



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
**POSGRADO EN CIENCIAS FÍSICAS**

ESTIMACIÓN DE PARÁMETROS EN SISTEMAS CUÁNTICOS

TESIS  
QUE PARA OPTAR POR EL GRADO DE:  
DOCTOR EN CIENCIAS (FÍSICA)

PRESENTA:  
MARCO ANTONIO RODRÍGUEZ GARCÍA

TUTOR PRINCIPAL:

DR. PABLO BARBERIS BLOSTEIN  
INSTITUTO DE INVESTIGACIONES EN MATEMÁTICAS APLICADAS Y EN  
SISTEMAS (IIMAS), UNAM

MIEMBROS DEL COMITÉ TUTOR:

DR. FRANÇOIS LEYVRAZ WALTZ  
INSTITUTO DE CIENCIAS FÍSICAS (ICF)

DR. CARLOS FRANCISCO PINEDA ZORRILLA  
INSTITUTO DE FÍSICA (IF), UNAM

CIUDAD DE MÉXICO, 2020



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

# Agradecimientos

El primer reconocimiento de la tesis es naturalmente para Pablo Barberis Blostein, por compartir conmigo su curiosidad y entusiasmo, por su paciencia en todas nuestras discusiones, darme la libertad de elegir mi propio camino y apoyarme enormemente durante todo este tiempo. También agradezco particularmente al Dr. Isaac Pérez Castillo y al Dr. Francisco Elohim Becerra por su gran apoyo a mi proyecto doctoral. Igualmente agradezco al Dr. Carlos Francisco Pineda Zorrilla y al Dra. Andrea Valdés Hernández por sus acertadas correcciones y sus sinceros comentarios a este trabajo. Gracias por formar parte de mi jurado.

En general, también quiero agradecer al Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas que me proporcionó un entorno muy agradable y estimulante para la mayor parte de mis estudios de posgrado.

Quiero agradecer a toda mi familia, en concreto a mis padres Lilia García y Marco Rodríguez, a mis hermanos Estefany y Jose Luis y mis abuelos Javier García y Natividad Ravelo por su compañía y apoyo que me han brindado durante toda mi vida. Gracias a mis amigos César Alberto Díaz Hernández y Sergio Arturo Ortiz Serafín por todas las experiencias y el tiempo compartido.

Muy especialmente agradezco a Edzna Claudia Silva Pérez por compartir conmigo esta melodía incierta, por todas las aventuras que hemos vivido juntos, por todo su soporte, empatía, ánimos infundidos, y cariño. Gracias.

Finalmente agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT) por apoyarme económicamente a través de su programa de becas de posgrado y a los proyectos PAPIIT de la Dirección General de Asuntos del Personal Académico (DGAPA), UNAM por el apoyo económico de los proyectos IG100518: "Estimación óptima de parámetros cuánticos y aplicaciones en información cuántica" y IG101421: "Mediciones en sistemas cuánticos de muchos cuerpos: nuevos métodos y perspectivas" que hicieron posible la realización de esta tesis.



# Índice general

<b>Agradecimientos</b>	<b>1</b>
<b>1. Introducción</b>	<b>9</b>
<b>2. Estimación de parámetros en sistemas clásicos</b>	<b>13</b>
2.1. Espacios de probabilidad . . . . .	13
2.2. Variables aleatorias . . . . .	14
2.3. Problema de estimación de parámetros . . . . .	18
2.4. Estimadores insesgados de varianza mínima . . . . .	19
2.5. Desigualdad de Crámer-Rao . . . . .	21
2.6. Estimador de máxima verosimilitud . . . . .	23
<b>3. Estimación de parámetros en sistemas cuánticos</b>	<b>27</b>
3.1. Estados cuánticos . . . . .	27
3.2. Observables cuánticos . . . . .	28
3.3. Problema de estimación de parámetros en sistemas cuánticos . . . . .	29
3.4. Estimación escalar . . . . .	30
3.4.1. Estimación multiparamétrica . . . . .	32
3.5. Estimación covariante . . . . .	34
3.6. Simetría temporal . . . . .	37
<b>4. Estimación de fase en qubits</b>	<b>39</b>
4.1. POVMs localmente óptimos . . . . .	41
4.1.1. Estimación de fase covariante en qubits . . . . .	42
4.2. Medidas entrelazadas . . . . .	47
4.3. Estimación cuántica adaptativa de estados (AQSE) . . . . .	48
4.4. Estimación adaptativa utilizando intervalos de confianza . . . . .	50
<b>5. Estimación de frecuencia</b>	<b>57</b>
5.1. Estimación de frecuencia con intervalos de confianza . . . . .	58
5.2. Estimación de frecuencia en qutrits . . . . .	62
<b>6. Estimación de fase en estados coherentes</b>	<b>67</b>
6.1. Estados coherentes . . . . .	67
6.2. Descripción del problema . . . . .	70

6.3.	Diseño Bayesiano óptimo de experimentos . . . . .	71
6.4.	Esquema adaptativo no-Gaussiano . . . . .	74
6.4.1.	Correcta elección del diseño óptimo . . . . .	76
6.5.	Desempeño del esquema adaptativo no-Gaussiano . . . . .	78
6.5.1.	Modelo para el rendimiento asintótico de la estrategia adaptativa . . . . .	81
<b>7.</b>	<b>Sensores con escalamiento Super-Heisenberg</b>	<b>85</b>
7.1.	Límite cuántico estándar y límite de Heisenberg . . . . .	85
7.2.	Protocolo para generar dinámica paramétrica no lineal . . . . .	87
7.3.	Mejoría en la estimación de parámetros . . . . .	91
7.4.	Información cuántica de Fisher para un ensamble post-medición . . . . .	91
7.5.	Sensores con escalamiento Super-Heisenberg . . . . .	93
<b>8.</b>	<b>Conclusiones</b>	<b>99</b>

# Lista de Símbolos

Para conveniencia del lector, se presenta una lista de los símbolos más utilizados en esta tesis:

$\mathcal{H}$  Espacio de Hilbert

$\mathcal{B}(\mathcal{H})$  Espacio de operadores acotados en  $\mathcal{H}$

$\mathcal{S}(\mathcal{H})$  Conjunto de operadores densidad en  $\mathcal{H}$

$\text{Spec}(A)$  Espectro del operador  $A$

$L(\theta)$  Operador de derivada simétrica logarítmica

$I_Q(\theta)$  Información de Fisher cuántica acerca de  $\theta$

$\mathcal{S}_\theta$  Modelo estadístico cuántico

$\mathbb{F}(\theta)$  Información de Fisher generalizada acerca de  $\theta$

$\mathcal{M}(\Theta, U)$  Espacio de POVMs covariantes respecto a una representación unitaria  $U$  de  $\mathcal{H}$  con salidas en  $\Theta$

$\mathcal{M}(\mathcal{H})$  Espacio de POVMs sobre un espacio de Hilbert  $\mathcal{H}$

$\mathcal{B}(\mathbb{R})$   $\sigma$ -álgebra de Borel de los números reales

$\mathcal{L}(\theta; x)$  Función de verosimilitud de  $\theta$  en función de una observación  $x$

$I_M(\theta)$  Información de Fisher asociada al POVM  $M$  acerca de  $\theta$

$I_X(\theta)$  Información de Fisher que contiene la variable aleatoria  $X$  acerca de  $\theta$





# Siglas

**AQSE** adaptive quantum state estimation. [46](#)

**CICR** cota inferior de Crámer-Rao. [19](#), [20](#)

**CICRC** cota inferior de Crámer-Rao Cuántica. [28](#), [29](#), [97](#)

**ECM** error cuadrático medio. [17](#), [35](#)

**EER** error estándar de la regresión. [80](#)

**EIC** estimación por intervalos de confianza. [56](#), [61](#)

**EIMVU** estimador insesgado de mínima varianza uniforme. [17](#)

**EMV** estimador de máxima verosimilitud. [21](#)

**IC** intervalo de confianza. [22](#), [49](#)

**IFC** información de Fisher cuántica. [28](#)

**MAP** máxima a posteriori. [71](#)

**MK II** Mark-II. [69](#)

**PNR** photon number resolution. [75](#)

**POVM** positive operator-valued measure. [9](#), [27](#)

**SQL** standard quantum limit. [84](#)



# Capítulo 1

## Introducción

Un problema fundamental para el desarrollo de las tecnologías cuánticas es el de determinar el estado en que se encuentra un sistema cuántico particular. No obstante, la naturaleza estocástica de la medición en la mecánica cuántica, además de la posible existencia de ruidos o errores sistemáticos en las mediciones, provoca que en general los estados de un sistema cuántico no se puedan determinar de forma perfecta. El marco teórico que describe este problema se conoce como la teoría de estimación cuántica y comenzó a desarrollarse a partir de los trabajos de Holevo [57] y Helstrom [54] en la década de 1970.

Uno de los objetivos de la teoría de estimación cuántica es la búsqueda de mediciones que produzcan la mejor estrategia de estimación para un estado de un sistema cuántico. Es frecuente suponer que los estados cuánticos se conocen parcialmente por un conjunto de parámetros a los cuales no se les puede asociar un observable cuántico. De este modo, el problema de estimación del estado cuántico recae en la estimación de los parámetros que determinan el sistema a partir del resultado de la medición de otras características del sistema. Bajo esta suposición, el problema es llamado estimación de parámetros en sistemas cuánticos. En particular, estimación de fases, constantes dieléctricas y frecuencias son ejemplos típicos de este tipo de problema. La estimación de parámetros en sistemas cuánticos es el eje central para el desarrollo de un gran espectro de aplicaciones como la preparación experimental de estados [30, 42], desarrollo de sensores [1], comunicación [6, 32, 54], metrología [43] y procesamiento de información [99], por mencionar algunas.

Aunque se tiene un largo historial sobre el trabajo de este problema, nuevas aplicaciones para la estimación cuántica de parámetros continúan apareciendo [1, 100, 119]. Por esta variedad de escenarios, hoy en día es de interés común para la comunidad científica obtener estrategias de estimación más precisas que sean factibles de implementar físicamente. En concreto, dos problemas centrales existen a la hora de querer implementar una estrategia de medición óptima. El primero recae en el hecho que en general la medida óptima para estimar un parámetro desconocido depende de este mismo, lo cual vuelve inviable su implementación física de forma directa, aunque enfoques Bayesianos y estrategias adaptativas pueden resolver esta dificultad si el sistema cuántico satisface ciertas condiciones de regularidad [38, 80]. El segundo corresponde al hecho que las mediciones óptimas pueden generar ambigüedad sobre el parámetro a estimar. Este problema se conoce como falta de identificabilidad y cuando se tiene, existe una misma descripción estadística para dos o más valores diferentes del parámetro [10, 11, 103].

El objetivo de esta tesis es ofrecer un trato sistemático al problema de estimación de paráme-

tros en sistemas cuánticos haciendo énfasis en cómo el problema de no identificabilidad es una limitante seria a la hora de intentar implementar distintas estrategias de estimación que son aparentemente óptimas, y resolver esta disyuntiva para casos relevantes que aparecen en aplicaciones concretas. Bajo este espíritu se hará especial énfasis en el desarrollo matemático de la teoría y en su interpretación física. En consecuencia, el método de presentación consistirá primero en establecer las propiedades generales del problema de estimación en sistemas clásicos, después generalizarlas al caso cuántico y finalmente proponer soluciones para el problema de estimación de fase en qubits y estados coherentes y estimación de frecuencia en qubits y qutrits.

El primer sistema que se estudia es el qubit. El qubit es un sistema cuántico de dimensión 2 el cual es frecuentemente utilizado para codificar información en forma de una fase sobre  $[0, 2\pi)$ . Cuando esta fase es desconocida se tiene un problema de estimación de fase. Para este problema se pueden encontrar condiciones iniciales donde las medidas óptimas no dependan del parámetro. Sin embargo, las distribuciones de probabilidad que se obtienen al utilizar estas medidas son periódicas con periodo  $\pi$ . Este hecho tiene como consecuencia que existan dos posibles valores de la fase que generan la misma descripción estadística, es decir, se tiene un problema de identificabilidad. Esta ambigüedad en los parámetros hace inasequible la implementación de estrategias de estimación basadas en mediciones óptimas sin tener conocimiento a priori acerca del parámetro. En este trabajo se resolvió el problema de identificabilidad para el problema de estimación de fase en qubits aprovechando la simetrías del sistema y haciendo uso de intervalos de confianza.

Las fases se pueden descomponer en un parámetro temporal y una frecuencia; cuando se tiene control sobre el parámetro temporal, el problema de estimación de fase se reduce a un problema de estimación de frecuencia. En este problema, el límite fundamental en la precisión de las estimaciones decrece conforme el parámetro temporal aumenta. Sin embargo, a partir de cierto valor del tiempo, el problema de identificabilidad sobre la frecuencia aparece, lo cual restringe el dominio temporal sobre el cual se puede trabajar y por ende el límite fundamental del error de las estimaciones. Para resolver esta disyuntiva, se extendió la estrategia de estimación de fase en qubits con intervalos de confianza al caso de estimación de frecuencia, obteniendo un esquema de estimación que permite ir aumentando el parámetro temporal para así, obtener mayor precisión en la estimación de la frecuencia. Además, esta idea permite su extensión a sistemas de mayor dimensión, tratando el ejemplo de estimación de frecuencia en un qutrit (sistema cuántico de dimensión 3).

Así mismo, el problema de estimación de fase se estudió para un estado coherente, esto es, un estado de un sistema de dimensión infinita que modela un láser perfecto. Para el problema de estimación de fase es posible encontrar teóricamente la medida que minimiza el error promedio de las estimaciones. A dicha medida se le conoce como medida canónica de fase y su implementación física es posible cuando el estado es un qubit. Sin embargo, para estados coherentes, no existe ninguna propuesta realista para poder implementar esta medida de forma experimental. Debido a esto se han propuesto distintos métodos de estimación para aproximar el desempeño de la medida canónica de fase [119]. En este trabajo se estudia un método con mediciones adaptativas basadas en conteo de fotones y retroalimentación capaz de aproximar de mejor manera el desempeño de medida canónica de fase que los métodos estándar expuestos en la literatura [31].

Por último, utilizando un caso pragmático, se muestra que la forma de codificar el parámetro puede influir en la precisión de cómo este se estima. Concretamente se estudió el caso donde una operación no lineal codifica una fase sobre un estado coherente, que aunque en principio, puede mejorar el desempeño de las estrategias de estimación, conlleva múltiples problemáticas para su implementación física. En esta tesis se discute una propuesta de como la codificación no lineal

de una fase sobre un estado coherente puede ser implementada y la correcta forma de medir la precisión de las estrategias de estimación aplicadas sobre esta propuesta. A continuación, el contenido particular de los capítulos de este trabajo es resumido:

En el **Capítulo 2** se da una breve introducción a la teoría de estimación de parámetros en sistemas clásicos. Esencialmente es un resumen de los resultados de la teoría que se puede encontrar en libros estándar del campo [19, 29, 67, 68]. Dentro del capítulo se introduce el concepto de modelo estadístico, de estimador y función de riesgo. La función de riesgo provee una forma de medir la precisión de un estimador. Finalmente el concepto de información de Fisher es introducido; cuyo inverso brinda una cota inferior a la precisión de estimadores insesgados llamada cota de Cramér-Rao.

El **Capítulo 3**, basado en el libro [57], generaliza la teoría de estimación de parámetros a sistemas cuánticos. En el capítulo se desarrollan a detalle las nociones fundamentales de esta teoría, y se construye el concepto de modelo estadístico cuántico, especificando los términos de estados y medidas cuánticas (observables). El concepto de POVM (del inglés [positive operator-valued measure](#)) es la descripción más general de una medida sobre un sistema cuántico. La optimización sobre el espacio de medidas de la información de Fisher define la información de Fisher cuántica, la cual es la cota superior a la información que se puede extraer de un POVM acerca de un parámetro que se desea estimar sobre un sistema cuántico. Para cerrar, se da especial énfasis al caso de estimación covariante, donde los estados y los POVMs son invariantes bajo la acción de un grupo que representa cierta simetría del sistema cuántico. En especial, se prueba que dentro de los POVMs covariantes se puede encontrar la medición con menor error de estimación para una sola copia del sistema, este POVM es llamado medida canónica. La medida canónica es de particular relevancia para problemas de estimación de fase.

En los siguientes capítulos 4, 5, 6 y 7 se exponen los resultados originales de esta tesis. Los resultados del **Capítulo 4** fueron publicados en el artículo [103]. En este capítulo se discute cómo la falta de identificabilidad en los parámetros es una limitante para producir estrategias de estimación eficientes basadas en POVMs localmente óptimos. Con el fin de resolver el problema de no identificabilidad, y probando que los POVMs covariantes son identificables, se deriva una nueva estrategia de estimación basada en intervalos de confianza construidos a partir de las estimaciones de mediciones covariantes. Finalmente, se demuestra que esta nueva propuesta es eficiente para cualquier condición inicial.

En el **Capítulo 5**, se extiende la propuesta de estimación con intervalos de confianza al problema de estimación de frecuencia en qubits. Se demuestra que la estrategia de estimación es viable para este problema. Para esta situación, a mayor tiempo de medición se obtiene en principio un menor error en las estimaciones. Sin embargo, esto conlleva la aparición del problema de no identificabilidad. Nuestra propuesta ofrece una manera de aumentar el tiempo de medición resolviendo la no identificabilidad que conlleva el incremento del parámetro temporal. Por último, al trabajar con qutrits, se prueba que el hecho de aumentar la dimensión del sistema cuántico puede mejorar aún más el desempeño de la estrategia de medición.

El **Capítulo 6** trata el problema de estimación de fase en un estado coherente. A diferencia del qubit, donde es necesario reducir el error de las estimaciones utilizando múltiples copias del sistema, en los estados coherentes es posible optimizar la precisión de las estrategias de estimación para el caso de una única copia del sistema dividiendo el pulso electromagnético en pulsos más pequeños. De acuerdo con la teoría descrita en el capítulo 3, la precisión de la medida covariante canónica de fase es el límite fundamental en la precisión de las estimaciones. Como no existe una

propuesta experimental para implementar este POVM, se presenta una estrategia adaptativa de estimación basada en la teoría de diseño óptimo de experimentos y conteo de fotones, cuyo desempeño puede aproximar al de la medida canónica [31]. Con la finalidad de mejorar las inferencias acerca de la fase desconocida, la estrategia adaptativa utiliza la fase de un oscilador local sobre el estado coherente como un diseño experimental, el cual puede controlarse antes de realizar cualquier medición sobre el sistema. La estrategia de estimación permite que en cada paso adaptativo se escoja el valor del diseño de forma óptima maximizando una función de costo. En este trabajo se prueba que las funciones de costo adecuadas para el problema de estimación de fase son la información mutua entre la distribución de las posibles salidas de la detección de fotones y la fase modelada como una variable aleatoria. Además, como resultado principal se encuentra una expresión analítica para el rendimiento de esta estrategia de estimación. Esta fórmula demuestra teóricamente que este protocolo de estimación supera a la mejor estrategia de estimación de fase estándar encontrada en la literatura [118].

Por último, en el **Capítulo 7** cuyos resultados fueron publicados en [62], se analiza el caso en que el generador de una fase para estados coherentes es un operador no lineal. Bajo este supuesto, se obtiene una información cuántica de Fisher que es función de una potencia sexta del número promedio de fotones, lo que trae como consecuencia la mejora de la cota inferior del error que puede tener un estimador de fase. No obstante, la implementación física de estos generadores es escasa. Por ello se introduce un circuito cuántico capaz de preparar generadores no lineales. Dentro de este circuito se pierde información tras realizar una medición a un estado auxiliar. Como resultado, se debe utilizar la información de Fisher generalizada como medida de información. Un análisis numérico de las potencias de esta cantidad, corrobora el correcto funcionamiento de este circuito. Las técnicas de este capítulo son de utilidad para la aplicación de sensores cuánticos donde los supuestos que se toman garantizan la identificabilidad de parámetros.

## Capítulo 2

# Estimación de parámetros en sistemas clásicos

La teoría de estimación de parámetros en sistemas cuánticos se puede tratar como una generalización de la teoría sobre sistemas clásicos. Por esta razón, los conceptos básicos utilizados de la teoría clásica, como lo son las variables aleatorias, distribuciones de probabilidad, estimadores y funciones de costo también son utilizados en el contexto cuántico. En este capítulo se dará una breve introducción a la teoría de estimación de parámetros en sistemas clásicos mostrando los conceptos básicos para poder generalizar la teoría de forma natural al caso de sistemas cuánticos.

### 2.1. Espacios de probabilidad

Una gran cantidad de fenómenos en la naturaleza se pueden considerar como aleatorios. Esta clase de fenómenos se pueden modelar matemáticamente utilizando el concepto de espacio de probabilidad. Un espacio de probabilidad es una terna  $(\Omega, \mathcal{F}, P)$ , donde  $\Omega$  es un conjunto llamado espacio muestral, que agrupa todos las posibles salidas de un fenómeno aleatorio,  $\mathcal{F}$  es una  $\sigma$ -álgebra de subconjuntos de  $\Omega$  cuyos elementos son interpretados como los eventos que pueden observarse y  $P$  una medida de probabilidad para cada elemento de  $\mathcal{F}$ . La definición precisa de estos conceptos es la siguiente:

**Definición 1** (Espacio muestral). El espacio muestral de un experimento aleatorio es el conjunto de todos los posibles resultados del experimento y es denotado por  $\Omega$ .

**Definición 2** ( $\sigma$ -álgebra). Sea  $\Omega$  un conjunto y  $\mathcal{F} \subset 2^\Omega$  una familia de subconjuntos de  $\Omega$ . Entonces  $\mathcal{F}$  es una  $\sigma$ -álgebra si y sólo si

- i)  $\emptyset \in \mathcal{F}$ ,
- ii) Si  $A, B \in \mathcal{F}$  entonces  $A \cup B \in \mathcal{F}$
- iii) Si  $A, B \in \mathcal{F}$  entonces  $A \setminus B \in \mathcal{F}$
- iv)  $\Omega \in \mathcal{F}$
- v) Para cada  $\{A_n\}_{n=1}^\infty \subset \mathcal{F}$ , se tiene que  $\cup_{n=1}^\infty A_n \in \mathcal{F}$ .

La  $\sigma$ -álgebra de un conjunto  $\Omega$  se interpreta como la colección de todos los eventos que se pueden observar en un fenómeno aleatorio.

**Definición 3** (Medida de probabilidad). Una función  $P$  definida sobre una  $\sigma$ -álgebra  $\mathcal{F}$  y con valores en el intervalo  $[0, 1]$  es una medida de probabilidad si  $P(\Omega) = 1$  y

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n), \quad (2.1)$$

donde  $\{A_n\}_{n=1}^{\infty} \subset \mathcal{F}$  y  $A_i \cap A_j = \emptyset$  para  $i$  y  $j$  distintos. El número  $P(A)$  representa una forma de medir la probabilidad de observar el evento  $A \in \mathcal{F}$ .

Aunque un espacio de probabilidad  $(\Omega, \mathcal{F}, P)$  es un modelo satisfactorio para describir cualquier experimento aleatorio, trabajar con conjuntos en abstracto suele ser en demasía complicado, por lo que es conveniente trasladar la manipulación matemática del espacio muestral a los números reales. Las herramientas matemáticas que sirven para realizar este proceso son las variables aleatorias.

## 2.2. Variables aleatorias

Considerando  $(\Omega, \mathcal{F}, P)$  un espacio de probabilidad, una variable aleatoria  $X : \Omega \rightarrow \mathbb{R}$  es una función medible del espacio muestral a los reales. Es decir,  $X$  es una función real tal que la imagen inversa de cualquier evento es un elemento de  $\mathcal{F}$ ,

$$\forall B \in \mathcal{B}(\mathbb{R}) \quad X^{[-1]}(B) \in \mathcal{F}, \quad (2.2)$$

donde  $\mathcal{B}(\mathbb{R})$  es la  $\sigma$ -álgebra de Borel de los números reales (la  $\sigma$ -álgebra más pequeña que incluye la topología de los reales) y  $X^{[-1]}(\cdot)$  es la notación de la imagen inversa de algún conjunto bajo  $X$ . De esta forma, dado un espacio de probabilidad  $(\Omega, \mathcal{F}, P)$ , una variable aleatoria induce una medida de probabilidad sobre los reales de la siguiente manera

$$P_X(B) = P\left(X^{[-1]}(B)\right), \quad \forall B \in \mathcal{B}(\mathbb{R}). \quad (2.3)$$

Entonces  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$  es un espacio de probabilidad sobre los números reales.

Dentro del espacio de probabilidad  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$  se puede definir una función de distribución  $F_X : \mathbb{R} \rightarrow [0, 1]$  para  $X$ , vía

$$F_X(x) = P_X[(-\infty, x]] = P(\{\omega \in \Omega \mid X(\omega) \leq x\}) = P\left(X^{[-1]}[(-\infty, x]]\right) = P(X \leq x). \quad (2.4)$$

De la definición (2.4) se puede comprobar que cada función de distribución  $F$ , es una función creciente, continua por la derecha,  $\lim_{x \rightarrow \infty} F(x) = 1$  y  $\lim_{x \rightarrow -\infty} F(x) = 0$ .

En general, toda función de distribución  $F$  se puede descomponer como la combinación convexa de una función de distribución constante por pedazos  $F_d$  y una función de distribución continua  $F_c$ . Esto es,

$$F(x) = \alpha F_d(x) + (1 - \alpha) F_c(x), \quad 0 \leq \alpha \leq 1. \quad (2.5)$$

De esta manera, una variable aleatoria  $X$  se puede clasificar en discreta, continua o mixta, dependiendo si su correspondiente función de distribución  $F_X$  se puede escribir solo como una función



constante por pedazos ( $F_c = 0$ ), como una distribución continua ( $F_d = 0$ ) o si sus partes discretas y continuas son distintas de cero respectivamente.

Cuando una variable aleatoria  $X$  es discreta, su correspondiente función de distribución  $F_X(x)$  tiene un conjunto de discontinuidades. Si  $x_1, x_2, \dots$ , son sus puntos de discontinuidad, entonces cada discontinuidad es de tamaño  $P(X = x_i) = P(X \leq x_i) - P(X < x_i)$  y la función de distribución  $F_X$  se puede reconstruir como

$$F_X(x) = \sum_{u \leq x} p(u), \quad (2.6)$$

donde  $p$  es una función definida como

$$p(x) = \begin{cases} P(X = x) & \text{si } x = x_1, x_2, \dots \\ 0 & \text{en otro caso} \end{cases} \quad (2.7)$$

llamada función de probabilidad.

Debido a que es posible obtener la función de distribución de una variable aleatoria discreta a partir de la ecuación (2.6), en muchos problemas prácticos se suele trabajar directamente con la función de probabilidad correspondiente a la distribución que se está estudiando.

Por otro lado, cuando  $X$  es una variable aleatoria continua, entonces su función de distribución es continua y se puede descomponer a su vez como la combinación convexa de una función de distribución singular  $F_s$  y una función de distribución absolutamente continua  $F_{ac}$  [75], es decir,

$$F_X(x) = F_c(x) = \beta F_s(x) + (1 - \beta) F_{ac}(x), \quad 0 \leq \beta \leq 1. \quad (2.8)$$

En particular, cuando  $F_s = 0$ , la variable aleatoria  $X$  es llamada absolutamente continua, ya que su función de distribución induce una medida sobre los reales  $P_F$  definida por

$$P_F[(a, b]] := F(b) - F(a) \text{ donde } a, b \in \mathbb{R}, \quad (2.9)$$

la cual es absolutamente continua con respecto a la medida de Lebesgue  $\lambda$  de los reales ( $\lambda[(a, b]] = b - a$ ). Esto es,

$$P_F(A) = 0 \forall A \in \mathcal{B}(\mathbb{R}) \text{ tal que } \lambda(A) = 0. \quad (2.10)$$

A consecuencia del teorema de Radon-Nikodym [75], existe una función  $f$  tal que  $f(\omega) \geq 0$  para cada  $\omega \in \Omega$  e integrable respecto a la medida de Lebesgue tal que para cualquier valor de  $x$  cumple con

$$F(x) = \int_{-\infty}^x f(u) du. \quad (2.11)$$

A la función  $f$  se le llama función de densidad.

