarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad |1\rangle \leftarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ 
17:   $|e_1\rangle \leftarrow \frac{1}{\sqrt{z_1 + z_2}}(z_1 |0\rangle + z_2 |1\rangle), \quad |e_2\rangle \leftarrow \frac{1}{\sqrt{z_3 + z_4}}(z_3 |0\rangle + z_4 |1\rangle)$ 
18:   $|e_3\rangle \leftarrow \frac{1}{\sqrt{z_5 + z_6}}(z_5 |0\rangle + z_6 |1\rangle), \quad |e_4\rangle \leftarrow \frac{1}{\sqrt{z_7 + z_8}}(z_7 |0\rangle + z_8 |1\rangle)$ 
19:   $P_1 \leftarrow |e_1\rangle\langle e_1|, \quad P_2 \leftarrow |e_2\rangle\langle e_2|, \quad P_3 \leftarrow |e_3\rangle\langle e_3|, \quad P_4 \leftarrow |e_4\rangle\langle e_4|$ 
20:   $A_1 \leftarrow 2P_1 - \mathbb{1}_2, \quad A_2 \leftarrow 2P_2 - \mathbb{1}_2$ 
21:   $B_1 \leftarrow 2P_3 - \mathbb{1}_2, \quad B_2 \leftarrow 2P_4 - \mathbb{1}_2$ 
22:   $E_{11} \leftarrow \text{Tr}[(A_1 \otimes B_1)\rho], \quad E_{12} \leftarrow \text{Tr}[(A_1 \otimes B_2)\rho]$ 
23:   $E_{21} \leftarrow \text{Tr}[(A_2 \otimes B_1)\rho], \quad E_{22} \leftarrow \text{Tr}[(A_2 \otimes B_2)\rho]$ 
24:   $S \leftarrow |E_{11} + E_{21} + E_{12} - E_{22}|$ 
25:  return S
26: end function

```

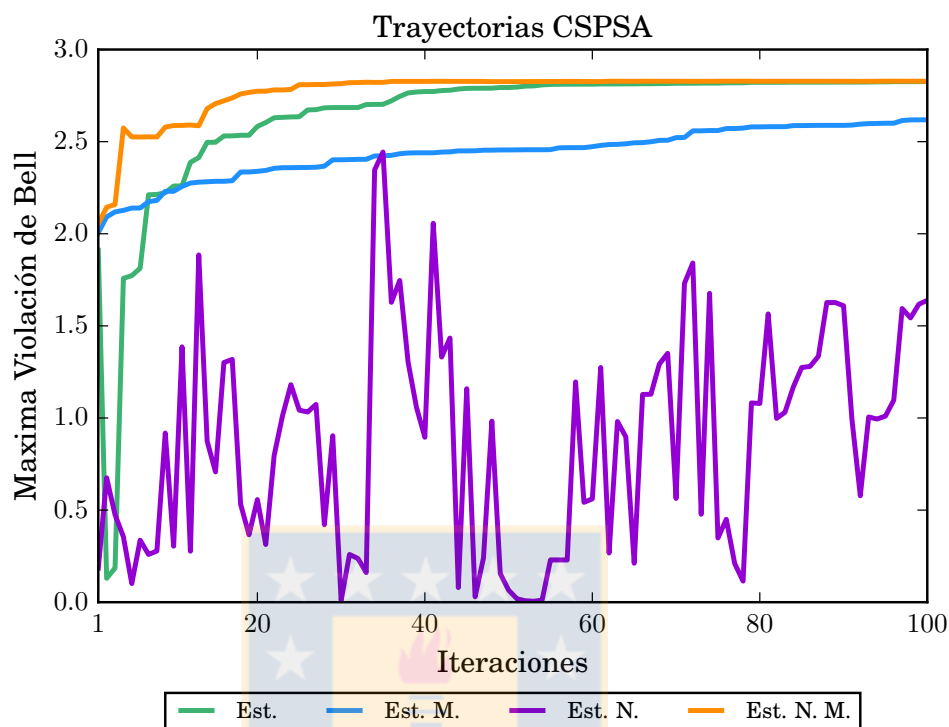


Figura 6.10: Trayectorias de CSPA para las ganancias estándar. El gráfico muestra la máxima violación de Bell versus el número de iteraciones del algoritmo de CSPA. Cada trayectoria representa el camino que sigue cada simulación para diferentes ganancias. La nomenclatura corresponde a Est.: Estándar, Est. M.: Estándar Modificado, Est. N.: Estándar para CSPAN, y Est. N. M.: Estándar Modificado para CSPAN. **Fuente:** Elaboración propia

Análogamente a la sección precedente, analizaremos el rendimiento y comportamiento de los algoritmos de CSPA y su versión normalizada CSPAN para la función objetivo CHSH siguiendo la parametrización desarrollada en (6.13 - 6.18). Consideraremos nuevamente el estado $|\phi_+\rangle$ y el Algoritmo 2. La Figura 6.10 muestra cuatro trayectorias para cada conjunto de ganancias. Como vemos, el conjunto de ganancias estándar modificadas solo presenta ruido hasta las 100 iteraciones, esto es debido a que este conjunto de parámetros tiene una convergencia más lenta que las otras tres ganancias. El ruido indica que no es conveniente utilizar esos parámetros debido a que posee muchísima variabilidad en su trayectoria.

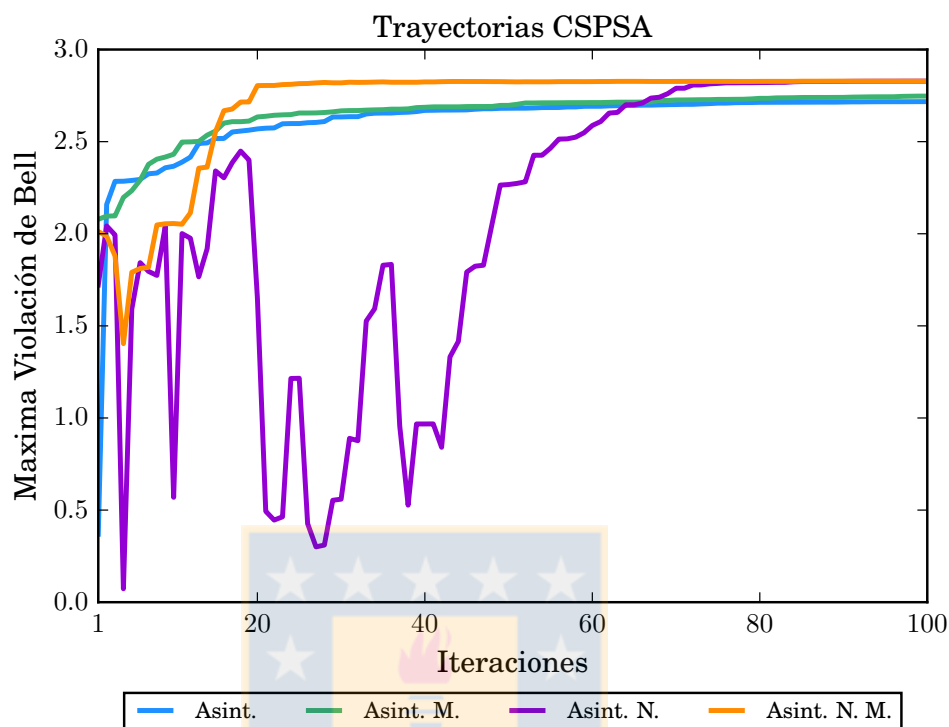


Figura 6.11: Trayectorias de CSPA para las ganancias asintóticas. Cada trayectoria representa el camino que sigue cada simulación para diferentes ganancias. La nomenclatura corresponde a Asint.: Asintótico, Asint. M.: Asintótico Modificado, Asint. N.: Asintótico para CSPAN, y Asint. N. M.: Asintótico Modificado para CSPAN. **Fuente:** Elaboración propia

La [Figura 6.11](#) muestra las trayectorias descritas para los algoritmos CSPA y CSPAN cuando son utilizadas las ganancias asintóticas. Las curvas presentadas son bastante similares en estabilidad a las mostradas en el caso de SPSA. A diferencia de las ganancias estándar, las asintóticas requieren menos iteraciones para alcanzar el máximo de la función. El gráfico muestra que, independientemente del guess y las ganancias usadas en este análisis, los algoritmos de CSPA y CSPAN convergen al máximo de la función, solo que a diferentes tasas de convergencia.

Notemos que la curva generada por los parámetros asintóticos en CSPAN es diferente a las producidas por las ganancias asintóticas modificadas aplicadas a CSPAN, esto se debe a que al normalizar el guess en el algoritmo la función objetivo cambia y, por tanto, las ganancias utilizadas no necesariamente reproducen las mismas curvas. De esta forma se puede observar que, para cada problema, se deben ajustar nuevamente las ganancias.

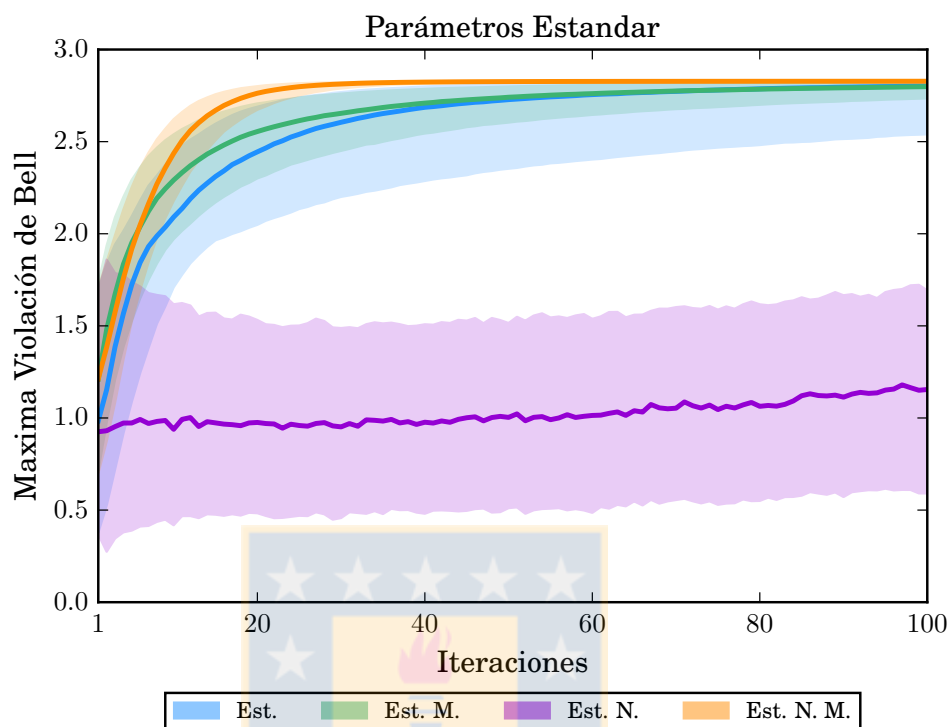


Figura 6.12: Simulación de CSPA y CPSAN usando ganancias estándar para un estado maximalmente entrelazado con 5×10^3 realizaciones. La figura muestra un gráfico de la máxima violación de Bell versus número de iteraciones. Las líneas continuas representan la mediana y el sombreado el rango intercuartil para cada ganancia. La nomenclatura corresponde a: Est.: Estándar, Est. M.: Estándar Modificado, Est. N.: Estándar para CPSAN, y Est. N. M.: Estándar Modificado para CPSAN. **Fuente:** Elaboración propia

Ya que este método también es estocástico en las trayectorias, se deben simular realizaciones independientes para estudiar el rendimiento de los algoritmos. En la [Figura 6.12](#) se muestran 5×10^3 realizaciones independientes. Las líneas centrales representan la mediana y el sombreado el rango intercuartil.

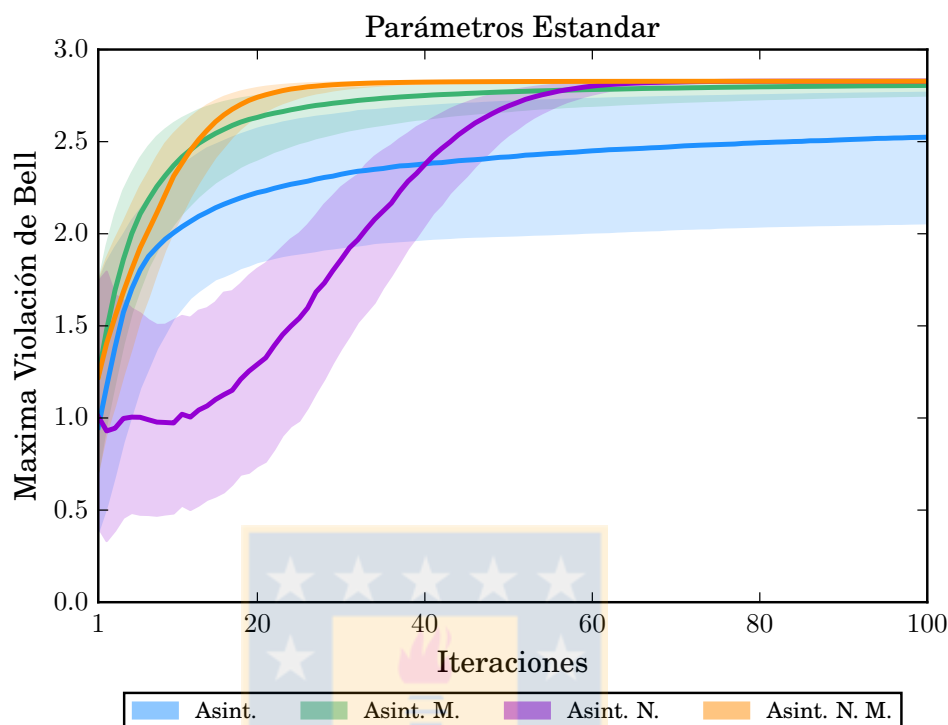


Figura 6.13: Simulación de CSPA y CSPAN usando ganancias asintóticas para un estado maximalmente entrelazado con 5×10^3 realizaciones. La figura muestra el gráfico de la máxima violación de Bell versus número de iteraciones. Las líneas continuas representan la mediana y el sombreado el rango intercuartil para cada ganancia. La nomenclatura corresponde a: Asint.: Asintótico, Asint. M.: Asintótico Modificado, Asint. N.: Asintótico para CSPAN, y Asint. N. M.: Asintótico Modificado para CSPAN. **Fuente:** Elaboración propia

La **Figura 6.13** muestra el gráfico de la máxima violación de Bell versus número de iteraciones para 5×10^3 realizaciones. Vemos que, pese a que la curva producida por CSPA con las ganancias asintóticas modificadas superan el valor de 2.0 en menos iteraciones, el método CSPAN con las ganancias asintóticas modificadas converge más rápidamente a $2\sqrt{2}$.

Siguiendo el análisis realizado para SPSA, se estudiará el rendimiento de los algoritmos de CSPA y CSPAN aplicado sobre los estados de Werner (6.9). La **Figura 6.14** muestra el valor promedio de las ganancias asintóticas utilizadas en los dos algoritmos complejos. Cada punto en el gráfico representa el promedio del máximo valor conseguido a las 200 iteraciones para 10 realizaciones independientes.

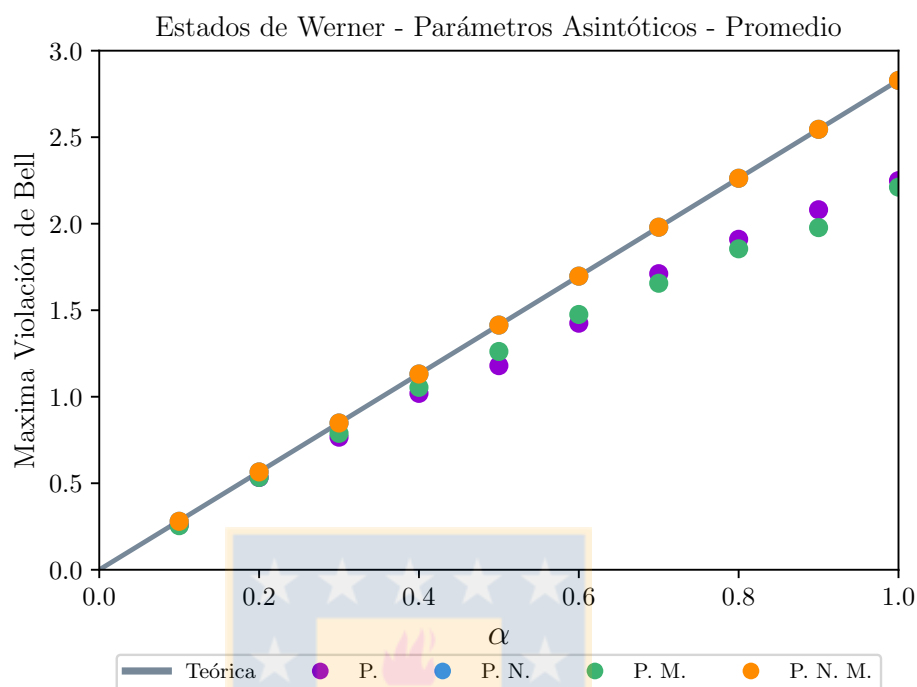