De forma análoga al caso de las distribuciones discretas y la función de probabilidad, la ecuación (2.11), trae como resultado que en las aplicaciones prácticas se suele trabajar directamente con la función de densidad correspondiente a la distribución bajo estudio.

Por otra parte cuando  $F_{ac} = 0$ , la medida  $P_F$  (2.9) es singular a la medida de Lebesgue  $\lambda$  en los reales y  $P_f(\{x\}) = 0$  para cada  $x \in \mathbb{R}$ . En otras palabras,

$$\exists A \in \mathcal{B}(\mathbb{R}) \text{ tal que } \lambda(A) = 0 = P_F(A^c) \quad (2.12)$$

y  $P_f(\{x\}) = 0$  para cada  $x \in \mathbb{R}$ . Las funciones de distribución suelen ser difíciles de construir <sup>1</sup>. En este trabajo, todas las variables aleatorias continuas serán absolutamente continuas ( $F_s = 0$  en la ecuación (2.8)).

De igual forma, la clasificación de las variables aleatorias también sirve para calcular sus momentos asociados. El  $k$ -ésimo momento de una variable aleatoria se define como

$$E[X^k] = \int_{-\infty}^{\infty} x^k dF_X(x), \quad (2.13)$$

donde  $\int_{-\infty}^{\infty} h(x)dF_X(x)$  es la integral de Riemann-Stieltjes de una función real  $h$  respecto a la función de distribución  $F_X$  la cual se calcula como

$$\int_{-\infty}^{\infty} h(x)dF_X(x) = \begin{cases} \sum_{i=1}^{\infty} h(x_i)p(x_i) & \text{si } X \text{ es discreta} \\ \int_{-\infty}^{\infty} h(x)f(x)dx & \text{si } X \text{ es continua.} \end{cases} \quad (2.14)$$

De particular importancia tenemos al primer momento de una variable aleatoria  $E[X]$ , también llamada esperanza o valor esperado que se interpreta como la media (promedio) de los posibles valores que puede tomar  $X$ , y la varianza de  $X$

$$\text{Var}[X] = E[X^2] - (E[X])^2 \quad (2.15)$$

que se interpreta como una medida de dispersión de los valores que puede tomar  $X$  con respecto a su media. Además para medir la correlación entre dos variables aleatorias  $X_1$  y  $X_2$  se utiliza el concepto de covarianza

$$\text{Cov}[X_1, X_2] = E[X_1 \cdot X_2] - E[X_1]E[X_2]. \quad (2.16)$$

Una familia  $\{X_t\}_{t \in T}$  de variables aleatorias definidas sobre el mismo espacio de probabilidad  $(\Omega, \mathcal{F}, P)$ , es independiente si y sólo si para cualquier  $A, B$  en la familia de eventos  $\mathcal{A} = \{X_t^{-1}(B(\mathbb{R}))\}_{t \in T}$  se tiene que

$$P(A \cap B) = P(A)P(B). \quad (2.17)$$

De este modo, si dos variables aleatorias  $X_1$  y  $X_2$  son independientes, entonces su covarianza es cero.

Los conceptos previos se pueden generalizar a variables aleatorias sobre  $\mathbb{R}^n$ , en este caso, una variable aleatoria  $X$  es una función medible del espacio muestral  $\Omega$  a  $\mathbb{R}^n$ , produciendo un espacio de probabilidad  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), P_X)$  donde  $\mathcal{B}(\mathbb{R}^n)$  es la  $\sigma$ -álgebra de Borel de  $\mathbb{R}^n$  ( $\sigma$ -álgebra más pequeña que contiene la topología usual de  $\mathbb{R}^n$ ) y  $P_X(B) = P(X^{-1}(B))$  para cada  $B \in \mathcal{B}(\mathbb{R}^n)$ . Así,  $P_X$  induce una función de distribución  $F_X : \mathbb{R}^n \rightarrow [0, 1]$ ,

$$\begin{aligned} F_X(x) &= F_X((x_1, x_2, \dots, x_n)) \\ &= P_X[(-\infty, x_1] \times (-\infty, x_2] \times \dots \times (-\infty, x_n)] \\ &= P[X^{-1}((-\infty, x_1] \times (-\infty, x_2] \times \dots \times (-\infty, x_n))] , x = (x_1, \dots, x_n) \in \mathbb{R}^n. \end{aligned} \quad (2.18)$$

Como  $X$  es una función de  $\Omega$  a  $\mathbb{R}^n$ , esta se puede representar como un "vector aleatorio"  $X = (X_1, \dots, X_n)$  en donde cada coordenada es una variable aleatoria de  $\Omega$  a  $\mathbb{R}$ . Entonces,  $F_X((x_1, x_2, \dots, x_n))$

<sup>1</sup>Un ejemplo de este tipo de distribuciones es la distribución de Cantor [75].

se puede interpretar como la probabilidad de que cada coordenada  $X_i, i = 1, \dots, n$  del vector aleatorio  $X = (X_1, \dots, X_n)$  sea menor que  $x_i, i = 1, \dots, n$  al mismo tiempo. Es decir,

$$F_X((x_1, x_2, \dots, x_n)) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n). \quad (2.19)$$

Este hecho es la razón por la cual a la función de distribución de  $X$  se le llama también distribución conjunta multivariada de las variables aleatorias  $X_1, \dots, X_n$ . Además el concepto de distribución multivariada trae como consecuencia que la lógica en el estudio de variables aleatorias sobre  $\mathbb{R}^n$  es análogo al caso  $\mathbb{R}$ .

Además de facilitar el cálculo de los momentos de una variable aleatoria, la importancia del concepto de la función de distribución radica en que dada una  $F_X$ , entonces existe un espacio de probabilidad  $(\Omega, \mathcal{F}, P)$  y una variable aleatoria  $X$  cuya función de distribución es justamente  $F_X$  [49]. Por lo tanto es suficiente dar una función de distribución particular para saber que existe cierto espacio de probabilidad y una variable aleatoria definida sobre él, cuya función de distribución es la especificada. Este punto de vista es el que suele adoptarse, de este modo, estudiar el fenómeno aleatorio de interés es equivalente a estudiar la función de distribución  $F_X$  correspondiente a la variable aleatoria que da su correcta descripción.

Por ejemplo, el fenómeno aleatorio de observar cara con probabilidad  $p$ , tras el lanzamiento de una moneda, se puede estudiar usando la distribución Bernoulli de una variable aleatoria  $X \in \{0, 1\}$ , la cual codifica el evento de observar cara con la salida 1 y el evento de observar cruz con la salida 0. La distribución Bernoulli se denota como  $Ber(p)$  y tiene función de probabilidad [104]

$$p(x) = \begin{cases} p & \text{si } x = 0 \\ 1 - p & \text{si } x = 1 \end{cases}. \quad (2.20)$$

Otro ejemplo es el fenómeno de contar el número de estrellas fugaces durante una noche con media  $\lambda$ , el cual se puede describir con una distribución de Poisson de una variable aleatoria  $X \in \mathbb{N} \cup \{0\}$ . La distribución de Poisson se denota por  $Poiss(\lambda)$  y tiene función de probabilidad [104]

$$p(x) = \frac{e^{-\lambda x}}{x!}, x \in \mathbb{N} \cup \{0\}. \quad (2.21)$$

Un tercer ejemplo, es el registro de la concentración de mercurio en el cuerpo de atunes en el pacífico. Este fenómeno se puede describir con una distribución normal  $\mathcal{N}(\mu, \sigma)$ , con media  $\mu$  y varianza  $\sigma^2$  y función de densidad [68]

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (2.22)$$

Además, las distribuciones (2.20), (2.21) y (2.22) aparecen frecuentemente en la teoría cuántica; por ejemplo, la distribución Bernoulli también sirve para modelar las salidas de la medición de un qubit; la distribución de Poisson modela la probabilidad de un detector de fotones y la distribución normal aparece como la distribución de probabilidad límite de una gran cantidad de esquemas de medición. No obstante, también es muy frecuente que se tenga incertidumbre acerca del modelo probabilístico que mejor describe el fenómeno de interés. Es decir, no se conoce con certeza la mejor función de distribución que describe el fenómeno aleatorio bajo estudio. En este caso, la estadística inferencial es de gran utilidad, pues esta área de las matemáticas nos permite reducir (tanto como sea posible) la incertidumbre acerca del modelo a partir de datos obtenidos de la medición de ciertas características del fenómeno aleatorio.

### 2.3. Problema de estimación de parámetros

La estadística inferencial se define como la rama de las matemáticas que utiliza conjuntos de datos para obtener inferencias sobre una población basadas en el cálculo de probabilidades. Aquí se postula que los datos observados de una característica del fenómeno aleatorio de interés, son valores que puede tomar un observable  $X$  modelado como una variable aleatoria, la cual tiene una función de distribución  $F_X$ . Esta distribución se conoce a lo más parcialmente, por lo que solamente se puede asegurar que pertenece a una familia de distribuciones  $\mathcal{P}$  llamada modelo estadístico.

El problema principal de la estadística inferencial es la estimación de la función desconocida  $F_X$  a partir de un conjunto de datos. Una vez determinada esta función de distribución se puede hacer cualquier tipo de inferencia sobre la distribución de los datos. Frecuentemente, se puede suponer que las distribuciones en  $\mathcal{P}$  están indexadas por un vector de parámetros  $\theta$ , el cual toma valores en un conjunto  $\Theta$  llamado espacio paramétrico. Es decir,

$$\mathcal{P} = \{F_{X;\theta}, \theta \in \Theta \subseteq \mathbb{R}^s, s \in \mathbb{N}\}. \quad (2.23)$$

De esta forma, se reemplaza el problema de estimación original por la estimación de los parámetros que identifica la distribución de interés. Este problema se conoce como *estimación de parámetros* [19, 67].

El objetivo de la estimación de parámetros es especificar un plausible valor para  $\theta$  (problema de estimación puntual), o al menos determinar una cierto subconjunto de  $\Theta$  del que podamos afirmar con cierta probabilidad que contiene o no a  $\theta$  (estimación por conjuntos de confianza o pruebas de hipótesis).

Formalmente, los problemas de estimación de parámetros requieren dos ingredientes:

- 1) Un valor  $\theta$  a ser estimado, definido sobre un conjunto  $\Theta \subseteq \mathbb{R}^s$  llamado espacio paramétrico.
- 2) Un observable  $X$ , llamado muestra de tamaño  $n \in \mathbb{N}$ , el cual es descrito por una variable aleatoria sobre  $\mathbb{R}^n$  con distribución  $F_{X;\theta}$  y que toma valores sobre un espacio muestral  $\mathcal{X}$ . De forma que los datos observados son salidas  $x$  de la variable aleatoria  $X$ .

En este sentido, la solución del problema es encontrar un estimador adecuado para  $\theta$ .

Un estimador  $\hat{\theta} : \mathcal{X} \rightarrow \Theta$  es una función de los datos al espacio paramétrico, cuyos valores de su imagen son utilizados como estimaciones de  $\theta$ . Como  $X$  es una variable aleatoria, los estimadores son a su vez variables aleatorias, entonces cada estimación  $d = \hat{\theta}(x)$  se considera como una salida de la variable aleatoria  $\hat{\theta}(X)$ . De esta forma, para cuantificar la "cercanía" a  $\theta$  de cada estimación  $d$  se utiliza una función de costo  $c$  que satisface:

$$c(\theta, d) \geq 0 \quad \forall \theta, d \quad (2.24)$$

y

$$c(\theta, d = \theta) = 0 \quad \forall \theta. \quad (2.25)$$

Entonces, la precisión de un estimador  $\hat{\theta}(X)$  se determina utilizando la función de riesgo

$$R(\hat{\theta}, \theta) = E_{\theta} [c(\theta, \hat{\theta}(X))] = \int_{\mathcal{X}} c(\theta, \hat{\theta}) dF_{X;\theta}, \quad (2.26)$$

la cual representa el valor esperado de la función de costo respecto a la distribución de la muestra condicionada a  $\theta$ .

Por su gran cantidad de aplicaciones, la función de riesgo más utilizada es el **error cuadrático medio (ECM)** [110], el cual se define como

$$ECM(\hat{\theta}, \theta) = E_{\theta} \left[ \|\hat{\theta} - \theta\|^2 \right]. \quad (2.27)$$

El error cuadrático medio, se puede interpretar como la diferencia entre un error aleatorio y un error sistemático del estimador. Esta interpretación viene del hecho que (2.27) se puede escribir como la diferencia entre la varianza del estimador (medida de dispersión de la distribución de estimaciones) y el cuadrado del sesgo (distancia de la estimación promedio con el valor real del parámetro), el cual se interpreta como la contribución de los errores sistemáticos que se tienen al medir,

$$\begin{aligned} ECM(\hat{\theta}, \theta) &= E_{\theta} \left[ (\hat{\theta} - \theta)^2 \right] \\ &= E_{\theta} \left[ \left( \sqrt{(\hat{\theta} - \theta)^T (\hat{\theta} - \theta)} \right)^2 \right] \\ &= E_{\theta} \left[ \left( (\hat{\theta} - \theta)^T (\hat{\theta} - \theta) \right) \right] \\ &= E_{\theta} \left[ \sum_i (\hat{\theta}_i - E_{\theta}[\hat{\theta}_i]) \right] + \text{Sesgo}(\hat{\theta})^T \text{Sesgo}(\hat{\theta}) \\ &= \text{Tr}(\text{Cov}(\hat{\theta})) + \left( \|\text{Sesgo}(\hat{\theta})\| \right)^2. \end{aligned} \quad (2.28)$$

En particular, si  $\theta \in \Theta \subset \mathbb{R}$ ,

$$ECM(\hat{\theta}, \theta) = \text{Var}(\hat{\theta}) + \text{Sesgo}(\hat{\theta})^2. \quad (2.29)$$

En general, no existe un estimador que minimice (2.27) simultáneamente para todos los valores de  $\theta \in \Theta$  [67]. Una forma de solucionar esta dificultad es restringir la búsqueda de estimadores con error cuadrático medio mínimo a la clase de estimadores insesgados, es decir, a la clase de aquellos estimadores  $\hat{\theta}$  tales que su sesgo es cero.

## 2.4. Estimadores insesgados de varianza mínima

Cuando se restringe la búsqueda de estimadores que minimicen el error cuadrático medio a la clase de estimadores insesgados, es posible encontrar un **estimador insesgado de mínima varianza uniforme (EIMVU)** para todo  $\theta \in \Theta$ . Las condiciones para encontrar el estimador de varianza mínima se establecen en el teorema de Rao-Blackwell [29, 67], pero antes de enunciar este teorema es necesario definir el concepto de estadística suficiente y completa.

**Definición 4** (Estadística suficiente y completa).

i) Una estadística  $T$  es una función de la muestra

$$T : x \in \mathcal{X} \rightarrow T(x) = t \in \mathcal{T}, \quad (2.30)$$

donde  $\mathcal{T}$  es la imagen de  $T$ . Si  $x$  es generado por una variable aleatoria  $X$ , la función  $T(X)$  es una variable aleatoria con función de distribución definida por la medida de probabilidad

$$P_{T;\theta}(B) = P_{X;\theta}(\{x \in \mathcal{X} \mid T(x) \in B\}). \quad (2.31)$$

ii) Una estadística  $T$  es suficiente para  $\mathcal{P}$  si para cada  $t$  y  $\theta$ , la distribución de  $X$  condicionada a  $T = t$  es independiente de  $\theta$ .

iii) Un modelo  $\mathcal{P}$  es completo si para cualquier función  $h : \mathcal{X} \rightarrow \mathbb{R}$ :

$$E_{\theta}[h(X)] = 0 \quad \forall \theta \in \Theta \quad (2.32)$$

implica

$$P_{X;\theta}(h(X) = 0) = 1 \quad \forall \theta \in \Theta \quad (2.33)$$

Una estadística  $T$  es completa si el modelo estadístico  $\{F_{T;\theta} \mid \theta \in \Theta\}$  es completo.

**Teorema 1** (Rao-Blackwell [67]). *Sea  $X$  una variable aleatoria cuya distribución pertenece al modelo estadístico  $\mathcal{P} = \{F_{X;\theta} \mid \theta \in \Theta\}$  y  $T$  una estadística suficiente y completa para  $\mathcal{P}$ . Si  $\hat{\theta}_T$  es una función tal que*

$$E_{\theta}[\hat{\theta}_T(T)] = \theta, \quad (2.34)$$

entonces  $\hat{\theta}_T(T)$  es el estimador de varianza mínima para  $\theta$ .

Además, como el error cuadrático medio es una función convexa del conjunto de estimadores, cuando el estimador de varianza mínima uniforme existe, este es único. De esta forma, el teorema de Rao-Blackwell nos dice que la búsqueda del mejor estimador insesgado debe restringirse a aquéllos que son funciones de la estadística suficiente  $T$ . En particular, cuando  $\mathcal{P}$  forma a una familia exponencial, el estimador de varianza mínima uniforme se puede obtener de forma directa.

**Definición 5** (Familia exponencial). *Sea  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  una función no negativa,  $\eta \in \mathbb{R}^s$ ,  $T_1, \dots, T_s$  funciones medibles de  $\mathbb{R}^n$  a  $\mathbb{R}$ , y*

$$A(\eta) = \log \left[ \int \exp \left[ \sum_{i=1}^s \eta_i T_i(x) \right] h(x) dx \right]. \quad (2.35)$$

Si  $A(\eta) < \infty$ , la familia de funciones

$$f_{\eta}(x) = \exp \left[ \sum_{i=1}^s \eta_i T_i(x) - A(\eta) \right] h(x), \quad x \in \mathbb{R}^n, \quad (2.36)$$

indexada por el conjunto  $\{\eta : A(\eta) < \infty\}$ ; es una familia de distribuciones de probabilidad, llamada familia exponencial de  $s$  parámetros en su forma canónica.

Más aún, si  $\eta$  es una función de algún espacio paramétrico  $\Theta \subset \mathbb{R}^s$  al conjunto  $\{\eta : A(\eta) < \infty\}$ , las funciones

$$f_{\theta}(x) = \exp \left[ \sum_{i=1}^s \eta_i(\theta) T_i(x) - A(\eta(\theta)) \right] h(x), \quad x \in \mathbb{R}^n, \quad (2.37)$$

indexadas por  $\theta \in \Theta$  es llamada familia exponencial de  $s$  parámetros.

**Corolario 1.** Si un modelo estadístico  $\mathcal{P} = \{\exp [\sum_{i=1}^s \eta_i(\theta)T_i(x) - A(\eta(\theta))] h(x), x \in \mathbb{R}^n, \theta \in \Theta \subset \mathbb{R}^s\}$  forma una familia exponencial, entonces, las conclusiones del teorema 1 se cumplen para  $\theta = (\theta_1, \dots, \theta_s)$  y  $T = (T_1, \dots, T_s)$ .

Cabe resaltar que una gran mayoría de modelos utilizados en la literatura, como lo son la distribución binomiales, Poisson, exponencial, Gamma y normal por mencionar algunas, forman familias exponenciales [29, 67, 68]. Desafortunadamente, incluso cuando se tenga un modelo donde se puede construir un EIMVU, calcular su varianza resulta una tarea bastante difícil en general. Por tal motivo, se han derivado desigualdades para acotar inferiormente a la varianza de cualquier estimador, que son sencillas de calcular. La desigualdad más importante es la **cota inferior de Crámer-Rao (CICR)** [67, 68].

## 2.5. Desigualdad de Crámer-Rao

La CICR acota a la varianza de una estimador por el inverso de la información de Fisher. La información de Fisher es una medida de información que nos dice cuánta información acerca de  $\theta$  hay en los datos. La información de Fisher se calcula a partir de la función de verosimilitud.

**Definición 6** (Verosimilitud).

Para una observación  $x \in \mathcal{X}$ , la función de verosimilitud  $\mathcal{L} : \Theta \rightarrow \mathbb{R}_+$  es definida por

$$\mathcal{L}(\theta; x) = f_\theta(x), \quad (2.38)$$

donde  $f_\theta$  son las densidades o funciones de probabilidad de  $X$  correspondientes a las distribuciones de  $\mathcal{P}$ .

Para que la información de Fisher esté bien definida, el modelo estadístico debe satisfacer algunas condiciones de regularidad [68]:

- C1** Las distribuciones en  $\mathcal{P}$  tienen soporte común  $\mathcal{A} = \{x \in \mathcal{X} \mid f_\theta(x) > 0\}$  e independiente de  $\theta$ .
- C2** El espacio paramétrico  $\Theta \subseteq \mathbb{R}^n$  es un conjunto abierto.
- C3** Para todo  $x \in \mathcal{A}$ , las derivadas parciales de la función de verosimilitud existen.

De esta forma, si las condiciones **C1** a **C3** se cumplen, la información de Fisher se define como la covarianza de la derivada logarítmica, es decir, la covarianza del gradiente del logaritmo de la función verosimilitud, la cual también es llamada función score.

**Definición 7** (Score e Información de Fisher).

- i) Para todo  $x \in \mathcal{A}$ , el vector

$$V(\theta; x) = \nabla_\theta l(\theta; x) = \left( \frac{\partial}{\partial \theta_1} l(\theta; x), \dots, \frac{\partial}{\partial \theta_n} l(\theta; x) \right)^T \quad (2.39)$$

es llamado función *score* o derivada logarítmica, donde  $l(\theta; x) = \log(\mathcal{L}(\theta; x))$ .

ii) La información de Fisher es una matriz de  $n \times n$  definida por

$$I_X(\theta) = \text{Cov}_\theta [V(\theta; X)]. \quad (2.40)$$

Los elementos de esta matriz son

$$I_X(\theta)_{ij} = \text{E}_\theta \left[ \frac{\partial}{\partial \theta_i} l(\theta; X), \frac{\partial}{\partial \theta_j} l(\theta; X) \right]. \quad (2.41)$$

Ya contando con la definición de la información de Fisher, la cota inferior de Crámer-Rao se puede enunciar como el siguiente teorema:

**Teorema 2** (Cota inferior de Crámer-Rao CICR [68]). *Supongamos que las condiciones de regularidad (2.5) se satisfacen para  $\mathcal{P}$  y que la información de Fisher es no singular. Si  $\hat{\theta}$  es un estimador insesgado para  $\theta$ , entonces*

$$\text{Cov}_\theta [\hat{\theta}] \geq I_X(\theta)^{-1}, \quad (2.42)$$

donde  $I_X(\theta)^{-1}$  es la matriz inversa de la información de Fisher.

En la ecuación (2.42) se tiene una desigualdad de matrices. La desigualdad  $A \geq B$ , para dos matrices  $A$  y  $B$  es equivalente a verificar que  $A - B \geq 0$ , es decir,  $A \geq B$  si y sólo si  $A - B$  es una matriz positiva semidefinida.

En particular, para el caso de parámetros escalares ( $\Theta \subseteq \mathbb{R}$ ), la CICR obtienen la forma

$$\text{Var}_\theta (\hat{\theta}) \geq \frac{1}{I_X(\theta)}, \quad (2.43)$$

con

$$I_X(\theta) = \text{E}_\theta \left[ \left( \frac{\partial}{\partial \theta} l(\theta; x) \right)^2 \right] = \text{Var}_\theta [V(\theta; x)]. \quad (2.44)$$

La información de Fisher tiene la ventaja de ser una cantidad independiente del estimador, por lo que la CICR es una cota universal para la clase de estimadores insesgados que evade la dificultad de no conocer en general las distribuciones de los estimadores y por tanto la adversidad de calcular  $\text{Var}(\hat{\theta})$ .

A los estimadores que saturan la CICR se les llama eficientes, esto es debido a que si se cuenta con una muestra de tamaño  $N$ ,  $X = (X_1, \dots, X_N)$  cuyos elementos son variables aleatorias independientes e idénticamente distribuidas, la función de distribución de  $X$  se puede escribir como

$$F_X((x_1, \dots, x_N)) = P(X_1 \leq x_1)P(X_2 \leq x_2) \cdots P(X_N \leq x_N)$$

y en consecuencia, la información de Fisher es igual a  $NI_{X_i}(\theta)$ , donde  $X_i$  es cualquier componente del vector aleatorio  $X$  [19]. Así, la varianza de un estimador eficiente decrece como el inverso de  $N$  conforme el tamaño de la muestra crece.

Adicionalmente, la cota de Crámer-Rao se puede extender al caso de estimadores sesgados, donde se obtiene una cota inferior para el error cuadrático medio [57].

$$ECM(\hat{\theta}, \theta) \geq \frac{[1 + \text{Sesgo}(\hat{\theta})']^2}{I_X(\theta)} + \text{Sesgo}(\hat{\theta})^2, \quad (2.45)$$

con  $\text{Sesgo}(\hat{\theta})' = \frac{\partial}{\partial \theta} \text{Sesgo}(\hat{\theta})$ .



## 2.6. Estimador de máxima verosimilitud

Es común pensar entonces que el objetivo final en la estimación puntual de parámetros es encontrar un estimador que sature la CICR. Sin embargo, no siempre es posible encontrar un estimador eficiente. Por ejemplo, los EIMVU son los mejores estimadores insesgados para una sola medición, pero estos solo saturan la CICR cuando el modelo estadístico es una familia exponencial completa [29]. De otro modo, no existe un estimador insesgado que sature la desigualdad (2.42). No obstante, debido a que la cota de Crámer-Rao decrece como función del inverso del tamaño de la muestra, en muchas aplicaciones se suele trabajar en el límite asintótico (cuando el tamaño de datos tiende a infinito). En este límite y bajo ciertas condiciones de regularidad adicionales, el EMV puede saturar la CICR, dicho de otra forma, el estimador de máxima verosimilitud es asintóticamente eficiente y por tanto será el mejor estimador que podamos obtener [45, 67, 68].

**Definición 8** (Estimador de máxima verosimilitud **EMV**).

Un estimador  $\hat{\theta}_{EMV}$  para  $\theta \in \Theta$  es llamado de máxima verosimilitud, si

$$\hat{\theta}_{EMV}(X) = \arg \max_{\theta \in \Theta} \mathcal{L}(\theta; X) \quad (2.46)$$

donde  $f_\theta$  son las densidades respecto a una medida común, para las distribuciones de  $\mathcal{P}$ .

Bajo las siguientes condiciones de regularidad sobre  $\mathcal{P}$ , el estimador  $\hat{\theta}_{EMV}$  es consistente y asintóticamente eficiente:

**C4** El modelo debe ser identificable:

$$\theta \neq \theta_0, \quad f_\theta \neq f_{\theta_0}. \quad (2.47)$$

Esta condición establece que la función de verosimilitud tiene un máximo global.

**C5** La información de Fisher existe para cualquier  $\theta \in \Theta$  y es semidefinida positiva.

**C6** Existe un conjunto abierto de  $\Theta$  que contiene al parámetro real, tal que la función de verosimilitud es al menos tres veces derivable con probabilidad 1.

**C7** Existe una función  $M_{jkl}$  tal que

$$\left| \frac{\partial^3}{\partial \theta_j \partial \theta_k \partial \theta_l} l(\theta; x) \right| \geq M_{jkl}(x), \quad (2.48)$$

y su valor esperado es finito.

Los modelos estadísticos que satisfacen **C4-C7** son llamados regulares.

**Teorema 3** (Consistencia y eficiencia del estimador de máxima verosimilitud [67]). *Sea  $X = (X_1, \dots, X_n)$  un observable de variables aleatorias independientes e idénticamente distribuidas con densidad (o función de probabilidad)  $f_\theta$ , con  $\theta = (\theta_1, \dots, \theta_m) \in \Theta \subset \mathbb{R}^m$  que satisfacen las condiciones **C4-C7**. Entonces, con probabilidad tendiendo a 1 cuando  $n \rightarrow \infty$ , la secuencia de estimadores  $\hat{\theta}_n = \hat{\theta}_{EMV}(X_1, \dots, X_n)$  cumple con:*

a)  $\hat{\theta}_n^{(j)}$  es consistente para estimar  $\theta_j$ . Es decir,

$$\lim_{n \rightarrow \infty} P \left( \hat{\theta}_n^{(j)} - \theta_j \right) \rightarrow 1, \quad (2.49)$$

donde  $\hat{\theta}_n^{(j)}$  es el  $j$ -ésimo componente del estimador  $\hat{\theta}_n$ .

b)  $\sqrt{n} \left( \hat{\theta}_n - \theta \right)$  es asintóticamente una distribución normal multivariada con vector cero de media y matriz de covarianza  $[I(\theta)]^{-1}$ , con  $I(\theta)$  la información de Fisher asociada a la densidad  $f_\theta$ .

c)  $\hat{\theta}_n^{(j)}$  es asintóticamente eficiente en el sentido que

$$\sqrt{n} \left( \hat{\theta}_n^{(j)} - \theta_j \right) \xrightarrow{d} \mathcal{N} \left( 0, [I(\theta)^{-1}]_{jj} \right)^2, \quad (2.50)$$

es decir, la distribución de  $\hat{\theta}_n^{(j)}$  tiende a una distribución normal con media  $\theta_j$  y varianza  $[nI(\theta)_{jj}]^{-1}$ .

En general, el EMV es un estimador con sesgo, sin embargo este decae con el inverso de  $\sqrt{n}$ , por lo que en el límite asintótico el error cuadrático medio del estimador de máxima verosimilitud es igual a su varianza. De esta manera, el teorema previo trae como consecuencia que la varianza de la distribución límite del estimador de máxima verosimilitud sea la cota de Crámer-Rao. Por lo que bajo las condiciones de regularidad (2.6) y en el límite asintótico, el EMV es el mejor estimador que se puede obtener. Por estas buenas propiedades, el estimador de máxima verosimilitud es de los más utilizados.

En particular, el estimador de máxima verosimilitud es muy útil para estimar intervalos de confianza. Un intervalo de confianza es un conjunto de valores que puede contener a  $\theta \in \Theta \subseteq \mathbb{R}$  con cierta probabilidad.

**Definición 9** (Intervalo de confianza IC).

Un intervalo a un nivel  $(1 - \alpha)$  de confianza para  $\theta$  es un intervalo de la forma  $(\hat{\theta}_1, \hat{\theta}_2)$  que satisface

$$P \left( \hat{\theta}_1 < \theta < \hat{\theta}_2 \right) = 1 - \alpha \quad (2.51)$$

en donde  $\hat{\theta}_1$  y  $\hat{\theta}_2$  son funciones de los datos y  $\alpha \in [0, 1]$ . Al parámetro  $\alpha$  se le llama nivel de significancia y se interpreta como el número de intervalos sobre 100 que no contienen el valor de  $\theta$ .

Esta definición se puede extender al caso de un espacio paramétrico que es un subconjunto de  $\mathbb{R}^n$ ,  $n > 1$ , como un conjunto de intervalos de confianza para cada componente de  $\theta$ .

Especialmente, una aplicación común de este concepto es cuando se desea estimar la media  $\theta$  de una distribución normal  $\mathcal{N}(\theta, \sigma^2)$  y con varianza  $\sigma^2$  conocida, a partir de una muestra  $X = (X_1, \dots, X_n)$ . Para este caso, el estimador de máxima verosimilitud coincide con la media aritmética, esto es,

$$\hat{\theta}_{EMV} = \bar{X} = \frac{X_1 + \dots + X_n}{n}. \quad (2.52)$$

---

<sup>2</sup>Una sucesión  $X_1, X_2, \dots, X_n$  de variables aleatorias converge en distribución, a una variable aleatoria  $X$  ( $X_n \xrightarrow{d} X$ ) si  $\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x)$ ,  $\forall x \in \mathbb{R}$ , donde  $F_{X_n}$  y  $F_X$  denotan las funciones de distribución de las variables aleatorias  $X_n$  y  $X$ , respectivamente.

Es fácil verificar que la variable aleatoria

$$Z = \frac{\bar{X} - \theta}{\frac{\sigma}{\sqrt{n}}}. \quad (2.53)$$

sigue una distribución normal estándar (distribución normal con media 1 y varianza 1) [104]. De esta forma, calculando puntos críticos  $x_{\alpha/2}$  y  $x_{-\alpha/2}$  (o equivalentemente sus versiones estandarizadas  $z_{\pm\alpha/2}$ ) tales que

$$P(\bar{X} \geq x_{\alpha/2}) = P(Z \geq z_{\alpha/2}) = \frac{\alpha}{2}, \quad (2.54)$$

se cumple para cualquier estimación  $\bar{x}$  que

$$P\left(\bar{x} + z_{-\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \theta \leq \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right). \quad (2.55)$$

Además, como la distribución normal es simétrica se tiene que  $z_{-\alpha/2} = -z_{\alpha/2}$  se tiene que el intervalo de confianza para  $\theta$  es

$$\left(\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right). \quad (2.56)$$

Un aplicación del resultado (2.56) se obtiene cuando un EMV  $\hat{\theta}_{EMV}$  es consistente y asintóticamente eficiente para un parámetro  $\theta \in \mathbb{R}$ . Para esta situación,  $\sigma/\sqrt{n}$  se puede estimar por el recíproco de la raíz cuadrada de  $n$  veces la información de Fisher y  $\bar{X}$  se puede cambiar por  $\hat{\theta}_{EMV}$ . De esta forma, dado el resultado de una muestra de tamaño  $n$ ,  $x_1, \dots, x_n$ , se puede estimar un intervalo a  $1 - \alpha$  de confianza para  $\theta$  por la siguiente fórmula

$$\left(\hat{\theta}_{EMV}(x_1, \dots, x_n) - z_{\alpha/2} \cdot \frac{1}{\sqrt{nI(\theta)}}, \hat{\theta}_{EMV}(x_1, \dots, x_n) + z_{\alpha/2} \cdot \frac{1}{\sqrt{nI(\theta)}}\right), \quad (2.57)$$

con  $I(\theta)$  la información de Fisher asociada a la variable aleatoria que sigue cualquier dato  $x_i$ ,  $i \in \{1, \dots, n\}$ .



## Capítulo 3

# Estimación de parámetros en sistemas cuánticos

Dentro de la teoría clásica, los espacios de probabilidad surgen de la necesidad de describir fenómenos aleatorios. No obstante, la aleatoriedad de los eventos es relativa al observador y depende de qué tanta información y control tenga sobre las condiciones del fenómeno bajo estudio. En principio, si se tuviese total control e información acerca de todas las condiciones iniciales del sistema en cuestión, es teóricamente posible describir este fenómeno con un modelo determinista. De esta manera, la aparente aleatoriedad observada en ciertos eventos no es más que una consecuencia de falta de información. En cambio, en el caso cuántico, los sistemas llevan consigo de forma inherente la causa de la aleatoriedad que se observa al medir un observable sobre ellos, por lo que los resultados de cada medición sobre un estado del sistema, son siempre descritos como una variable aleatoria independientemente del control que se tenga sobre las condiciones del fenómeno bajo estudio.

### 3.1. Estados cuánticos

En la teoría de la mecánica cuántica, todo sistema tiene asociado un espacio de Hilbert  $\mathcal{H}$ , con dimensión  $\dim(\mathcal{H}) \leq \infty$ . Los operadores densidad que actúan sobre este espacio, son usados para representar los estados en que puede encontrarse el sistema. Un operador densidad es un operador lineal  $\rho : \mathcal{H} \rightarrow \mathcal{H}$  de tipo traza, tal que  $\rho \geq 0$  y  $\text{Tr}(\rho) = 1$ . Físicamente, un estado se puede interpretar como la clase de equivalencia de todas las posibles preparaciones del sistema que producen las mismas distribuciones de probabilidad para cualquier posible medición sobre el sistema [57].

El conjunto de operadores densidad sobre un espacio de Hilbert  $\mathcal{H}$  es denotado por  $S(\mathcal{H})$ . Este conjunto es un conjunto convexo, es decir, que para cualesquiera  $\rho_1, \rho_2 \in S(\mathcal{H})$  y cualquiera  $\alpha \in [0, 1]$ , la combinación convexa  $\alpha\rho_1 + (1-\alpha)\rho_2 \in S(\mathcal{H})$ . La combinación convexa de dos operadores densidad representa una preparación aleatoria del sistema entre dos estados diferentes.

Se puede probar que el conjunto  $S(\mathcal{H})$  es un conjunto compacto (con respecto a su topología  $w^*$ ) [66], por lo que  $S(\mathcal{H})$  puede ser descrito como la envolvente convexa de sus puntos extremales [57]. Los puntos extremales de  $S(\mathcal{H})$  son llamados *estados puros*, mientras que el resto son llamados

---

<sup>1</sup>La topología  $w^*$  de  $S(\mathcal{H})$  es la topología con el menor número de conjuntos abiertos que hace a todos los funcionales lineales definidos sobre el dual de  $S(\mathcal{H})$  continuos [101].

estados mixtos. Además, cualquier  $\rho \in \mathcal{S}(\mathcal{H})$  puede ser diagonalizado, esto es,

$$\rho = \sum_{i=1}^k p_i |\psi_i\rangle \langle \psi_i|, \quad (3.1)$$

con  $p_i \geq 0$ ,  $\sum_{i=1}^k p_i = 1$ , y  $\{|\psi_i\rangle\}$ ,  $i = 1, \dots, k \leq \infty$  una base ortonormal adecuada. Se puede ver fácilmente que los estados puros son los proyectores de rango uno  $\rho = |\psi\rangle \langle \psi|$ , con  $|\psi\rangle \in \mathcal{H}$ .

Un sistema compuesto de dos subsistemas  $A$  y  $B$  con espacios de Hilbert  $\mathcal{H}_A$  y  $\mathcal{H}_B$  se representa por el producto tensorial  $\mathcal{H}_A \otimes \mathcal{H}_B$  de los espacios de Hilbert de los subsistemas  $A$  y  $B$ . Cuando el espacio de Hilbert  $\mathcal{H}_A$  ( $\mathcal{H}_B$ ) tiene una base ortonormal  $\{|u\rangle_1^A, \dots, |u\rangle_{d_A}^A\}$  ( $\{|u\rangle_1^B, \dots, |u\rangle_{d_B}^B\}$ ), el espacio  $\mathcal{H}_A \otimes \mathcal{H}_B$  tiene una base ortonormal

$$\left\{ |u\rangle_1^A \otimes |u\rangle_1^B, \dots, |u\rangle_1^A \otimes |u\rangle_{d_B}^B, \dots, |u\rangle_2^A \otimes |u\rangle_2^B, \dots, |u\rangle_2^A \otimes |u\rangle_{d_B}^B, \dots, |u\rangle_1^A \otimes |u\rangle_{d_A}^B, \dots, |u\rangle_{d_A}^A \otimes |u\rangle_{d_B}^B \right\}. \quad (3.2)$$

De este modo, cada elemento  $|z\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$  puede ser escrito como  $|z\rangle = \sum_{i,j} z^{i,j} |u\rangle_i^A \otimes |u\rangle_j^B$  y el producto tensorial de dos vectores  $|x\rangle^A = \sum_k x^k |u\rangle_k^A$  y  $|y\rangle^B = \sum_j y^j |u\rangle_j^B$ ,  $x^k, y^j \in \mathbb{C}$  definirse como  $|x\rangle^A \otimes |y\rangle^B = \sum_k \sum_j x^k y^j |u\rangle_k^A \otimes |u\rangle_j^B$ .

El producto tensorial  $X_A \otimes X_B$  de un operador  $X_A$  en  $\mathcal{H}_A$  y un operador  $X_B$  en  $\mathcal{H}_B$  se define como un operador en  $\mathcal{H}_A \otimes \mathcal{H}_B$  tal que

$$X_A \otimes X_B(|u\rangle \otimes |v\rangle) = X_A(|u\rangle) \otimes X_B(|v\rangle), \quad \forall |u\rangle \in \mathcal{H}_A \quad |v\rangle \in \mathcal{H}_B. \quad (3.3)$$

La traza de este producto tensorial satisface la relación

$$\text{Tr}[X_A \otimes X_B] = \text{Tr}[X_A] \text{Tr}[X_B]. \quad (3.4)$$

Si los sistemas  $A$  y  $B$  son independientes y se encuentran en los estados  $\rho_A \in \mathcal{S}(\mathcal{H}_A)$  y  $\rho_B \in \mathcal{S}(\mathcal{H}_B)$ , entonces el estado del sistema compuesto se representa por el producto tensorial  $\rho_A \otimes \rho_B$ . Cuando un operador densidad  $\rho \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$  puede ser descrito como una combinación convexa de productos tensoriales

$$\rho = \sum_i p_i \rho_A^i \otimes \rho_B^i, \quad p_i \geq 0, \quad \sum_i p_i = 1, \quad \rho_A^i \in \mathcal{S}(\mathcal{H}_A), \quad \rho_B^i \in \mathcal{S}(\mathcal{H}_B) \quad (3.5)$$

es llamado estado separable. Por otro lado, cuando el estado  $\rho \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$  no tiene la forma (3.5) se le llama estado entrelazado.

Cuando se tienen  $n$  sistemas idénticos con espacio de Hilbert  $\mathcal{H}$ , el sistema compuesto se escribe como  $\underbrace{\mathcal{H} \otimes \dots \otimes \mathcal{H}}_{n\text{-veces}}$  y es denotado  $\mathcal{H}^{\otimes n}$  por brevedad. En particular, si todos los sistemas son

independientes, y cada sistema se encuentra en un estado  $\rho$ , el estado del sistema compuesto es  $\rho^{\otimes n} = \underbrace{\rho \otimes \dots \otimes \rho}_{n\text{-veces}}$ .

## 3.2. Observables cuánticos

Como ya se mencionó, al realizar una medición sobre un sistema cuántico se produce como resultado la observación de un evento con cierta probabilidad. De esta forma, cada medida  $M$  que pueda efectuarse sobre el sistema con posibles salidas sobre un conjunto  $\mathcal{X} \subseteq \mathbb{R}$  define una variable

aleatoria, la cual, dado el estado  $\rho$  del sistema, asigna una regla para predecir las probabilidades de los posibles eventos.

La descripción más general de una medida cuántica es dada por el concepto de **POVM** (del acrónimo en inglés **positive operator-valued measure**)

**Definición 10** (POVM [57, 66]). Sea  $\mathcal{X}$  un espacio muestral,  $\mathcal{F}$  una  $\sigma$ -álgebra de subconjuntos de  $\mathcal{X}$  y  $\mathcal{B}(\mathcal{H})$  el conjunto de operadores acotados y positivos de un espacio de Hilbert  $\mathcal{H}$ . Un POVM  $M$  con salidas en el conjunto  $\mathcal{X}$  es un mapeo  $M : \mathcal{F} \rightarrow \mathcal{B}(\mathcal{H})$ , el cual asocia a cada evento  $B \in \mathcal{F}$  con un operador  $M(B) \in \mathcal{B}(\mathcal{H})$  que satisface lo siguiente

- i)  $M(B) \geq 0, \forall B \in \mathcal{F}$ ,
- ii)  $M(\Omega) = I$ ,
- iii)  $M(\cup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} M(B_i)$ , donde  $B_i \cap B_j = \emptyset, i \neq j$ .

Equivalentemente, un POVM se puede pensar como un mapeo

$$M : \mathcal{X} \rightarrow \mathcal{B}(\mathcal{H}), \quad (3.6)$$

tal que

$$\sum_{x \in \mathcal{X}} M(x) = I \quad (3.7)$$

Cabe resaltar que los operadores auto adjuntos (definidos como los observables cuánticos en los libros básicos de mecánica cuántica) son un caso particular de un POVM, pues para cualquier operador auto adjunto  $A$ , si tomamos  $\mathcal{X} = \text{Spec}(A)$ , el espectro de  $A$ , se obtiene un POVM cuyos elementos son los operadores  $A(\lambda) = E_\lambda$ , los proyectores a los eigenespacios asociados al eigenvalor  $\lambda \in \text{Spec}(A)$ .

Así, cuando un POVM  $M$  se especifica, se puede dar un único mapeo afín entre cualquier operador densidad  $\rho$  y cualquier evento  $B$ , vía la regla de Born [14]

$$P(B | \rho) = \text{Tr}[M(B)\rho]. \quad (3.8)$$

En efecto, las condiciones *i) – iii)* de la definición (10) garantizan la normalización, positividad y aditividad de eventos disjuntos de la probabilidad para cualquier estado  $\rho \in \mathcal{S}(\mathcal{H})$ . Además, a consecuencia de la ecuación (3.8) los POVMs también son llamados medidas u observables cuánticos.

### 3.3. Problema de estimación de parámetros en sistemas cuánticos

Suponiendo que el sistema está en un estado  $\rho_0$ , el objetivo de la teoría de estimación en sistemas cuánticos es estimar  $\rho_0$  a partir del resultado de mediciones del sistema [50]. En muchas ocasiones, se tiene el conocimiento a priori de que  $\rho_0$  pertenece a un subconjunto particular de estados parametrizados, pero del cual se tiene incertidumbre acerca del parámetro. Bajo esta suposición, podemos reemplazar el problema original por la estimación de parámetros que identifican el estado bajo estudio. Este problema es llamado estimación de parámetros en sistemas cuánticos. Esta situación se modela como una familia de estados parametrizados (llamado modelo estadístico cuántico)

$$\mathcal{S}_\theta = \{\rho_\theta; \theta \in \Theta \subset \mathbb{R}^s\}, \quad (3.9)$$

donde  $\theta \in \Theta$  es un vector de parámetros desconocido, y  $\Theta$  es el conjunto de valores plausibles de  $\theta$  llamado espacio paramétrico. Entonces, el objetivo de la teoría de estimación de parámetros en sistemas cuánticos es producir la mejor estimación para  $\theta$  bajo la premisa de que el estado del sistema se encuentra en alguno de los estados de  $\mathcal{S}_\theta$ .

La construcción de una familia de estados (3.9) es consecuencia de la evolución dinámica sobre los estados, pues para cualquier estado inicial  $\rho$ , que evolucione a partir de un proceso dinámico dependiente de un parámetro  $\theta$  se obtiene un operador densidad  $\rho_\theta \in \mathcal{S}_\theta$  como resultado. Así, suponiendo que el sistema se encuentra en  $\rho_\theta$ , al obtener datos a partir de mediciones sobre el sistema, se puede obtener una estimación de  $\theta$  realizando un análisis clásico de estos.

En este sentido, la implementación práctica de la teoría de estimación de parámetros en sistemas cuánticos se contempla en dos pasos: asumiendo que el sistema puede prepararse en  $n \in \mathbb{N}$  estados idénticos  $\rho_\theta$ , el primero consiste en la apropiada elección de un POVM  $M^n$  que se realizarán sobre el sistema  $\mathcal{H}^{\otimes n} = \underbrace{\mathcal{H} \otimes \cdots \otimes \mathcal{H}}_{n\text{-veces}}$ , y el segundo, en procesar los resultados de estas mediciones (datos) a través de un estimador  $\hat{\theta}_n$ . A la dupla  $(M^n, \hat{\theta}_n)$  se le conoce como estrategia de estimación.

En concreto, cuando se fija un POVM  $M$ , el problema se reduce al caso clásico de estimar una distribución en la familia  $\{F_{M;\theta}, \theta \in \Theta\}$ . De este modo, se obtiene un modelo estadístico clásico para cada POVM diferente. Por tal motivo, para buscar la estrategia de estimación óptima (aquella que minimice la función de riesgo) es imprescindible realizar un proceso de optimización sobre la elección del POVM, Aunque tal proceso depende altamente de la dimensión del espacio paramétrico.

### 3.4. Estimación escalar

Primero consideremos el problema de estimar un estado que pertenece a la familia

$$\mathcal{S}_\theta = \{\rho_\theta; \theta \in \Theta \subset \mathbb{R}\}, \quad (3.10)$$

y supongamos que el desempeño de una estrategia de medición se cuantifica con el error cuadrático medio (2.29). Entonces dado un POVM  $M^n$  con salidas en un conjunto  $\mathcal{X}$  y aplicando el estimador de máxima verosimilitud  $\hat{\theta}_{n,EMV}$ , sabemos, de la desigualdad de Crámer-Rao (2.43), que bajo las condiciones de regularidad 2.6 el error cuadrático medio de este estimador es aproximadamente  $(I_{M^n}(\theta))^{-1}$  en el límite asintótico, donde  $I_{M^n}(\theta)$  es la información de Fisher acerca de  $\theta$  para la distribución de probabilidad obtenida a partir de la regla de Born

$$P(B | \theta) = \text{Tr}[M^n(B)\rho_\theta], \quad B \text{ un conjunto de Borel sobre } \mathcal{X}. \quad (3.11)$$

Sin embargo, como fue mencionado, la diferencia principal al caso de estimación clásica es que cada POVM produce una distribución diferente y por tanto diferentes cotas de Crámer-Rao, por lo que la maximización de la información de Fisher sobre el conjunto de POVMs nos daría la menor cota de Crámer-Rao.

Esta información de Fisher es conocida como [información de Fisher cuántica \(IFC\)](#) y es denotada por  $I_Q(\theta)$ , además a la mínima cota inferior de Cramér-Rao se le llama la [cota inferior de Crámer-Rao Cuántica \(CICRC\)](#) [16, 44, 53, 55, 57]. De esta manera, si se utiliza un POVM con información de Fisher cuántica acerca de  $\theta$  y que produzca un modelo estadístico regular, el estimador de máxima verosimilitud alcanza la CICRC en el límite asintótico. Por este motivo, una gran parte



del trabajo en la estimación de parámetros en sistemas cuánticos trata el problema de encontrar POVMs con información cuántica de Fisher. A estos POVMs se les llama óptimos.

Formalmente, la información cuántica de Fisher  $I_Q(\theta)$  puede ser definida por la derivada simétrica logarítmica [16, 54, 57], por lo que la cota inferior de Crámer-Rao puede ser descrita por el siguiente teorema.

**Teorema 4** (Cota inferior de Cramér-Rao cuántica (CICRC) [16, 53, 85]). *Dada una estrategia de estimación  $(M^n, \hat{\theta}_n)$ ,  $n \in \mathbb{N}$  para  $\theta \in \Theta$ , insesgada, es decir*

$$E_\theta \left[ \left( M^n, \hat{\theta}_n \right) \right] = \theta, \forall \theta \in \Theta, \quad (3.12)$$

entonces la siguiente desigualdad se cumple

$$\text{Var}_\theta \left[ \left( M^n, \hat{\theta} \right) \right] \geq \frac{1}{I_Q(\theta)}, \quad (3.13)$$

donde

$$I_Q(\theta) = \text{Tr} \left[ \rho_\theta L(\theta)^2 \right], \quad (3.14)$$

con  $L(\theta)$  el operador Hermitiano de derivada simétrica logarítmica definido implícitamente por

$$\frac{\partial \rho_\theta}{\partial \theta} = \frac{1}{2} (\rho_\theta L(\theta) + L(\theta) \rho_\theta) \quad (3.15)$$

nombrado score cuántico.

Además, el siguiente corolario, probado en [85], confirma la intuición física acerca del significado de  $I_Q$

**Corolario 2.** *Sea  $\mathcal{M}(\mathcal{H})$  el espacio de POVMs sobre el espacio de Hilbert  $\mathcal{H}$  de un sistema cuántico. Denotando como  $I_M(\theta)$  a la información de Fisher para un POVM  $M \in \mathcal{M}(\mathcal{H})$ , se cumple que*

$$I_Q(\theta) = \max_{M \in \mathcal{M}(\mathcal{H})} I_M(\theta). \quad (3.16)$$

Un caso muy particular, pero de gran relevancia es cuando  $\rho_\theta = |\psi_\theta\rangle \langle \psi_\theta|$  es un estado puro. En este caso, el score cuántico se puede calcular fácilmente [96]. Ya que,  $\rho_\theta = \rho_\theta^2$ , implica,

$$\frac{\partial \rho_\theta}{\partial \theta} = \frac{\partial}{\partial \theta} (\rho_\theta^2) = \frac{\partial \rho_\theta}{\partial \theta} \rho_\theta + \rho_\theta \frac{\partial \rho_\theta}{\partial \theta}. \quad (3.17)$$

De esta manera, (3.15) se puede reescribir como

$$L(\theta) = 2 \frac{\partial \rho_\theta}{\partial \theta}, \quad (3.18)$$

y por tanto la información cuántica de Fisher para estados puros puede ser escrita como

$$I_Q(\theta) = \text{Tr} \left[ |\psi_\theta\rangle \langle \psi_\theta| L(\theta)^2 \right] = 4 \text{Tr} \left[ \left( \frac{\partial}{\partial \theta} (|\psi_\theta\rangle \langle \psi_\theta|) \right)^2 |\psi_\theta\rangle \langle \psi_\theta| \right]. \quad (3.19)$$

Y si además la evolución paramétrica de cualquier  $\rho_\theta \in \mathcal{S}_\theta$  se obtiene a partir de una evolución unitaria sobre un estado  $\rho$ ,

$$\rho_\theta = e^{-i\theta A} \rho e^{i\theta A} \in \mathcal{S}_\theta, \quad (3.20)$$

donde  $A$  es un operador auto adjunto sobre  $\mathcal{H}$  llamado generador. Así, de acuerdo con la ecuación (3.19), la información cuántica de Fisher se puede escribir como

$$\begin{aligned} I_Q(\theta) &= 4 \left[ \text{Tr} [A^2 \rho_\theta] - \text{Tr} [A \rho_\theta]^2 \right] \\ &= 4 \text{Var}_\rho [A]. \end{aligned} \quad (3.21)$$

En [16] se mostró que el POVM  $M = \{M_\theta(j) \mid j \in \text{Spec}(L(\theta))\}$  construido por los proyectores  $M_\theta(j)$  a los eigenspacios del operador  $L(\theta)$  es óptimo. A pesar de esto, en general, este POVM depende del parámetro desconocido que se desea estimar (son localmente óptimos). Por lo tanto, estrategias de estimación basadas en esta clase de POVMs no son prácticas pues es necesario conocer de antemano el valor del parámetro que se desea estimar [38].

Para tratar este problema, dos enfoques han sido adoptados. El primero consiste en esquemas de estimación adaptativos basados en POVMs localmente óptimos. En este tipo de esquemas, en principio, uno puede construir en el límite asintótico el POVM localmente óptimo sin conocer a priori el valor del parámetro [9, 38, 60, 85]. No obstante, un conjunto de condiciones de regularidad para cada modelo estadístico involucrado en el esquema deben ser satisfechas [38].

En particular, para el esquema de estimación adaptativa de estados cuánticos (AQSE del acrónimo en inglés, adaptive quantum state estimation) [38, 85], es necesario que cada modelo estadístico involucrado en la estrategia sea regular para garantizar que el estimador de máxima verosimilitud sea consistente y eficiente, y por tanto, alcanzar la cota de Crámer-Rao cuántica. Sin embargo, para problemas de estimación con parámetros periódicos, las funciones de verosimilitud producidas por POVMs localmente óptimos son no identificables [10, 22]. Es decir, para este caso los modelos estadísticos son no regulares y por consiguiente, no hay razón matemática para asegurar que este métodos adaptativo (y otros similares) puedan funcionar para este caso, lo que limita su uso para una gran cantidad de posibles aplicaciones.

El segundo enfoque consiste en encontrar condiciones específicas donde el POVM sea independiente de  $\theta$  [22, 111], por lo que en principio, si uno aplica una sucesión grande de estos POVMs de forma independiente y usando el estimador de máxima verosimilitud es posible alcanzar la cota de Crámer-Rao cuántica. A pesar de esto, para lograr este objetivo, se debe garantizar que el POVM óptimo producirá un modelo estadístico regular. De lo contrario el EMV no será consistente, lo que implica que la cota de Cramér-Rao no pueda ser alcanzada por el POVM  $M$ .

### 3.4.1. Estimación multiparamétrica

El proceso de estimación, cuando se tiene el caso de estimar un parámetro de un estado desconocido que pertenece a una familia

$$\mathcal{S}_\theta = \{\rho_\theta; \theta \in \Theta \subset \mathbb{R}^n, n > 1 \in \mathbb{N}\}, \quad (3.22)$$

es análogo al caso escalar. Primero se escoge un POVM  $M^n$  apropiado, el cual se efectúa sobre el espacio de Hilbert  $\mathcal{H}^{\otimes n}$ , y después se procesan los datos obtenidos con un estimador  $\hat{\theta}_n$ . Suponiendo que el estimador es insesgado, la matriz de covarianzas estará acotada por la desigualdad de Crámer-Rao (2.42), es decir, por  $I_{M^n}(\theta)^{-1}$ , donde  $I_{M^n}(\theta)$  es una matriz no singular definida por (2.40) y  $I_{M^n}(\theta)^{-1}$  su matriz inversa.