Figura 6.14: Promedio de las simulaciones realizadas con \mathbb{C} SPSA y \mathbb{C} SPSAN aplicada estados de Werner para parámetros asintóticos. Cada punto representa el promedio del máximo valor conseguido a las 200 iteraciones, donde se simularon 10 realizaciones. La nomenclatura corresponde a: P.: Promedio, P.N.: Promedio de \mathbb{C} SPSAN, P.M.: Promedio ganancias modificadas, P.N.M.: Promedio ganancias modificadas aplicadas en \mathbb{C} SPSAN. **Fuente:** Elaboración propia

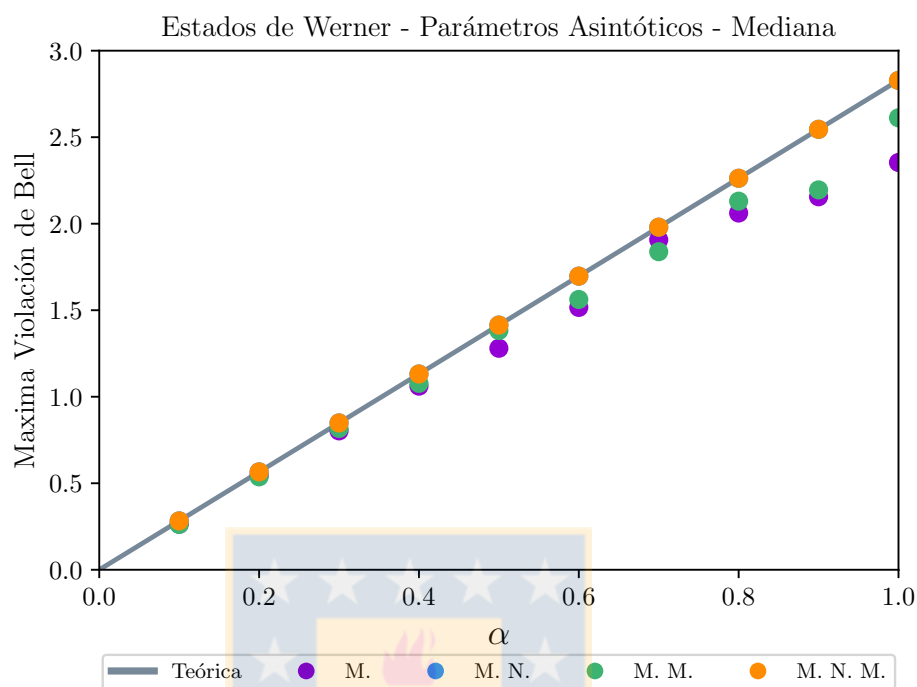


Figura 6.15: Mediana de las simulaciones realizadas con \mathbb{C} SPSA y \mathbb{C} SPSAN aplicada estados de Werner para parámetros asintóticos. Cada punto representa la mediana del máximo valor conseguido a las 200 iteraciones, donde se simularon 10 realizaciones. La nomenclatura corresponde a: M.: Mediana, M.N.: Mediana de \mathbb{C} SPSAN, M.M.: Mediana de ganancias modificadas, M.N.M.: Mediana de ganancias modificadas aplicadas en \mathbb{C} SPSAN. **Fuente:** Elaboración propia

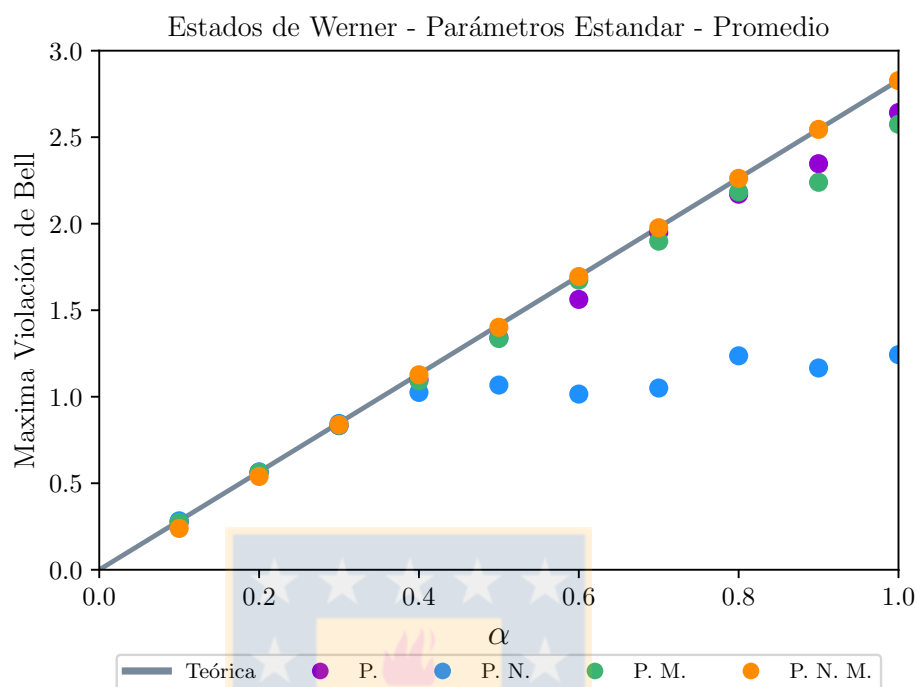


Figura 6.16: Promedio de las simulaciones realizadas con \mathbb{C} SPSA y \mathbb{C} SPSAN aplicada estados de Werner para parámetros estándar. Cada punto representa el promedio del máximo valor conseguido a las 200 iteraciones, donde se simularon 10 realizaciones. La nomenclatura corresponde a: P.: Promedio, P.N.: Promedio de \mathbb{C} SPSAN, P.M.: Promedio ganancias modificadas, P.N.M.: Promedio ganancias modificadas aplicadas en \mathbb{C} SPSAN. **Fuente:** Elaboración propia

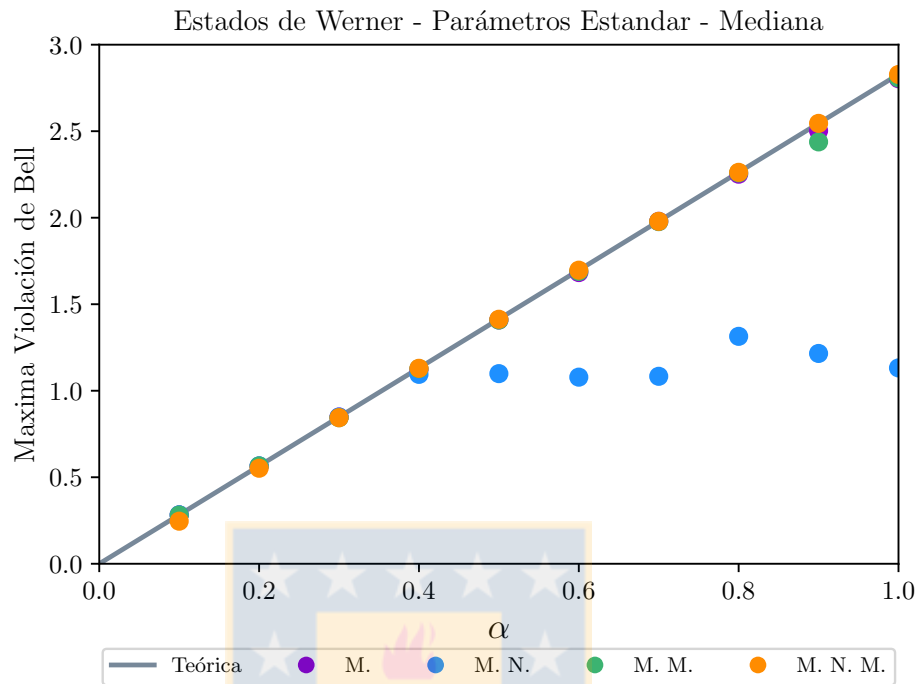


Figura 6.17: Mediana de las simulaciones realizadas con CSPSA y CSPSAN aplicada estados de Werner para parámetros estándar. Cada punto representa la mediana del máximo valor conseguido a las 200 iteraciones, donde se simularon 10 realizaciones. La nomenclatura corresponde a: M.: Mediana, M.N.: Mediana de CSPSAN, M.M.: Mediana de ganancias modificadas, M.N.M.: Mediana de ganancias modificadas aplicadas en CSPSAN. **Fuente:** Elaboración propia

De las [Figura 6.14](#), [Figura 6.15](#), [Figura 6.16](#) y [Figura 6.17](#) se observar que, para en todos los casos, tanto la mediana como el promedio de las ganancias modificadas aplicadas sobre CSPSAN reproducen los resultados predichos por la teoría. De acuerdo a este resultado se tiene evidencia suficiente para afirmar que las ganancias modificadas aplicadas a CSPSA normalizado poseen un mejor rendimiento que las otras configuraciones estudiadas.

Finalmente se revisarán los algoritmos aplicados sobre un estado descrito en descomposición de Schmidt, como fue mostrado en la eq. (6.11).

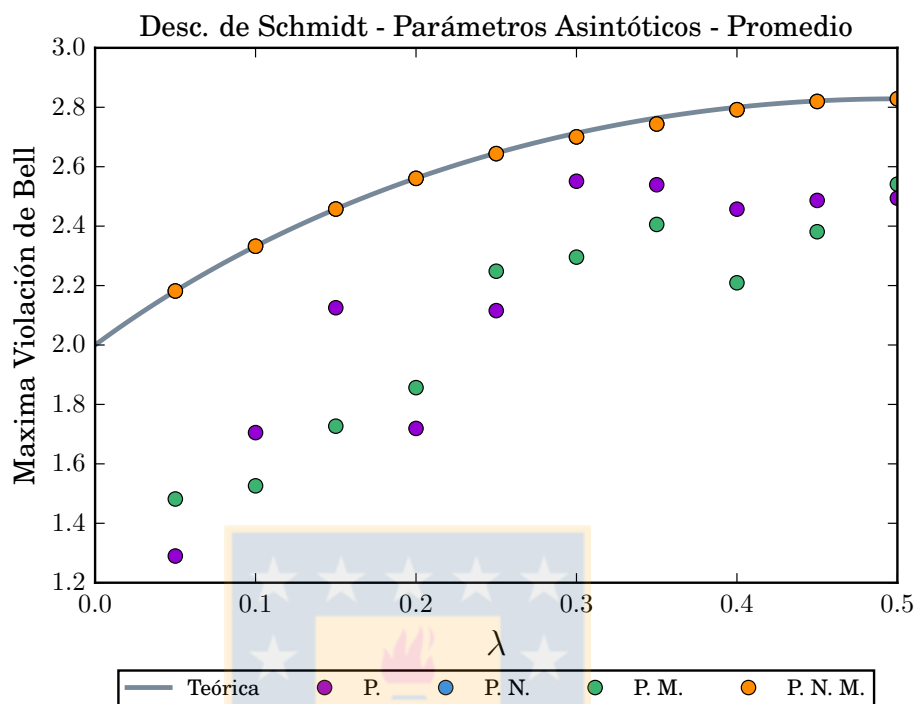


Figura 6.18: Promedio de las simulaciones realizadas con \mathbb{C} SPSA y \mathbb{C} SPSAN aplicada a estados descritos en descomposición de Schmidt para ganancias Asintóticas. Cada punto representa el promedio del máximo valor conseguido a las 200 iteraciones, donde se simularon 10 realizaciones. La nomenclatura corresponde a: P.: Promedio, P.N.: Promedio de \mathbb{C} SPSAN, P.M.: Promedio ganancias modificadas, P.N.M.: Promedio ganancias modificadas aplicadas en \mathbb{C} SPSAN. **Fuente:** Elaboración propia

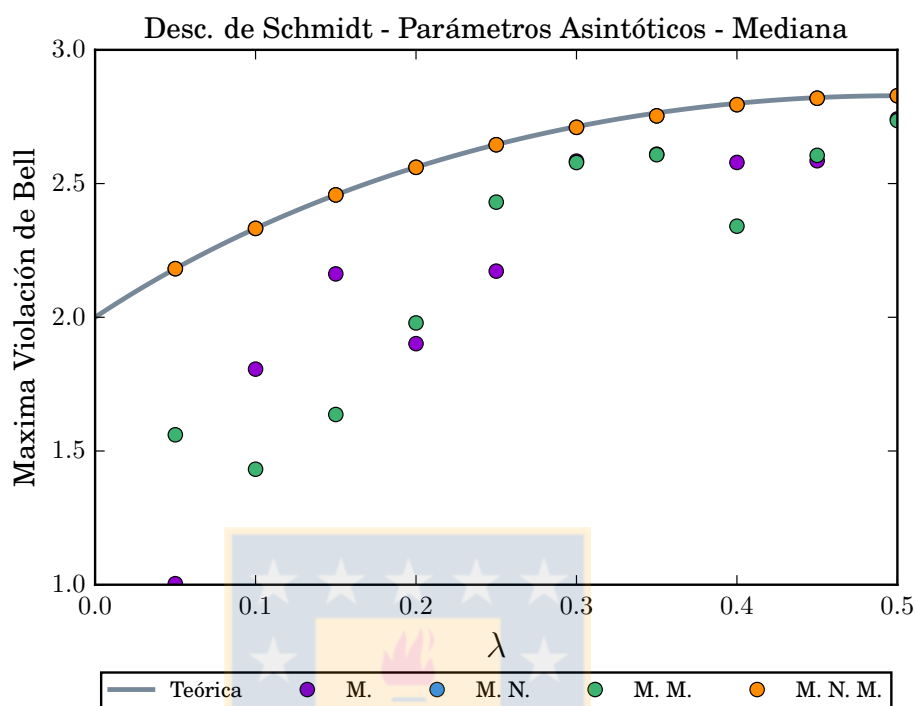


Figura 6.19: Mediana de las simulaciones realizadas con CSPA y CSPAN aplicada a estados descritos en descomposición de Schmidt para ganancias asintóticas. La nomenclatura corresponde a: M.: Mediana, M.N.: Mediana de CSPAN, M.M.: Mediana de ganancias modificadas, M.N.M.: Mediana de ganancias modificadas aplicadas en CSPAN. **Fuente:** Elaboración propia

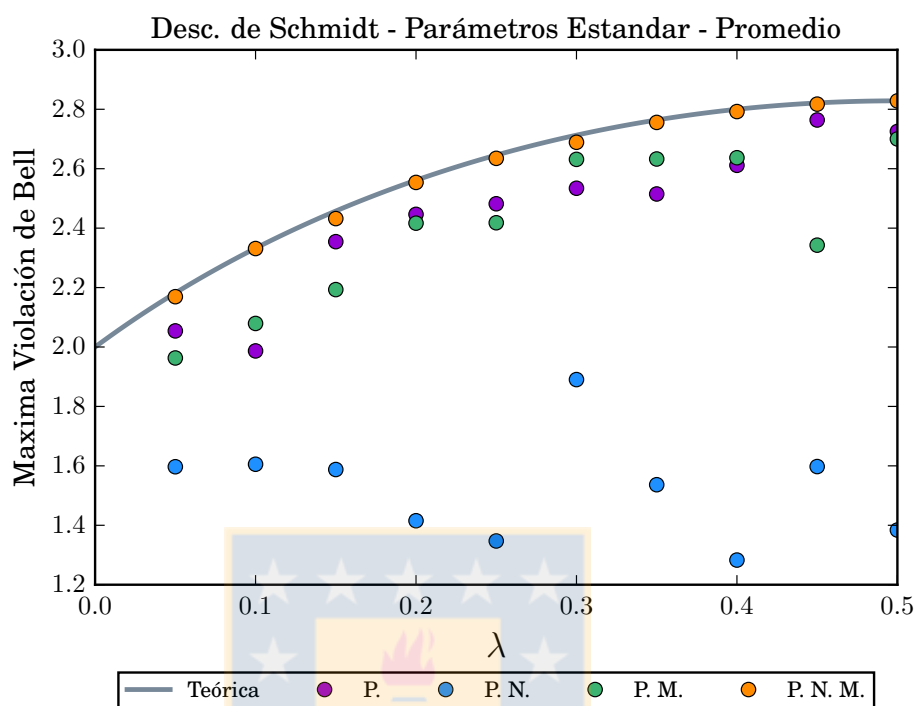


Figura 6.20: Promedio de las simulaciones realizadas con CSPSA y CSPSAN aplicada a estados descritos en descomposición de Schmidt para ganancias estándar. La nomenclatura corresponde a: P.: Promedio, P.N.: Promedio de CSPSAN, P.M.: Promedio ganancias modificadas, P.N.M.: Promedio ganancias modificadas aplicadas en CSPSAN. **Fuente:** Elaboración propia

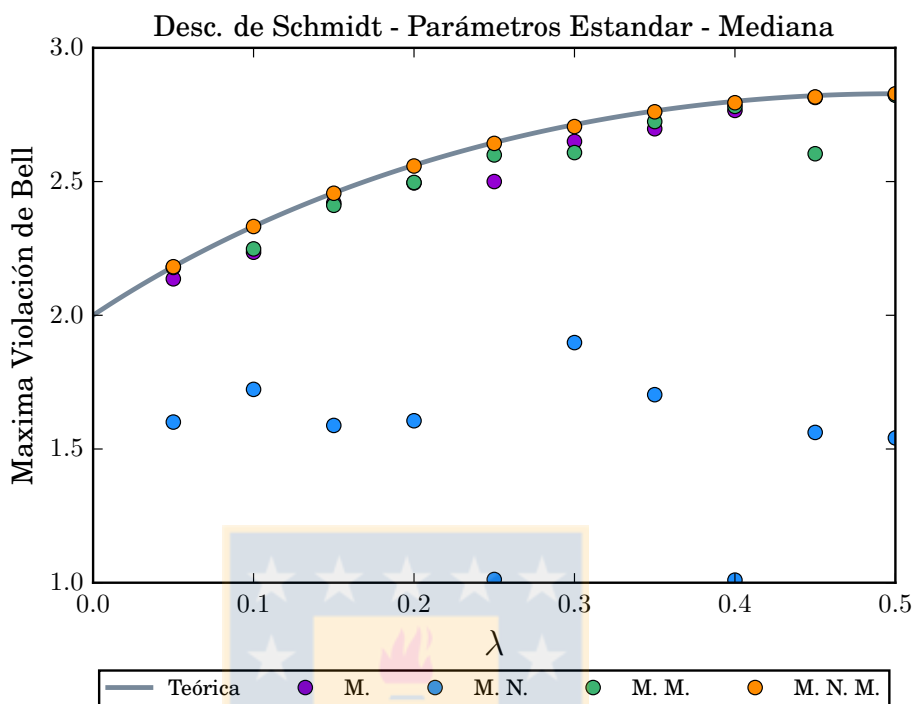


Figura 6.21: Mediana de las simulaciones realizadas con CSPSA y CSPSAN aplicada a estados descritos en descomposición de Schmidt para ganancias estándar. Cada punto representa el promedio del máximo valor conseguido a las 200 iteraciones, donde se simularon 10 realizaciones. La nomenclatura corresponde a: M.: Mediana, M.N.: Mediana de CSPSAN, M.M.: Mediana de ganancias modificadas, M.N.M.: Mediana de ganancias modificadas aplicadas en CSPSAN. **Fuente:** Elaboración propia

De las [Figura 6.18](#), [Figura 6.19](#), [Figura 6.20](#) y [Figura 6.21](#) se observa que las ganancias estándar y asíntóticas modificadas aplicadas a CSPSAN reproducen la curva que predice la teoría, tanto en el cálculo de mediana como en el promedio. De esta forma podemos afirmar que el método que nos lleva al mejor rendimiento es CSPSAN con ganancias modificadas.

6.4. Estimación de la desigualdad CHSH para estados cuánticos desconocidos

A continuación se reportarán las simulaciones realizadas con los algoritmos SPSA, CSPSA y CSPSAN, utilizando las ganancias mostradas en la [Tabla 5.2](#). Cada gráfico muestra la mediana y el rango intercuartil para 10^3 estados puros generados uniformemente donde cada uno de ellos fue simulado utilizando el mismo guess, es decir, inician desde el mismo punto. Los gráficos de error relativo vs número de iteraciones se obtienen de acuerdo a la expresión

$$E_{\text{rel}} = \frac{|\mathcal{S}_{\text{teo}} - \mathcal{S}_{\text{sim}}|}{\mathcal{S}_{\text{teo}}}, \quad (6.21)$$

donde \mathcal{S}_{sim} es el valor estimado de la desigualdad mediante algún proceso de optimización y \mathcal{S}_{teo} es el valor máximo de la desigualdad que puede lograr el estado simulado. El valor exacto se calcula mediante el procedimiento mostrado en la [Sección 4.5](#),

$$E_{11} + E_{12} + E_{21} - E_{22} = 2[(1 - \gamma)^2 + K^2]^{1/2} + 2\gamma, \quad (6.22)$$

con $K = 2(c_1c_2 + c_3c_4 + \dots)$, donde c_i son los coeficientes de Schmidt de la descomposición. Para el caso de la desigualdad CHSH $\gamma = 2$ y, por lo tanto

$$E_{11} + E_{12} + E_{21} - E_{22} = 2[(1 + K^2)]^{1/2}, \quad (6.23)$$

6.4.1. Simulaciones SPSA