Como ya se mencionó, la información de Fisher  $I_{M^n}(\theta)$  es una función del POVM escogido. Por tal motivo, con el fin de obtener la cota de Crámer-Rao más pequeña, se debe realizar nuevamente una optimización sobre el espacio de POVMs. Sin embargo, en este caso no existe una única desigualdad de Crámer-Rao cuántica como en el caso escalar. Este resultado fue probado por [53, 57] y expresado en el siguiente teorema

**Teorema 5.** *Considerando una estrategia de estimación insesgada  $(M^n, \hat{\theta}_n)$ ,  $n \in \mathbb{N}$  para  $\theta \in \Theta$ . Las siguientes desigualdades matriciales se cumplen*

$$\text{Cov}_\theta \left[ (M^n, \hat{\theta}) \right] \geq I_x(\theta)^{-1}, \quad x = r, s, \quad (3.23)$$

donde

$$I_r(\theta) = \text{Tr} \left[ \rho_\theta \tilde{L}(\theta) \tilde{L}(\theta)^\dagger \right], \quad (3.24)$$

con  $\tilde{L}(\theta)$  es el operador Hermitiano de derivada derecha logarítmica definido por

$$\tilde{L}(\theta) = \rho_\theta^{-1} \frac{\partial \rho_\theta}{\partial \theta} \quad (3.25)$$

y

$$I_s(\theta) = \text{Tr} \left[ \rho_\theta L(\theta)^2 \right], \quad (3.26)$$

con  $L(\theta)$  el operador de derivada simétrica logarítmica.

En efecto, se han encontrado varios casos donde la desigualdad basada en la derivada derecha logarítmica ofrece una cota de Crámer-Rao más pequeña que usando la derivada simétrica logarítmica [54, 57], por lo que no hay una relación de orden entre estas dos desigualdades. Aunado a este problema, en general no existe una estrategia de estimación que sature (3.23) [51]. Por este motivo, en muchas ocasiones se suele adoptar el enfoque de minimizar el error cuadrático medio directamente para cada  $i$ -ésimo componente del parámetro. Para esto, se busca minimizar la cantidad escalar  $\text{Tr} \left( W \cdot \text{Cov}_\theta \left[ M, \hat{\theta} \right] \right)$  sobre el espacio de POVMs, con  $W$  una matriz positiva definida adecuada de pesos. Esta cantidad representa la suma de los errores cuadráticos medio de todas las componentes del parámetro para una estrategia insesgada.

Incluso adoptando este enfoque se tiene el problema de que, en general, el POVM que minimiza este criterio de optimalidad depende del parámetro que se desea estimar, volviéndola inviable [38], por lo que al igual que en el caso escalar, se tiene la necesidad de utilizar estrategias adaptativas. Aunque si el sistema tiene presente simetrías es posible encontrar una solución teórica al problema de minimizar el error cuadrático medio (y varios otros criterios de optimización).

Las simetrías del sistema describen el formalismo covariante de estimación utilizando la teoría de grupos. Con esta estructura matemática se puede tomar ventaja de las simetrías del sistema para resolver el problema de minimizar funciones de costo invariantes bajo la operación del grupo (como el ECM) sobre el espacio de POVMs. En la siguiente sección se dará una introducción a este formalismo.

### 3.5. Estimación covariante

Consideremos el caso donde un grupo  $G$  finito o continuo (supondremos para un grupo de Lie localmente compacto para el caso continuo) actúa sobre el espacio paramétrico  $\Theta$ , esto es, existe un mapeo  $\varphi : G \times \Theta \rightarrow \Theta$  tal que

$$\begin{aligned} \forall \theta \in \Theta, \varphi(e, x) &= e \cdot \theta = \theta, \text{ donde } e \text{ es el neutro de } G \\ \forall \theta \in \Theta, g, h \in G, \varphi(g \cdot h, x) &= g \cdot (h \cdot \theta). \end{aligned} \quad (3.27)$$

Las condiciones de arriba implican que para cada elemento  $g \in G$ , la función  $\varphi_g = \varphi(g, \cdot) : \Theta \rightarrow \Theta$  es una función biyectiva y por tanto existe un morfismo de grupos entre  $G$  y el grupo simétrico  $S_\Theta$  formado por todas las funciones biyectivas de  $\Theta$  en sí mismo, bajo la composición de funciones.

El problema de estimación covariante se caracteriza por el hecho de que el espacio paramétrico es invariante bajo la acción del grupo, es decir, para cualquier  $g \in G$  se tiene que  $g\Theta = \Theta$ , donde

$$g\Theta := \{g\theta \mid \theta \in \Theta\}. \quad (3.28)$$

En general, dada una acción de un grupo  $G$  sobre un conjunto  $C$ , se dice que dos elementos  $c, d \in C$  son equivalentes si existe  $g \in G$  tal que  $gc = d$ . De esta forma, una órbita  $O_c$  de  $G$  para cada  $c \in C$  se define como el conjunto de todos los elementos de  $C$  equivalentes a  $c$ , es decir,

$$O_c = \{gc \mid g \in G\}. \quad (3.29)$$

En particular, para el problema de estimación covariante, la órbita para cualquier elemento  $\theta \in \Theta$  resultante de la acción de  $G$  es el conjunto

$$O_\theta = \{y \in \Theta \mid \exists g \in G : g \cdot \theta = y\} \subset \Theta. \quad (3.30)$$

Se puede notar que las órbitas para dos elementos distintos  $\theta_1$  y  $\theta_2$  en  $\Theta$ , o bien, son iguales o son disjuntas, por consiguiente las órbitas bajo la acción de  $G$  forman una partición del espacio paramétrico. Por este motivo, el espacio paramétrico puede considerarse como la unión de un cierto número de órbitas disjuntas,

$$\Theta = \bigcup_{x \in X} O_x, \quad (3.31)$$

donde  $X = \Theta/G$ . Además, dado un punto arbitrario  $x_0 \in O_{x_0}$ , existe una biyección entre su órbita y su grupo estabilizador

$$G_{x_0} = \{g \in G \mid g \cdot \theta_0 = x_0\}, \quad (3.32)$$

lo que introduce la identificación  $O_{x_0} = G/G_{x_0}$  y entonces  $\Theta = \bigcup_{x \in X} G/G_{x_0}$  [34].

De acuerdo al teorema de Wigner [117], la acción del grupo  $G$  sobre  $\Theta$  se puede trasladar al espacio  $\mathcal{S}(\mathcal{H})$  de estados cuánticos por el automorfismo

$$\rho \mapsto U_g \rho U_g^\dagger = \mathcal{A}_g(\rho), \quad (3.33)$$

donde  $\{U_g \in \mathcal{B}(\mathcal{H}) \mid g \in G\}$  es una representación proyectiva de  $G$  en  $\mathcal{H}$ . El automorfismo  $\mathcal{A}_g$  representa una simetría del conjunto  $\mathcal{S}(\mathcal{H})$  con respecto al grupo  $G$ . Por ende el problema de

estimación covariante se puede caracterizar por familias  $\mathcal{S} = \{\rho_\theta \mid \theta \in \Theta\}$  que son invariantes bajo esta simetría, es decir, para cualquier elemento  $g \in G$  se tiene que

$$g\mathcal{S} := \{\mathcal{A}_g(\rho_\theta) \mid \theta \in \Theta\} = \mathcal{S}. \quad (3.34)$$

Consecuentemente, el modelo estadístico cuántico se puede escribir como la unión de un cierto número de conjuntos de estados disjuntos

$$\mathcal{S} = \cup_{x \in X} \{\mathcal{A}_g(\rho_\theta) \mid \theta \in \Theta\}, \quad (3.35)$$

con  $X = \Theta/G$ .

Dentro de este contexto, un POVMs  $M$  sobre  $\mathcal{H}$  que toma valores en  $\Theta$  es llamado *covariante* con respecto a la representación unitaria  $U$ , si para cada evento  $B$  y  $g \in G$ , se satisface la siguiente ecuación

$$M(Bg^{-1}) = U_g^\dagger M(B) U_g, \quad (3.36)$$

donde  $Bg = \{\theta \mid \theta = g\theta', \theta' \in B\}$ . Esta definición es equivalente a que la distribución de probabilidad producida por  $M$  sea invariante bajo  $G$ ,

$$P(B \mid \rho) = P(gB \mid \mathcal{A}_g(\rho)) \quad \forall \rho \in \mathcal{S}(\mathcal{H}). \quad (3.37)$$

Al conjunto de POVMs con salidas en  $\Theta$  lo denotaremos como  $\mathcal{M}(\Theta)$  y al conjunto de POVMs covariantes con respecto a la representación unitaria en cuestión,  $\mathcal{M}(\Theta, U)$ .

Aunque hay trabajos donde se consideran acciones de grupo en general [27], el caso de estudio más común para la estructura de  $\mathcal{M}(\Theta)$  es el de acciones transitivas [54, 57]. En este caso, el espacio paramétrico es compuesto por una única órbita

$$\Theta = G/G_0. \quad (3.38)$$

En particular para esta situación, los POVMs covariantes pueden expresarse de forma relativamente simple.

Una medida  $\mu(dg)$  sobre la  $\sigma$ -álgebra de Borel de subconjuntos de  $G$   $\mathcal{B}(G)$ , es invariante por la izquierda (derecha) si

$$\mu(gA) = \mu(A) \quad (\mu(Ag) = \mu(A)), \quad \forall g \in G, \quad (3.39)$$

donde  $gA = \{y \in G \mid y = g \cdot a, a \in A\}$  y  $Ag = \{y \in G \mid y = a \cdot g, a \in A\}$ . Un grupo es unimodular si existe una medida  $\mu$  invariante, es decir, que sea invariante por la izquierda y por la derecha. Adicionalmente, si el grupo unimodular es compacto, entonces  $\mu(g) \leq \infty$  y esta se puede normalizar a uno,  $\mu(G) = 1$ . Para describir la estructura de los POVMs covariantes nos restringiremos a grupos unimodulares compactos.

Si también  $G_0$  es unimodular, entonces existe una medida  $\nu$  invariante sobre la  $\sigma$ -álgebra de Borel de subconjuntos de  $\Theta$   $\mathcal{B}(\Theta)$

$$\nu(Bg) = \nu(B), \quad \forall B \in \mathcal{B}(\Theta) \text{ y } g \in G \quad (3.40)$$

y si en adición  $G_0$  es compacto se puede construir una relación entre  $\mu$  y  $\nu$  de la forma

$$\nu(B) = \mu\left(\theta^{[-1]}(B)\right), \quad (3.41)$$

donde  $\theta^{[-1]}(B) = \{g \mid g\theta_0 = \theta \in B\}$ , para cualquier  $B \in \mathcal{B}(\Theta)$ . Entonces, para cualquier función  $f$  integrable en  $\Theta$  se cumple la relación

$$\int_G f(g\theta_0)\mu(dg) = \int_\Theta f(\theta)\nu(d\theta). \quad (3.42)$$

Estos conceptos son de necesarios para expresar la estructura de los POVMs covariantes, la cual es enunciada en el siguiente teorema.

**Teorema 6 ([57]).** *Sea  $G$  un grupo localmente compacto unimodular, y  $\{U_g \mid g \in G\}$  una representación proyectiva de  $G$  sobre  $\mathcal{H}$ . Si  $K \subseteq G$  es un subgrupo compacto, entonces un POVM  $M = \{M(d\theta)\}_{\theta \in \Theta}$  con salidas en  $\Theta = G/K$  es covariante si y sólo si tiene la forma*

$$M(d\theta) = U_{g(\theta)}P_0U_{g(\theta)}^\dagger\nu(d\theta), \quad (3.43)$$

donde  $g(\theta) \in G$  es cualquier elemento representativo de la clase de equivalencia  $\theta \in \Theta$ ,  $\nu(d\theta)$  es una medida de grupo invariante sobre  $\theta$ , y  $P_0$  un operador Hermitiano positivo sobre  $\mathcal{H}$ , que conmuta con cada elemento de  $\{U_g \in \mathcal{B}(\mathcal{H}) \mid g \in K\}$  y satisface

$$\int_G U_gP_0U_g^\dagger\mu(dg) = I. \quad (3.44)$$

La clase de POVMs covariantes  $\mathcal{M}(\Theta, U)$  tiene ventajas cuando sus salidas son tomadas como las estimaciones del parámetro y la función de costo bajo consideración es invariante bajo la acción de  $G$ ,  $c(\hat{\theta}, \theta) = c(g\hat{\theta}, \theta), \forall g \in G$ . Para este caso se obtiene que

$$\begin{aligned} E_{g\theta, M} [c(\hat{\theta}, g\theta)] &= \int_\Theta c(\hat{\theta}, g\theta) \text{Tr} [M(d\hat{\theta})\rho_{g\theta}] \\ &= \int_\Theta c(g\hat{\theta}', g\theta) \text{Tr} [M(d\hat{\theta}')\rho_\theta] \\ &= E_{\theta, M} [c(\hat{\theta}, \theta)], \forall M \in \mathcal{M}(\Theta, U) \text{ y } \hat{\theta}' := g^{-1}\hat{\theta}. \end{aligned} \quad (3.45)$$

Por lo que es posible encontrar una medida óptima dentro de  $\mathcal{M}(\Theta, U)$  trabajando en el enfoque Bayesiano, es decir, si tomamos la esperanza de  $c(\hat{\theta}, \theta)$  con respecto a una medida de probabilidad a priori  $\pi$  sobre  $\theta$ . En particular, cuando  $\pi$  es una distribución uniforme,  $\pi$  es equivalente a la medida de Haar  $\nu_\Theta$  sobre  $\Theta$  (medida invariante normalizada sobre  $\Theta$ ). El siguiente teorema sintetiza este resultado

**Teorema 7 ([55]).** *Si  $\Theta$  es compacto,*

$$\begin{aligned} &\min_{M \in \mathcal{M}(\Theta)} \int_\Theta E_{\theta, M} [c(\hat{\theta}, \theta)] \nu_\Theta(d\theta) \\ &= \min_{M \in \mathcal{M}(\Theta, U)} \int_\Theta E_{\theta, M} [c(\hat{\theta}, \theta)] \nu_\Theta(d\theta). \end{aligned} \quad (3.46)$$

donde  $\nu_\Theta$  es la medida de Haar sobre  $\theta$ .

Bajo las condiciones del teorema previamente expuesto, es posible restringir la búsqueda de mediciones óptimas a la clase  $\mathcal{M}(\Theta, U)$  de POVMs. En especial, para el caso de grupos de estabilidad triviales  $K = \{e\}$  (con  $e$  la identidad del grupo  $G$ ), existe una biyección entre los elementos de  $\Theta$  y  $G$ ; es decir,  $\Theta \cong G$ , lo que tiene como consecuencia que el valor esperado  $E_{\theta, M} [c(\hat{\theta}, \theta)]$  sea una constante para todo  $\theta \in \Theta$ . De este hecho, se puede restringir la optimización de la esperanza de funciones de costo invariantes sobre el conjunto de POVMs covariantes. Un ejemplo típico de esta situación es el problema de estimación de fase. En este caso, el grupo  $G$  es isomorfo a  $\Theta = [0, 2\pi)$ , logrando obtener la optimización del [ECM](#) dentro del conjunto de POVMs covariantes.

### 3.6. Simetría temporal

El principio de relatividad Galileana en sistemas cuánticos nos dice que *la estadística de cualquier experimento es la misma en cualquier sistema de referencia* [57]. Es decir, si consideramos que nuestro sistema está en un estado base  $\rho$  y entonces se le realiza una medición  $M$ , la función de distribución resultante debe ser invariante ante transformaciones del grupo Galileano,

$$P(M(B) | \rho) = P(M(gB) | \rho_g), \quad (3.47)$$

donde  $gB$  es el evento  $B$  tras la aplicación de  $g$  y  $\rho_g$  es el nuevo estado que resulta tras la transformación de  $g$ , donde  $g$  es un elemento del grupo Galileano. En otros términos, todo POVM  $M$  debe satisfacer la propiedad de covarianza respecto al grupo de transformaciones Galileano, el cual incluye traslaciones sobre todas las coordenadas del sistema de referencia (incluyendo la coordenada temporal).

En particular, cuando  $g$  es una traslación temporal, la distribución de salidas de una medición  $M$  debe ser invariante. Es decir,

$$P(M(B) | \rho) = P(M(B+t) | \rho_t), \quad (3.48)$$

con  $B+t = \{\lambda+t | \lambda \in B\}$  es el desplazamiento temporal de  $B$ ,  $t$  unidades de tiempo.

La ecuación (3.48) refleja la covarianza de  $M$ , la cual es consecuencia de la existencia de una simetría temporal en el sistema. De acuerdo al teorema de Wigner [117], esta simetría temporal se puede representar como un grupo de operadores unitarios  $\{V_t\}_{t \in \mathbb{R}}$ , donde cada elemento del grupo, de acuerdo con el teorema de Stone [48], tiene la forma

$$V_t = e^{-itH}, \quad (3.49)$$

donde  $H$  es un operador autoadjunto en el espacio de Hilbert  $\mathcal{H}$  del sistema llamado Hamiltoniano.

En general, un Hamiltoniano de un sistema cuántico se puede escribir como

$$H = \frac{P^2}{2m} + V(Q), \quad (3.50)$$

donde  $P$  es el operador de momento y  $V$  el operador potencial, función del operador de posición  $Q$ . Entonces, de acuerdo con (3.49), la familia de estados

$$\rho_t = e^{-itH} \rho e^{itH}, \quad t \in \mathbb{R}, \quad (3.51)$$

corresponde a la evolución temporal, determinada por el Hamiltoniano del sistema, de un estado base  $\rho$  para diferentes valores del parámetro de evolución  $t$ . Además por la propiedad de covarianza, cada elemento de un POVM  $M$  debe evolucionar temporalmente como

$$M(B + t) = e^{-itH} M(B) e^{itH}, \quad t \in \mathbb{R}, \quad \text{y } B \text{ un conjunto de Borel.} \quad (3.52)$$

Esta simetría es de especial importancia ya que todos los sistemas cuánticos cerrados la cumplen, lo que trae como consecuencia que en múltiples escenarios como la estimación de fase y estimación de frecuencia, la dependencia paramétrica sea causada por esta simetría.



## Capítulo 4

# Estimación de fase en qubits

Un qubit es un sistema descrito por un espacio de Hilbert  $\mathcal{H}$  2-dimensional. Para este sistema, cada Hamiltoniano se puede representar como una matriz hermitiana de  $2 \times 2$ . Una base para el espacio de matrices hermitianas de  $2 \times 2$  es la matriz identidad  $I$ , junto con las matrices de Pauli

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (4.1)$$

Bajo la representación de esta base, cualquier Hamiltoniano del sistema puede expresarse como

$$H = c_0 I + \vec{c} \cdot \vec{\sigma} \quad (4.2)$$

con  $c_0 \in \mathbb{R}$ ,  $\vec{c} = c_1 \hat{x} + c_2 \hat{y} + c_3 \hat{z} \in \mathbb{R}^3$  y  $\vec{\sigma} = \sigma_1 \hat{x} + \sigma_2 \hat{y} + \sigma_3 \hat{z}$ , con  $\sigma_i$ , para  $i = 1, 2, 3$ , las matrices de Pauli (4.1). Por consiguiente, los elementos de la representación unitaria correspondiente a la simetría temporal (3.49) se pueden escribir de la siguiente forma: [82, 88]

$$V_t = e^{i\gamma} e^{-i|\vec{c}|t\vec{n} \cdot \frac{\vec{\sigma}}{2}}, \quad (4.3)$$

donde  $\vec{n} = c_x/|\vec{c}|\hat{x} + c_y/|\vec{c}|\hat{y} + c_z/|\vec{c}|\hat{z} \in \mathbb{R}^3$ , y  $\gamma = c_0 t$  una fase global que puede ser omitida.

Debido a que  $V_t$  produce una evolución periódica, es decir,  $V_{t+\frac{2\pi}{|\vec{c}|}} = V_t$ , la evolución temporal es equivalente a la producida por los operadores unitarios

$$U_\theta = e^{-i\theta\vec{n} \cdot \frac{\vec{\sigma}}{2}}, \quad (4.4)$$

donde  $\theta = -|\vec{c}|t \pmod{2\pi}$  es un parámetro de fase del operador exponencial. De esta forma, la acción del grupo  $\Theta = [0, 2\pi)$  con la suma módulo  $2\pi$ , sobre un estado base  $\rho \in \mathcal{S}(\mathcal{H})$  del qubit produce estados de la forma

$$\rho_\theta = U_\theta \rho U_\theta^\dagger, \quad (4.5)$$

y una familia parametrizada de operadores densidad  $\mathcal{S} = \{\rho_\theta; \theta \in \Theta = [0, 2\pi)\}$  sobre el espacio de Hilbert  $\mathcal{H}$  del sistema. El problema de estimación de fase sobre qubits surge cuando se tiene desconocimiento sobre el parámetro  $\theta$  de un estado  $\rho_\theta \in \mathcal{S}$  y se desea estimar a partir de los resultados de mediciones sobre el sistema. El objetivo final de este problema es encontrar la mejor estrategia para estimar la fase  $\theta \in \Theta$ .

Para tratar este problema es conveniente utilizar la representación de la esfera de Bloch. En esta representación, el estado base se escribe como

$$\rho = \frac{1}{2} (I + \vec{a} \cdot \vec{\sigma}) , \quad \|\vec{a}\| \leq 1 , \quad (4.6)$$

y la acción de la transformación  $U_\theta$  se puede representar como una rotación del vector  $\vec{a}$  un ángulo  $\theta$  sobre la esfera de Bloch alrededor de otro vector  $\vec{n}$ , llamado eje. Así, un el estado  $\rho_\theta$  es totalmente determinado por

$$\vec{a}(\theta) = \cos(\theta)\vec{a} + (\sin(\theta)(\vec{n} \times \vec{a})) + \left( 2(\vec{n} \cdot \vec{a}) \sin^2 \left( \frac{\theta}{2} \right) \right) \vec{n} . \quad (4.7)$$

Usando esto, se puede probar que el score cuántico (3.15) tiene la siguiente forma [22]:

$$L(\theta) = (\vec{n} \times \vec{a}(\theta)) \cdot \vec{\sigma} , \quad (4.8)$$

lo que implica, de acuerdo a la ecuación (3.14), que la información cuántica de Fisher sea

$$I_Q = \|a\|^2 - (\vec{a} \cdot \vec{n})^2 . \quad (4.9)$$

De esta manera, si el estado inicial de un qubit es un estado puro ( $\|\vec{a}\| = 1$ ) y el vector  $\vec{a}$  es ortogonal al eje de rotación  $\vec{n}$ , uno obtiene el valor máximo para la información de Fisher cuántica,  $I_Q^{\max} = 1$ . Por tal motivo, a estas condiciones se les llama condiciones iniciales óptimas.

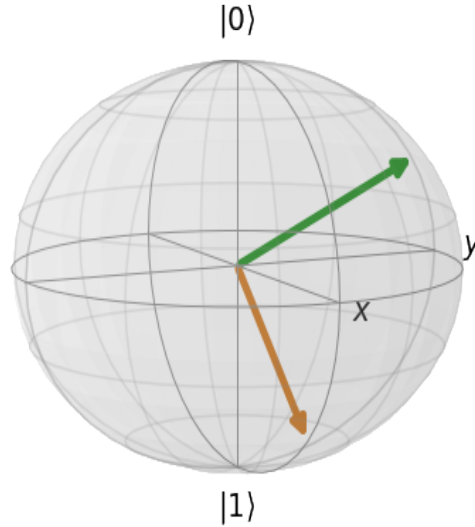


Figura 4.1: Representación de la esfera de Bloch de las condiciones iniciales óptimas. El eje de rotación  $\vec{n}$  es representado por la flecha verde y el vector  $\vec{a}$  por la flecha anaranjada. Los vectores que apuntan a los polos de la esfera corresponden a los vectores  $|0\rangle = (0, 0, 1)$  y  $|1\rangle = (0, 0, -1)$ .

A continuación discutiremos diferentes estrategias de estimación para  $\theta$ , resaltando los problemas que se tiene al utilizar algunas de las más comunes que se encuentran en la literatura, así como una propuesta propia que da solución en gran medida a los problemas previamente expuestos.

## 4.1. POVMs localmente óptimos

Los POVMs localmente óptimos para este problema de estimación se construyen directamente de la ecuación (4.8), pues los operadores de proyección a los eigenspacios del operador (4.8) están dados por la expresión:

$$M_g(0) = \frac{1}{2} \left( I + \frac{(\vec{n} \times \vec{a}(g))}{\|\vec{n} \times \vec{a}(g)\|} \cdot \vec{\sigma} \right), \quad M_g(1) = \frac{1}{2} \left( I - \frac{(\vec{n} \times \vec{a}(g))}{\|\vec{n} \times \vec{a}(g)\|} \cdot \vec{\sigma} \right), \quad g \in \Theta. \quad (4.10)$$

Entonces, la familia de POVMs  $M_g = \{M_g(0), M_g(1)\}$ , con salidas  $\mathcal{X} = \{0, 1\}$ , produce la información de Fisher:

$$I_M(\theta; g) = \frac{I_Q \cos^2(\theta - g)}{1 - I_Q \sin^2(\theta - g)}. \quad (4.11)$$

Por lo que  $I_M(\theta, g) = I_Q(\theta)$  cuando  $g = \theta$  para cualquier  $\theta \in \Theta$ . Esto es, el POVM  $M = \{M_\theta(0), M_\theta(1)\}$  es localmente óptimo.

Ahora, un caso muy especial es cuando  $\vec{a} \perp \vec{n}$ , pues para esta situación se tiene que

$$I_M(\theta; g) = \frac{\cos^2(\theta - g)}{1 - \sin^2(\theta - g)} = 1 = I_Q^{\max}. \quad (4.12)$$

Es decir, cuando el estado inicial es preparado en las condiciones iniciales óptimas, el POVM  $M$  es óptimo, independiente del valor de  $\theta \in \Theta$ . Y así, cualquier POVM  $M_g$  es óptimo.

No obstante, contrario a lo que la intuición puede dictar, el POVM  $M$  no produce una estrategia óptima de estimación en el sentido que el estimador de máxima verosimilitud para una muestra de medidas independientes del POVM  $M$  no puede saturar el límite de Cramér-Rao cuántico, incluso en el límite asintótico. Esto es debido a que  $M$  produce funciones de verosimilitud no-identificables (con dos máximos globales), lo que trae como consecuencia que el estimador de máxima verosimilitud no sea consistente.

Cuando una secuencia de  $N$  sistemas cuánticos idénticos descritos por un espacio de Hilbert  $\mathcal{H}$  son preparados en el mismo estado  $\rho_\theta$ , se describe el sistema compuesto por el  $N$  producto tensorial  $\underbrace{\mathcal{H} \otimes \cdots \otimes \mathcal{H}}_{N\text{-veces}}$  y sus estados son descritos por el  $N$  producto tensorial de los estados  $\underbrace{\rho_\theta \otimes \cdots \otimes \rho_\theta}_{N\text{-veces}}$ . Entonces, la función de verosimilitud producida por una muestra  $\vec{x} = (x_1, \dots, x_N) \in \mathcal{X}^N$  obtenida por la una secuencia de  $N$  POVMs  $M_g$  independientes es

$$\begin{aligned} \mathcal{L}(\theta | \vec{x}; g) &= \text{Tr}[(M_g(x_1) \otimes \cdots \otimes M_g(x_N)) \rho_\theta \otimes \cdots \otimes \rho_\theta] = \prod_{i=1}^N \text{Tr}[M_g(x_i) \rho_\theta] = \prod_{i=1}^N p(x_i | \theta; g) \\ &= p(0 | \theta; g)^m p(1 | \theta; g)^{N-m}, \end{aligned} \quad (4.13)$$

donde  $m$  es el número de 0's en  $\vec{x}$ , y

$$p(x | \theta; g) = \begin{cases} \text{Tr}[M_\theta(1) \rho_\theta] = \frac{1}{2} [1 + \sin(\theta - g) \sqrt{I_Q}] , & x = 1, \\ \text{Tr}[M_\theta(0) \rho_\theta] = \frac{1}{2} [1 - \sin(\theta - g) \sqrt{I_Q}] , & x = 0. \end{cases} \quad (4.14)$$

En consecuencia, la estimación de máxima verosimilitud para  $\theta$ , denotada como  $\hat{\theta}_{\text{EMV}}(\vec{x})$ , es:

$$\hat{\theta}_{\text{EMV}}(\vec{x}) = \arg \max_{\theta \in \Theta} \mathcal{L}(\theta | \vec{x}; g) = \arcsin \left( \frac{1}{\sqrt{I_Q}} \left[ 1 - \frac{2m}{N} \right] \right) + g, \quad (4.15)$$

la cual regresa dos valores dentro de  $\Theta = [0, 2\pi)$ . Esto prueba que la información de Fisher no es suficiente para caracterizar el error de una estrategia de estimación.

Ahora consideremos el caso cuando se tiene un POVM con  $\infty > k \geq 3$  número de salidas. Aquí, cualquier elemento de un POVM  $M = \{M(i)\}$  con salidas en un conjunto  $\mathcal{X} = \{1, 2, \dots, k\}$  puede ser escrito como  $M(i) = f_0^{(i)} I + \vec{f}^{(i)} \cdot \vec{\sigma}$ , con  $\sum_{i=1}^k f_0^{(i)} = 1$  y  $\sum_{i=1}^k \vec{f}^{(i)} = \vec{0}$ . Cuando un qubit está en un estado  $\rho_\theta$ , la probabilidad de medir la salida  $i$  condicionada a  $\theta$  es

$$p(i | \theta) = \text{Tr} [\rho_\theta M(i)] = f_0^{(i)} + \vec{a}(\theta) \cdot \vec{f}^{(i)}. \quad (4.16)$$

Una condición necesaria y suficiente para la identificabilidad de las funciones de verosimilitud, para cualesquiera  $\theta_1, \theta_2 \in \Theta$ , es que la ecuación  $p(i | \theta_1) = p(i | \theta_2), i = 1, \dots, k$  admita una única solución en el espacio paramétrico  $\Theta$  [77].

Las condiciones para tener identificabilidad expuestas en el párrafo anterior se pueden satisfacer fácilmente utilizando POVMs con múltiples salidas. Sin embargo, estos POVMs no necesariamente tienen que ser óptimos. En particular, fuera de las condiciones iniciales óptimas, no existe un POVM  $M$  independiente del parámetro que sature la CICRC [4]. Este hecho se encuentra en [22] donde se muestra una familia de POVMs con más de dos elementos, la cual produce funciones de verosimilitud identificables, pero sólo en las condiciones iniciales óptimas es localmente óptima. No obstante, de acuerdo al teorema 7 si es posible construir un POVM que para una medida minimiza el error cuadrático medio en cualquier condición inicial y sin el problema de identificabilidad utilizando el enfoque covariante.

#### 4.1.1. Estimación de fase covariante en qubits

La estimación de fase es un problema de estimación covariante con grupo estabilizador trivial, pues el conjunto  $G = [0, 2\pi)$  junto con la suma módulo  $2\pi$  actúa sobre el mismo implicando que el conjunto  $\Theta$  sea isomorfo a  $G$ . Así,  $U = \{U_\theta\}_{\theta \in [0, 2\pi)}$  es una  $G$  representación unitaria de  $G$  sobre  $\mathcal{H}$ . Por lo que al usar (3.43), el POVM  $M$  para estimar  $\theta$  tiene la forma

$$M(d\hat{\theta}) = e^{-i\hat{\theta}\vec{n} \cdot \frac{\vec{\sigma}}{2}} P_o e^{i\hat{\theta}\vec{n} \cdot \frac{\vec{\sigma}}{2}} \frac{d\hat{\theta}}{2\pi}, \quad (4.17)$$

donde  $P_o$  es un operador positivo que satisface la ecuación (3.44).

El generador de la representación unitaria es el operador  $J := \vec{n} \cdot \frac{\vec{\sigma}}{2}$ , el cual tiene la siguiente descomposición espectral

$$J = \frac{1}{2} \left| \frac{1}{2} \right\rangle \left\langle \frac{1}{2} \right| - \frac{1}{2} \left| -\frac{1}{2} \right\rangle \left\langle -\frac{1}{2} \right|, \quad (4.18)$$

donde usamos el hecho que  $|\pm\frac{1}{2}\rangle \langle \pm\frac{1}{2}| = \frac{1}{2} (I \pm \vec{n} \cdot \vec{\sigma})$ . Además, en la base  $\{|\frac{1}{2}\rangle, |-\frac{1}{2}\rangle\}$ , uno puede expresar  $M(d\hat{\theta})$  como

$$M(d\hat{\theta}) = \frac{d\hat{\theta}}{2\pi} \sum_{m,n \in \{-\frac{1}{2}, \frac{1}{2}\}} e^{i\hat{\theta}(n-m)} p_{nm} |m\rangle \langle n|, \quad (4.19)$$

donde  $p_{mn} = \langle m|P_o|n\rangle$ . Por consiguiente cualquier POVM covariante es caracterizado por los números  $0 \leq p_{nm} \leq 1$  y como consecuencia de la ecuación (3.11)

$$P(\hat{\theta} \in B|\theta) = \int_B \frac{d\hat{\theta}}{2\pi} \sum_{m,n \in \{-\frac{1}{2}, \frac{1}{2}\}} e^{i(n-m)(\hat{\theta}-\theta)} p_{mn} \langle n|\rho|m\rangle. \quad (4.20)$$

Notemos que la medida (4.20) es una distribución de probabilidad  $2\pi$ -periódica, por lo que se deben considerar los momentos de la variable aleatoria  $e^{i\hat{\theta}}$  en lugar de los momentos de  $\theta$ . El primer momento para una distribución circular  $p(\theta)$  se define como

$$\mathbb{E} [e^{i\theta}] = \int_0^{2\pi} e^{i\theta} p(\phi) d\theta, \quad (4.21)$$

y de esta expresión podemos obtener un estimador para  $\theta$  como  $\hat{\theta} = \text{Arg} (\mathbb{E} [e^{i\hat{\theta}}])$ , el cual es equivalente al estimador de máxima verosimilitud restringido a un intervalo entre  $[0, 2\pi]$ .

En este caso, el estimador  $\hat{\theta}$  será una variable aleatoria periódica con periodo  $2\pi$  por lo que el error cuadrático medio no será la forma correcta de cuantificar las dispersiones de las estimaciones, ya que cada estimación  $\hat{\theta}(x) = \hat{\theta}(x) + 2\pi n$  con  $n \in \mathbb{N}$  deben considerarse "idénticas". Sin embargo, a mayor  $n$  el error cuadrático medio también será más grande. En cambio, si uno toma en cuenta la periodicidad del problema se puede utilizar a la varianza de Holevo [57] para cuantificar de manera correcta la dispersión de las estimaciones

$$V_M^H(\hat{\theta}) = \mu^{-2} - 1, \quad (4.22)$$

donde  $M \in \mathcal{M}(\Theta, U)$  y  $\mu = |\mathbb{E} [e^{i\hat{\theta}}]|$ . Cuando  $\hat{\theta}$  es un estimador sesgado, entonces  $\mu = \mathbb{E} [\cos(\hat{\theta} - \theta)]$  [10]. Un estimador periódico  $\hat{\theta}$  es insesgado si  $e^{i\hat{\theta}} \propto \mathbb{E} [e^{i\hat{\theta}}]$ . Para distribuciones simétricas y altamente concentradas alrededor de  $\theta$ ,  $V_M^H(\hat{\theta}) \sim \text{ECM}_M(\hat{\theta})$ . Como consecuencia, la varianza de Holevo de un estimador periódico insesgado esta acotada inferiormente por

$$V_M^H(\hat{\theta}) \geq \frac{1}{I_Q(\theta)}.$$

Como la información de Fisher cuántica alcanza sus máximos valores para familias de estados puros, restringiremos nuestra atención a estados iniciales puros. Además en este caso, uno puede encontrar el POVM covariante que minimiza la varianza de Holevo usando el espectro de  $J$  para cualquier parámetro de desplazamiento [57, 58]. En este trabajo, se adaptó la prueba presentada en [57] al problema de estimación de fase en qubits.

**Teorema 8.** Sea  $\rho_\theta = U_\theta \rho U_\theta^\dagger$ , con  $\rho = |\psi\rangle\langle\psi|$  y

$$M_*(d\hat{\theta}) = \frac{d\hat{\theta}}{2\pi} \sum_{m,n \in \text{Spec}(J)} e^{i\hat{\theta}(n-m)} \frac{P_m \rho P_n}{\sqrt{\text{Tr}[P_m \rho] \text{Tr}[P_n \rho]}}, \quad (4.23)$$

donde  $P_m = |m\rangle\langle m|$  es el proyector asociado al eigenvalor  $m$  de  $J$ . Entonces, para cualquier  $M \in \mathcal{M}(\Theta, U)$ ,

$$V_M^H(\hat{\theta}) \geq V_{M_*}^H(\hat{\theta}). \quad (4.24)$$

*Demostración.* En notación de Dirac,  $\rho = |\psi\rangle\langle\psi|$  de modo que la ecuación (4.23) toma la forma

$$M_*(d\hat{\theta}) = \frac{d\hat{\theta}}{2\pi} \sum_{m,n \in \text{Spec}(J)} e^{i\hat{\theta}(n-m)} p_{mn}^* |m\rangle\langle n|, \quad (4.25)$$

con  $p_{mn}^* = \frac{\langle m|\psi\rangle\langle\psi|n\rangle}{|\langle m|\psi\rangle||\langle\psi|n\rangle|}$ . Por otro lado, usando (4.19), se tiene que

$$E_{M,\theta}[e^{ik\hat{\theta}}] = \sum_{\substack{m,n: \\ |m-n|=k}} p_{mn} \langle\psi|m\rangle\langle n|\psi\rangle e^{-i(n-m)\theta} \quad (4.26)$$

y como  $|p_{mn}| \leq 1$ ,

$$\left| E_{M,\theta}[e^{ik\hat{\theta}}] \right| \leq \sum_{\substack{m,n: \\ |m-n|=k}} \left| \langle\psi|m\rangle\langle n|\psi\rangle e^{-i(n-m)\theta} \right| \leq \sum_{\substack{m,n: \\ |m-n|=k}} \left| \langle\psi|m\rangle||\langle n|\psi\rangle| e^{-i(n-m)\theta} \right|. \quad (4.27)$$

donde la igualdad es alcanzada si y sólo si  $p_{mn} = p_{mn}^*$ . Haciendo  $k = 1$  se obtiene el resultado deseado.  $\square$

Es importante señalar que la prueba anterior es independiente de la estructura del generador  $J$ . Mientras el espectro de  $J$  sea numerable el resultado anterior es válido. En general, para el problema de estimación de fase sobre estados puros con generadores con espectro a lo más numerable, al POVM  $M_*$  descrito por la ecuación 4.23 se le llama medida canónica de fase y es el que minimiza la varianza de Holevo.

En particular, si  $\langle m|\psi\rangle$  es una constante para todo  $m \in \text{Spec}(J)$ , el POVM  $M_*$  contendrá una información acerca de  $\theta$  igual a la información cuántica de Fisher [58]. En qubits, esta condición se cumple cuando el estado inicial es preparado en la condición inicial  $\vec{a} \cdot \vec{n} = 0$ . En este contexto, se puede probar que  $M_*$  maximiza la información de Fisher sobre el conjunto de POVMs covariantes.

**Teorema 9.** Sea  $\rho_\theta = e^{-i\theta\vec{n}\cdot\frac{\vec{\sigma}}{2}} \rho e^{-i\theta\vec{n}\cdot\frac{\vec{\sigma}}{2}}$ , con

$$\rho = |\psi\rangle\langle\psi| = \frac{1}{2} [I + \vec{a} \cdot \hat{n}], \quad \|\vec{a}\| = 1.$$

Entonces, para cualquier  $M \in \mathcal{M}(\Theta, U)$ ,

$$I_M(\theta) \leq I_{M_*}(\theta), \quad (4.28)$$

donde  $M_*(d\hat{\theta})$  es definido de acuerdo a (4.23).

*Demostración.* Escribiendo explícitamente la ecuación (4.20) obtenemos

$$p(d\hat{\theta}) = \frac{d\hat{\theta}}{2\pi} \left[ 1 + 2\text{Re} \left( e^{-(\hat{\theta}-\theta)} \cdot C \right) \right], \quad (4.29)$$

donde  $C = p_{\frac{1}{2}-\frac{1}{2}} \langle -\frac{1}{2} | \psi \rangle \langle \psi | \frac{1}{2} \rangle$  es un número complejo con módulo  $|C|$  y fase  $\varphi$ . Entonces,

$$p(d\hat{\theta}) = \frac{d\hat{\theta}}{2\pi} \left[ 1 + 2|C| \cos \left( (\hat{\theta} - \theta) + \varphi \right) \right]. \quad (4.30)$$

Por lo tanto la información de Fisher,

$$I_M = \frac{1}{2\pi} \int_{\Theta} d\hat{\theta} \frac{4|C|^2 \sin^2 \left( (\hat{\theta} - \theta) + \varphi \right)}{1 + 2|C| \cos \left( (\hat{\theta} - \theta) + \varphi \right)} = \frac{4|C|^2}{2\pi} \int_{\Theta} dx \frac{\sin^2(x)}{1 + 2|C| \cos(x)}. \quad (4.31)$$

Observando que (4.31) crece monótonamente conforme el valor de  $|C|$  aumenta. Como  $|\langle \pm \frac{1}{2} | \psi \rangle| = \sqrt{\frac{1}{2}(1 \pm (\vec{a} \cdot \vec{n}))}$ , entonces  $|C| = |\leq \frac{1}{2} \sqrt{I_Q}$ , y el máximo valor  $I_M$  es alcanzado por  $M_*$ .  $\square$

Para encontrar una expresión explícita de  $p(d\hat{\theta} | \theta)$ , se debe notar que

$$\langle \pm \frac{1}{2} | \langle \pm \frac{1}{2} | \psi \rangle \langle \psi | \mp \frac{1}{2} \rangle \langle \mp \frac{1}{2} | = \frac{1}{4} [2(\vec{a} \cdot \vec{\sigma}) + i(\vec{a} \times \vec{n}) \cdot \vec{\sigma} - (\vec{n} \cdot \vec{\sigma})(\vec{n} \cdot \vec{a})], \quad (4.32)$$

por lo que la ecuación (4.23) puede ser reescrita como sigue

$$M_*(d\hat{\theta}) = \frac{d\hat{\theta}}{2\pi} \left[ I + \frac{1}{\sqrt{I_Q}} \left[ (\vec{a} \cdot \vec{\sigma} - (\vec{n} \cdot \vec{\sigma})(\vec{n} \cdot \vec{a})) \cos(\hat{\theta}) - \sin(\hat{\theta})(\vec{a} \times \hat{n}) \cdot \vec{\sigma} \right] \right]. \quad (4.33)$$

De este modo, la medición  $M_*$  produce la siguiente distribución de probabilidad

$$p(d\hat{\theta} | \theta) = \frac{d\hat{\theta}}{2\pi} \left[ 1 + \sqrt{I_Q} \cos(\hat{\theta} - \theta) \right], \quad (4.34)$$

por lo que, el POVM  $M_*$  genera un modelo estadístico identificable independiente de  $\theta$ . Para ilustrar este hecho, se simuló numéricamente números aleatorios con distribución (4.34). En la Fig. 4.2 se muestra gráficamente una función de verosimilitud particular producida por (4.34).

Además, de acuerdo a la expresión (4.34), la información de Fisher dada por  $M_*$  es:

$$I_{M_*} = \frac{1}{2\pi} \int_0^{2\pi} \frac{I_Q \sin^2(\hat{\theta} - \theta)}{1 + \sqrt{I_Q} \cos(\hat{\theta} - \theta)} d\hat{\theta} = 1 - \sqrt{1 - I_Q}. \quad (4.35)$$

Cuando  $\vec{a} \cdot \vec{n} = 0$ ,  $I_Q = 1$  y como consecuencia  $I(\theta; M_*) = I_Q^{\max} = 1$ . Esto es un resultado esperado pues bajo esta condición inicial óptima,  $|\langle \frac{1}{2} | \psi \rangle| = |\langle -\frac{1}{2} | \psi \rangle|$ . Por lo tanto, en esta preparación óptima, si el estimador de máxima verosimilitud es utilizado en  $N$  copias independientes del sistema, uno puede alcanzar la cota cuántica de Crámer-Rao cuántica en el límite asintótico ( $N \rightarrow \infty$ ). No obstante, bajo esta condición inicial óptima es posible encontrar otros POVMs covariantes

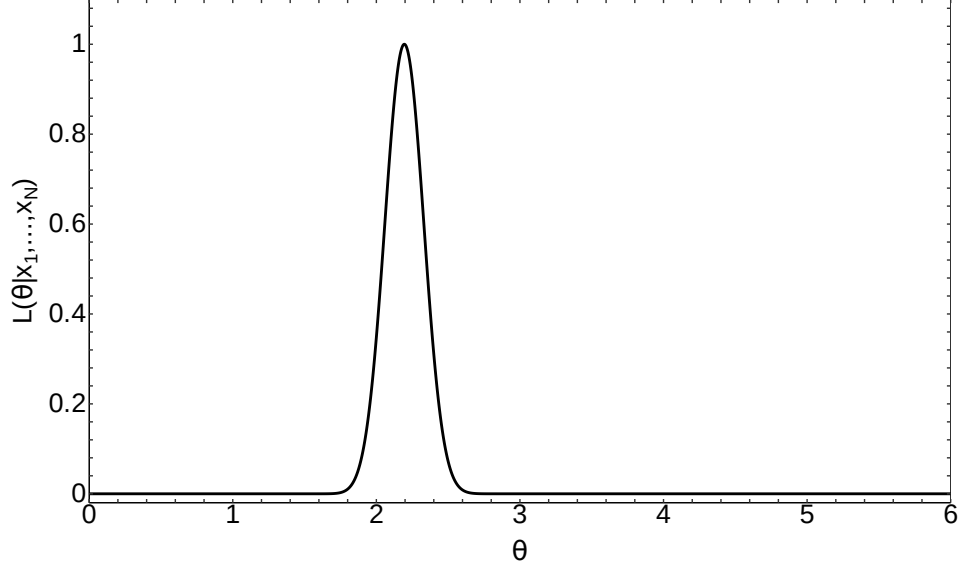


Figura 4.2: Función de verosimilitud correspondiente a la densidad de probabilidad (4.34) para la medida  $M_*$  y  $N = 64$ . Se consideró como valor del parámetro  $\theta = 2$  y se utilizó la condición inicial  $\vec{a} \cdot \vec{n} = 0$ . Para mayor claridad se escaló la figura para que el máximo fuese igual a uno.

óptimos. Para encontrarlos basta con caracterizar todos los POVMs covariantes para este problema de estimación.

Cada operador positivo y acotado  $P_o$  puede ser expresado en la representación de la esfera de Bloch, esto es,  $P_o = d_0 + \vec{d} \cdot \vec{\sigma}$ , donde  $(d_0, \vec{d})$  son números reales. Evaluando (3.44), se observa que  $P_o$  forma un POVM covariante si  $d_0 = 1$  y  $\vec{d}$  es ortogonal a  $\vec{n}$ . Por lo tanto cualquier POVM covariante tiene la forma

$$M(d\hat{\theta}) = \frac{d\hat{\theta}}{2\pi} \left[ I + \vec{d} \cdot \vec{\sigma} \cos(\hat{\theta}) - \sin(\hat{\theta})(\vec{d} \times \hat{n}) \cdot \vec{\sigma} \right]. \quad (4.36)$$

Así, cuando un qubit está en el estado  $\rho_\theta$ , las salidas del POVM  $M(d\hat{\theta})$  siguen la densidad de probabilidad

$$p(d\hat{\theta}|\theta) = \frac{d\hat{\theta}}{2\pi} \left[ 1 + (\vec{a} \cdot \vec{d}) \cos(\hat{\theta} - \theta) - \vec{a} \cdot (\vec{d} \times \vec{n}) \sin(\hat{\theta} - \theta) \right]. \quad (4.37)$$

De esta expresión, uno puede calcular la información de Fisher asociada al POVM covariante  $M$ . En particular, el caso  $\vec{d} = \pm(\vec{n} \times \vec{a})$  es de sumo interés, pues la información de Fisher para este caso es

$$I_{M|\vec{d}=\pm(\vec{n} \times \vec{a})}(\theta) = \frac{1}{2\pi} \int_0^{2\pi} \frac{I_Q^2 \cos^2(\hat{\theta} - \theta)}{1 \pm I_Q \sin(\hat{\theta} - \theta)} d\hat{\theta}. \quad (4.38)$$

Notando que, en general,  $I_{M_*} \geq I_{M|\vec{d}=\pm(\vec{n} \times \vec{a})}(\theta)$ , pero para este caso particular  $\vec{a} \cdot \vec{n} = 0$ , el POVM  $M|\vec{d}=\pm(\vec{n} \times \vec{a})$  es capaz de alcanzar el máximo de información de Fisher,  $I_Q^{\max} = 1$ . Entonces, el POVM dado por la ecuación (4.36) es equivalente a la medida  $M_*$ . Cabe señalar que el POVM



$M_*$ , en general, es difícil de implementar experimentalmente. Sin embargo, recientemente se ha podido realizar en qubits utilizando retroalimentación cuántica [79]. Además es posible aproximar este POVM utilizando un POVM de  $k \geq 8$  elementos que discretiza la ecuación (4.25) [22].

En resumen, se ha mostrado que las estrategias de estimación basadas en POVMs localmente óptimos con dos salidas no pueden saturar la CICRC pues producen funciones de verosimilitud no identificables, este es uno de los puntos centrales de este capítulo y sus implicaciones se detallarán en la sección 4.3. POVMs discretos con más de dos salidas pueden resolver el problema de identificabilidad pero estos siguen teniendo el problema de ser localmente óptimos. Posteriormente, se probó que el POVM covariante  $M_*$  resuelve el problema de identificabilidad y minimiza el error cuadrático medio para una medida, aunque solamente en las condiciones iniciales óptimas tiene información de Fisher cuántica. Entonces, según estos resultados, no existe un conjunto de mediciones independientes e idénticas que pueden alcanzar la CICRC en general.

## 4.2. Medidas entrelazadas

Es posible alcanzar la cota cuántica de Crámer-Rao para un estado inicial no ortogonal al eje de rotación si no restringimos el proceso de estimación a un conjunto de mediciones independientes e idénticas. De hecho, se puede probar que es posible alcanzar la CICRC para mediciones entrelazadas independientemente de la condición inicial.

De acuerdo a [58], al estimar  $\theta \in [0, 2\pi)$  para una familia de estados  $\rho_\theta^{\otimes N}$  en el espacio de Hilbert  $\mathcal{H}^{\otimes N}$  con  $\rho_\theta = U_\theta \rho U_\theta^\dagger$  y  $\rho$  un qubit inicial, se nota que  $\rho_\theta^{\otimes N}$  puede ser escrito como  $\rho_\theta^{\otimes N} = e^{-i\theta J^{(N)}} \rho^{\otimes N} e^{i\theta J^{(N)}}$ , donde

$$J^{(N)} = \left( \hat{n} \cdot \frac{\vec{\sigma}}{2} \right) \otimes \cdots \otimes I + \cdots + I \otimes \cdots \otimes \left( \hat{n} \cdot \frac{\vec{\sigma}}{2} \right), \quad (4.39)$$

de este modo, la familia de estados  $\rho_\theta^{\otimes N}$  en  $\mathcal{H}^{\otimes N}$  es covariante con respecto a la representación unitaria  $e^{-i\theta J^{(N)}}$ . Por lo tanto, análogo al caso de un único estado inicial, el POVM que minimiza la varianza de Holevo es expresado en términos del espectro de  $J^{(N)}$ , por lo que el POVM  $M_*^{(N)}$  que minimiza la varianza de Holevo tiene la forma

$$M_*^{(N)}(d\hat{\theta}) = \frac{d\hat{\theta}}{2\pi} \left[ \sum_{\substack{\lambda, \lambda' \in \\ \text{Spec}(J^{(N)})}} e^{i\hat{\theta}(\lambda' - \lambda)} \frac{P_\lambda \rho^{\otimes N} P_{\lambda'}}{\sqrt{\text{Tr}[P_\lambda \rho^{\otimes N}] \text{Tr}[P_{\lambda'} \rho^{\otimes N}]}} \right], \quad (4.40)$$

con  $P_\lambda$  el operador de proyección de  $J^{(N)}$  asociado al eigenvalor  $\lambda$ .

La diferencia aquí, en comparación con el caso de una única medición, es que el espectro de  $J^{(N)}$  es degenerado. Entonces

$$P_* = \int_{\Theta} M(d\hat{\theta}) \neq I,$$

y por tanto  $M_*^{(n)}$  no es una resolución de la identidad en el espacio de Hilbert. No obstante, este no es un problema real pues se puede extender a un POVM agregando el operador  $I - P_*$ , el cual opera en el complemento ortogonal del subespacio  $\mathcal{H}_* := P_*(\mathcal{H}^{\otimes N})$ . Así, el POVM óptimo es

$$M_*^{(N)} := P_* \oplus I - P_*. \quad (4.41)$$

En particular, si  $N = 1$ ,  $M_*^{(N)} = M_*$ .

Como se establece en [58],  $M_*^{(N)}$  puede alcanzar la cota cuántica de Crámer-Rao para cualquier estado inicial en el límite asintótico. Además, el error escala como el cuadrado del número de qubits (escala de Heisenberg). Desafortunadamente,  $M_*^{(N)}$  es una medida entrelazada. Es decir, para cada conjunto  $B \in \mathcal{B}(\Theta)$ , el operador  $M(B)/\text{Tr}[M(B)]^1$  define un operador densidad entrelazado sobre  $\mathcal{H}^{\otimes N}$  [113]. De este modo, si uno quiere construir (4.41), es necesario tener  $N$  copias del estado inicial y realizar operaciones cuánticas sobre todas ellas al mismo tiempo. Por tal motivo, la implementación de este tipo de POVMs es un gran reto experimental.

### 4.3. Estimación cuántica adaptativa de estados (AQSE)

Como se mencionó previamente, la estimación cuántica de parámetros produce estimadores que dependen del parámetro que uno quiere estimar. Una forma de tratar este problema es utilizar el siguiente esquema adaptativo llamado estimación cuántica adaptativa de estados (AQSE de sus siglas en inglés [adaptive quantum state estimation](#)) [38, 85, 121], el cual consiste en lo siguiente: supongamos que uno tiene un conjunto óptimo de POVMs  $\{M_g\}_{g \in \Theta}$  y se empieza con una suposición inicial arbitraria  $g_0$  para el parámetro  $g$  del POVM. Entonces, uno aplica la media  $M_{g_0}$ , la cual es óptima en  $g_0$ ,  $M_{g_0}$ . Suponiendo que se observa el dato  $x_1$ , uno aplica el estimador de máxima verosimilitud a la función de verosimilitud  $\mathcal{L}_1(\theta | x_1; g_0) = p(x_1 | \theta; g_0)$  para obtener la estimación  $\hat{\theta}_1(x_1) = g_1$ , y se repite este proceso de forma iterativa. De este modo, para  $n \geq 2$ , uno aplica el POVM  $M_{g_{n-1}=\hat{\theta}_{n-1}(x_1, \dots, x_{n-1})}$  obteniendo la salida  $x_n$ . Donde  $\hat{\theta}_{n-1}$  es la estimación del paso previo usando los datos  $x_1, \dots, x_{n-1}$ . Así, la función de verosimilitud para el  $n$ -ésimo paso es

$$\mathcal{L}_n(\theta | x_1, \dots, x_{n-1}; g_{n-1}) = \prod_{i=1}^n p(x_i | \theta; g_{i-1}), \quad (4.42)$$

con  $x_i$  el dato observado en el paso  $i$ . Aplicando el estimador de máxima verosimilitud, uno obtiene el  $n$ -ésimo valor del parámetro  $g$  del POVM,  $g_n = \hat{\theta}_n(x_1, \dots, x_n)$ . Aunque este método relaja la condición de mediciones idénticas, este no garantiza la obtención de un modelo estadístico regular, lo que impacta directamente al desempeño del estimador [28].