En la presente sección se mostrarán las simulaciones realizadas utilizando el algoritmo de SPSA para 10^3 estados puros uniformemente generados. En las siguientes figuras se mostrará el rendimiento del algoritmo bajo las diferentes ganancias presentadas anteriormente.

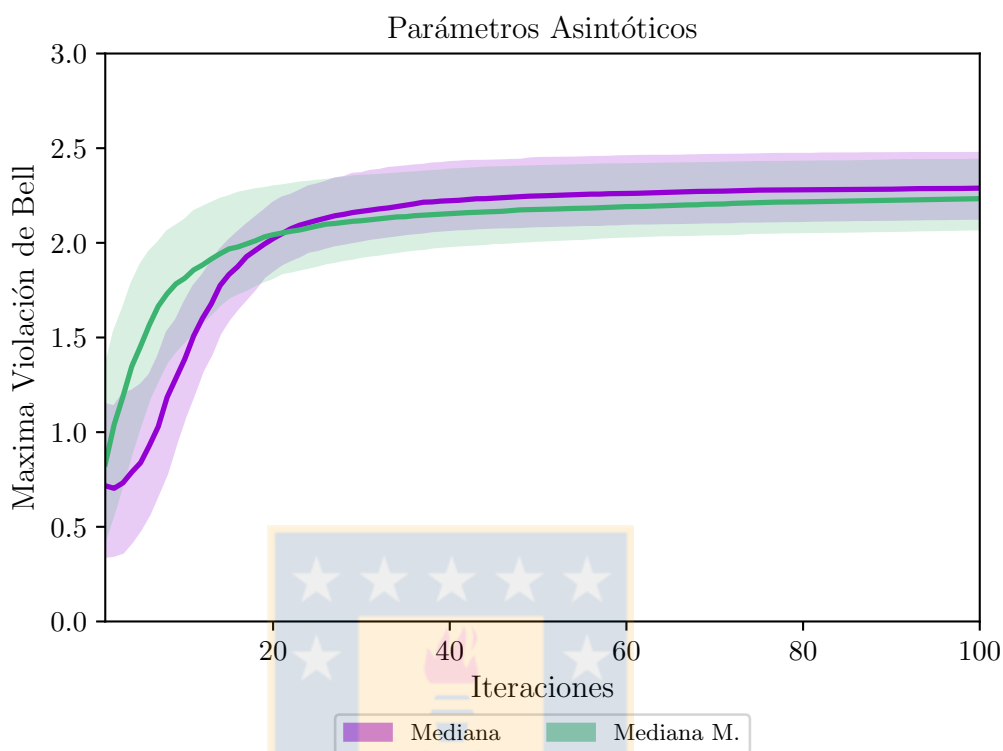


Figura 6.22: Simulaciones de SPSA para las ganancias asintóticas para 10^3 estados generados uniformemente. El gráfico indica la máxima violación de Bell versus número de iteraciones. Las líneas continuas representan la mediana y el sombreado el rango intercuartil, según corresponda.

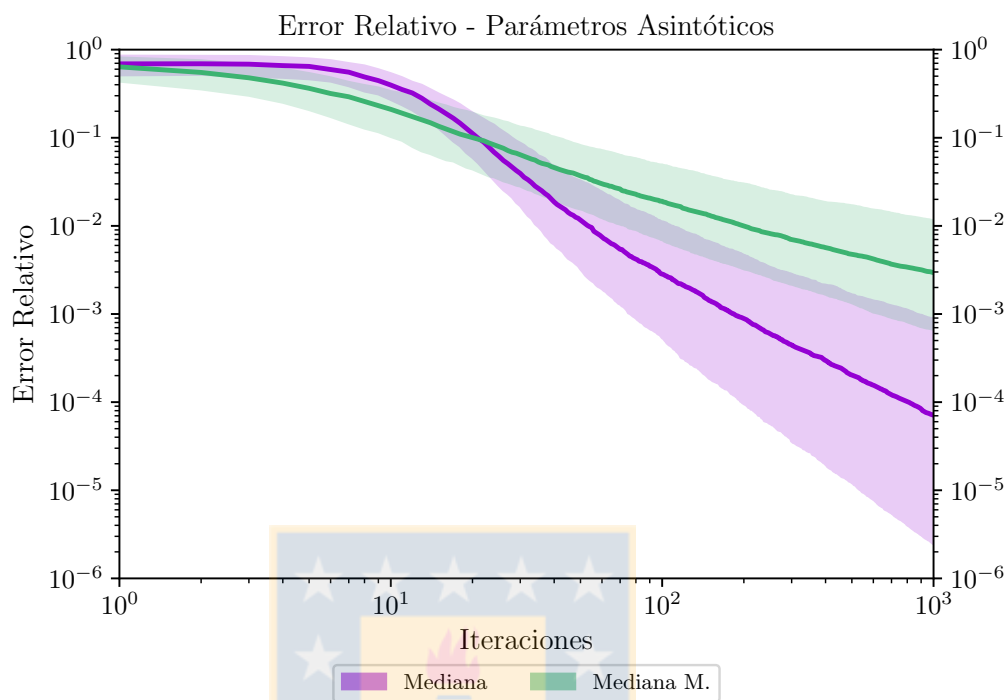


Figura 6.23: Error relativo de SPSA de las ganancias asintóticas para 10^3 estados puros generados uniformemente. El gráfico presenta el error relativo versus número de iteraciones en escala logarítmica. Las líneas continuas indican la mediana y el sombreado el rango intercuartil, según corresponda.

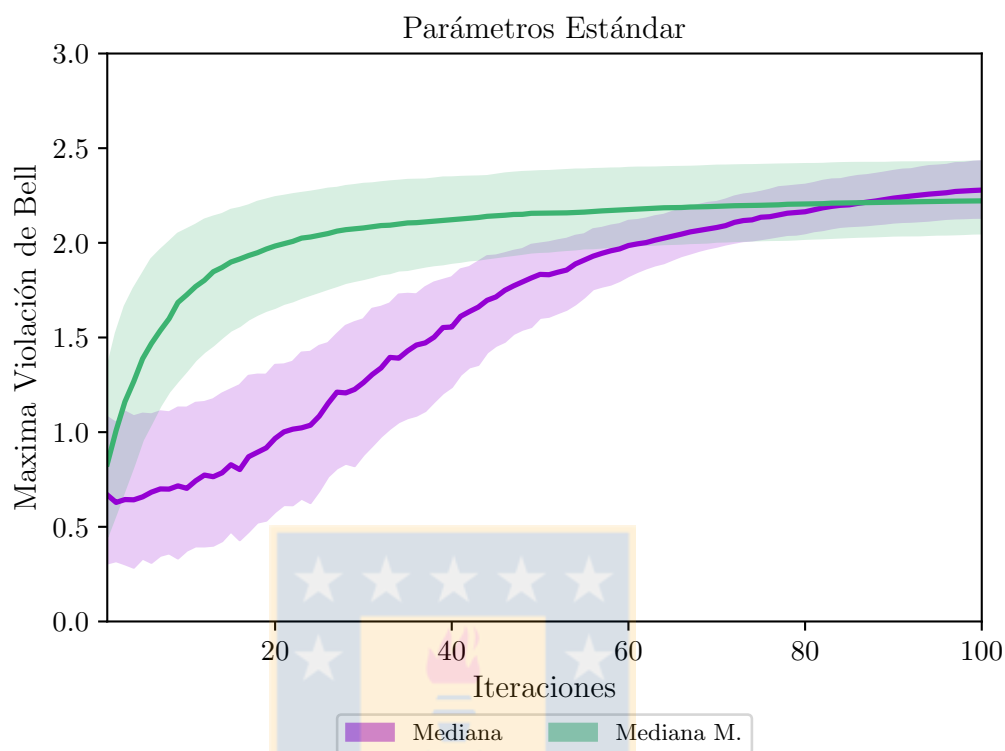


Figura 6.24: Simulaciones de SPSA para las ganancias estándar sobre 10^3 estados puros generados uniformemente. El gráfico indica la máxima violación de Bell versus número de iteraciones. Las líneas continuas representan la mediana y el sombreado el rango intercuartil en cada caso, según corresponda.

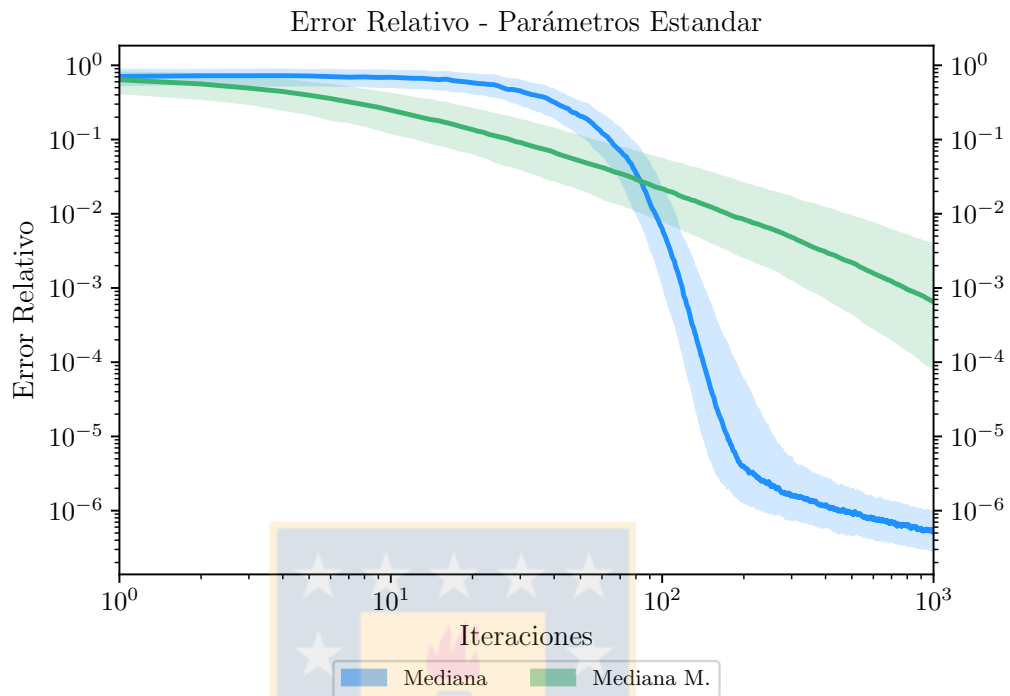


Figura 6.25: Error relativo de SPSA de las ganancias estándar para 10^3 estados puros generados uniformemente. El gráfico presenta el error relativo versus número de iteraciones en escala logarítmica. Las líneas continuas indican la mediana y el sombreado el rango intercuartil en cada caso, según corresponda.

6.4.2. Simulaciones CSPAy CSPAN

En la presente sección se mostrarán las simulaciones realizadas utilizando algoritmo de CSPAy su versión que incluye normalización CSPAN para 10^3 estados puros uniformemente generados. El objetivo es investigar el rendimiento en mediana para todo tipo de estados, esto es, estados con diferente cantidad de entrelazamiento. En las siguientes figuras se mostrará la mediana y los rangos intercuantiles de los algoritmos bajo las diferentes ganancias presentadas la [Tabla 5.2](#).

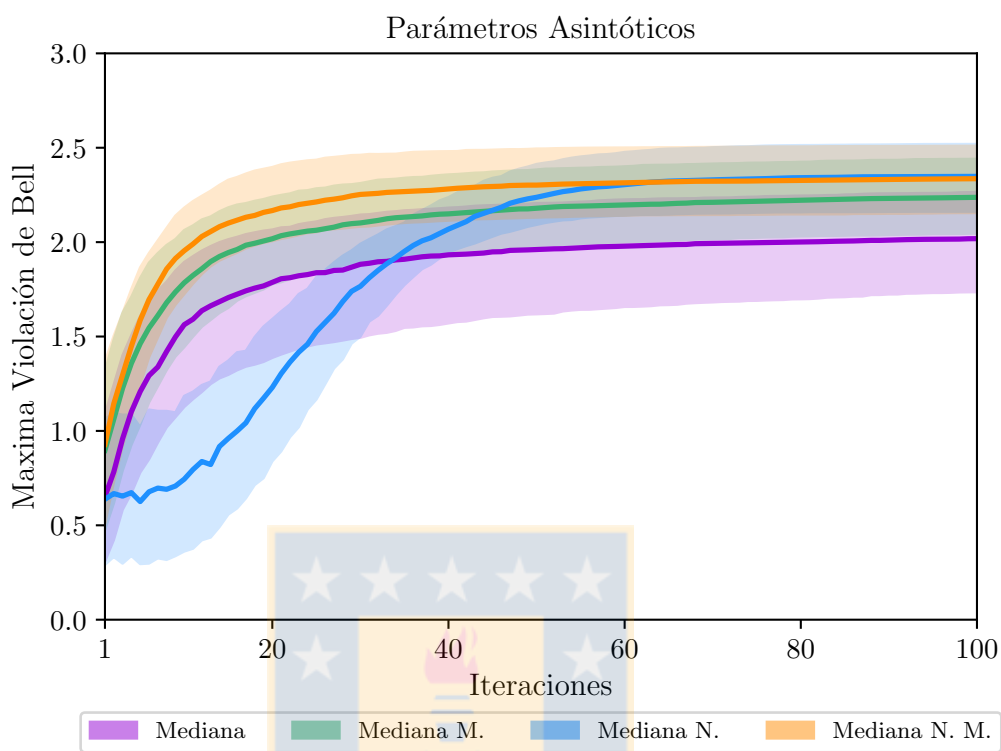


Figura 6.26: Simulaciones de CSPA y CSPSAN para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente. El gráfico indica la máxima violación de Bell versus número de iteraciones. Las líneas continuas representan la mediana y el sombreado el rango intercuartil en cada caso, según corresponda.

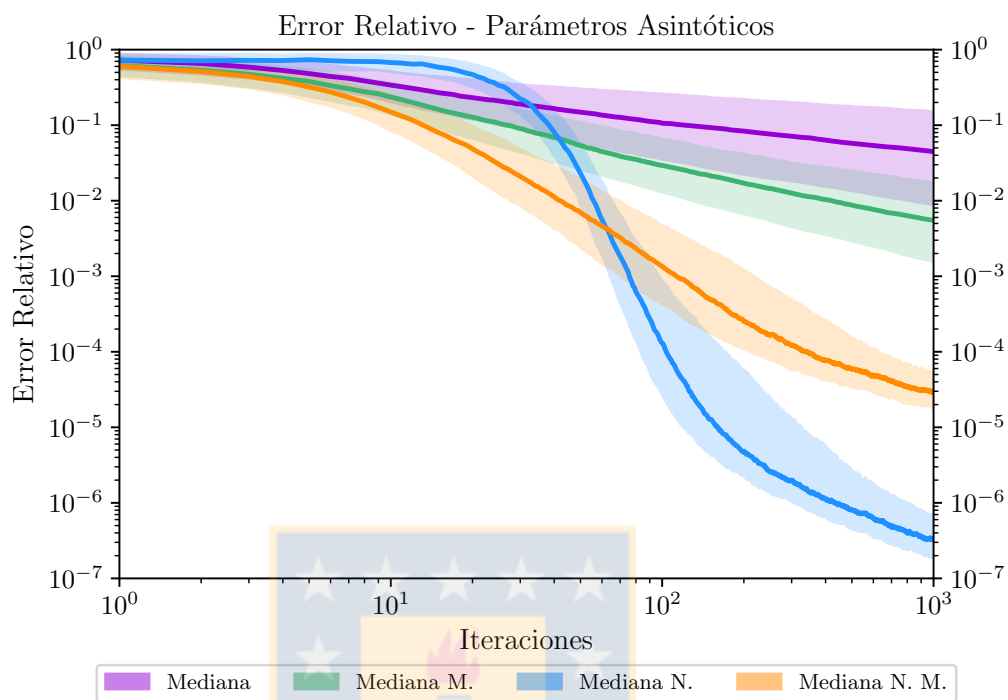


Figura 6.27: Error relativo de \mathbb{C} SPSA y \mathbb{C} SPSAN para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente. El gráfico presenta el error relativo versus número de iteraciones en escala logarítmica. Las líneas continuas indican la mediana y el sombreado el rango intercuartil en cada caso, según corresponda.

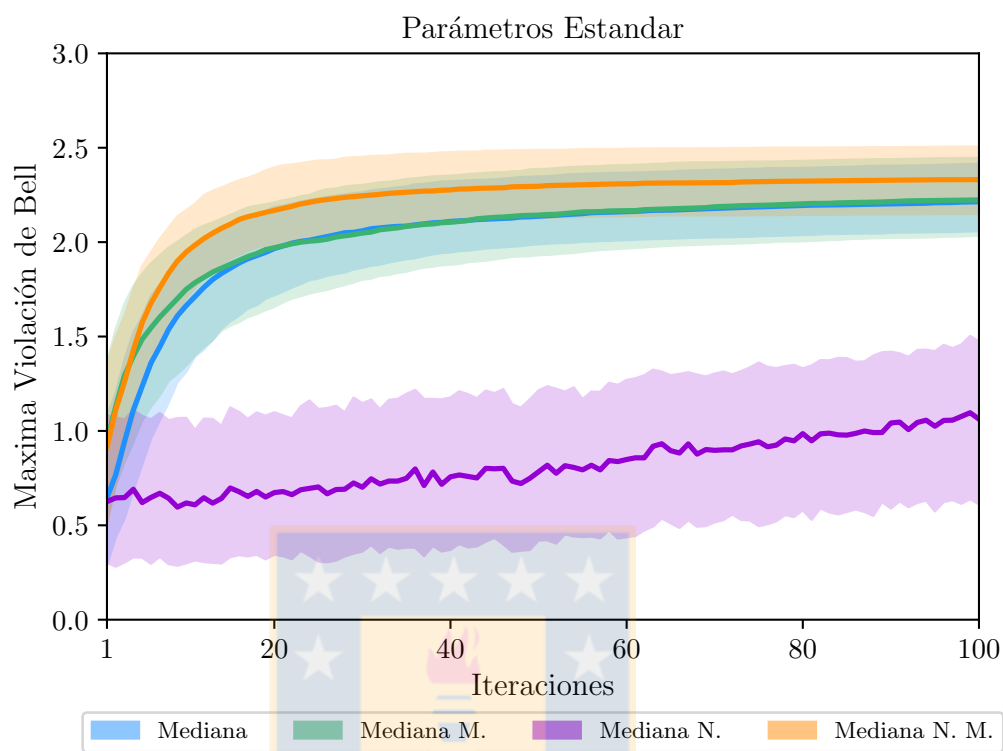


Figura 6.28: Simulaciones de CSPA y CSPAN para las ganancias estándar sobre 10^3 estados generados uniformemente. El gráfico indica la máxima violación de Bell versus número de iteraciones. Las líneas continuas representan la mediana y el sombreado el rango intercuartil en cada caso, según corresponda.

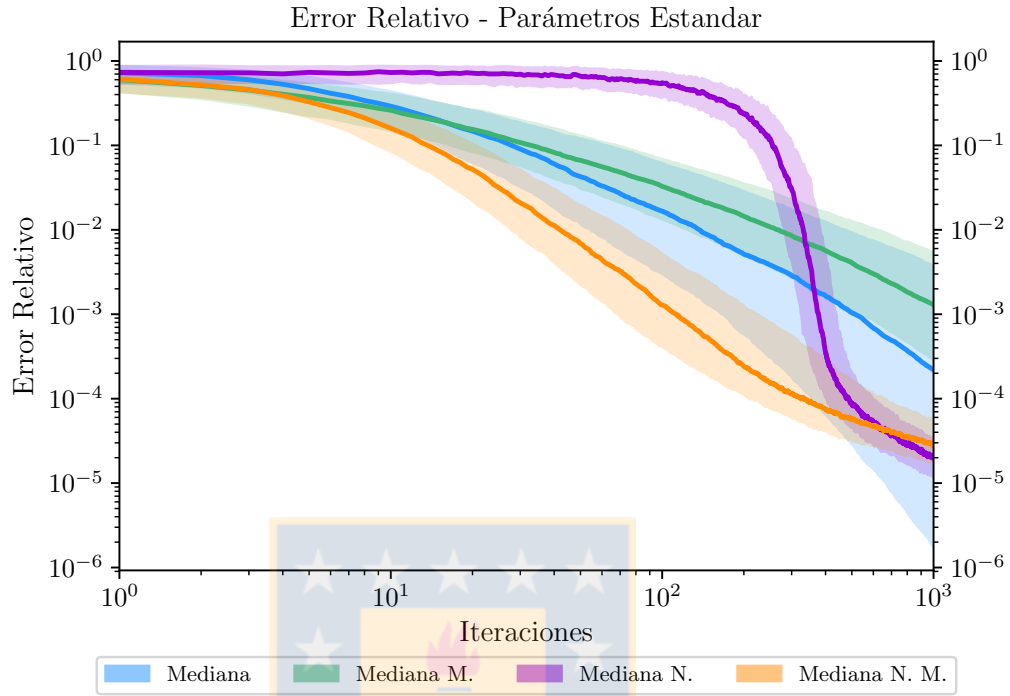


Figura 6.29: Error relativo de CSPA y CSPAN para las ganancias estándar sobre 10^3 estados puros generados uniformemente. El gráfico presenta el error relativo versus número de iteraciones en escala logarítmica. Las líneas continuas indican la mediana y el sombreado el rango intercuartil en cada caso, según corresponda.

6.5. Simulaciones con ensemble finito

En las secciones anteriores se mostraron los resultados de las simulaciones realizadas con estadística exacta. A continuación se mostrará el procedimiento que se utilizó para el desarrollo de simulaciones para un ensemble finito.

Primero, recordemos la definición de valor esperado mostrada en la expresión eq. (4.43).

$$E_{ij} = \text{Tr} \{ \rho (A_i \otimes B_j) \}, \quad (6.24)$$

donde $i, j = 1, 2$ son las etiquetas de los observables para Alice y Bob. Se definen los operadores de proyección para Alice y Bob de acuerdo a

$$\begin{aligned}
P_{a,i} &:= |a_i\rangle\langle a_i|; & P_{a,i}^\perp &:= |a_i^\perp\rangle\langle a_i^\perp| \\
P_{b,j} &:= |b_j\rangle\langle b_j|; & P_{b,j}^\perp &:= |b_j^\perp\rangle\langle b_j^\perp|.
\end{aligned} \tag{6.25}$$

Luego,

$$A_i = +P_{a,i} - P_{a,i}^\perp; \quad B_j = +P_{b,j} - P_{b,j}^\perp \tag{6.26}$$

Con esto, la eq. (6.24) puede ser rescrita como

$$E_{ij} = \text{Tr} \{ \rho (+P_{a,i} - P_{a,i}^\perp) \otimes (+P_{b,j} - P_{b,j}^\perp) \} \tag{6.27}$$

$$E_{ij} = \text{Tr} \{ \rho (P_{a,i} \otimes P_{b,j} - P_{a,i} \otimes P_{b,j}^\perp - P_{a,i}^\perp \otimes P_{b,j} + P_{a,i}^\perp \otimes P_{b,j}^\perp) \} \tag{6.28}$$

$$\begin{aligned}
E_{ij} &= \text{Tr} \{ \rho (P_{a,i} \otimes P_{b,j}) \} - \text{Tr} \{ \rho (P_{a,i} \otimes P_{b,j}^\perp) \} \\
&\quad - \text{Tr} \{ \rho (P_{a,i}^\perp \otimes P_{b,j}) \} + \text{Tr} \{ \rho (P_{a,i}^\perp \otimes P_{b,j}^\perp) \},
\end{aligned} \tag{6.29}$$

donde, por ejemplo, $\text{Tr} \{ \rho (P_{a,i} \otimes P_{b,j}) \}$ representa la probabilidad de obtener el valor propio +1 asociado al proyector $P_{a,i}$ de Alice en conjunto con el valor propio +1 asociado al proyector $P_{b,j}$ de Bob. Por simplicidad definimos cada uno de estos términos como

$$\begin{aligned}
\mathcal{P}_{pp} &= \text{Tr} \{ \rho (P_{a,i} \otimes P_{b,j}) \}, & \mathcal{P}_{pm} &= \text{Tr} \{ \rho (P_{a,i} \otimes P_{b,j}^\perp) \} \\
\mathcal{P}_{mp} &= \text{Tr} \{ \rho (P_{a,i}^\perp \otimes P_{b,j}) \}, & \mathcal{P}_{mm} &= \text{Tr} \{ \rho (P_{a,i}^\perp \otimes P_{b,j}^\perp) \}.
\end{aligned} \tag{6.30}$$

Consideremos las probabilidades de obtener las correlaciones mostradas anteriormente $\mathcal{P}_{pp}, \mathcal{P}_{pm}, \mathcal{P}_{mp}, \mathcal{P}_{mm}$. Estas probabilidades deben satisfacer

$$\mathcal{P}_{pp} + \mathcal{P}_{pm} + \mathcal{P}_{mp} + \mathcal{P}_{mm} = 1. \quad (6.31)$$