Para ilustrar este hecho, es necesario desarrollar con mayor detalle el problema de no identificabilidad para el POVM  $M_g$  previamente discutido. Con este fin, en el panel A de la Fig 4.3 se muestra el resultado de la función de verosimilitud (4.14) para  $g = 1.5$ ,  $\theta = 2$  y  $N = 64$ . En efecto, se observa que esta función de verosimilitud es no identificable ya que presenta dos máximos globales dentro del intervalo  $[0, 2\pi)$ . En el panel B de la misma figura se muestra el resultado de utilizar AQSE comenzando con un valor inicial de  $g_0 = 1.5$ . En este caso, se observa un único máximo global pero para un valor incorrecto de  $\theta$ . Este hecho es resultado de tener un modelo con funciones de verosimilitud no identificables y aunque AQSE puede en principio eliminar una posible degeneración del máximo global, no es posible controlar la obtención de estimaciones únicamente alrededor del máximo correcto.

De este modo, parece bastante natural esperar que si conocemos alrededor de qué máximo pertenece el parámetro y restringimos la verosimilitud a este conjunto, se puede asegurar la identificabilidad para la función de verosimilitud, obteniendo la convergencia del método adaptativo

<sup>1</sup>El hecho que  $M(B)$  se un operador acotado y positivo implica que  $M(B)$  tenga traza finita para todo  $B$  conjunto de Borel.

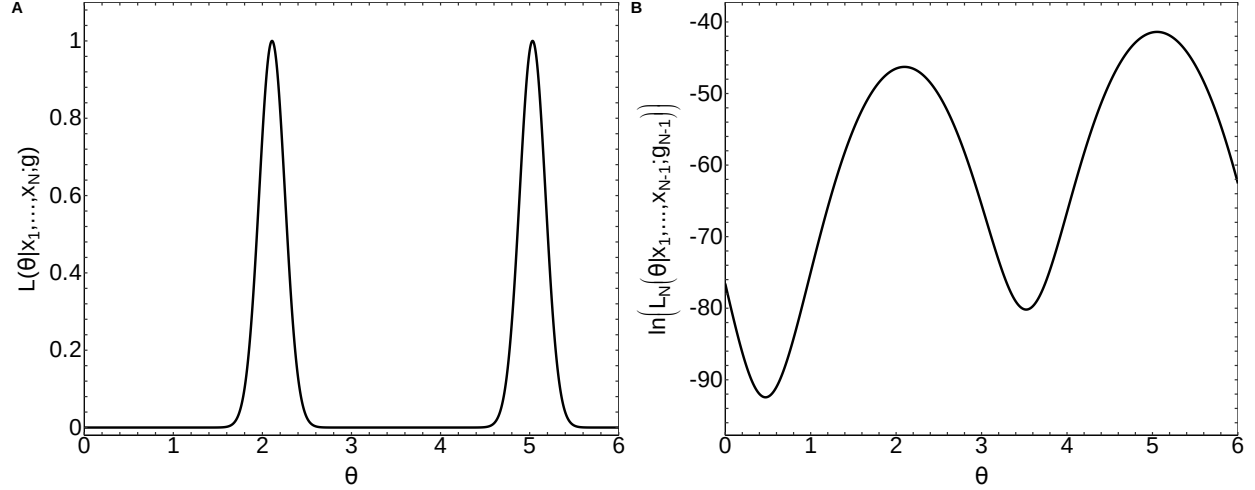


Figura 4.3: Gráfica de las funciones de verosimilitud para el POVM de dos salidas  $P(g)$  dado por la Eq. (4.10) y considerando  $\theta = 2$  y  $N = 64$ . Panel (A) muestra la función de verosimilitud dada por la fórmula (4.13). Aquí se consideró  $g = 1.5$ , obteniendo 32 salidas igual a 1 de las posible  $N = 64$ . Para mayor claridad en la figura se reescalaron para tener los máximos igual a 1. Panel (B) muestra la función de verosimilitud logarítmica para el esquema AQSE dado por la Eq. (4.42) usando 64 pasos adaptativos. La existencia de este tipo de trayectorias trae como consecuencia un aumento en el número de estimaciones incorrectas. Esto explica por qué AQSE no alcanza el límite cuántico de Crámer-Rao. En ambos casos, la condición inicial fue tomada como  $\vec{a} \cdot \vec{n} = 0.5$ .

al límite cuántico de Crámer-Rao. Con el fin de analizar este hecho, se comparó el comportamiento de AQSE en dos diferentes escenarios: funciones de verosimilitud sin restricción y con restricción en su dominio. En ambos casos se tomó  $\theta = \pi$ . Para el caso sin restricción el espacio parametral es  $\Theta = [0, 2\pi)$ , mientras que para el caso restringido se tomó  $\theta \in \Theta = [\frac{\pi}{2}, \frac{3\pi}{2})$ , dominio en el cual la función de verosimilitud es identificable. En la Fig. 4.4 se muestra el resultado de estos esquemas para las condiciones iniciales extremales  $\vec{a} \cdot \vec{n} = 0.5$  (panel A) y  $\vec{a} \cdot \vec{n} = 0$  (panel B). Estas condiciones iniciales se eligieron porque corresponden al mayor y menor valor de la diferencia  $|I_Q - I_{M^*}|$ , respectivamente. Para ambos casos, se realizó un bootstrap<sup>2</sup> con  $10^4$  repeticiones para cada punto.

De estos resultados se puede observar que en el caso sin restricción, la carencia de identificabilidad afecta directamente al escalamiento de la varianza de Holevo como función del número de pasos adaptativos. Además estos resultados muestran que AQSE no alcanza la cota cuántica de Crámer-Rao. Por otro lado, para el caso donde se restringe el espacio parametral, se encontró que el estimador de máxima verosimilitud alcanza la cota cuántica de Crámer-Rao alrededor de los 128 pasos para ambas condiciones iniciales, por lo que es factible concluir que el deficiente rendimiento para AQSE es debido a la no identificabilidad para  $\theta$  en las funciones de verosimilitud.

Finalmente, en la Fig. 4.4, se comparó también el desempeño de AQSE con la inferencia covariante óptima. En este caso, la secuencia de  $M_*$  medidas tiene un mejor escalamiento cuando  $N$  es pequeño (menor que 4). Por lo que, en adición se puede concluir que el POVM covariante óptimo

<sup>2</sup>La técnica de bootstrap es un método de remuestreo que se utiliza para aproximar la distribución de una variable aleatoria [33].

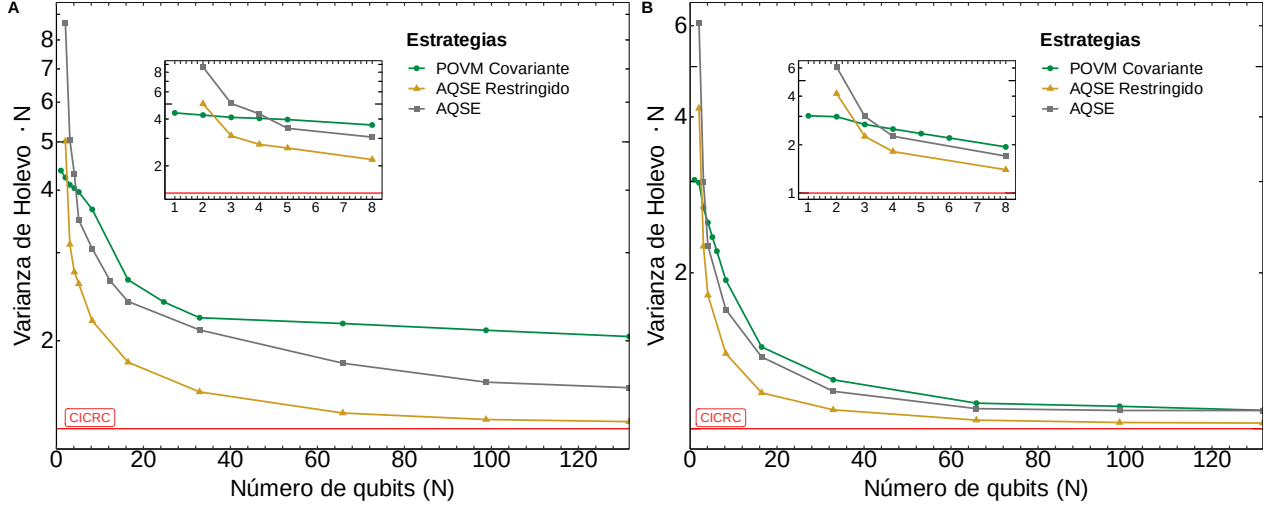


Figura 4.4: Resultados numéricos de la varianza de Holevo como función del número de pasos adaptativos para las diferentes estrategias de estimación en dos diferentes condiciones iniciales:  $\vec{a} \cdot \vec{n} = 0.5$ , correspondiente al menor valor de la información de Fisher cuántica (panel A) y  $\vec{a} \cdot \vec{n} = 0$  (panel B) correspondiente al máximo valor de la información de Fisher cuántica. En ambos casos se utilizó una secuencia de  $10^4$  repeticiones para calcular la varianza de Holevo en cada punto. Como se puede observar, el AQSE restringido presenta un desempeño mucho mejor que AQSE. Como punto de referencia, se incluyó el desempeño de la estrategia de estimación basada en el POVM covariante  $M_*$ . Se puede notar que ambas condiciones iniciales AQSE no satura la desigualdad de Crámer-Rao cuántica, mientras que para la condición inicial óptima, como era de esperar, la secuencia de POVMs covariantes alcanza el límite cuántico de Crámer-Rao (panel B). Además, como  $M_*$  minimiza la varianza de Holevo para una copia del sistema se observa un mejor desempeño de la estimación covariante cuando  $N$  es pequeño (inset de ambos paneles).

es la mejor estrategia basada en mediciones independientes en el régimen de  $N$  pequeños. Esto es de esperar pues el POVM covariante óptimo se construyó como aquel que minimiza el error cuadrático medio para  $N = 1$ .

En la siguiente sección, proponemos un esquema de estimación que utiliza la inferencia covariante para evadir el problema de identificabilidad sobre las funciones de verosimilitud en AQSE y que entonces, nos permita saturar la desigualdad de Crámer-Rao en cualquier condición inicial.

#### 4.4. Estimación adaptativa utilizando intervalos de confianza

De la discusión previa, es claro que si se tiene un conocimiento previo sobre alrededor de qué máximo está realmente el parámetro que se desea estimar y se restringe el espacio parametral a este conjunto, podemos producir funciones de verosimilitud identificables, obteniendo a su vez la convergencia de AQSE al límite cuántico de Crámer-Rao. Con esto en mente, proponemos el siguiente esquema de estimación, el cual nos permitirá obtener ese conocimiento sobre a cuál conjunto pertenece el parámetro, y entonces tener un modelo estadístico regular el cual puede ser aplicado al esquema adaptativo.

Nuestro esquema consiste en dos pasos principales. En el primer paso uno utiliza una secuencia de POVMs covariantes óptimos para producir un **intervalo de confianza (IC)** con el estimador de máxima verosimilitud  $\hat{\theta}_{\text{EMV}}$ . El IC nos da un rango de los posibles valores donde el parámetro desconocido tenga la mayor probabilidad de pertenecer. En el segundo paso, uno aplica AQSE restringiendo el espacio parametral al intervalo de confianza. Además, uno puede ir actualizando el centro del intervalo de confianza utilizando la estimación del paso anterior.

Supongamos que se tienen  $N$  copias del estado  $\rho$ , entonces cada medida sobre el sistema puede ser efectuada de forma consecutiva sobre cada copia. Denotaremos al resultado de cada medición como  $x_i$ . En la primer parte de nuestro esquema se efectúan  $N_1$  mediciones covariantes  $\mathcal{M}_1$  dadas por (4.25), esto es

$$\mathcal{M}_1 := M_*(dx_k), \quad (4.43)$$

y entonces se construye el estimador de máxima verosimilitud

$$\hat{\theta}_{\text{EMV}} = \arg \max_{\theta \in \Theta} \prod_{i=1}^{N_1} p(dx_i | \theta) = \arg \max_{\theta \in \Theta} \prod_{i=1}^{N_1} \frac{dx_i}{2\pi} \left( 1 + \sqrt{I_Q} \cos(x_i - \theta) \right) \quad (4.44)$$

para obtener la estimación  $\hat{\theta}_{\text{EMV}}(x_1, \dots, x_{N_1})$ . El intervalo a un nivel  $1 - \alpha$  de confianza se estima utilizando la ecuación (2.57):

$$\text{IC}(\hat{\theta}_{\text{EMV}}(\vec{x})) = \left( \hat{\theta}_{\text{EMV}}(\vec{x}) - z_{\alpha/2} \cdot I_{M_*}^{-\frac{1}{2}}, \hat{\theta}_{\text{EMV}}(\vec{x}) + z_{\alpha/2} \cdot I_{M_*}^{-\frac{1}{2}} \right), \quad (4.45)$$

donde  $I_{M_*}$  es la información de Fisher de  $M_*$  dada por la ecuación (4.35) y  $z_{\alpha/2}$  es un punto crítico determinado por el nivel de confianza y definido por la ecuación (2.54). Además, uno puede obtener el mínimo tamaño de muestra  $N_1$  necesario para obtener un intervalo de confianza de tamaño  $2E$ , a partir de

$$N_1 \geq \frac{z_{\alpha/2}^2}{I_{M_*} E^2}. \quad (4.46)$$

En el segundo paso de nuestro esquema se realizan  $N_2 = N - N_1$  medidas del POVM

$$\mathcal{M}_2 := \mathcal{P}(\hat{\theta}_{\text{EMV}}(x_1, \dots, x_k)). \quad (4.47)$$

Resumiendo, nuestro esquema está dado por:

$$\hat{\theta}_{\text{EMV}} = \arg \max_{\theta \in \Theta} \prod_{i=1}^{N_1} p(dx_i | \theta) \prod_{k=N_1}^{N_2} p(x_{k+1} | \theta; \hat{\theta}_{\text{EMV}}(x_1, \dots, x_k)), \quad (4.48)$$

con  $\Theta = \text{IC}(\hat{\theta}_{\text{EMV}}(x_1, \dots, x_k))$ . Así, la ecuación (4.48) es el estimador de nuestra propuesta.

Antes de discutir algunos resultados numéricos obtenidos para esta propuesta es importante discutir las nuevas posibles fuentes de errores asociadas a este nuevo método. Suponga que el nivel de confianza  $0 \leq C_l \leq 1$  se fija para un error marginal  $E$  dado. Entonces, la varianza de Holevo puede escribirse como

$$V^H(\hat{\theta}_{\text{EMV}}) = C_l \Delta_1 + (1 - C_l) \Delta_2, \quad (4.49)$$

donde se identifican dos tipos de errores  $\Delta_1$  y  $\Delta_2$ , lo cuales corresponden al error asociado a AQSE cuando el IC incluye correctamente el valor del parámetro, y cuando no lo hace, respectivamente. Durante el resto del trabajo, llamaremos a los intervalos de confianza que no incluyen al parámetro como intervalos malos. En adición, se tomará  $E < \pi/2$ , a modo de garantizar que las funciones de verosimilitud usadas en AQSE sean identificables. De este modo, el error  $\Delta_1$ , cuando los IC incluyen el valor del parámetro está acotado inferiormente por

$$\Delta_1 \geq \frac{1}{I_Q N_2 + N_1 I_{M^*}}. \quad (4.50)$$

Para caracterizar  $\Delta_2$ , primero se debe notar que existen dos tipo de intervalos de confianza malos. Para ver esto, consideremos  $\theta_A$  el valor del parámetro y  $\theta_B = \theta_A + \pi$ . Entonces, la función de verosimilitud producida por AQSE tiene dos máximos locales, uno alrededor de  $\theta_A$ , y otro alrededor de  $\theta_B$ . El primer tipo de intervalo de confianza malo corresponde a aquellos donde se incluye el segundo máximo alrededor de  $\theta_B$ . Así, si uno aplica AQSE restringido a este intervalo, el estimador  $\theta_{EMV}$  tenderá a  $\theta_B$ . El segundo tipo de intervalo de confianza malo aparece cuando el intervalo no incluye a  $\theta_B$ . En este caso, la estimador de máxima verosimilitud producido por AQSE tiende al punto del intervalo que esté más cerca de cualquiera de los máximos alrededor de  $\theta_A$  o  $\theta_B$ . Cuando la estimación es más cercana a  $\theta_A$ , al actualizar el centro del intervalo con las estimaciones posteriores, uno puede, en principio, obtener un intervalo de confianza que capture el valor de parámetro eventualmente. De esta manera, se obtendrán mejores resultado cuando se actualizan los centros de los intervalos de confianza.

Para obtener un intervalo de confianza malo del primer tipo, es necesario que la mayoría de los datos obtenidos de las mediciones covariantes sean números cercanos a  $\theta_B$ , por lo que de acuerdo a la ecuación (4.34), para una muestra covariante de tamaño  $N$ , se sigue que la probabilidad de obtener una función de verosimilitud con un máximo global alrededor de  $\theta_B$  es

$$\left[ \int_{\theta_B - \epsilon}^{\theta_B + \epsilon} p(d\hat{\theta} | \theta) \right]^N = \left[ \int_{\theta_B - \epsilon}^{\theta_B + \epsilon} \frac{d\hat{\theta}}{2\pi} \left( 1 + \sqrt{I_Q} \cos(\hat{\theta} - \theta) \right) \right]^N. \quad (4.51)$$

Entonces, para  $N$  suficientemente grande, esta probabilidad es despreciable. De esta forma, para niveles de confianza cercanos a 1, y un error marginal  $E$  suficientemente pequeño, uno esperaría que el número de intervalos de confianza malos vaya desapareciendo conforme el número de pasos adaptativos se incrementa. Para ilustrar este hecho, en la Tabla 4.1, se muestra el número de intervalos de confianza malos en función del número de pasos adaptativos realizados en la parte correspondiente a AQSE. Estos resultados fueron producidos por una simulación bootstrap de 10000 repeticiones por punto. Aquí se eligió el error marginal como  $\pi/4$ , y un nivel de confianza de 0.99. En este caso, con el fin de obtener el intervalo de confianza deseado, se requirió realizar 11 y 22 medidas covariantes para las condiciones iniciales  $\vec{a} \cdot \vec{n} = 0$  y  $\vec{a} \cdot \vec{n} = 0.5$ , respectivamente. Para estos valores, el número de intervalos de confianza malos del tipo uno es despreciable ya que al fijar  $\epsilon = E$  en (4.51), las probabilidades de obtener intervalos de este tipo es de  $2.3 \times 10^{-10}$  y  $2.3 \times 10^{-18}$  para cada un de las condiciones iniciales.

La distribución para las estimaciones de AQSE basadas en intervalos de confianza, puede ser aproximada como la mezcla de 3 distribuciones normales,  $\mathcal{N}\left(\theta, \frac{1}{I_Q N_2 + N_1 I_{M^*}}\right)$ ,  $\mathcal{N}\left(\theta + E, \frac{1}{I_{M^*, M}(\theta)}\right)$ , y  $\mathcal{N}\left(\theta - E, \frac{1}{I_{M^*, M}(\theta)}\right)$ , donde  $M$  es la secuencia de medidas en AQSE y  $E$  es el error marginal. Esto implica que la cota inferior de la varianza de Holevo cuando el centro de los intervalos de

Número de pasos en AQSE	# Intervalos malos ( $\vec{a} \cdot \vec{n} = 0$ )	# Intervalos malos ( $\vec{a} \cdot \vec{n} = 0.5$ )
0	111	125
4	50	56
8	19	20
16	8	1
32	6	1
48	2	0

Cuadro 4.1: Número de intervalos de confianza malos como función del número de pasos adaptativos para la parte de ASQE donde el parámetro real  $\theta$  fue  $\pi$ .

confianza no es actualizado, este dada por

$$V^H \left( \hat{\theta}_{\text{EMV}} \right) \geq \frac{1}{I_Q N_2 + N_1 I_{M_*}} + (1 - C_l) E^2. \quad (4.52)$$

El segundo término del lado derecho de la ecuación (4.52) es el término dominante cuando  $N_2 \rightarrow \infty$ . Esto tiene una simple interpretación: si el parámetro se encuentra fuera del intervalo de confianza, al realizar más pasos en AQSE no se disminuye este error. Desafortunadamente, no es posible dar una expresión analítica para la cota inferior a la varianza de Holevo para el caso en que el centro de los intervalos de confianza se van actualizando. No obstante, como se verá en los resultados numéricos, si uno fija el nivel de confianza cercano a 1, un error marginal  $E$  suficientemente pequeño, y los centros de los intervalos de confianza se van actualizando, la varianza de Holevo tiende al límite cuántico de Crámer-Rao.

Para evaluar el desempeño de nuestro método se realizó una simulación Monte Carlo con un error marginal de  $E = \pi/4$  y un nivel de confianza de  $C_l = 0.99$ . Además, comparamos los resultados numéricos con el caso ideal ( $C_l = 1$ ) y una estrategia de estimación donde el centro de los intervalos de confianza no son actualizados. Los resultados son resumidos en las figuras 4.5 y 4.6 para las condiciones iniciales  $\vec{a} \cdot \vec{n} = 0$  y  $\vec{a} \cdot \vec{n} = 0.5$ .

En la figura 4.5 se consideró la condición inicial  $\vec{a} \cdot \vec{n} = 0$ , entonces  $I_{M_*} = 1$ , y  $N_1 = 11$ . La línea vertical discontinua en  $N_1 = 11$  es la separación entre las inferencias covariantes y AQSE. Por supuesto, para la parte covariante ( $N \leq 11$ ), ambas estrategias presentan los mismo resultados. sin embargo, para  $N > 11$ , si el centro de los intervalos de confianza no son actualizados, de acuerdo a la ecuación (4.52) el segundo término es el dominante y el error de la estimación comienza a crecer rápidamente conforme el número de mediciones se incrementa. Si por otro lado, el centro de los intervalos de confianza son actualizados, entonces el error de la estimación tiende a la cota cuántica de Crámer-Rao. De forma similar en la figura 4.6, se tiene  $N_1 = 22$ ,  $\vec{a} \cdot \vec{n} = 0.5$  y  $I_{M_*} = 0.75$ . De esta figura se observa el mismo comportamiento cualitativo, mostrando que nuestra estrategia de estimación tiende a la cota cuántica de Crámer-Rao conforme el número de pasos se incrementa en cualquier condición inicial.

Finalmente en la figura 4.7 se muestra el desempeño de todas las estrategias de estimación discutidas en esta sección considerando las condiciones iniciales  $\vec{a} \cdot \vec{n} = 0.5$  y  $\vec{a} \cdot \vec{n} = 0$ , correspondientes a los casos extremos donde la distancia entre  $I_Q$  y  $I_{M_*}$  es la mayor y la menor respectivamente. Obviamente, el comportamiento de estas estrategias de estimación es independiente de  $\theta$ , ya que la varianza de Holevo es invariante bajo traslaciones módulo  $2\pi$  sobre el espacio parámé-



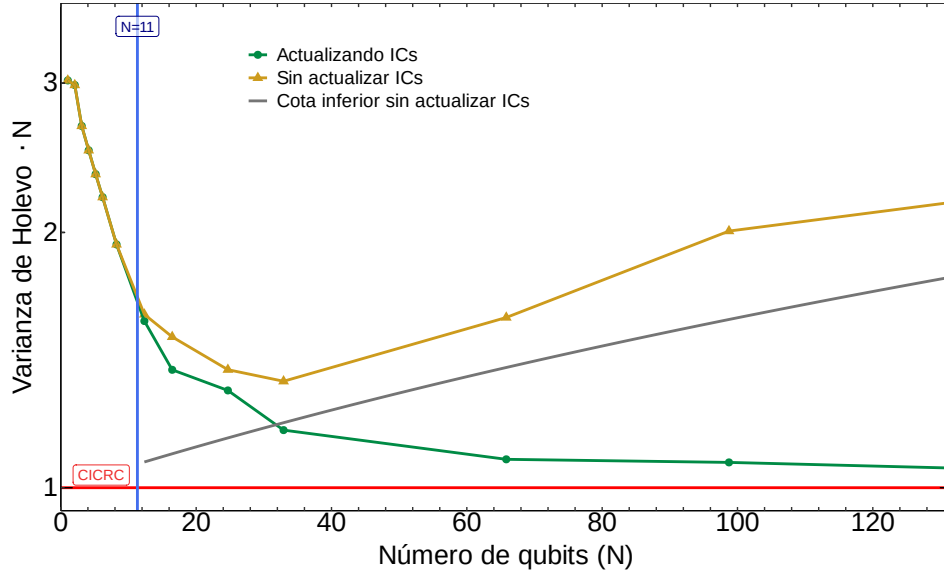


Figura 4.5: Varianza de Holevo como función de  $N$  para nuestro esquema de estimación. La estrategia óptima (curva verde con círculos), en la cual actualizamos el centro de los intervalos, satura la desigualdad de Crámer-Rao. La línea horizontal roja corresponde a tomar un nivel de confianza de 100 %. En adición, se muestra la estrategia alternativa donde los centros de los intervalos no son actualizados (curva amarilla con triángulos). Este resultado es acotado inferiormente por la CICRC, cuya fórmula está dada por (4.52). Se eligió la condición inicial  $\vec{a} \cdot \vec{n} = 0$ . La línea vertical azul en  $N = 11$  separa la inferencia covariante de la inferencia producida por AQSE en nuestro esquema.

trico. Además, para todas las estrategias, el comportamiento es cualitativamente el mismo para cualquier condición inicial. Como es de esperar, la estrategia basada en mediciones entrelazadas es la mejor de todas, en el sentido que alcanza la cota cuántica de Crámer-Rao con el mejor escalamiento. La segunda mejor estrategia es AQSE restringido, pero esta no es práctica pues se tiene que saber de antemano a qué vecindad pertenece el parámetro y producir un estimador de máxima verosimilitud consistente. Nuestro método, aunque es la tercer mejor estrategia, es la única propuesta realista, tanto de forma matemática como experimentalmente para cualquier condición inicial, mejorando el desempeño de AQSE y aproximando el error de las estimaciones a la cota cuántica de Crámer-Rao.



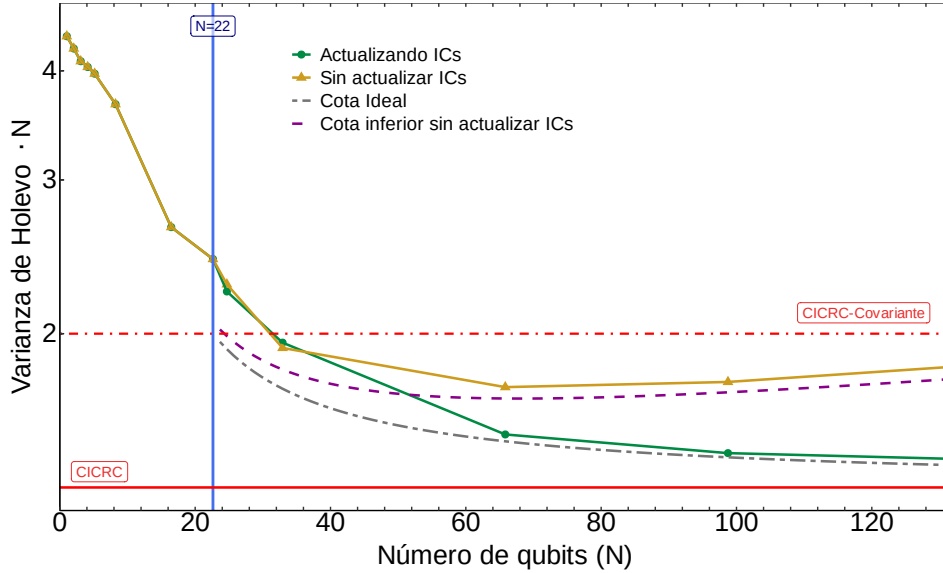


Figura 4.6: Varianza de Holevo como función de  $N$  para nuestro esquema de estimación. La estrategia óptima (curva verde con círculos), en la cual el centro de los intervalos es actualizado, tiende a la cota cuántica de Cramér-Rao conforme  $N$  se incrementa. El desempeño de esta estrategia está acotado por línea discontinua gris, la cual corresponde a la ecuación (4.52) para un nivel de confianza de 100%. También se muestra el desempeño de la estrategia alternativa, en donde los centros de los intervalos no son actualizados (curva amarilla con triángulos). En este caso, se eligió la condición inicial  $\vec{a} \cdot \vec{n} = 0.5$ . La línea vertical azul en  $N = 22$  marca el punto de separación entre las inferencias covariantes y el desempeño de las inferencias de la parte correspondiente a AQSE en nuestro esquema. Finalmente, la línea roja horizontal discontinua corresponde a la cota de Crámer-Rao asociada a los POVMs covariantes.