De esta forma podemos separar el intervalo $[0, 1]$ en cuatro zonas: $0 \leq z_1 < \mathcal{P}_{pp}$, $\mathcal{P}_{pp} \leq z_2 < \mathcal{P}_{pp} + \mathcal{P}_{mp}$, $\mathcal{P}_{pp} + \mathcal{P}_{mp} \leq z_3 < \mathcal{P}_{pp} + \mathcal{P}_{mp} + \mathcal{P}_{mm}$ y $\mathcal{P}_{pp} + \mathcal{P}_{mp} + \mathcal{P}_{mm} < z_4 \leq 1$, como se ilustra en la siguiente figura.

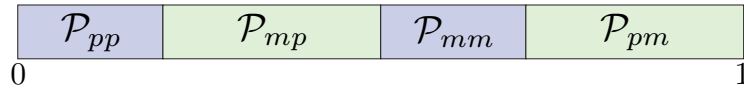


Figura 6.30: Diagrama de las zonas para emular detecciones experimentales.

Fuente: Elaboración propia

Supongamos que se genera un ensemble de N números al azar entre 0 y 1, con los cuales se realizará un conteo de cuantos coinciden en cada zona. Así, η_{pp} números se encuentran en z_1 , η_{mp} en z_2 , η_{mm} en z_3 y η_{pm} en z_4 . De esta forma podemos estimar la probabilidad de que estos números aleatorios coincidan en las diferentes zonas,

$$\tilde{\mathcal{P}}_{pp} = \frac{\eta_{pp}}{N}, \quad \tilde{\mathcal{P}}_{pm} = \frac{\eta_{pm}}{N}, \quad \tilde{\mathcal{P}}_{mp} = \frac{\eta_{mp}}{N}, \quad \tilde{\mathcal{P}}_{mm} = \frac{\eta_{mm}}{N}. \quad (6.32)$$

Finalmente, con estas probabilidades y los valores propios asociados a ellas, podemos calcular el valor estimado \tilde{E}_{ij} para esta simulación de detecciones,

$$\tilde{E}_{ij} = \frac{\eta_{pp} - \eta_{pm} - \eta_{mp} + \eta_{mm}}{N} = \tilde{\mathcal{P}}_{pp} - \tilde{\mathcal{P}}_{pm} - \tilde{\mathcal{P}}_{mp} + \tilde{\mathcal{P}}_{mm}. \quad (6.33)$$

donde, para un ensemble suficientemente grande, tenemos

$$E_{ij} = \lim_{N \rightarrow \infty} \tilde{E}_{ij}. \quad (6.34)$$

En las siguientes secciones se reportarán todas las simulaciones realizadas con los algoritmos de SPSA, CSPA y CSPAN para las distintas ganancias mostradas anteriormente.

6.5.1. Ensemble con $N = 10^2$

A continuación se mostrarán las simulaciones realizadas utilizando los métodos de SPSA, CSPSA, CSPSAN para una muestra de tamaño $N = 10^2$.

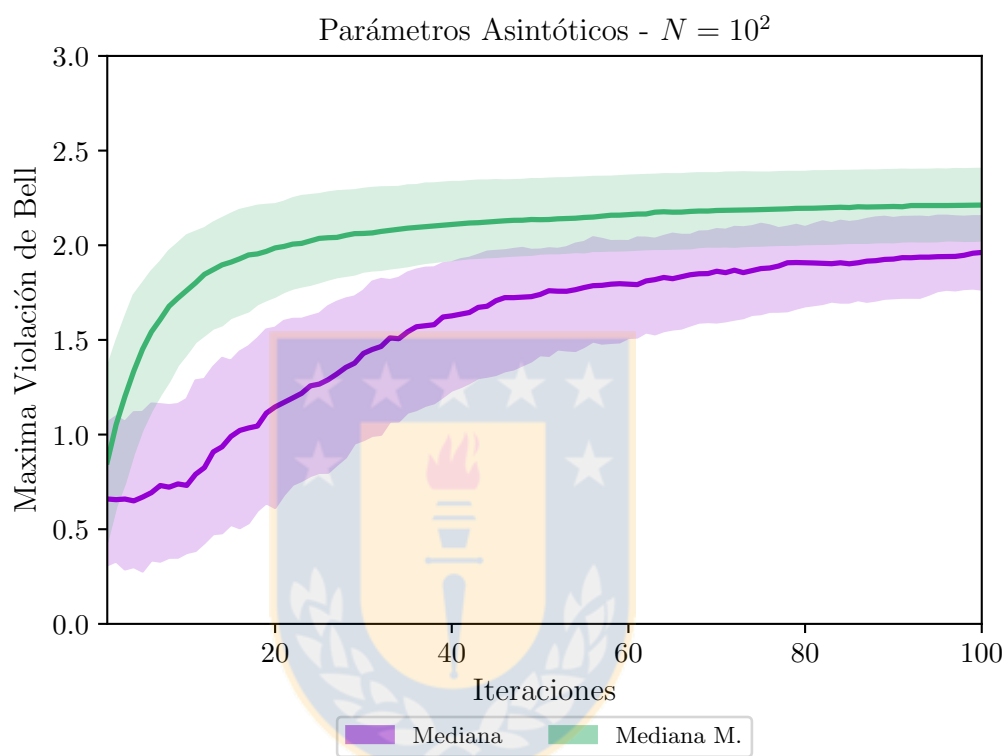


Figura 6.31: Simulaciones de SPSA para las ganancias asintóticas sobre 10^3 estados generados uniformemente con un tamaño de muestra $N = 10^2$.

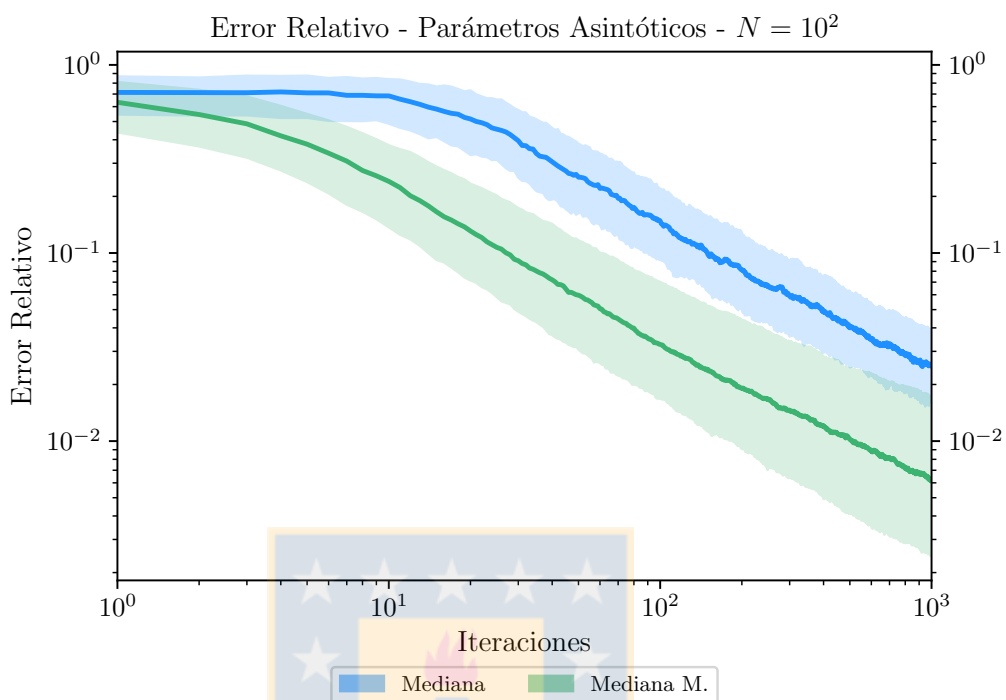


Figura 6.32: Error relativo de SPSA para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^2$.

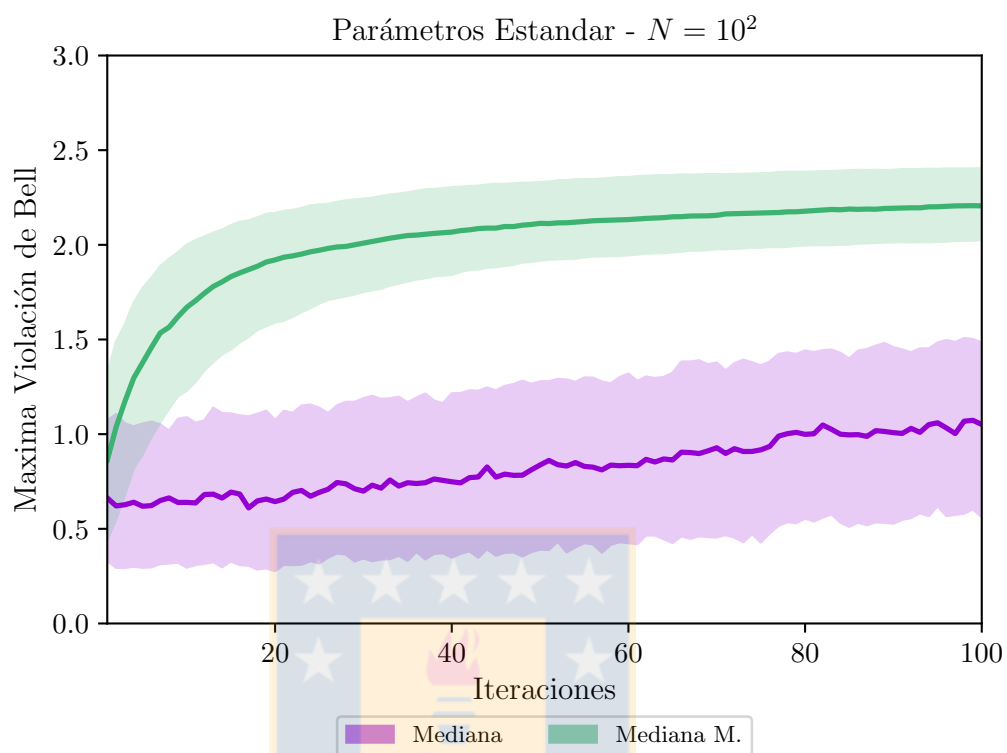


Figura 6.33: Simulaciones de SPSA para las ganancias estándar sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^2$.

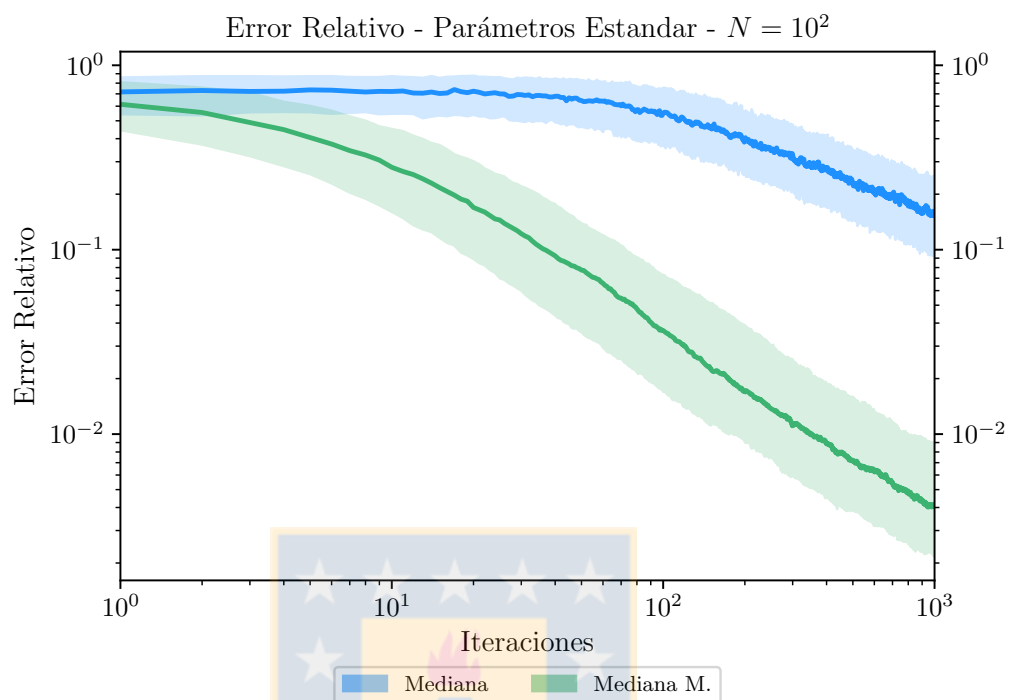


Figura 6.34: Error relativo de SPSA para las ganancias estándar sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^2$.

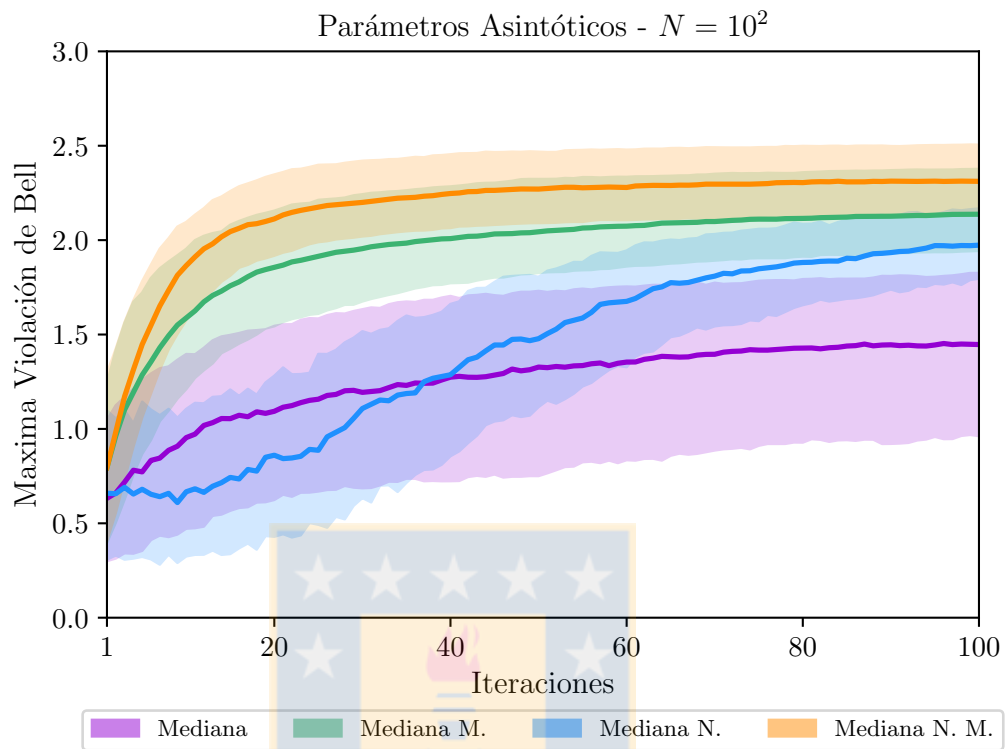


Figura 6.35: Simulaciones de \mathbb{CSPSA} y \mathbb{CSPSAN} para las ganancias asintóticas sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^2$.

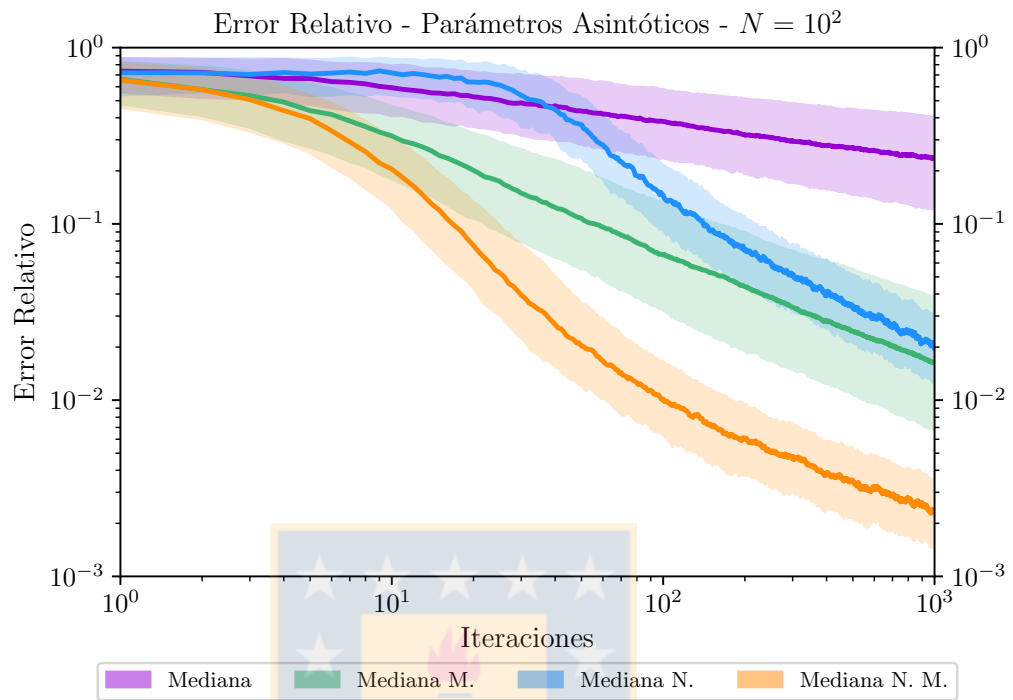


Figura 6.36: Error relativo de \mathbb{C} SPSA y \mathbb{C} SPSAN para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^2$.

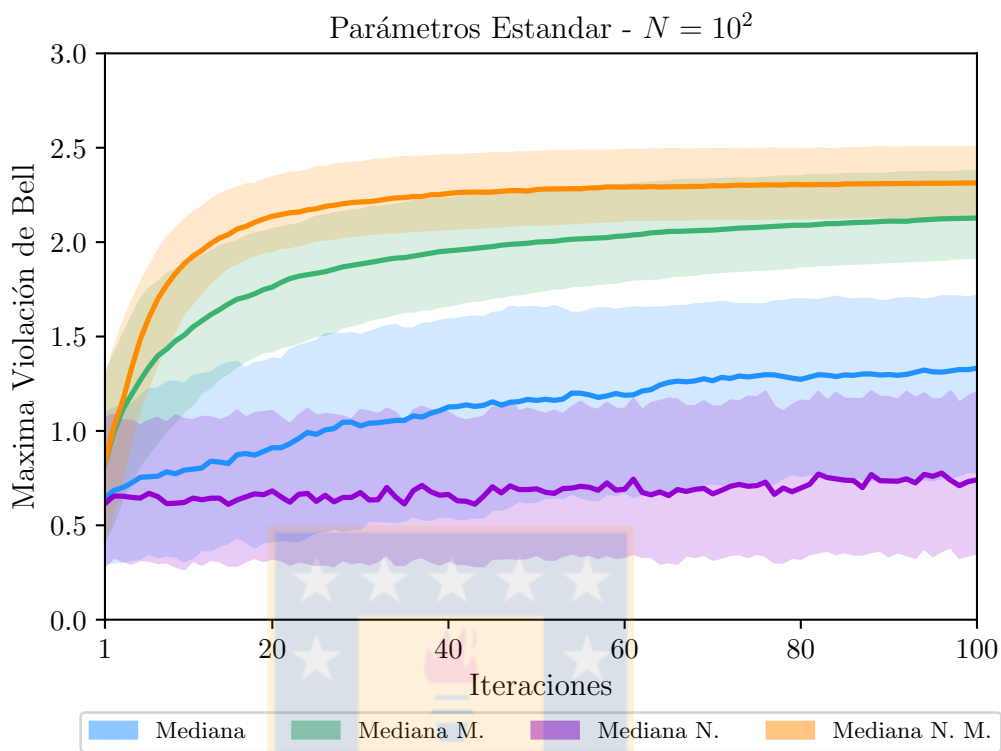


Figura 6.37: Simulaciones de CSPSA y CSPSAN para las ganancias estándar sobre 10^3 estados generados uniformemente y $N = 10^2$ experimentos.

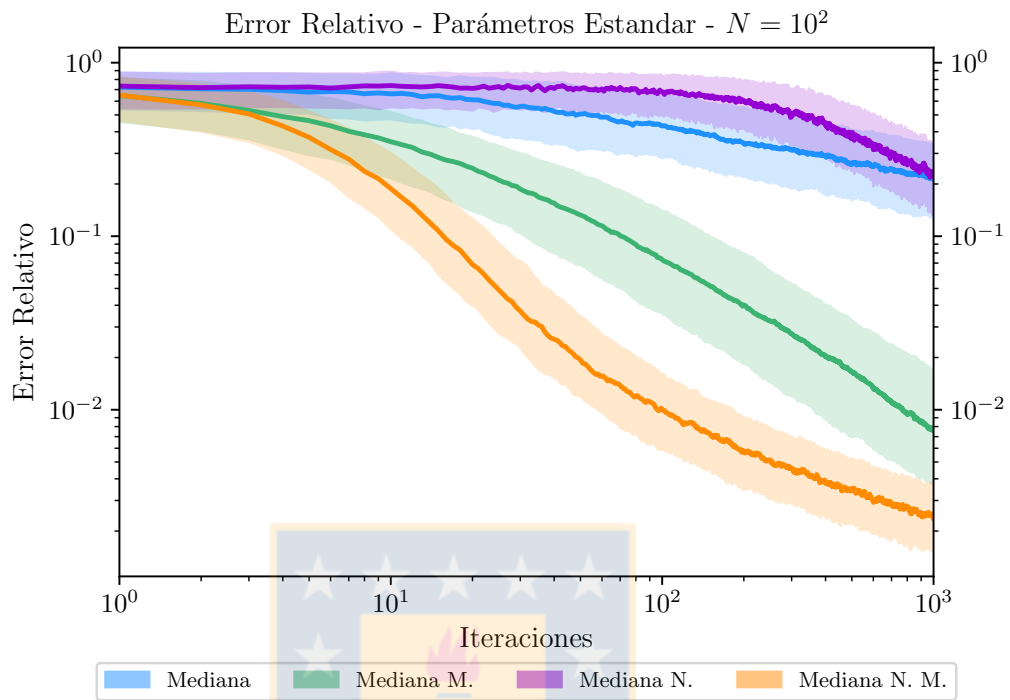


Figura 6.38: Error relativo de \mathbb{C} SPSA y \mathbb{C} SPSAN para las ganancias estándar sobre 10^3 estados puros generados uniformemente y $N = 10^2$ experimentos.

6.5.2. Ensemble con $N = 10^3$

A continuación se mostrarán las simulaciones realizadas utilizando los métodos de SPSA, \mathbb{C} SPSA, \mathbb{C} SPSAN para una muestra de tamaño $N = 10^3$.

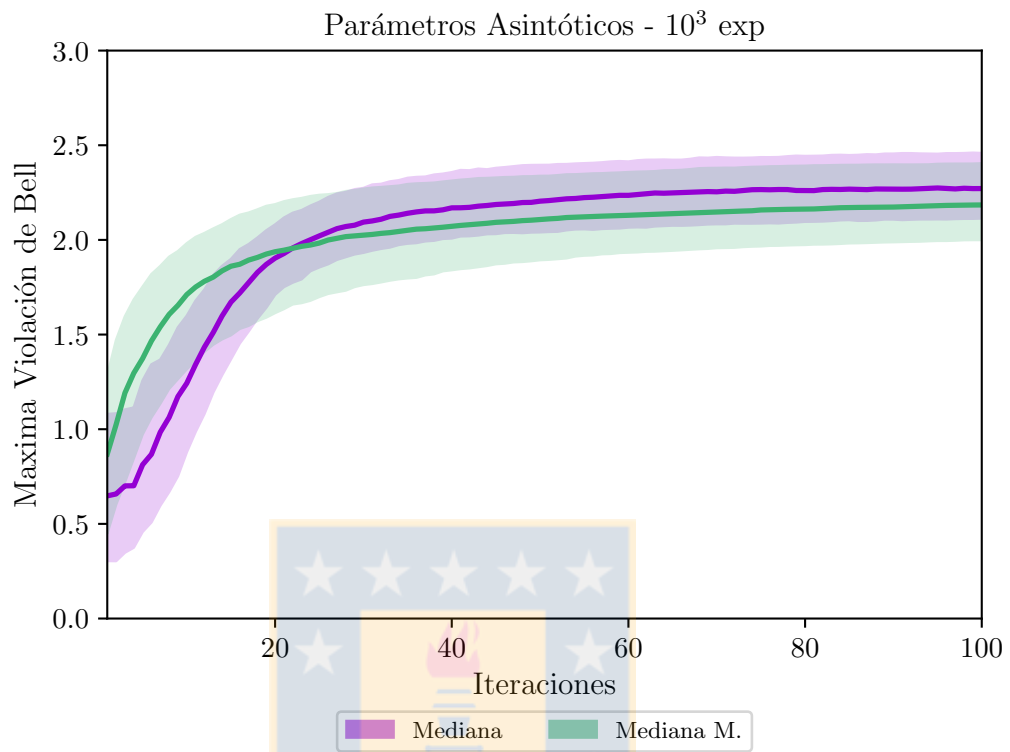


Figura 6.39: Simulaciones de SPSA para las ganancias asintóticas sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^3$.

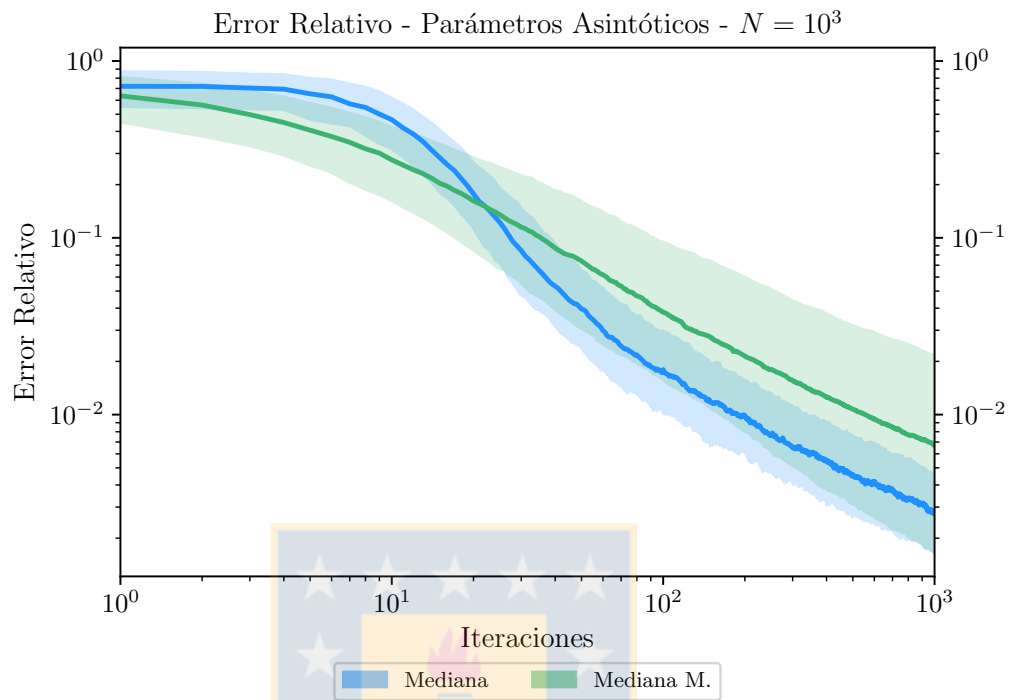


Figura 6.40: Error relativo de SPSA para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^3$.

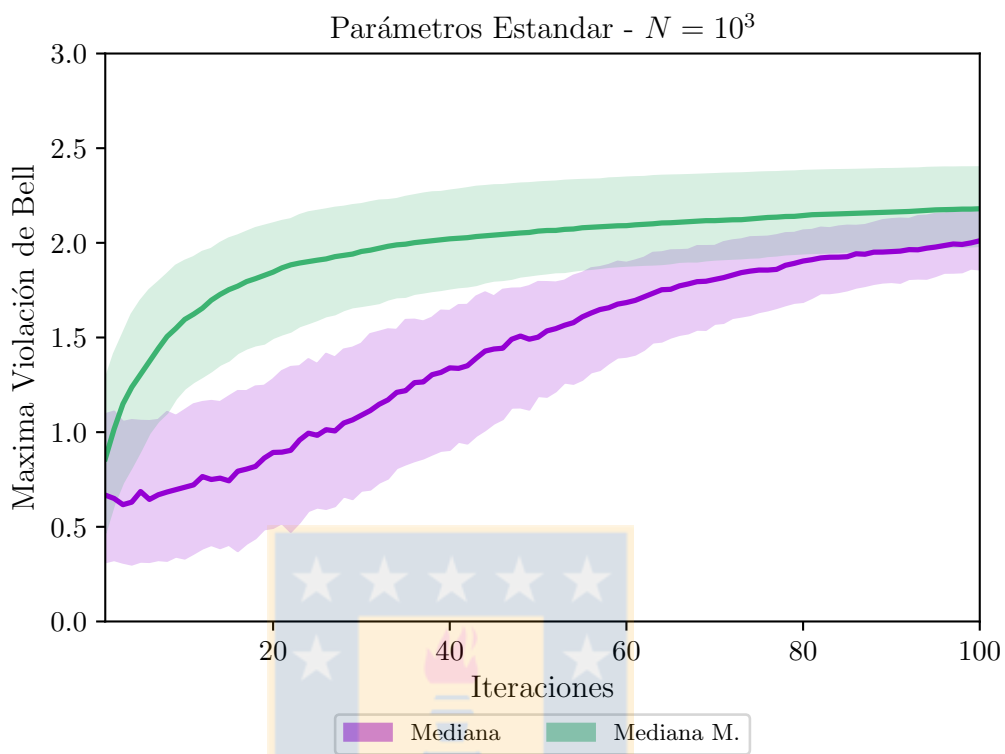


Figura 6.41: Simulaciones de SPSA para las ganancias estándar sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^3$.

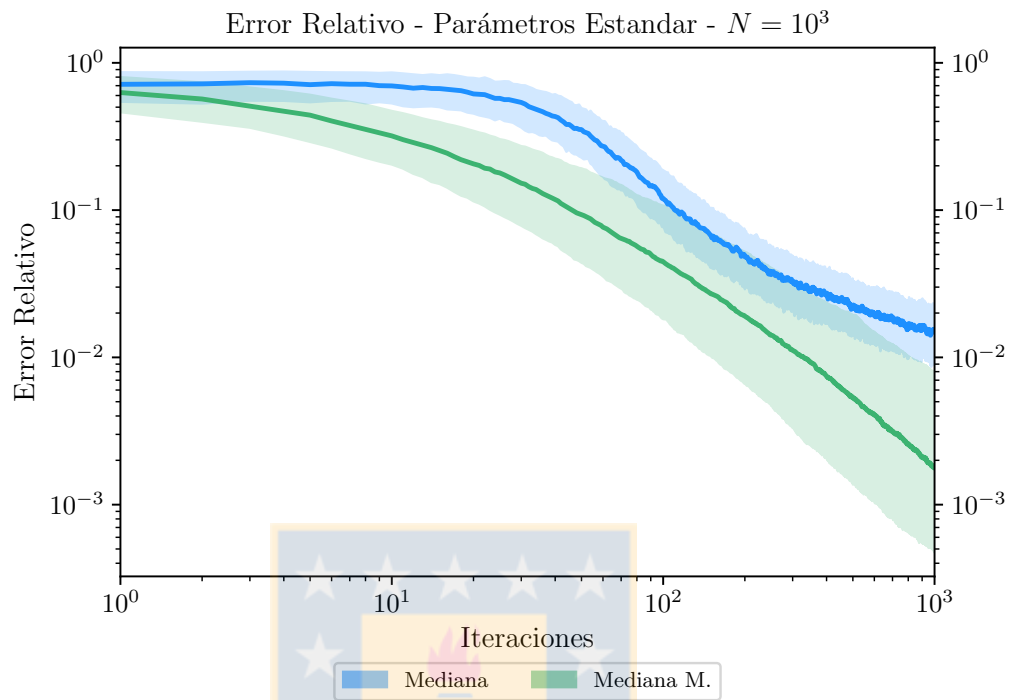


Figura 6.42: Error relativo de SPSA para las ganancias estándar sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^3$.

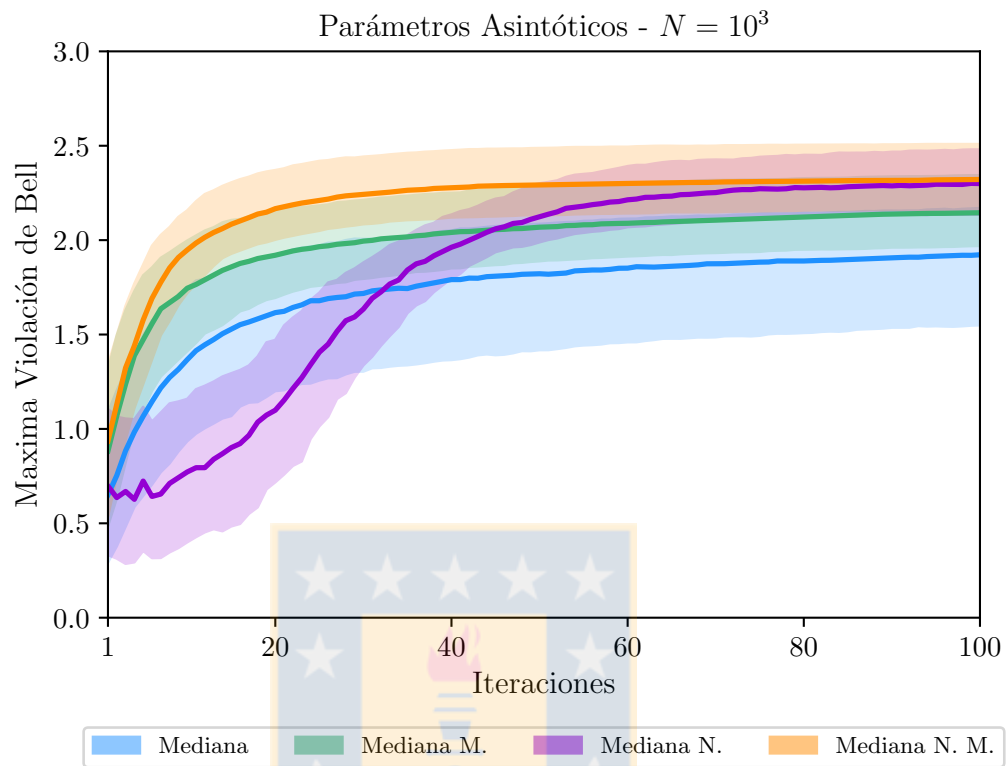


Figura 6.43: Simulaciones de CSPA y CSPSAN para las ganancias asintóticas sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^3$.

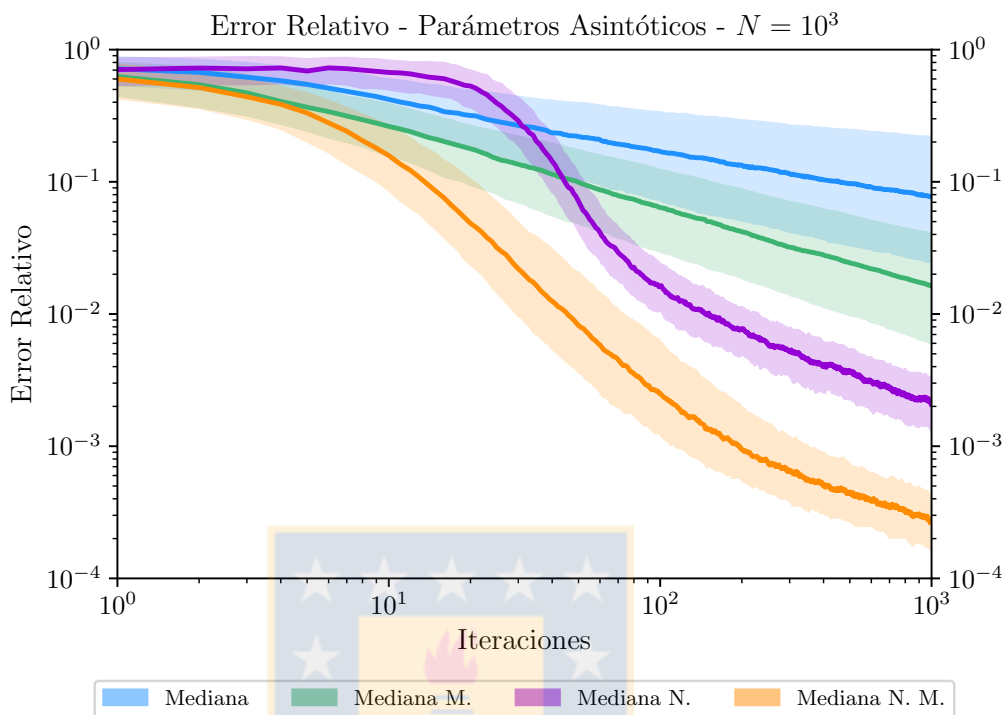


Figura 6.44: Error relativo de \mathbb{C} SPSA y \mathbb{C} SPSAN para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^3$.

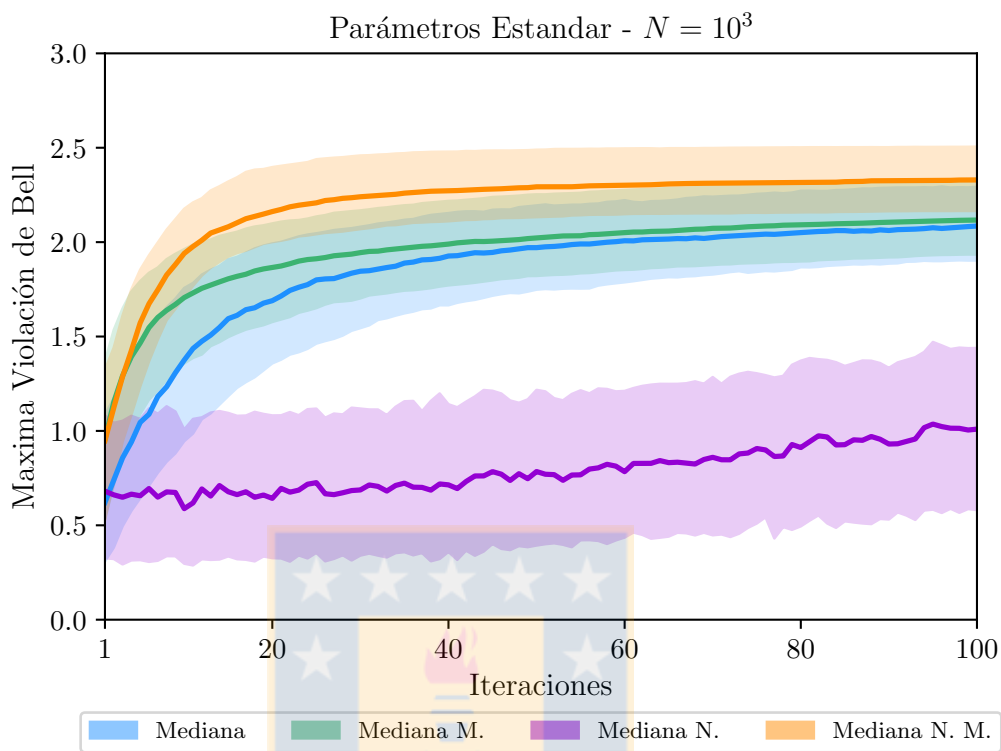


Figura 6.45: Simulaciones de CSPSA y CSPSAN para las ganancias estándar sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^3$.

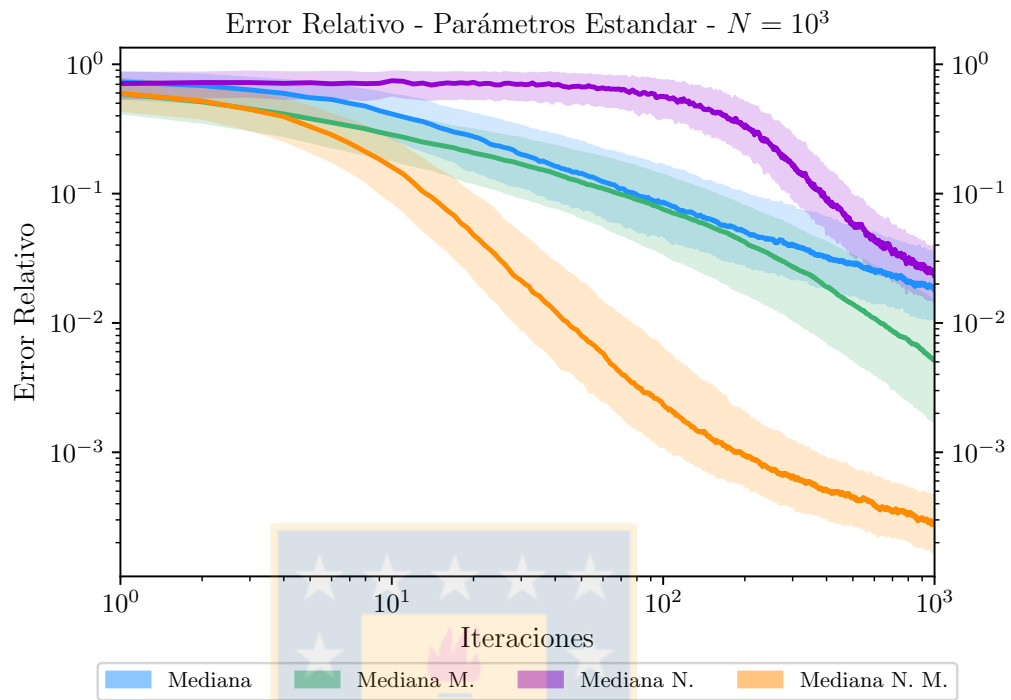


Figura 6.46: Error relativo de \mathbb{C} SPSA y \mathbb{C} SPSAN para las ganancias estándar sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^3$.

6.5.3. Ensemble con $N = 10^4$

A continuación se mostrarán las simulaciones realizadas utilizando los métodos de SPSA, \mathbb{C} SPSA, \mathbb{C} SPSAN para una muestra de tamaño $N = 10^4$.

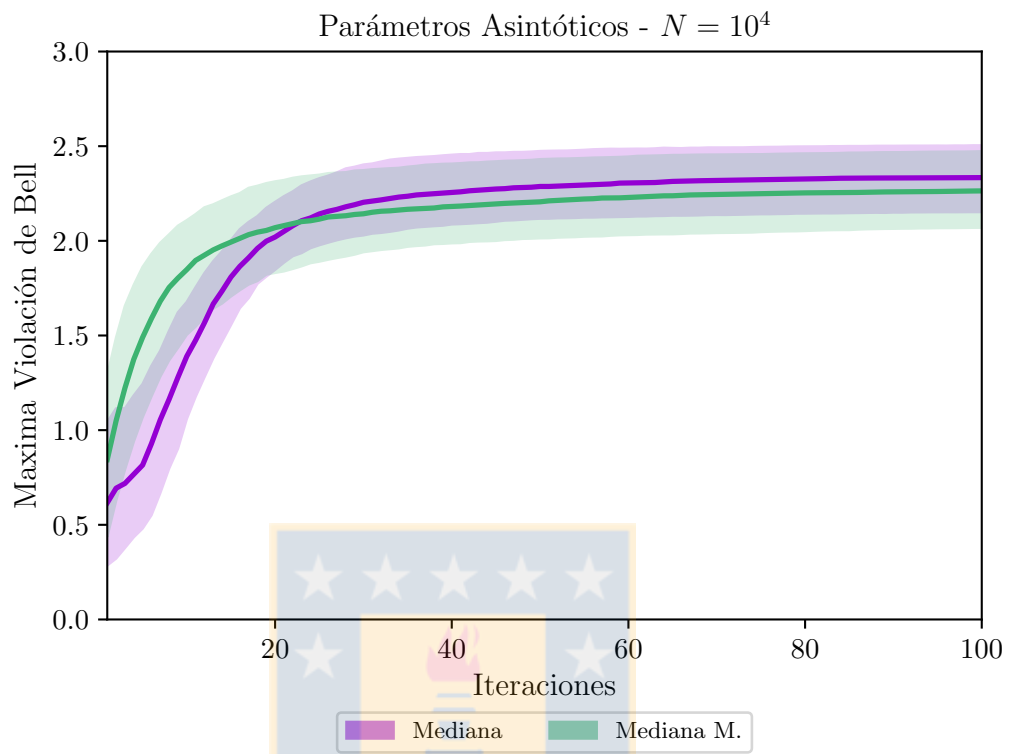


Figura 6.47: Simulaciones de SPSA para las ganancias asintóticas sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^4$.

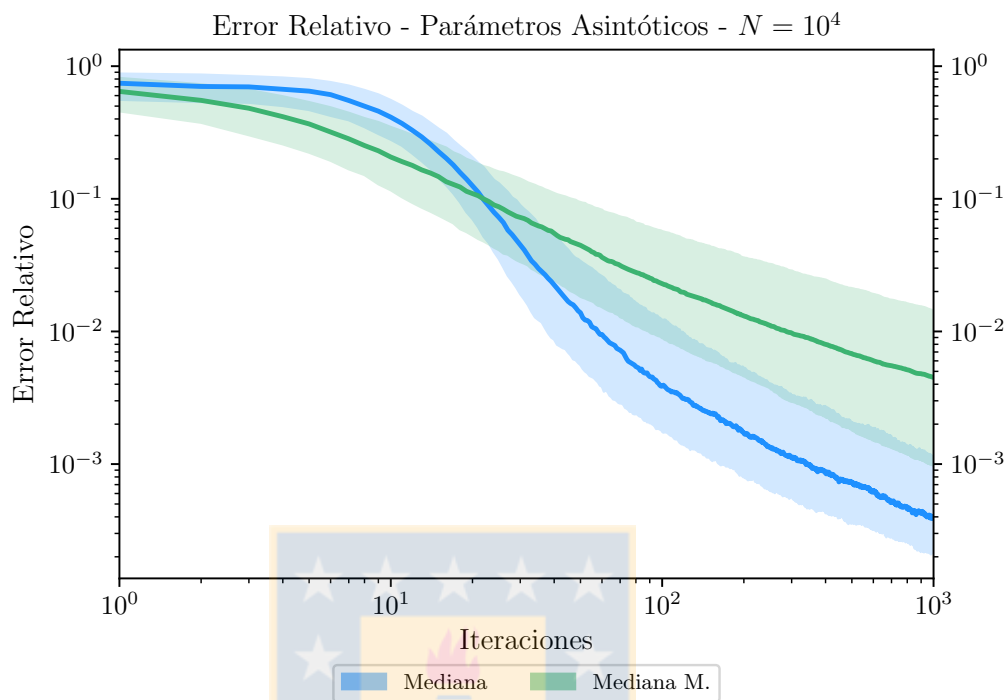


Figura 6.48: Error relativo de SPSA para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^4$.

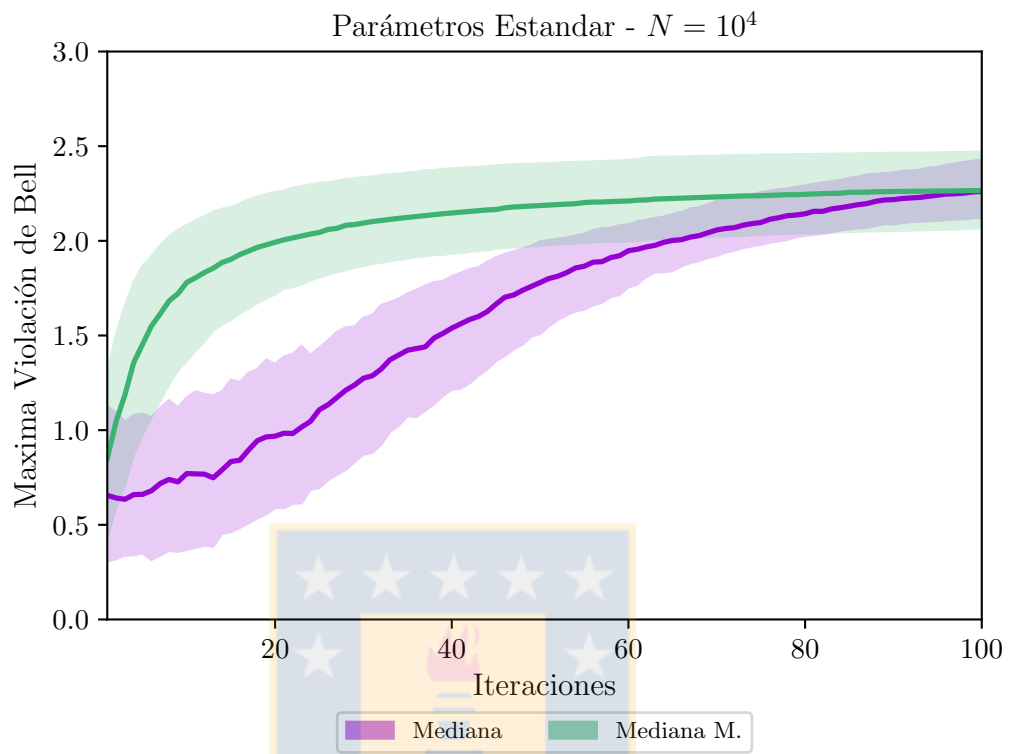


Figura 6.49: Simulaciones de SPSA para las ganancias estándar sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^4$.

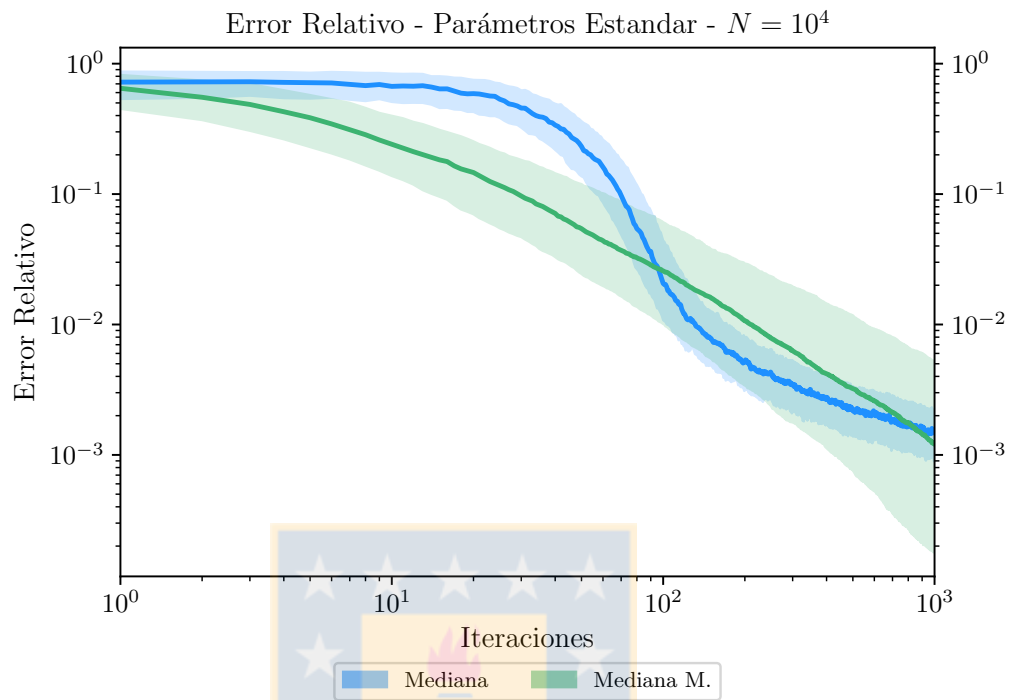


Figura 6.50: Error relativo de SPSA para las ganancias estándar sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^4$.

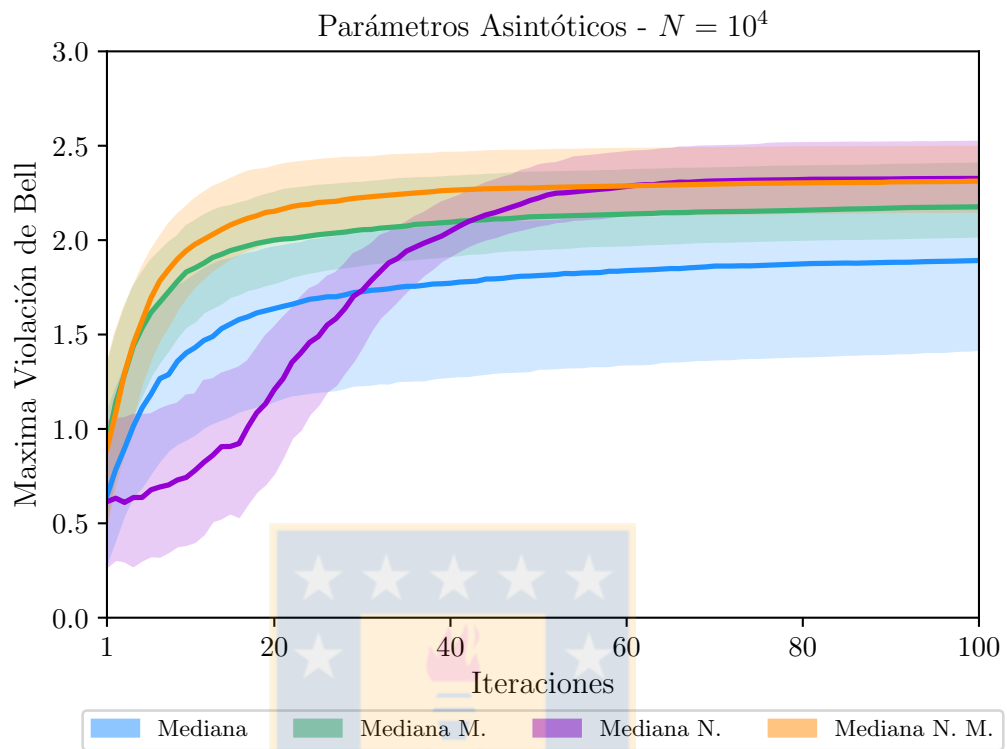


Figura 6.51: Simulaciones de CSPA y CSPSAN para las ganancias asintóticas sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^4$.

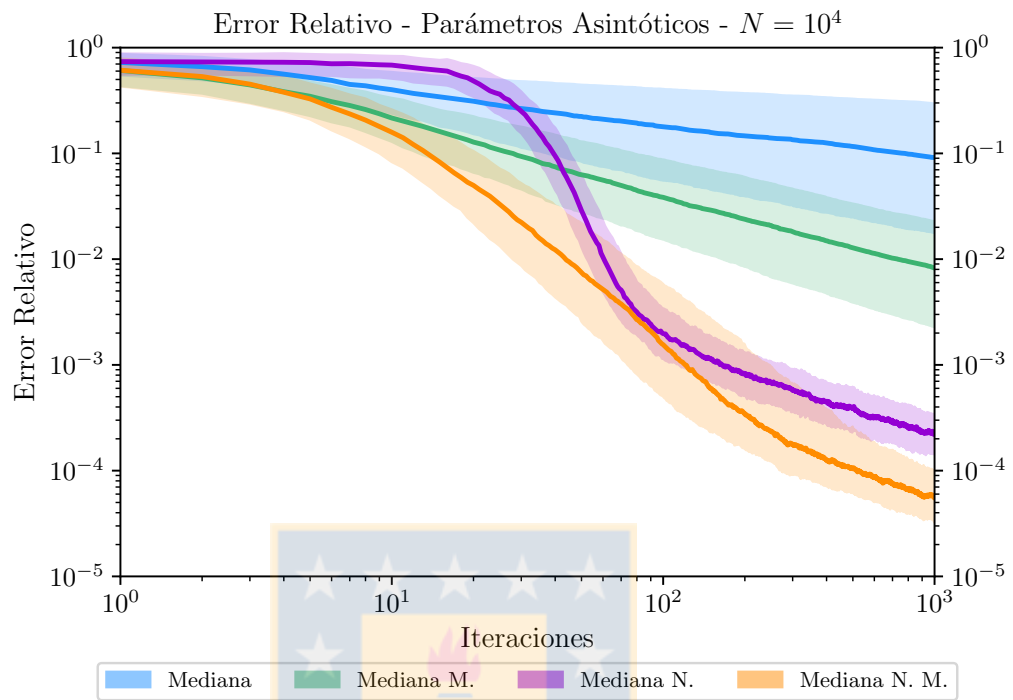


Figura 6.52: Error relativo de \mathbb{C} SPSA y \mathbb{C} SPSAN para las ganancias asintóticas sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^4$.

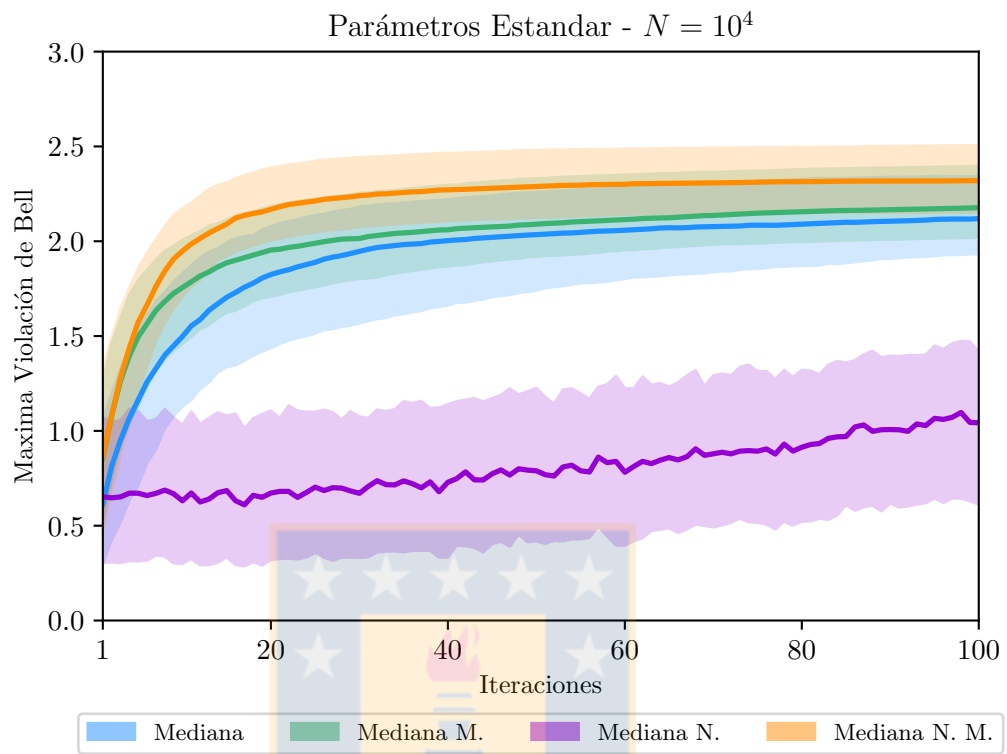


Figura 6.53: Simulaciones de CSPA y CSPAN para las ganancias estándar sobre 10^3 estados generados uniformemente con un ensemble de $N = 10^4$.