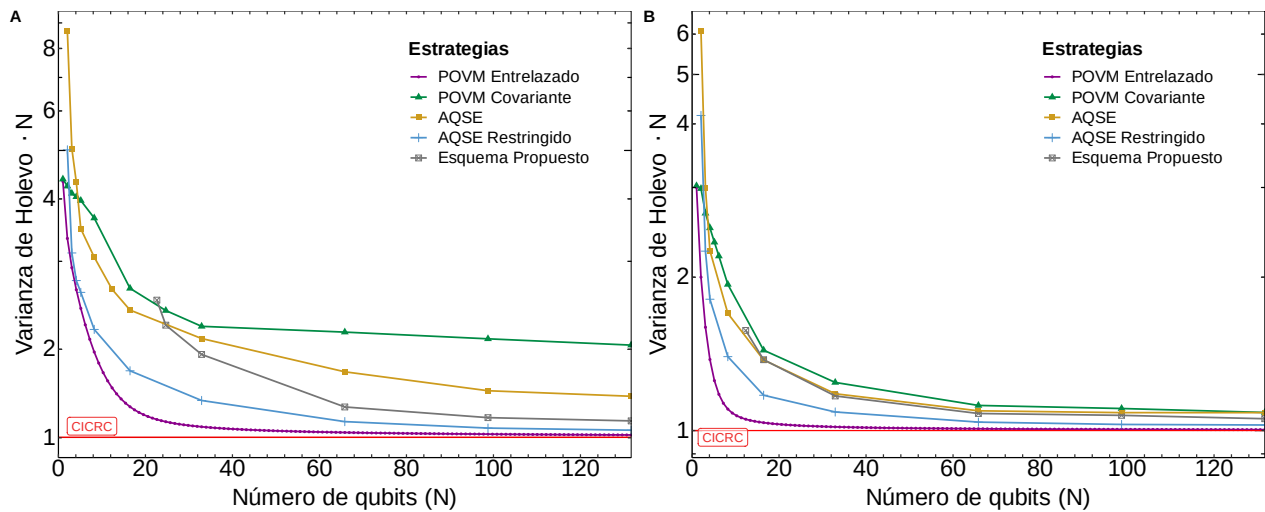


Figura 4.7: Varianza de Holevo como función del número de estados iniciales para las diferentes estrategias de medición. Para calcular la varianza de Holevo se realizó un bootstrap sobre las estimaciones con  $10^4$  repeticiones en cada punto. La curva de medidas entrelazadas fue calculada analíticamente. La curva de nuestra propuesta se realizó con un nivel de confianza de 99 % en los intervalos de confianza. En el AQSE restringido se supuso un espacio paramétrico de longitud  $\pi$  el cual incluía el valor del parámetro real  $\theta$  y producía un modelo estadístico regular para las POVMs involucrados en el método. La mejor estrategia corresponde a la de medidas entrelazadas, seguido del AQSE restringido. Estas dos estrategias son imprácticas experimentalmente. La tercer mejor estrategia es nuestra propuesta. El eje  $y$  está en escala logarítmica.

## Capítulo 5

# Estimación de frecuencia

En el capítulo anterior fue descrito el problema de estimación de fase en qubits. En dicho problema, la evolución paramétrica es producida por la conjugación de los operadores (4.3). De dicha ecuación se puede observar que el parámetro  $|\vec{c}|$  tiene unidades de 1/tiempo, por lo que este parámetro se interpreta como una frecuencia que denotaremos como  $\omega$ . Así, la evolución paramétrica para el estado de un qubit  $\rho$  es de la forma

$$\rho_\theta = e^{-i\omega t \frac{\vec{n} \cdot \vec{\sigma}}{2}} \rho e^{i\omega t \frac{\vec{n} \cdot \vec{\sigma}}{2}}. \quad (5.1)$$

Cuando se tiene control sobre el parámetro temporal  $t$ , la incertidumbre paramétrica recae sobre la frecuencia  $\omega$ . Este problema se conoce como estimación de frecuencia en qubits. Un ejemplo típico de este problema aparece en partículas de espín 1/2 interactuando con un campo magnético. Para este sistema, el Hamiltoniano es de la forma

$$H = -\mu \vec{\sigma} \cdot B, \quad (5.2)$$

donde  $\mu$  es el momento magnético de la partícula y  $\Delta = \mu B / \hbar$ .

Aquí el POVM localmente óptimo se puede implementar con espectroscopía de Ramsey sobre átomos de dos niveles, los cuales juegan el papel de los qubits. La espectroscopía de Ramsey se aplica de la siguiente manera [23, 89]: se comienza preparando  $N$  átomos con dos niveles (denotados por  $|0\rangle$  y  $|1\rangle$ ) en el estado base  $|0\rangle \langle 0|$ . Después se le aplica una rotación de  $\pi/2$  para cambiar el estado del sistema a  $|\psi_1\rangle \langle \psi_1|$ , con  $|\psi_1\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ . Después de esto, durante un tiempo  $t$ , se aplica un campo magnético que induce una oscilación en la transición de los niveles de los átomos de dos niveles con frecuencia  $\omega_0$ . Este proceso cambia el estado del sistema a  $|\psi_2\rangle \langle \psi_2|$ , donde  $|\psi_2\rangle = (|0\rangle + e^{-i\omega t} |1\rangle)/\sqrt{2}$ ,  $\omega = \omega_0 - \Delta$  y  $\omega_0$  la frecuencia de transición natural entre los niveles de los átomos. Finalmente, se aplica otra rotación de  $\pi/2$  para producir un estado final  $|\psi_3\rangle \langle \psi_3|$  con  $|\psi_3\rangle = \sin(\omega t/2) |0\rangle + \cos(\omega t/2) |1\rangle$ . Una vez que el sistema se encuentra en el estado  $|\psi_3\rangle \langle \psi_3|$ , cada átomo se mide utilizando dispersión y detectando su fluorescencia con un fotomultiplicador. Entonces, el parámetro  $\omega$  se puede estimar a partir de muestras de variables aleatorias Bernoulli con función de probabilidad

$$\begin{aligned} p(1 | \omega; t) &= \cos^2 \left[ \frac{\omega t}{2} \right] = \frac{1 + \cos(\omega t)}{2}, \\ p(0 | \omega; t) &= 1 - p(1 | \omega). \end{aligned} \quad (5.3)$$

Físicamente, el número de 1's en los datos es un registro del número de detecciones de los átomos en el estado  $|1\rangle \langle 1|$ .

La información de Fisher para la función de probabilidad (5.3) es igual a  $t^2$ , la cual coincide con la información de Fisher cuántica, por lo que la CICR para  $N$  mediciones independientes del sistema es

$$\text{Var}_\omega(\hat{\omega}) \geq \frac{1}{Nt^2}, \quad (5.4)$$

donde  $t$  es el tiempo de una única medición [13, 23, 61]. De esta ecuación, se puede pensar que al incrementar el tiempo de medición  $t$  es posible obtener errores para la estimación de  $\omega$  tan pequeños como se desee, sin embargo, este hecho no es verdad.

El parámetro  $\omega t$  es  $2\pi$ -periódico, por lo tanto si uno conoce el orden de  $\omega$  es posible escoger  $t \leq 2\pi/\omega$  para que  $0 \leq \omega t \leq 2\pi$ . Este hecho muestra que no es posible a primera instancia escoger tiempos de medición arbitrarios para reducir la cota (5.4). Además, estrategias de estimación como la espectroscopía de Ramsey restringen aún más el tiempo de medición. Como la función de probabilidad (5.3) es una función lineal de  $\cos^2(\omega t/2)$ , para obtener una única estimación del parámetro  $\omega t$  utilizando el estimador de máxima verosimilitud dentro del intervalo  $[0, 2\pi)$  se debe acotar el tiempo por  $t \leq \pi/\omega$ , pues de otro modo la función de verosimilitud producida por la ecuación (5.3) tendrá un problema de no identificabilidad sobre  $\omega$ , el cual se irá acentuando como función de  $t$ .

Por comodidad en los cálculos redefiniremos las escalas temporales como  $t = \tilde{t}/\omega_N$ , donde  $\omega_N$  es un valor plausible para el máximo de  $\omega$ . De esta forma se obtiene que  $0 \leq \tilde{t}\tilde{\omega} \leq 2\pi$ , con  $\tilde{\omega} = \omega/\omega_N$ . En el resto del capítulo, se describirá cómo es posible resolver el problema de no identificabilidad acerca de  $\omega$  utilizando el método de [estimación por intervalos de confianza \(EIC\)](#) descrito en (4.4). Solucionar este problema abre la posibilidad de aumentar el tiempo de medición y reducir el error en la estimación de  $\omega$ . Por simplicidad de notación, se continuará utilizando  $t$  y  $\omega$ , en vez de  $\tilde{t}$  y  $\tilde{\omega}$ .

## 5.1. Estimación de frecuencia con intervalos de confianza

El problema de estimar un parámetro  $\omega$  codificado en un qubit dentro de una fase  $\theta = \omega t$   $2\pi$ -periódica con  $t \in \mathbb{R}$  es llamado estimación cuántica de frecuencia en qubits [69]. La forma más general de codificar  $\omega$  en un qubit, es dado por el automorfismo

$$\rho_{t\omega} = e^{-it\omega \frac{\vec{\sigma}}{2} \cdot \vec{n}} \rho e^{it\omega \frac{\vec{\sigma}}{2} \cdot \vec{n}}, \quad (5.5)$$

donde  $t\omega \in [0, 2\pi)$ ,  $\vec{\sigma}$  es el vector de matrices de Pauli (4.1), y  $\rho = \frac{1}{2}(I + \vec{\sigma} \cdot \vec{a})$  el estado base que representa las condiciones iniciales en el vector de Bloch  $\vec{a}$ . El objetivo es estimar  $\omega$  a partir de mediciones sobre el sistema. Con este fin, se utiliza un estimador  $\hat{\omega}$ , es decir, una función cuya imagen es el conjunto  $\Omega$  de posibles valores de  $\omega$ .

Sustituyendo  $\theta = \omega t$  en el score cuántico (4.8), se observa que la información cuántica de Fisher para  $\omega$  asociada al estado (5.5) es

$$I_Q = t^2 \left[ \|a\|^2 - (\vec{a} \cdot \vec{n})^2 \right]. \quad (5.6)$$

Cantidad que llega a su valor máximo  $t^2$  en las condiciones iniciales  $\vec{n} \cdot \vec{a} = 0$ . Además, en estas condiciones iniciales, el POVM  $M_\omega$  construido como los proyectores a los eigenestados de la derivada simétrica logarítmica, satura (5.4) para cualquier  $\omega \in \Omega$ , es decir, es globalmente óptimo.

Explícitamente, los elementos de  $M_\omega$  son de la forma

$$M_g(0) = \frac{1}{2} (I + \vec{n} \times \vec{a}(g) \cdot \vec{\sigma}), \quad M_g(1) = \frac{1}{2} (I - \vec{n} \times \vec{a}(g) \cdot \vec{\sigma}), \quad (5.7)$$

con  $\vec{a}(g) = \cos(g)\vec{a} + \sin(g)\vec{n} \times \vec{a}$ ,  $g \in \Omega$ . Entonces, las salidas del POVM  $M_\omega$  tienen una distribución Bernoulli con función de probabilidad

$$p(x | \omega; t) = \begin{cases} \text{Tr} [\rho_{\omega,t} M_g(0)] = \frac{1}{2} [1 + \sin(t(\omega - g))] & \text{si } x = 0 \\ \text{Tr} [\rho_{\omega,t} M_g(1)] = \frac{1}{2} [1 - \sin(t(\omega - g))] & \text{si } x = 1. \end{cases} \quad (5.8)$$

Por tanto, la información de Fisher de este POVM es

$$I_{M_\omega}(\omega) = t^2 \frac{\cos^2(t(\omega - g))}{1 - \sin^2(t(\omega - g))} = t^2 = I_Q, \quad \forall g, \omega \in \Omega. \quad (5.9)$$

No obstante, al igual que en el problema de estimación de fase, dado un conjunto de datos de tamaño  $N$  obtenidos tras efectuar una secuencia de  $N$  medidas  $M_\omega$  independientes, la función de verosimilitud

$$\mathcal{L}(\omega) = p(0 | \omega; t)^m (1 - p(0 | \omega; t))^{N-m}, \quad (5.10)$$

es no identificable (podemos tener varios máximos globales) para cualquier condición inicial, lo que produce estimadores no consistentes.

La diferencia principal con respecto al caso de estimación de fase es que, en el problema de estimación de frecuencia, el número de máximos globales en la función de verosimilitud (5.10) aumenta por cada cierto incremento del tiempo  $t$ . En particular, el máximo valor de  $t$  para garantizar que las funciones de verosimilitud (5.10) sean identificables es  $t = 1/2$ . En el panel A de la Fig. 5.1 mostramos este comportamiento para  $g = \omega = 2$  y  $N = 64$ . Cada incremento de  $1/2$  para  $t$  produce un nuevo máximo en la función de verosimilitud (paneles C-D de la Fig. 5.1).

Si uno conociera cerca de qué máximo global se encuentra la frecuencia que se desea estimar, al restringir el espacio paramétrico a esta región, la función de verosimilitud sería identificable y por tanto, se obtendrían cotas de Crámer-Rao menores. El método de estimación por intervalos de confianza descrito en el capítulo anterior se puede adaptar como la herramienta para localizar la región donde se encuentra la frecuencia para valores de  $t > 1/2$ . De esta forma, calculando los intervalos de confianza para  $\theta = \omega$  con la ecuación (2.57) se resuelve el problema de no identificabilidad y pérdida de información acerca del  $\omega$  a medida que  $t$  aumenta.

El método consiste en realizar una estrategia óptima sobre todo el espacio paramétrico hasta que uno obtiene suficiente resolución acerca de la fase para incrementar  $t$ , y luego repetir el proceso. De esta manera, la estrategia de estimación queda adaptada a este problema como sigue: Uno comienza suponiendo que  $\omega \in \Omega$  y fijando el tiempo de medición  $t$  al valor máximo donde podemos asegurar que la función de verosimilitud sea identificable. Entonces, usando una estimación de máxima verosimilitud  $\hat{\omega}_{EMV}$  producida por una estrategia  $M$ , se produce un intervalo de confianza  $IC$ . Después, fijando  $IC = \Omega$ , uno incrementa  $t$  al máximo valor donde (5.10) sea identificable y se repite el proceso anterior.

Por su fácil implementación experimental, se escogió  $M$  como AQSE junto con la familia de POVMs  $\{M_g\}_{g \in \Omega}$  [38, 92], la cual es localmente óptima. Así, para el  $n$ -ésimo paso adaptativo, la función de verosimilitud para  $\omega$  es

$$\mathcal{L}_n(\omega; x_1, \dots, x_{n-1}; g_{n-1}) = \prod_{i=1}^n p(x_i | \theta; g_{i-1}) \quad (5.11)$$

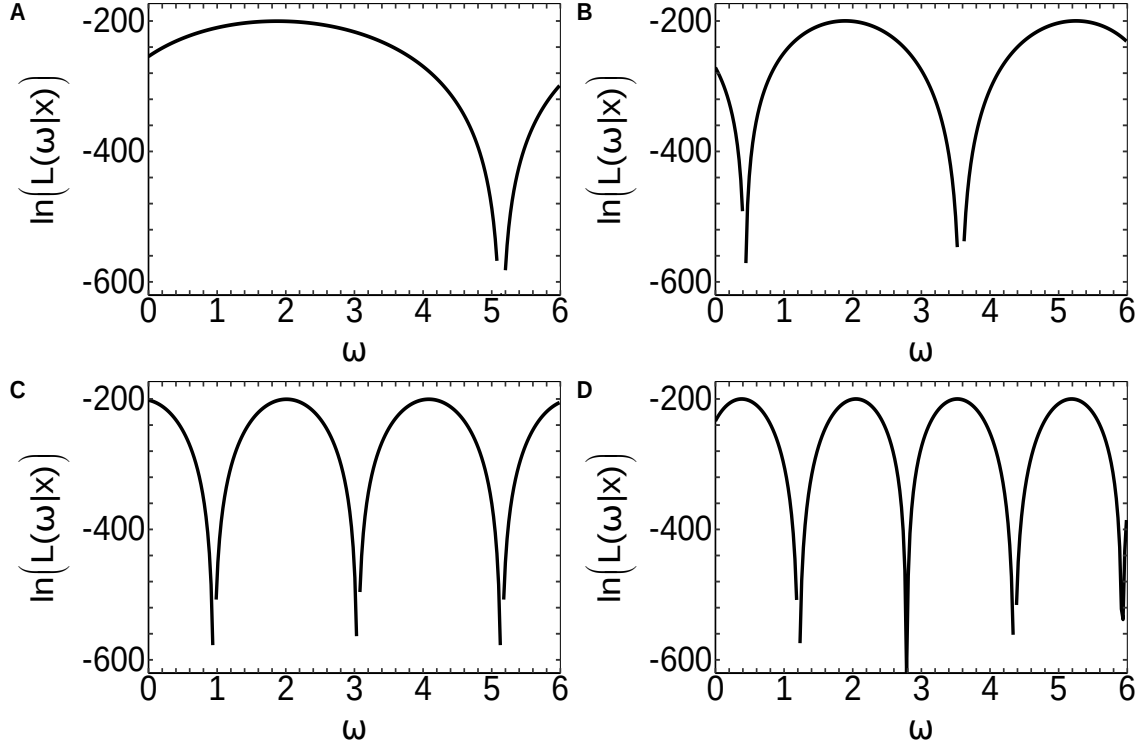


Figura 5.1: Funciones de verosimilitud producidas por  $M_\omega$ . **A**: verosimilitud cuando  $t = \frac{1}{2}$ . En este caso, la función es identificable. **B-D**: funciones de verosimilitud para  $t = 1, \frac{3}{2}, 2$ . Aquí, las funciones no son no identificables. Para todos los casos, se generó una secuencia de 64 datos provenientes del POVM  $M_\omega$  (5.7) y se supuso que  $\omega = 2$ .

done  $x_i$  es el dato observado en el paso  $i$  y  $g_{n-1}$  la estimación de máxima verosimilitud del paso anterior. Así, la estimación correspondiente a este paso es  $g_n = \hat{\theta}_n(x_1, \dots, x_n)$ , obteniendo la cota de Cramér-Rao para el estimador de máxima verosimilitud

$$\text{Var}_\omega [\hat{\omega}_{EMV}] \geq \frac{1}{nI_{M_\omega}(\omega)} = \frac{1}{nt^2}. \quad (5.12)$$

La función de verosimilitud (5.11) tiene dos máximos separados a una distancia de  $|\pi/t|$ . Por tanto, cada vez que  $t$  se incrementa  $1/2$ , un nuevo máximo aparece en la función de verosimilitud sobre el dominio de  $[0, 2\pi)$ , produciendo un estimador inconsistente para AQSE. No obstante, si uno inicia con un tiempo de  $t = 1/2$  y se va incrementando  $1/2$  sobre el tiempo en cada cierto número de pasos adaptativos, al usar intervalos de confianza con un error marginal  $E \leq \pi/t$  es posible producir estimadores consistentes para  $t \geq 1/2$  al usar AQSE.

Para un nivel de confianza  $C_l \in [0, 1]$ , en cada paso adaptativo, se tienen dos contribuciones al error del estimador. El primero se da cuando los intervalos de confianza contienen al parámetro en cuestión, y la segunda contribución es debida al caso contrario. Cuando los intervalos de confianza incluyen al parámetro, el estimador de máxima verosimilitud de AQSE converge en distribución a una normal con media  $\omega$  y varianza igual al inverso de la información de Fisher [19, 38]. Entonces,

la varianza para el estimador está acotada por  $1/I_M(\omega)$ , donde  $I_M(\omega)$  es la información de Fisher asociada a la estrategia  $M$ .

Por otro lado, si los intervalos de confianza no contienen al parámetro, como  $I(\omega, M_g(\omega)) = t^2$  para cualquier  $g \in [0, 2\pi)$ , AQSE produce un estimador de máxima verosimilitud normalmente asintótico con varianza  $1/I(\omega, M_g(\omega))K$  ( $K$  es el número de pasos adaptativos), y media alrededor de algún punto dentro del intervalo de confianza. Entonces, en el mejor de los casos, la distribución asintótica para el estimador es una mezcla de 3 distribuciones normales,  $\mathcal{N}(\omega, \frac{1}{I_M(\omega)})$ ,  $\mathcal{N}(\omega + E, \frac{1}{I_M(\omega)})$ , y  $\mathcal{N}(\omega - E, \frac{1}{I_M(\omega)})$ , donde  $M$  es el POVM producido por la secuencia de AQSE y  $E$  es el error marginal. De este modo, la varianza  $\sigma^2$  para esta distribución final puede ser calculada como

$$\begin{aligned} \sigma^2 &= \frac{(1 - C_l)}{2}(\omega + E)^2 + \frac{(1 - C_l)}{2}(\omega - E)^2 + C_l\omega^2 \\ &\quad - \left( C_l\omega + \frac{(1 - C_l)}{2}(\omega + E) + \frac{(1 - C_l)}{2}(\omega - E) \right)^2 + \frac{1}{I_M(\omega)} \\ &= \frac{1}{I_M(\omega)} + (1 - C_l)E^2. \end{aligned} \quad (5.13)$$

Entonces, como el error de los intervalos malos en los subsecuentes pasos adaptativos es despreciable, se obtiene como resultado que el error de estimación para  $N$  pasos adaptativos esté acotado inferiormente por

$$\text{Var}_\omega [\hat{\omega}_{EMV}] \geq \frac{1}{\sum_{i=0}^N N_i I(\omega; t_i)} + (1 - C_l)E^2, \quad (5.14)$$

donde  $t_0$  es el tiempo de medida para el paso inicial y  $\text{Var}_\omega [\hat{\omega}_{EMV}]$  es la varianza de Holevo [57]. Cabe resaltar que la segunda contribución es dominante cuando  $N \gg 1$  pues si el parámetro se encuentra fuera del intervalo de confianza, hacer más mediciones no va a reducir este error. Así, con el fin de observar pequeñas contribuciones del segundo término de (5.14), se recomienda utilizar niveles de confianza cercanos a 1.

Además, usando (4.46) uno puede deducir el mínimo de pasos adaptativos  $K$  en AQSE para obtener un intervalo de confianza cuya longitud sea menor al error marginal  $E$ ,

$$K_{min} \geq \frac{z_{\alpha/2}^2}{E^2 I(\omega; t_i)}. \quad (5.15)$$

Lo cota inferior (5.14) se muestra como función del número de mediciones realizadas  $N$  en la Fig. 5.2. Se muestran diferentes cotas para diferentes niveles de confianza ( $C_l = 0.95, 0.99$ , y  $0.999$ ). La línea discontinua roja representa la cota de Crámer-Rao cuántica para  $t = 1/2$ , indicando el punto de referencia donde el método adaptativo con intervalos de confianza (ECI) podría producir una mejora. Además, esta figura muestra cómo el segundo término de (5.14) domina cuando  $N \gg 1$ , reafirmando que para obtenerse una mejora consistente se requieren niveles de confianza cercanos a uno.

Como la consistencia del EMV es una propiedad asintótica, es necesario invertir una cantidad suficientes de pasos adaptativos  $K_0$  al tiempo inicial  $t_0$  para obtener una estimación adecuada para la construcción del primer intervalo de confianza. Posteriormente, para los tiempos  $t > t_0$ , se puede utilizar el mínimo de pasos adaptativos  $K_{min}$  en AQSE calculados por (5.15) ya que la distribución del EMV es una distribución normal. De esta forma, el método adaptativo con

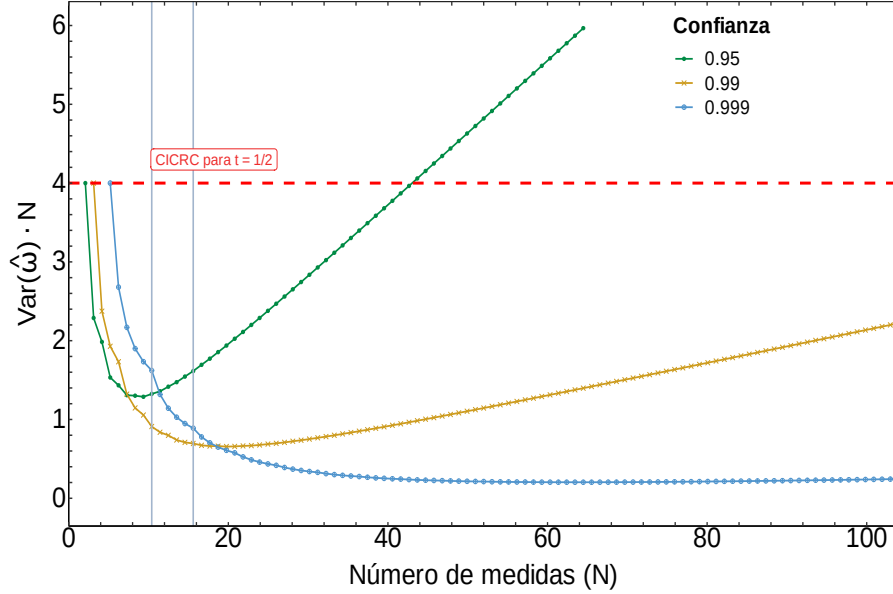


Figura 5.2: Cota inferior (5.14) en la varianza del estimador de máxima verosimilitud producido por ECI, donde los niveles de confianza utilizados fueron  $C_l = 0.95, 0.99$  y  $0.999$ . Las líneas grises verticales en  $N = 10$  y  $N = 15$  indican los dos primeros incrementos del tiempo para el caso de  $C_l = 0.999$ . Para realizar los incrementos temporales subsecuentes se necesita una cantidad menor de medidas, llegando a realizar eventualmente un incremento temporal en cada paso.

intervalos de confianza (ECI) amoldado al problema de estimación de frecuencia se puede resumir con el algoritmo 1.

Para probar la posible implementación de EIC, se realizó una simulación de Monte Carlo en la cual se compara el rendimiento de error de la estimación con la cota analítica (5.14). En la Fig. 5.3 se presentan los resultados de esta simulación, donde EIC 1 fue aplicado con  $K_0 = 16$  al estado (5.5). La línea punteada representa los resultados de ECI con  $C_l = 0.999$ , los puntos circulares representan el rendimiento de AQSE sin aumentar  $t$ . La curva discontinua representa la cota inferior calculada por (5.14), la línea sólida la cota inferior ideal, donde  $C_l = 1$ , y la línea discontinua con trazos largos representa la cota de Crámer-Rao cuántica (5.4) para  $t = 1/2$ . Estos resultados numéricos muestran que EIC puede superar la cota de Crámer-Rao particular para  $t = 1/2$  cuando la confianza con la que se trabaja es cercana a uno. Sin embargo, cuando el número de mediciones  $N \gg 1$ , la segunda parte de la ecuación (5.14) domina, incrementado el error de la estimación.

## 5.2. Estimación de frecuencia en qutrits

En los últimos años, se ha encontrado que los qutrits son capaces de producir mejoras en diversos protocolos de información cuántica [65]. En particular, al usar qutrits es posible mejorar el límite en la precisión de la secuencia de estimadores para  $\omega$  producida por el algoritmo 1.

Un qutrit es un estado de un sistema descrito por un espacio de Hilbert  $\mathcal{H}$  tridimensional.