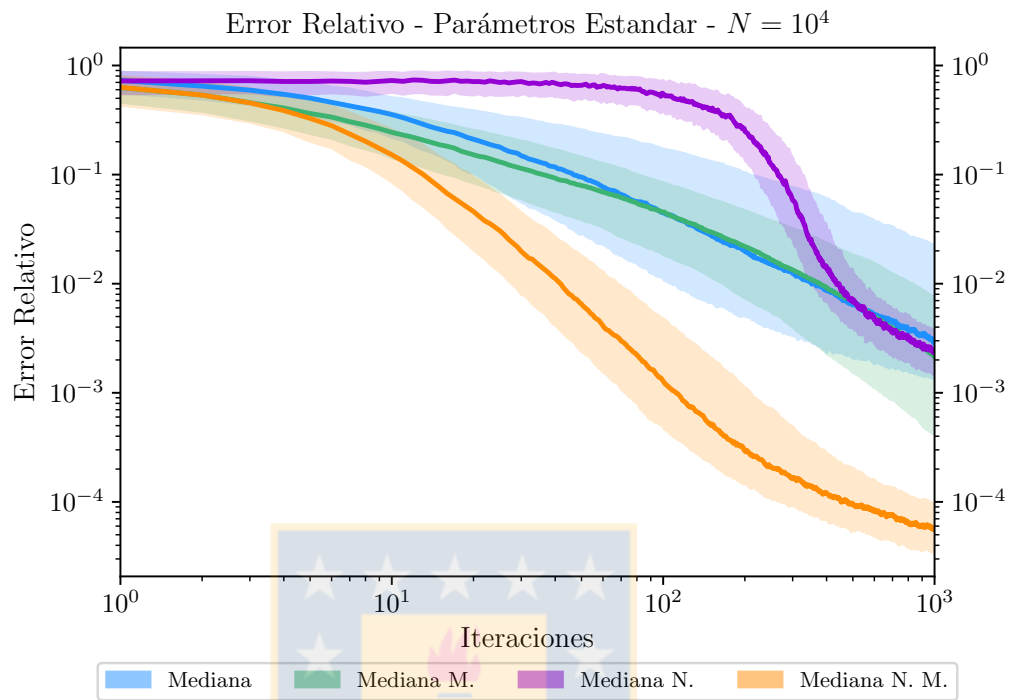


Figura 6.54: Error relativo de \mathbb{CSPSA} y \mathbb{CSPSAN} para las ganancias estándar sobre 10^3 estados puros generados uniformemente con un ensemble de $N = 10^4$.

6.6. Resumen de resultados

Debido a la gran cantidad de simulaciones que fueron realizadas, en esta sección se compararán los mejores resultados obtenidos para cada valor de N , en término de la rapidez de convergencia y en el error relativo.

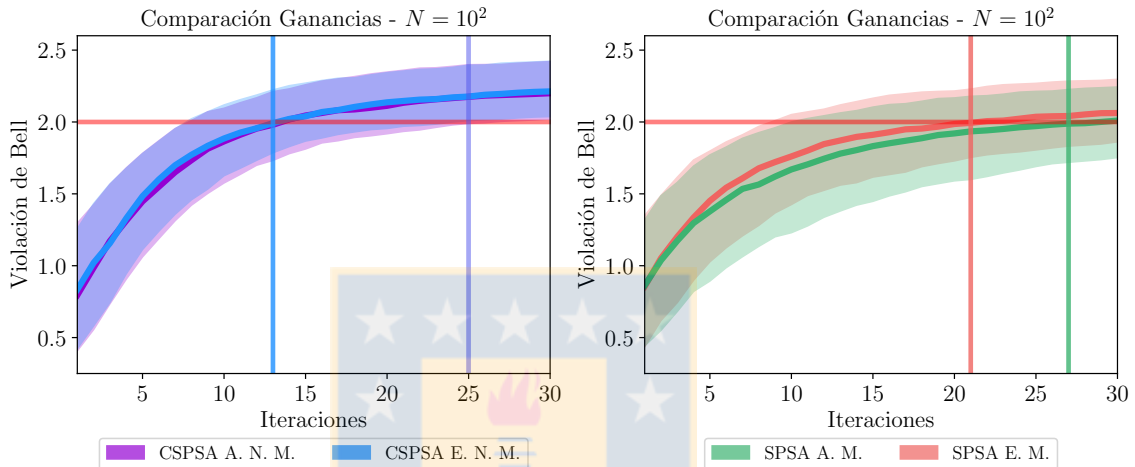


Figura 6.55: Rapidez de convergencia para un ensemble de $N = 10^2$. Violación de Bell vs iteraciones. El gráfico de la izquierda muestra la comparación las ganancias estándar modificadas y asíntóticas modificadas utilizadas en el algoritmo CSPSAN. La figura de la derecha muestra la respectiva comparación para SPSA para las mismas ganancias. Las líneas continuas representan la mediana y el sombreado el rango intercuartil, en ambos casos.

En la [Figura 6.55](#) se muestra la comparación de rapidez de convergencia para los algoritmos de CSPSAN y SPSA. Se observa que, para CSPSAN, las ganancias describen las mismas curvas, donde la mediana llega al valor 2.0 a las 13 iteraciones considerando además que todo el rango intercuartil supera este valor en las 25 iteraciones. Esto significa que, al menos, el 75% de las realizaciones viola la desigualdad CHSH a las 25 iteraciones. Para SPSA se observa que la mediana de las ganancias asíntóticas modificadas viola la desigualdad de Bell a las 21 iteraciones, mientras que en el caso de las ganancias estándar modificadas esto ocurre cerca de las 27 iteraciones. En este caso, el rango intercuartil de cada curva no viola la desigualdad como en el caso de CSPSAN, por lo que menos realizaciones logran una violación de Bell para la misma cantidad de iteraciones. Al menos para el ensemble de $N = 10^2$, CSPSAN exhibe un mejor rendimiento que SPSA.

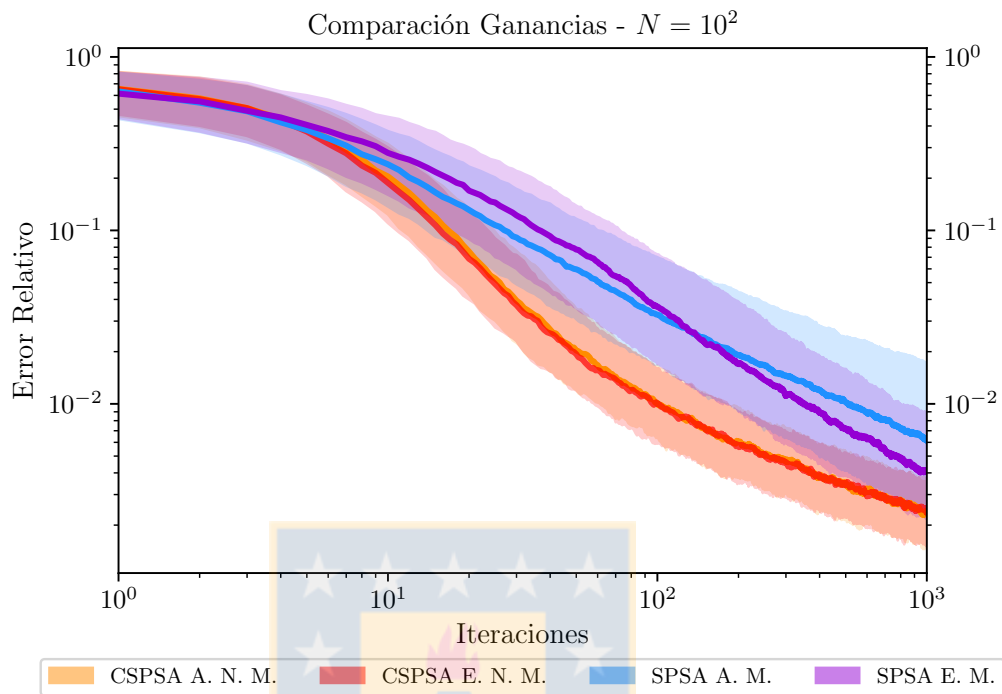


Figura 6.56: Error relativo para un ensemble de $N = 10^2$. El gráfico muestra las curvas logradas para las ganancias estándar modificadas y asíntóticas modificadas para los algoritmos CSPA y SPSA. Las líneas continuas representan la mediana y el sombreado el rango intercuartil para cada caso.

El panorama para la comparación del error relativo es muy similar. En la [Figura 6.56](#) se muestra las curvas descritas por las distintas ganancias usadas en SPSA y CSPA normalizado. Vemos que las ganancias asíntóticas y estándar modificadas reproducen los mismos resultados, al igual que en el caso de la rapidez de convergencia. Vemos que para estas ganancias un 50 % de las realizaciones consiguen un error relativo más pequeño. Por otra parte la dispersión de las realizaciones es menor en CSPA que en SPSA.

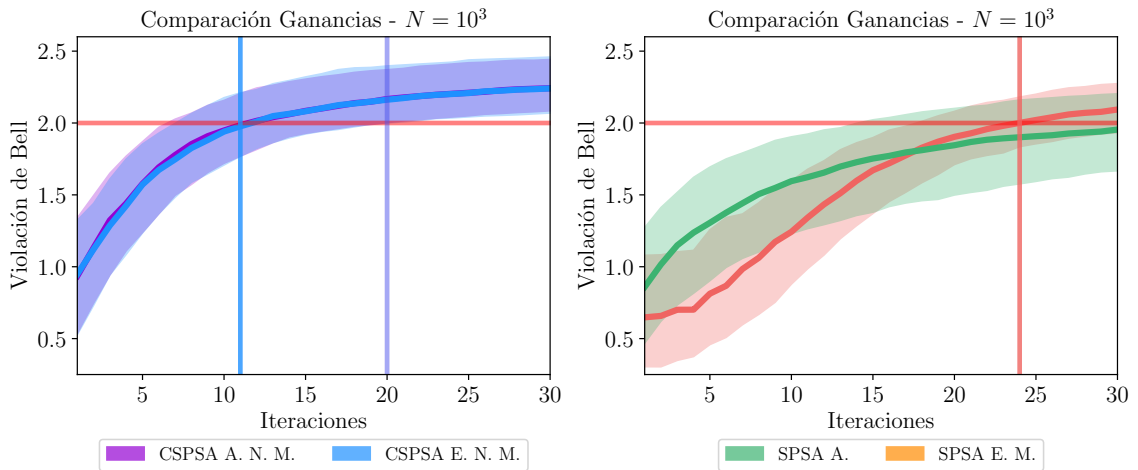


Figura 6.57: Rapidez de convergencia para un ensemble de $N = 10^3$. Violación de Bell vs iteraciones. Las líneas continuas representan la mediana y el sombreado el rango intercuartil, en ambos casos.

En la [Figura 6.57](#) se muestra la comparación de rapidez de convergencia para los algoritmos de C \mathcal{S} PSAN y SPSA. Se observa que, para C \mathcal{S} PSAN, las ganancias describen las mismas curvas, donde la mediana llega al valor 2.0 a las 11 iteraciones considerando además que todo el rango intercuartil supera este valor en las 20 iteraciones. Esto significa que, al menos, el 75 % de las realizaciones viola la desigualdad CHSH a las 20 iteraciones. Para SPSA se observa que la mediana de las ganancias asintóticas viola la desigualdad de Bell a las 24 iteraciones, mientras que en el caso de las ganancias estándar modificadas esto ocurre cerca de las 32 iteraciones. En este caso, el rango intercuartil de cada curva no viola la desigualdad como en el caso de C \mathcal{S} PSAN, por lo que menos realizaciones logran una violación de Bell para la misma cantidad de iteraciones. Para el ensemble de $N = 10^3$, C \mathcal{S} PSAN exhibe un mejor rendimiento que SPSA.

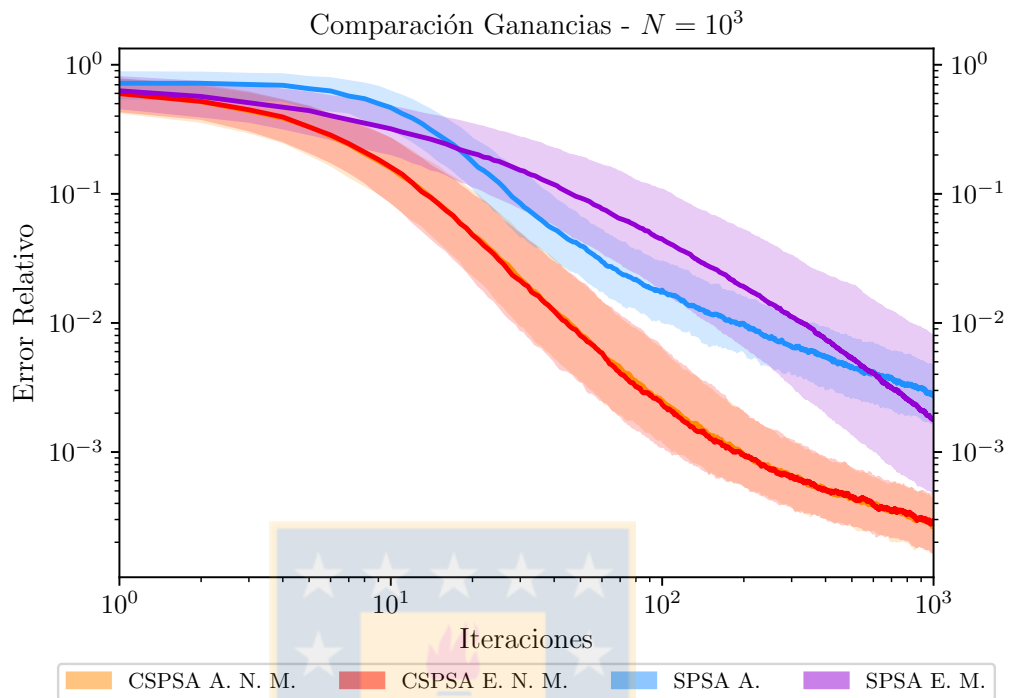


Figura 6.58: Error relativo para un ensemble de $N = 10^3$. El gráfico muestra las curvas logradas para las ganancias estándar modificadas y asíntoticas modificadas para CSPA y ganancias estándar modificadas y asíntoticas SPSA. Las líneas continuas representan la mediana y el sombreado el rango intercuartil para cada caso.

El panorama para la comparación del error relativo es muy similar. En la [Figura 6.58](#) se muestra las curvas descritas por las distintas ganancias usadas en SPSA y CSPA normalizado. Vemos que las ganancias asíntoticas y estándar modificadas reproducen los mismos resultados, al igual que en el caso de la rapidez de convergencia. Vemos que la mediana llega cerca de las 10^{-4} en el error relativo. A diferencia de lo mostrado para el caso del ensemble de $N = 10^2$, el 75% de las realizaciones para estas ganancias alcanza mejores resultados que los obtenidos por SPSA usando ganancias estándar modificadas.

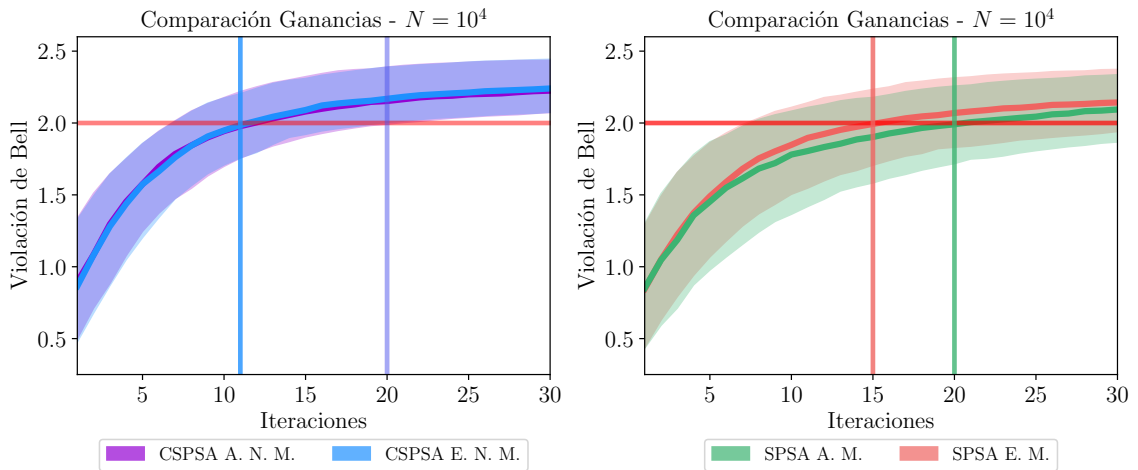


Figura 6.59: Rapidez de convergencia para un ensemble de $N = 10^4$. Violación de Bell vs iteraciones. Las líneas continuas representan la mediana y el sombreado el rango intercuartil, en ambos casos.

En la [Figura 6.59](#) se muestra la comparación de rapidez de convergencia para los algoritmos de C_{SPSA}N y SPSA. Se observa que, para C_{SPSA}N, las ganancias describen las mismas curvas, donde la mediana llega al valor 2.0 a las 11 iteraciones considerando además que todo el rango intercuartil supera este valor en las 20 iteraciones. Esto significa que, al menos, el 75 % de las realizaciones viola la desigualdad CHSH a las 20 iteraciones. Para SPSA se observa que la mediana de las ganancias asíntóticas modificadas viola la desigualdad de Bell a las 15 iteraciones, mientras que en el caso de las ganancias estándar modificadas esto ocurre cerca de las 20 iteraciones. En este caso, el rango intercuartil de cada curva no viola la desigualdad como en el caso de C_{SPSA}N, por lo que menos realizaciones logran una violación de Bell para la misma cantidad de iteraciones. Para el ensemble de $N = 10^4$, C_{SPSA}N exhibe un mejor rendimiento que SPSA.

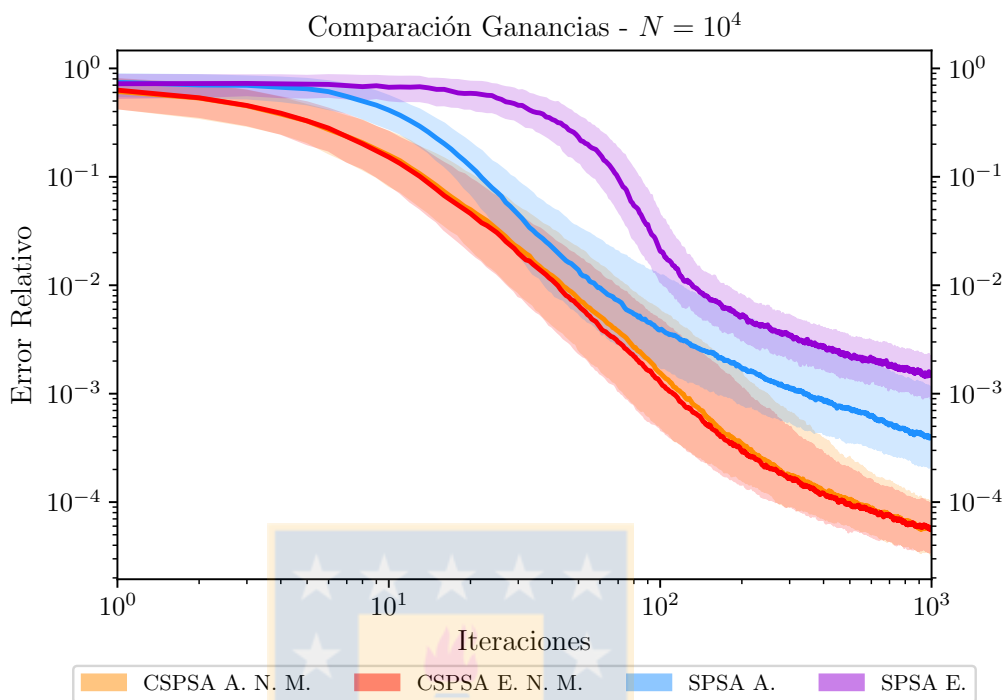


Figura 6.60: Error relativo para un ensemble de $N = 10^4$. El gráfico muestra las curvas logradas para las ganancias estándar modificadas y asíntoticas modificadas para CSPA y ganancias estándar y asíntoticas SPSA. Las líneas continuas representan la mediana y el sombreado el rango intercuartil para cada caso.

El panorama para la comparación del error relativo es muy similar. En la [Figura 6.60](#) se muestra las curvas descritas por las distintas ganancias usadas en SPSA y CSPA normalizado. Vemos que las ganancias asíntoticas y estándar modificadas reproducen los mismos resultados, al igual que en el caso de la rapidez de convergencia. Vemos que para estas ganancias un 50 % de las realizaciones consiguen un error relativo más pequeño. Por otra parte la dispersión de las realizaciones es menor en CSPA que en SPSA. Vemos que la mediana llega mas abajo de 10^{-4} en el error relativo. A diferencia de lo mostrado para el caso del ensemble de $N = 10^2$, al menos el 75 % de las realizaciones para estas ganancias alcanza mejores resultados que los obtenidos por SPSA usando ganancias estándar modificadas. En este caso CSPA en su versión normalizada alcanza aproximadamente un orden de magnitud menos que SPSA, es decir, es un orden de magnitud más preciso.

Los siguientes gráficos muestran el error relativo en función del ensemble N para las simulaciones mostradas anteriormente. Cada marca representa la mediana del error relativo obtenido a las 10^3 iteraciones. Las líneas continuas muestran los ajustes realizados para cada conjunto de valores en un gráfico log-log de acuerdo a la expresión,

$$E_{\text{rel}} = 10^{\beta} N^{\alpha} \quad (6.35)$$

donde α es la pendiente y β es el valor donde el ajuste intercepta el eje vertical.

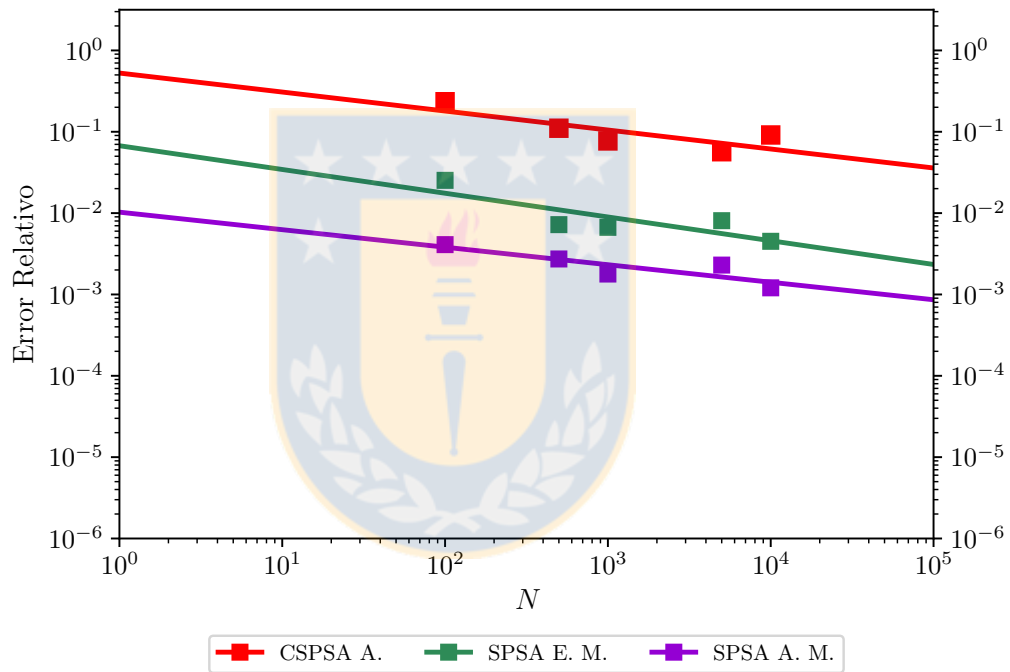


Figura 6.61: Gráfico error relativo vs ensemble N . Cada marca representa la mediana del error relativo obtenido a las 10^3 iteraciones. Las líneas continuas representan los ajustes lineales en un gráfico log-log para cada conjunto de datos.

	CSPSAA.	SPSA E. M.	SPSA A. M.
α	-0,2334	-0,2919	-0,2154
β	-0,2782	-1,1716	-1,9883

Tabla 6.1: Tabla comparativa con parámetros de ajuste de la [Figura 6.61](#). **Fuente:** Elaboración propia.

La [Tabla 6.1](#) muestra los valores para la pendiente α e intersección con el eje vertical β para las regresiones lineales realizadas para cada combinación combinación algoritmo-ganancia. Se puede observar que las pendientes están entre $-0,2$ y $-0,3$ lo que indica que las pendientes de las rectas son aproximadamente paralelas entre si en el rango de valores del ensemble con los que trabajamos. También se puede realizar una interpolación y calcular cuanto debiese ser la mediana del error si utilizamos un ensemble más grande. Por ejemplo, SPSA las ganancias asintóticas modificadas logra un error relativo de $8,6040 \times 10^{-4}$ para un ensemble de $N = 10^5$. Finalmente, podemos ver que el algoritmo de SPSA utilizando las ganancias asintóticas modificadas proporciona los mejores resultados del conjunto de simulaciones comparado.

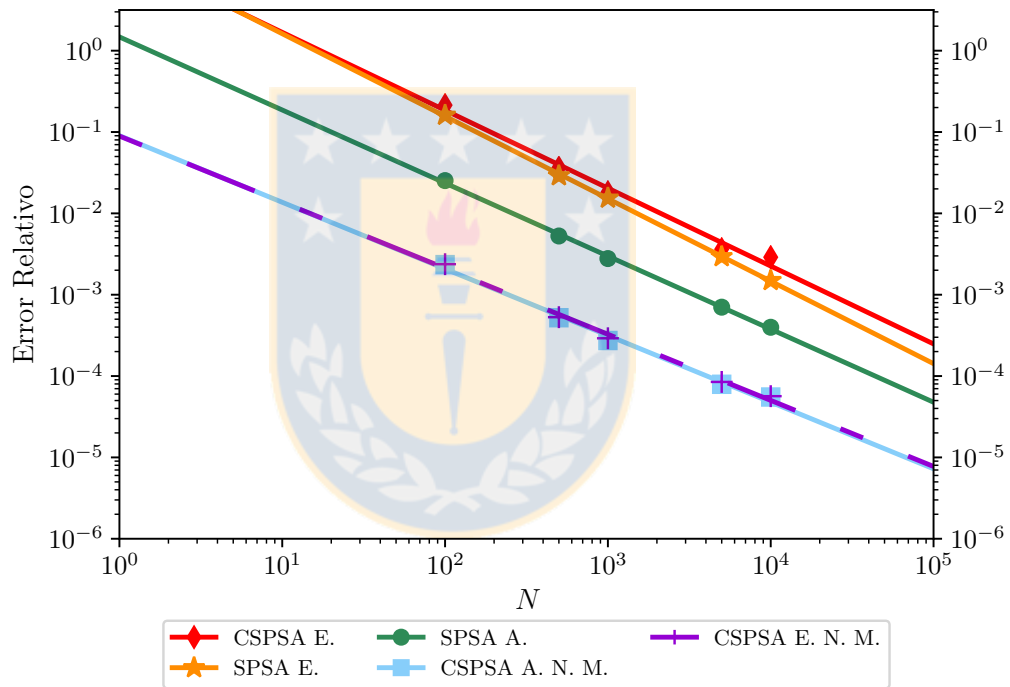


Figura 6.62: Gráfico error relativo vs ensemble N . Cada marca representa la mediana del error relativo obtenido a las 10^3 iteraciones. Las líneas continuas representan los ajustes lineales en un gráfico log-log para cada conjunto de datos.

	SPSA E.	CSPA E.	SPSA A.	CSPSAN A. M.	CSPSAN E. M.
α	-1,0139	-0,9575	-0,8981	-0,8190	-0,8124
β	1,2223	1,1826	0,1690	-1,0446	-1,0493

Tabla 6.2: Tabla comparativa con parámetros de ajuste de la [Figura 6.62](#). **Fuente:** Elaboración propia.

Análogamente, la [Tabla 6.2](#) muestra los valores para la pendiente e intersección con el eje vertical pero para otra combinación de algoritmo-ganancia. Para este caso las pendientes están entre $-0,80$ y $-1,02$. Podemos ver que las pendientes son aproximadamente paralelas para el intervalo entre $N = 10^0$ y $N = 10^5$. Realizando una interpolación, podemos ver que el error relativo para SPSA con ganancias asintóticas con $N = 10^5$ es de $4,7698 \times 10^{-5}$, mientras que el error relativo para CSPSAN con ganancias asintóticas y estándar modificadas es de $7,2510 \times 10^{-6}$ y $7,7393 \times 10^{-6}$. De este resultado podemos observar que CSPSAN con las ganancias asintóticas logran un menor error relativo si es comparada con SPSA y las ganancias asintóticas.



Capítulo 7

Conclusión

Se ha propuesto resolver el problema de la máxima violación de la desigualdad CHSH para estados cuánticos desconocidos mediante dos algoritmos iterativos llamados *Simultaneous Perturbation Stochastic Approximation* (SPSA) y *Complex Simultaneous Perturbation Stochastic Approximation* (CSPSA). Ambos métodos son bastante versátiles, puesto que pueden ser implementados para funciones cuyos gradientes no pueden ser obtenidos de manera exacta. Tanto SPSA como CSPSA están definidos por medio de parámetros que controlan la convergencia, los cuales son llamados ganancias. El vector inicial donde comienza la secuencia iterativa es llamado guess. Así, ambos algoritmos permiten la convergencia asintótica al máximo de la función pese a que el estado es desconocido.

Este problema se afrontó desde un punto de vista numérico, por lo que se realizaron exhaustivas simulaciones numéricas para estudiar el rendimiento medio de los algoritmos en la violación de la desigualdad CHSH. Para comparar se utilizó como medida de tendencia central la mediana y como medida de dispersión el rango intercuartil. Los casos estudiados fueron clasificados en dos grandes grupos: el primero es el método exacto, donde se consideró un experimento numérico en el cual la función es evaluada sin errores, y en el segundo se considera una fuente de error en las simulaciones numéricas, la cual consiste en utilizar un ensamble de tamaño finito para la evaluación de la función, emulando así un proceso de detección.

Hemos mostrado que los algoritmos de SPSA y CSPSA pueden ser utilizados en el problemas de la desigualdad CHSH aplicado a estados cuánticos de distinta naturaleza tales como: estados entrelazados, estados de Werner, estados en descomposición de Schmidt y estados cuánticos puros desconocidos generados uniformemente. Para cada uno de estos casos, las simulaciones reproducen las predicciones dadas por la teoría. Sin embargo, para los métodos SPSA y CSPSA utilizando las ganancias estándar y asintótica modificadas, se requieren menos de 50 iteraciones para lograr una violación de la desigualdad CHSH. También se observó la gran precisión de los

métodos, los cuales permiten obtener errores relativos del orden de 10^{-6} para las mismas ganancias nombradas anteriormente, donde estos resultados fueron obtenidos en el marco de simulaciones con método exacto.

Parte importante del desarrollo de nuestra investigación fue la implementación de una situación hipotética de laboratorio donde se utilizó un tamaño de ensemble finito para estimar los valores esperados de los observables involucrados en la desigualdad CHSH. Estas simulaciones se realizaron sin considerar errores producto de la implementación experimental, por ejemplo, los defectos tecnológicos de los detectores, entre otros. Las simulaciones muestran que CSPA en su versión normalizada consigue mejores resultados si es comparado con su versión sin normalizar y con SPSA. En primer lugar converge más rápidamente en media que SPSA, logrando una ventaja de 15 iteraciones para algunos casos y cerca de 20 iteraciones de ventaja si se compara el rango intercuartil completo. Para la comparación de los errores relativos se tiene que CSPAN consigue precisiones del orden de 10^{-5} , mientras que SPSA obtiene 10^{-4} , para un tamaño de ensamble de $N = 10^4$ por medición. De esta forma, CSPAN tiene ventaja en precisión respecto a SPSA para el mismo tamaño de ensamble.

Otro punto importante de lo investigado es el rol que juegan las ganancias en las simulaciones con estos métodos. Se observó que las ganancias recomendadas en la literatura de optimización estocástica no proporcionan buenos resultados para las funciones objetivos consideradas, teniendo que realizar una búsqueda exhaustiva de ganancias que proporcionen los mejores resultados. Es por este motivo que para cualquier modificación en la función objetivo, es necesario realizar una nueva búsqueda de ganancias, por lo que desde un punto de vista práctico es una desventaja de los algoritmos. Otra desventaja de estos métodos es la cantidad de mediciones totales que se realizan para una simulación completa. Para cada iteración se realizan 2 evaluaciones de la función. Si se fijan en k las iteraciones totales, entonces se deben hacer $2 \cdot k$ mediciones diferentes. Si este mismo proceso se hace para un tamaño de ensemble finito N , la cantidad de recursos totales son $2 \cdot 4 \cdot N \cdot k$. A pesar de esto la principal ventaja de este método es que no se requiere postprocesamiento para obtener los resultados de las simulaciones.

La implementación exitosa de estos algoritmos al problema de las desigualdad CHSH nos permitió tener nociones para extender la investigación a otro tipo de sistemas como la desigualdad de Mermin para 3-qubits, desigualdades para observables con tres salidas y sistemas multipartitos. También estos algoritmos pueden ser implementados en otros problemas en Información Cuántica como métodos tomográficos, medidas de entrelazamiento, estimación de la pureza de estados, entre otros.

Apéndice A

Tomografía de estados cuánticos

La **tomografía de estados cuánticos** es un proceso experimental que permite inferir un estado cuántico ρ desconocido por medio de una serie de mediciones. Este proceso requiere una gran cantidad de copias del estado que deseamos caracterizar.

El conjunto $1/\sqrt{2}, X/\sqrt{2}, Y/\sqrt{2}, Z/\sqrt{2}$ forman una base ortonormal para las matrices hermíticas con respecto al producto punto de Hilbert-Schmidt. De este modo, el estado ρ puede ser expandido como

$$\rho = \frac{1}{2} (\text{Tr}\{\rho\}\mathbb{1} + \text{Tr}\{X\rho\}X + \text{Tr}\{Y\rho\}Y + \text{Tr}\{Z\rho\}Z). \quad (\text{A.1})$$

Recordemos que la expresión $\text{Tr}\{A\rho\}$ puede ser interpretada como el valor esperado del observable A . Por ejemplo, para estimar $\text{Tr}\{Z\rho\}$ debemos medir el observable Z m veces, obteniendo z_1, \dots, z_m cuyos valores pueden ser $+1$ o -1 . Así, $\sum z_i/m$ es una aproximación del valor real de $\text{Tr}\{A\rho\}$. De acuerdo al teorema del límite central, para un m suficientemente grande, los resultados están aproximadamente en distribución Gaussiana, donde el valor esperado es precisamente $\text{Tr}\{Z\rho\}$ y con desviación estándar $\Delta(Z)/\sqrt{m}$ como es mostrado en la eq. (2.58). Así, procediendo de la misma forma que para el observable Z se debe estimar $\text{Tr}\{X\rho\}$ y $\text{Tr}\{Y\rho\}$ para poder obtener una estimación del estado ρ .

Siguiendo el mismo procedimiento, podemos generalizar naturalmente esta situación a sistemas cuánticos de N qubits. Para una tomografía de un estado de N qubits, se requieren estimar $D^2 = 4^N - 1$ parámetros. La elección usual para la base de operadores es usar el conjunto de operadores de Pauli generalizados, construido como producto tensorial de las matrices de Pauli. Así ρ se puede descomponer como,

$$\rho = \sum_{\vec{v}} \frac{\text{Tr}\{\sigma_{v_1} \otimes \sigma_{v_2} \otimes \cdots \otimes \sigma_{v_n} \rho\} \sigma_{v_1} \otimes \sigma_{v_2} \otimes \cdots \otimes \sigma_{v_n}}{2^n} \quad (\text{A.2})$$

donde v_i pueden tomar los valores 0, 1, 2, 3 para indicar los operadores $\mathbb{1}, X, Y, Z$ en la posición i .



Apéndice B

Códigos de Matlab

En el siguiente apéndice se mostrarán las líneas de código utilizados para las simulaciones realizadas en los capítulos anteriores.

B.0.1. Códigos SPSA

La función *CHSHang* depende de dos parámetros z y Rho (ρ), los cuales representan los valores de ángulos y el estado, respectivamente. La parametrización elegida para el caso real es de acuerdo a la eq. (6.2). Este script permite calcular, de acuerdo a las variables declaradas, el valor de la función S . Esta función es la que será utilizada como función objetivo en *SPSAang*.

```
function [S] = CHSHang(z,Rho)
%
sigmax = [0 1; 1 0]; % pauli matrix x
sigmay = [0 -1i; 1i 0]; % pauli matrix y
5 sigmaz = [1 0; 0 -1]; % pauli matrix z

% Construcción de los operadores de Alice y Bob

A1 = sin(z(1))*cos(z(5))*sigmax + sin(z(1))*sin(z(5))*sigmay
10 + cos(z(1))*sigmaz;
A2 = sin(z(2))*cos(z(6))*sigmax + sin(z(2))*sin(z(6))*sigmay
+ cos(z(2))*sigmaz;
B1 = sin(z(3))*cos(z(7))*sigmax + sin(z(3))*sin(z(7))*sigmay
+ cos(z(3))*sigmaz;
15 B2 = sin(z(4))*cos(z(8))*sigmax + sin(z(4))*sin(z(8))*sigmay
+ cos(z(4))*sigmaz;
```



```

%
%z(1,...,4) -> Theta, z(5,...,8) -> phi
20 % Valores esperados de los operadores

E_11 = trace(kron(A1,B1)*Rho);
E_12 = trace(kron(A1,B2)*Rho);
E_21 = trace(kron(A2,B1)*Rho);
25 E_22 = trace(kron(A2,B2)*Rho);

% Evaluación de S

S = abs(E_11 + E_21 + E_12 - E_22);
30 end

```

SPSAang es una función que depende de 9 parámetros: *Dim* es la dimensión del espacio, *Nit* es el número de iteraciones para la secuencia, *z* es el vector que contiene los ángulos iniciales y *rho* representa el estado a evaluar. Por otro lado *a, b, s, t, A* son parámetros denominados *ganancias* y deben ser ajustados, para este problema, de acuerdo a los mostrados en la table 5.1. En cada iteración, el valor de *S* se acerca a su máximo permitido, convergiendo asintóticamente en media a dicho valor.

```

function [Salida] = SPSAang(Dim,Nit,z,rho,a,b,s,t,A)

Salida = zeros(1,Nit);
for k = 1:Nit
5 % Vector con dirección aleatorias +1 y -1

Delta_k = (-1).^round(randi([0,1],Dim,1));

10 % Ganancias

alpha_k = a/(k + 1 + A)^s; beta_k = b/(k + 1)^t;

% Declaración de variables
15 z1 = z + beta_k*Delta_k; z2 = z - beta_k*Delta_k;
chsh1 = CHSHang(z1,rho); chsh2 = CHSHang(z2,rho);

% Estimación del gradiente
20 g_k = ((chsh1 - chsh2)/(2*beta_k)).*Delta_k;

```

```

% Actualización del siguiente punto inicial
25 z = z + alpha_k*g_k;
Salida(k) = CHSHang(z,rho);
end

end

```

B.0.2. Código CSPA

La función *CHSH* depende de dos parámetros; z y Rho , los cuales representan el vector complejo y el estado cuántico, respectivamente. Para este caso la parametrización usada en la mostrada en la eq. (6.18). Con este script y de acuerdo a las variables declaradas, calcula el valor de la función S . Esta función es la que se utiliza en la función *CSPA*.

```

function [S] = CHSH(z,Rho)

E_1 = z(1:2); E_2 = z(3:4);
E_3 = z(5:6); E_4 = z(7:8);
5
e_1 = E_1./norm(E_1); e_2 = E_2./norm(E_2);
e_3 = E_3./norm(E_3); e_4 = E_4./norm(E_4);

p_1 = e_1*e_1'; p_2 = e_2*e_2';
10 p_3 = e_3*e_3'; p_4 = e_4*e_4';

% Operadores para Alice y Bob

A_1 = 2*p_1 - eye(2); A_2 = 2*p_2 - eye(2);
15 B_1 = 2*p_3 - eye(2); B_2 = 2*p_4 - eye(2);

% Valor esperado de la combinación de operadores

E_11 = trace(kron(A_1,B_1)*Rho);
20 E_12 = trace(kron(A_1,B_2)*Rho);
E_21 = trace(kron(A_2,B_1)*Rho);
E_22 = trace(kron(A_2,B_2)*Rho);

% Evaluación de S
25 S = abs(E_11 + E_21 + E_12 - E_22);

```

```
end
```

```
function [Salida] = CSPSA(Dim,N,z,rho,a,b,s,r,A)
Salida = zeros(1,N);
for K =1:N
a_k = a/(K + A)^s; c_k = b/(K)^r;
5 l = randi(4,Dim,1);
D = (1i).^l);
%
%g = G(z,C,D,rho);
F1 = z + c_k*D;
10 F2 = z - c_k*D;
f1 = CHSH(F1, rho);
f2 = CHSH(F2, rho);
g = (f1 - f2)./(2*c_k*conj(D)); %aproximaci n del gradiente
z = z + a_k*g;
15 E_1 = z(1:2)/norm(z(1:2));
E_2 = z(3:4)/norm(z(3:4));
E_3 = z(5:6)/norm(z(5:6));
E_4 = z(7:8)/norm(z(7:8));
z = [E_1;E_2;E_3;E_4];
20 Salida(K) = CHSH(z,rho);
end
end
```

Bibliografía

- [1] David Avis, Sonoko Moriyama, and Masaki Owari. From bell inequalities to tsirelson’s theorem. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E92-A(5):1254–1267, 2009.
- [2] John Stewart Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1(3):195–200, 1964.
- [3] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an Unknown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels. *Physical Review Letters*, 70(13):1895–1899, 1993.
- [4] D. Bohm and Y. Aharonov. Discussion of experimental proof for the paradox of Einstein, Rosen, and Podolsky. *Physical Review*, 108(4):1070–1076, 1957.
- [5] D. Boschi, S. Branca, F. De Martini, L. Hardy, and S. Popescu. Experimental realization of teleporting an unknown pure quantum state via dual classical and einstein-podolsky-rosen channels. *Physical Review Letters*, 80(6):1121–1125, 1998.
- [6] Carlton M Caves, Christopher A Fuchs, and Rüdiger Schack. Unknown quantum states: The quantum de Finetti representation. *J. Math. Phys.*, 43(9):4537–4559, 2002.
- [7] Robert J Chapman, Christopher Ferrie, and Alberto Peruzzo. Experimental Demonstration of Self-Guided Quantum Tomography. *Phys. Rev. Lett.*, 117(4), 2016.
- [8] B S Cirel’son. Quantum generalizations of Bell’s inequality. *Letters in Mathematical Physics*, 4(2):93–100, 1980.
- [9] John F Clauser, Michael A Horne, Abner Shimony, and Richard A Holt. Proposed Experiment to Test Local Hidden-Variable Theories. 23(15):880–884, 1969.
- [10] P A M Dirac. A new notation for quantum mechanics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 35(3):416–418, 1939.

- [11] A Einstein, B Podolsky, and N Rosen. Can Quantum-Mechanical Description of Reality Be Considered Complete?, 1935.
- [12] A K Ekert. Quantum Cryptography based on Bell theorem. *Physical Review Letters*, 67(6):661–663, 1991.
- [13] Christopher Ferrie. Self-guided quantum tomography. *Phys. Rev. Lett.*, 113(19), 2014.
- [14] N Gisin and A Peres. Maximal violation of Bell’s inequality for arbitrarily large spin. *Phys. Lett. A*, 162(1):15–17, 1992.
- [15] Richard Jozsa. Entanglement and Quantum Computation. jul 1997.
- [16] B P Kibble, I A Robinson, and J H Belliss. A realization of the SI watt by the NPL moving-coil balance. *Metrologia*, 27(4):173–192, 1990.
- [17] Michael A Nielsen and Isaac L Chuang. *Quantum Computation and Quantum Information*. 2010.
- [18] Payman Sadegh and James C Spall. Optimal random perturbations for stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Trans. Automat. Contr.*, 43(10):1480–1484, 1998.
- [19] S Schlamminger, D Haddad, F Seifert, L S Chao, D B Newell, R Liu, R L Steiner, and J R Pratt. Determination of the Planck constant using a watt balance with a superconducting magnet system at the National Institute of Standards and Technology. *Metrologia*, 51(2), 2014.
- [20] J C Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *Automatic Control, IEEE Transactions on*, 37(3):332–341, 1992.
- [21] A. Utreras-Alarcón. An enhancement to self-guided quantum tomography.
- [22] A. Utreras-Alarcón, M. Rivera-Tapia, S. Niklitschek, and A. Delgado. Stochastic optimization on complex variables and tomography of pure states.
- [23] Stephanie Wehner. Tsirelson bounds for generalized Clauser-Horne-Shimony-Holt inequalities. *Physical Review A - Atomic, Molecular, and Optical Physics*, 73(2), 2006.
- [24] Reinhard F Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40(8):4277–4281, 1989.
- [25] W Wirtinger. *Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen*, volume 97. 1927.