---

**Algoritmo 1:** Estimación con intervalos de confianza (EIC)
 

---

```

 $t_0 \leftarrow \frac{1}{2};$ 
 $\Omega_0^{AQSE} \leftarrow [0, 2\pi);$ 
 $\hat{\omega}_0^{AQSE} \leftarrow \text{rand}([0, 2\pi));$ 
 $K_0 \leftarrow$  número de medidas iniciales;
for  $k = 1$  to  $K_0$  do
  |  $\Omega_0 \ni \hat{\omega}_k^{AQSE} \leftarrow$  Estimado a partir de  $M_L(\hat{\omega}_{k-1}^{AQSE}, t_n);$ 
end
 $\hat{\omega}_1 \leftarrow \hat{\omega}_{K_0}^{AQSE};$ 
 $\Omega_1 \leftarrow IC(\hat{\omega}_1);$ 
 $t_1 \leftarrow t_0 + \frac{1}{2};$ 
for  $n = 1$  to  $N - 1$  do
  |  $\hat{\omega}_0^{AQSE} \leftarrow \hat{\omega}_n;$ 
  | for  $k = 1$  to  $K_{min}$  do
  | |  $\Omega_n \ni \hat{\omega}_k^{AQSE} \leftarrow$  Estimado a partir  $M_L(\hat{\omega}_{k-1}^{AQSE}, t_n);$ 
  | end
  |  $\hat{\omega}_{n+1} \leftarrow \hat{\omega}_{K_{min}}^{AQSE};$ 
  |  $t_{n+1} \leftarrow t_n + \frac{1}{2};$ 
  |  $\Omega_{n+1} \leftarrow IC(\hat{\omega}_{n+1});$ 
end
return  $\hat{\omega}_{N-1}$ 

```

---

Cualquier ket  $|\psi\rangle \in \mathcal{H}$  se puede escribir como la combinación lineal

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle + \gamma |2\rangle, \quad \alpha, \beta, \gamma \in \mathbb{C}, \quad (5.16)$$

donde  $|0\rangle$ ,  $|1\rangle$ , y  $|2\rangle$  son los elementos de una base de  $\mathcal{H}$ . De este modo, el operador densidad  $\rho = |\psi\rangle\langle\psi|$  describe al estado de cualquier qutrit puro.

Tomando como condición inicial a un qutrit  $\rho \in \mathcal{S}(\mathcal{H})$ , la forma más general de codificar una frecuencia  $\omega$  sobre este estado es vía la conjugación del operador unitario

$$U(\omega, k_1, k_2) = e^{-i\omega[k_1|0\rangle\langle 0| - k_2|1\rangle\langle 1|]}, \quad k_1, k_2 \in \mathbb{R}. \quad (5.17)$$

En especial, si  $\rho$  es un estado puro se tiene que

$$\begin{aligned} \rho_{\omega; k_1, k_2} &= U(\omega, k_1, k_2) \rho U^\dagger(\omega, k_1, k_2) \\ &= |\psi(\omega)\rangle\langle\psi(\omega)|, \end{aligned} \quad (5.18)$$

donde  $|\psi(\omega)\rangle = \alpha e^{-ik_1\omega} |0\rangle + \beta e^{-ik_2\omega} |1\rangle + \gamma |2\rangle$ .

De acuerdo con (3.18), el score cuántico correspondiente a este estado es:

$$\begin{aligned} L(\omega) &= 2 \left[ i\alpha\bar{\beta} (k_2 - k_1) e^{i\omega(k_2 - k_1)} |0\rangle\langle 1| - i\alpha\bar{\gamma} k_1 e^{-i\omega k_1} |0\rangle\langle 2| \right. \\ &\quad \left. - i\beta\bar{\gamma} k_2 e^{-ik_2\omega} |1\rangle\langle 2| + h.c. \right]. \end{aligned} \quad (5.19)$$

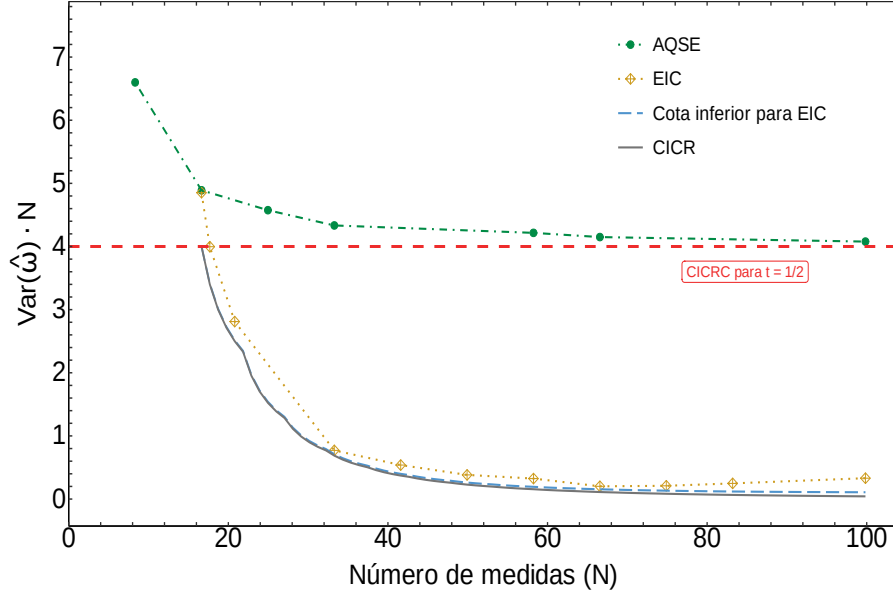


Figura 5.3: Varianza de Holevo como función del número de mediciones. El rendimiento de EIC para el sistema del qubit es presentado para  $K_0 = 16$  y  $C_t = 0.999\%$  de nivel de confianza (curva amarilla punteada con romboides). La cota inferior al error de EIC (curva discontinua azul) y la CICR de EIC (curva gris continua) se utilizan como referencia. Como complemento, también se muestra el desempeño de AQSE (curva verde discontinua con círculos) para  $t = 1/2$ . El cálculo de la varianza de Holevo del estimador de máxima verosimilitud en AQSE y EIC, se realizó a partir de un bootstrap de 10000 estimaciones en cada punto.

Por lo que la información cuántica de Fisher es

$$I_Q = 4 \left[ |\beta|^2 k_2^2 + |\alpha|^2 k_1^2 - (|\alpha|^2 k_1 + |\beta|^2 k_2)^2 \right]. \quad (5.20)$$

Entonces, cuando  $t := k_1 = -k_2$  y  $|\alpha| = |\beta| = \frac{1}{\sqrt{3}}$  la información cuántica de Fisher (5.20) alcanza su máximo valor

$$I_Q^{max} = \frac{8}{3} t^2. \quad (5.21)$$

En particular, si  $t = 1$ , entonces  $\omega \in [0, 2\pi)$  y el problema es análogo a la aplicación de la evolución paramétrica  $e^{-\frac{i}{2}\omega\sigma_z \cdot \hat{z}}$  sobre un qubit. Sin embargo, en este caso, se tiene un término adicional  $\gamma|2\rangle$ , el cual trae como consecuencia la obtención de una mayor información de Fisher cuántica. Además, este término adicional provoca que el POVM localmente óptimo  $M_\omega$  genere funciones de verosimilitud identificables (Fig. 5.4).

Para estas condiciones iniciales, los elementos de  $M_\omega$  son

$$M_\omega = \{|a\rangle \langle a|, |b\rangle \langle b|, |c\rangle \langle c|\}, \quad (5.22)$$

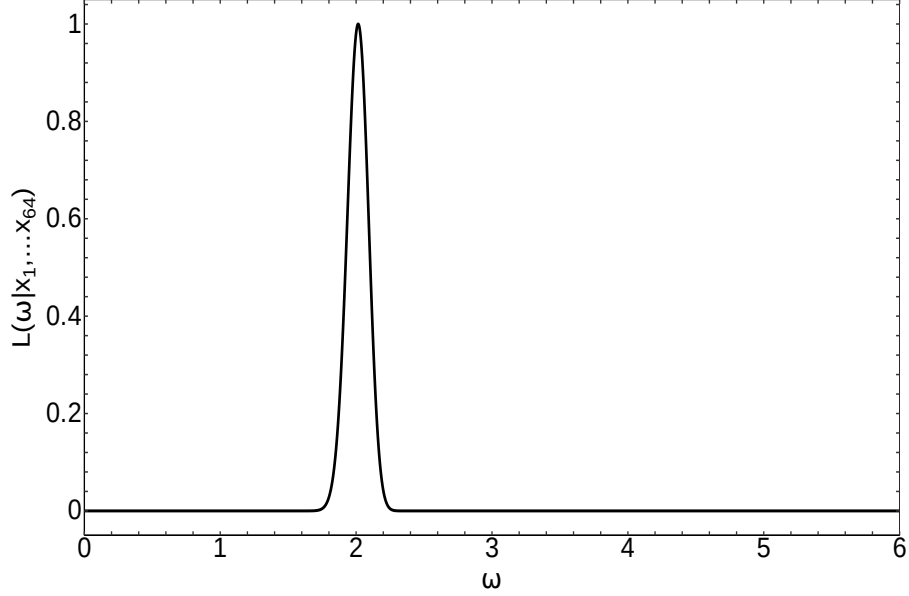


Figura 5.4: Función de verosimilitud identificable para una muestra de tamaño 64 obtenida del resultado de una secuencia de mediciones independientes  $M_\omega$ . La muestra fue numéricamente generada para un valor de  $\omega = 2$ .

donde

$$\begin{aligned}
 |a\rangle &= \left( \frac{\sqrt{5}}{2\sqrt{3}}, -\frac{(2^{\frac{3}{2}}\sqrt{3}i + 1)e^{2it\omega}}{2\sqrt{15}}, -\frac{(\sqrt{2}\sqrt{3}i - 2)e^{it\omega}}{2\sqrt{3}\sqrt{5}} \right) \\
 |b\rangle &= \left( \frac{\sqrt{5}}{2\sqrt{3}}, \frac{(2^{\frac{3}{2}}\sqrt{3}i - 1)e^{2it\omega}}{2\sqrt{3}\sqrt{5}}, \frac{(\sqrt{2}\sqrt{3}i + 2)e^{it\omega}}{2\sqrt{3}\sqrt{5}} \right) \\
 |c\rangle &= \left( \frac{1}{\sqrt{6}}, \frac{e^{2it\omega}}{\sqrt{6}}, -\frac{2e^{it\omega}}{\sqrt{6}} \right).
 \end{aligned} \tag{5.23}$$

De modo que al aplicar la regla de Born (3.11) la función de probabilidad para las salidas de  $M_\omega$  es

$$\begin{aligned}
 p(1 | \omega) &= \frac{2^{\frac{7}{2}} \sin(2t\omega - 2tg)}{25\sqrt{3}} + \frac{\sqrt{2} \sin(2t\omega - 2tg)}{25 \cdot 3^{\frac{3}{2}}} - \frac{\cos(2t\omega - 2tg)}{18} + \frac{2^{\frac{5}{2}} \sin(t\omega - tg)}{25\sqrt{3}} \\
 &\quad - \frac{\sqrt{2} \sin(t\omega - tg)}{25\sqrt{3}} - \frac{2^{\frac{3}{2}} \sin(t\omega - tg)}{53^{\frac{3}{2}}} + \frac{2 \cos(t\omega - tg)}{9} + \frac{1}{3} \\
 p(2 | \omega) &= \frac{\cos(2t\omega - 2tg)}{9} - \frac{4 \cos(t\omega - tg)}{9} + \frac{1}{3} \\
 p(3 | \omega) &= 1 - (p(1 | \omega) + p(2 | \omega)).
 \end{aligned} \tag{5.24}$$

La función de verosimilitud producida por (5.24) es identificable para todo  $\omega \in \Omega$ , por lo tanto AQSE es una estrategia de estimación óptima para  $t = 1$ . Sin embargo, la falta de resolución

sobre  $\omega$  aparece cuando  $t > 1$ . En particular, cada incremento de una unidad en  $t$  produce un máximo global adicional para la función de verosimilitud de  $M_\omega$ . Por tal motivo, AQSE no puede alcanzar la cota cuántica de Crámer-Rao para tiempos mayores a uno, lo que produce la necesidad de aplicar EIC para poder mejorar el error del AQSE al ir incrementando  $t$ . Para este caso, EIC se puede implementar aplicando el algoritmo 1 sobre el estado (5.18), con  $t_0 = 1$  un incremento para  $t$  de una unidad en cada paso del algoritmo.

En la Fig. 5.5 se presenta el comportamiento del error en la estimación de la frecuencia usando EIC, cuando esta es codificada en un qutrit (5.18). La línea punteada amarilla con rombos representa el escalamiento para EIC con  $C_l = 0.999$ , la línea discontinua verde con puntos circulares, el rendimiento de AQSE y la curva azul discontinua, la cota inferior (5.14). También, se agrega la cota inferior ideal (línea solida gris) y la cota cuántica de Crámer-Rao (5.21) cuando  $t = 1$  (línea discontinua roja) como comparación. Al igual que en el caso del qubit, estos resultados muestran que ECI puede superar la cota cuántica de Crámer-Rao cuando el nivel de confianza es cercano a uno, aunque en el límite  $N \gg 1$ , el error comienza a incrementar de acuerdo con (5.14). Este análisis numérico muestra que al codificar  $\omega$  en un qutrit y al aplicar EIC se puede disminuir el error de la estimación de la frecuencia a niveles más bajos que en el sistema del qubit.

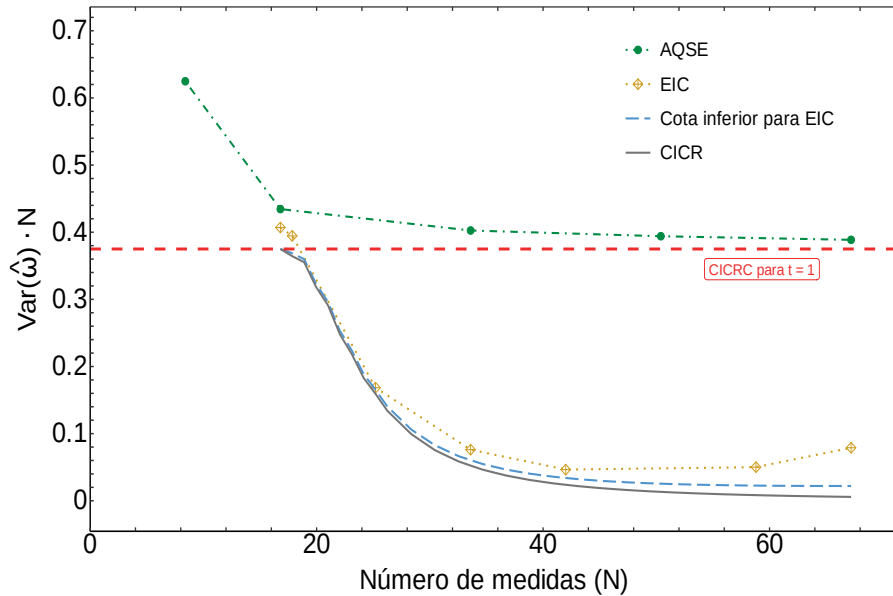


Figura 5.5: Varianza de Holevo como función del número de medidas  $N$ . El rendimiento en el error para EIC con  $K_0 = 16$  y  $C_l = .999\%$  para la estimación de una frecuencia  $\omega$  se muestra en la curva punteada amarilla con rombos. También se muestra el desempeño de AQSE (curva discontinua verde con círculos). La cota inferior ideal (línea solida gris) y la cota cuántica de Cramér-Rao (5.21) cuando  $t = 1$  (línea discontinua roja) sirven como comparación. Para calcular la varianza de Holevo las distribuciones de los estimadores de máxima verosimilitud para AQSE y ECI, se aproximaron usando un bootstrap de  $10^4$  estimaciones en cada punto.

## Capítulo 6

# Estimación de fase en estados coherentes

Como ya se ha mencionado, el POVM  $M(\theta)$  cuyos elementos son los proyectores a los eigenspacios del score cuántico (3.15) es localmente óptimo. Para tratar este problema se pueden utilizar métodos adaptativos de estimación como AQSE [38]. No obstante, para el problema de estimación de fase, el POVM  $M(\theta)$  produce funciones de verosimilitud no identificables y en consecuencia, incluso al utilizar métodos adaptativos se obtienen estimadores no consistentes.

En el capítulo 4 se dio una solución a esta problemática introduciendo un esquema adaptativo de estimación mixto con intervalos de confianza basado en la medida canónica de fase y AQSE cuando la fase es codificada en qubits. La medida canónica de fase minimiza el error de las estimaciones cuando se tiene una única copia del sistema y soluciona el problema de no identificabilidad. Para el caso de qubits se conocen técnicas para su implementación física [22, 79]. Sin embargo, la realización experimental de esta medida en sistemas cuánticos de mayor dimensión sigue siendo un problema abierto. Por esta razón, para implementar un esquema adaptativo con intervalos de confianza análogo al descrito en el capítulo 4 en sistemas cuánticos de mayor dimensión, se debe encontrar un sustituto a la medida canónica de fase que sea posible de realizar experimentalmente. Esta cuestión recae en la búsqueda de POVMs que minimizan la varianza de Holevo para una sola copia del sistema. A este problema se le llama “*estimación de disparo único*”. El caso pragmático de esta situación es la estimación de fase en estados coherentes. En esta sección se discutirán las técnicas para la estimación óptima de fase en estados coherentes en el caso de disparo único.

### 6.1. Estados coherentes

De acuerdo con la sección 3.6, la dinámica de un sistema cuántico se describe a partir del operador Hamiltoniano. El Hamiltoniano para el oscilador armónico tiene la forma

$$H = \frac{1}{2m} [p^2 + m^2\omega^2 q^2], \quad (6.1)$$

donde  $p$  y  $q$  son los operadores de momento y posición respectivamente. El Hamiltoniano  $H$  puede ser escrito como

$$\begin{aligned} H &= \hbar\omega \left( a^\dagger a + \frac{1}{2} \right) \\ &= \hbar\omega \left( \hat{n} + \frac{1}{2} \right) \end{aligned} \quad (6.2)$$

donde

$$a = \frac{1}{\sqrt{2m\hbar\omega}} [\omega q + ip], \quad a^\dagger = \frac{1}{\sqrt{2m\hbar\omega}} [\omega q - ip] \text{ y } \hat{n} = a^\dagger a. \quad (6.3)$$

En particular, el operador no acotado  $\hat{n}$  es esencialmente autoadjunto<sup>1</sup>, por lo que su extensión autoadjunta a todo  $\mathcal{H}$  dada por

$$\begin{aligned} a|0\rangle &= 0, \quad |n\rangle = \frac{1}{\sqrt{n!}} (a^\dagger)^n |0\rangle, \quad n = 0, 1, \dots \\ D(\hat{n}) &= \left\{ |\psi\rangle \in \mathcal{H} : \sum_{n=0}^{\infty} n^2 |\langle n|\psi\rangle|^2 < \infty \right\} \\ \hat{n}|\psi\rangle &= \sum_{n=0}^{\infty} n |n\rangle \langle n|\psi\rangle; \quad |\psi\rangle \in D(\hat{n}). \end{aligned} \quad (6.4)$$

define una base ortonormal  $\{|n\rangle\}$  (llamada base de Fock) sobre el espacio de Hilbert del sistema  $\mathcal{H}$ . Es decir,

$$\sum_{n=0}^{\infty} |n\rangle \langle n| = I, \text{ y } |\psi\rangle = \sum_{n=0}^{\infty} |n\rangle \langle n|\psi\rangle, \quad \forall |\psi\rangle \in \mathcal{H}, \quad (6.5)$$

con  $|n\rangle$  los eigenvectores de la extensión autoadjunta de  $\hat{n}$  a todo  $\mathcal{H}$  con eigenvalor asociado  $n \in \{0\} \cup \mathbb{N} = \{0, 1, 2, \dots\}$ . Además, debido a que se cumplen las siguiente relaciones

$$\begin{aligned} a|n\rangle &= \sqrt{n} |n-1\rangle, \\ a^\dagger|n\rangle &= \sqrt{n+1} |n+1\rangle, \\ \hat{n}|n\rangle &= n|n\rangle \end{aligned} \quad (6.6)$$

los operadores  $a$ ,  $a^\dagger$  y  $\hat{n}$  reciben el nombre operador de aniquilación, operador de creación y operador de número respectivamente [82].

Los kets de los estados coherentes se definen como los eigenestados del operador (no autoadjunto) de aniquilación [57]

$$a|\alpha\rangle = \alpha|\alpha\rangle, \quad \alpha \in \mathbb{C}. \quad (6.7)$$

Estos estados son utilizados para describir los estados de luz que tienen un número indefinido de fotones pero una fase bien definida, un ejemplo de esto son los rayos láser [40]. Estos estados cumplen con propiedades de gran interés práctico. A continuación se mencionarán algunas que serán relevantes para este trabajo.

Los estados coherentes también pueden ser generados a partir del estado vacío ( $|0\rangle$ ),

$$|\alpha\rangle = D(\alpha)|0\rangle, \quad (6.8)$$

donde

$$D(\alpha) = \exp\left(-\frac{1}{2}|\alpha|^2\right) \exp(\alpha a^\dagger) \exp(\alpha^* a), \quad \alpha \in \mathbb{C} \quad (6.9)$$

<sup>1</sup>Un operador no acotado sobre un espacio de Hilbert  $\mathcal{H}$  con dominio  $D(A) \subset \mathcal{H}$  es esencialmente autoadjunto si tiene una única extensión autoadjunta a todo  $\mathcal{H}$  [101].

es conocido como el operador de desplazamiento. También se pueden producir estados comprimidos si se aplica un operador de compresión previo al operador de desplazamiento

$$|\alpha, r\rangle = D(\alpha)S(r)|0\rangle, \quad (6.10)$$

donde

$$S(r) = \exp\left[\frac{1}{2}(r^*a^2 - r(a^\dagger)^2)\right] \quad r \in \mathbb{C}. \quad (6.11)$$

De igual forma, escribiendo los ket  $|\alpha\rangle$  en la representación de la base de Fock

$$|\alpha\rangle = \exp\left(-\frac{1}{2}|\alpha|^2\right) \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle, \quad (6.12)$$

se puede calcular la distribución de probabilidad de fotones de forma sencilla,

$$p(n) = \text{Tr}[|n\rangle\langle n| |\alpha\rangle\langle\alpha|] = \frac{|\alpha|^{2n} \exp(-|\alpha|^2)}{n!}, \quad (6.13)$$

es decir,  $p(n)$  es la función de probabilidad de una distribución de Poisson con parámetro  $|\alpha|^2$ , el cual corresponde con el número promedio de fotones

$$\begin{aligned} \bar{n} &= \sum_{i=1}^{\infty} ip(i) \\ &= \sum_{i=1}^{\infty} i \text{Tr}[|i\rangle\langle i| |\alpha\rangle\langle\alpha|] \\ &= \text{Tr}[\hat{n} |\alpha\rangle\langle\alpha|] \\ &= E_{|\alpha\rangle\langle\alpha|}[\hat{n}] = |\alpha|^2. \end{aligned} \quad (6.14)$$

Otra propiedad interesante es que los estados coherentes minimizan respecto a  $S(\mathcal{H})$ , la relación de incertidumbre de Heisenberg:

$$\text{Var}_\rho[q] \text{Var}_\rho[p] \geq \frac{1}{4}, \quad (6.15)$$

donde la varianza de un operador  $A$  sobre  $\mathcal{H}$  se define como

$$\text{Var}_\rho[A] = E_\rho[(A - E_\rho[A])^2]. \quad (6.16)$$

Es decir, si  $\rho = |\alpha\rangle\langle\alpha|$  con  $\alpha \in \mathbb{C}$ , entonces  $\text{Var}_\rho[q] \text{Var}_\rho[p] = 1/4$ .

Por último, señalemos que los estados coherentes forman un POVM con elementos  $M(d^2\alpha) = \frac{|\alpha\rangle\langle\alpha|}{\pi}$  no ortogonales entre sí [57], pues

$$\int M(d^2\alpha) = I, \quad (6.17)$$

y

$$\text{Tr}[|\alpha\rangle\langle\alpha| |\beta\rangle\langle\beta|] = \exp[-|\alpha - \beta|^2]. \quad (6.18)$$

Este POVM es la descripción matemática de la detección heterodina [3].

De esta forma, ya habiendo establecido los estados base sobre los cuales codificaremos un parámetro a estimar y sus propiedades más relevantes, se procederá a describir el problema de estimación de fase para el caso de un solo disparo y diferentes protocolos de estimación para el mismo.

## 6.2. Descripción del problema

De acuerdo con el teorema de Stone [48], se tiene una correspondencia uno a uno entre el Hamiltoniano de un sistema y una familia de operadores unitarios  $\{V_t\}_{t \in \mathbb{R}}$  que determina la dinámica del sistema. Como el Hamiltoniano (6.1), es diagonal en la base de Fock, los elementos de  $\{V_t\}_{t \in \mathbb{R}}$  se pueden escribir como

$$V_t = \sum_{n=0}^{\infty} e^{-i\omega(n+1/2)t} |n\rangle \langle n|. \quad (6.19)$$

En este caso, de acuerdo con (3.51), la familia de estados

$$\begin{aligned} \rho_t &= e^{-i\omega(n+1/2)t} \rho e^{i\omega(n+1/2)t} \\ &= e^{-i\omega t} \rho e^{i\omega t}, \quad t \in \mathbb{R}, \end{aligned} \quad (6.20)$$

corresponde a la evolución temporal del estado base  $\rho$  para diferentes valores del parámetro de evolución  $t$ . Además por la propiedad de covarianza, cada elemento de un POVM  $M$  debe evolucionar temporalmente como

$$M(B+t) = e^{-i\omega t} M(B) e^{i\omega t}, \quad t \in \mathbb{R}, \quad B \text{ un conjunto de Borel.} \quad (6.21)$$

Debido a que la dinámica es periódica en el sentido que  $V_{t+\frac{2\pi}{\omega}} = V_t$ , se tiene que la evolución temporal de un estado coherente  $\rho = |\alpha\rangle \langle \alpha|$  es equivalente a

$$\rho_\theta = e^{-i\theta \hat{n}} |\alpha\rangle \langle \alpha| e^{i\theta \hat{n}}, \quad (6.22)$$

donde  $\theta = (-\omega t) \text{ mód } 2\pi$  es un parámetro en  $[0, 2\pi) = \Theta$ . Como los eigenvalores de  $\hat{n}$  son enteros, entonces  $\exp(i\theta \hat{n}) = \exp(i(\theta + 2\pi k)\hat{n})$  y la familia  $\{U_\theta = e^{-i\theta \hat{n}}\}_{\theta \in \Theta}$  es una representación unitaria del grupo aditivo  $[0, 2\pi)$  módulo  $2\pi$ .

Este grupo es isomorfo al grupo de rotaciones sobre el círculo unitario  $\mathbb{T}$ , por lo que  $\theta$  en (6.22) es una fase. Por ende, si  $\theta$  es un parámetro desconocido, se tiene un problema de estimación de fase sobre estados coherentes con cota de cuántica de Crámer-Rao (ver (3.19)) de

$$\text{Var}_\theta \left[ \left( M, \hat{\theta} \right) \right] \geq \frac{1}{4|\alpha|^2}. \quad (6.23)$$

De acuerdo con el teorema 8, el POVM que minimiza la varianza de Holevo en el problema de estimación de disparo único es la medida canónica de fase:

$$M_*(d\theta) = \frac{d\theta}{2\pi} \sum_{n,m=0}^{\infty} |n\rangle \langle m| e^{i(n-m)\theta}. \quad (6.24)$$

Además este POVM, puede saturar la cota de cuántica de Crámer-Rao en el límite de altas energías  $|\alpha| \gg 1$  [57]. Fuera de este límite, este POVM solo es de varianza mínima. Físicamente, esto significa que la varianza del POVM (6.24)

$$V_{can} = \left| e^{-\alpha^2} \sum_{n=0}^{\infty} \frac{\alpha^{2n+1}}{n! \sqrt{n+1}} \right|^2 - 1 \quad (6.25)$$



es el límite fundamental de precisión para el caso de estimación con una sola medida [58, 118]. Sin embargo, a la fecha no hay un método experimental que implemente de forma exitosa este POVM. Por ejemplo, en el trabajo [97], se puede obtener la distribución de las salidas para la medición canónica a partir de la distribución conjunta de un POVM en el espacio fase, el cual requiere realizar el doble de mediciones sobre el sistema, y más recientemente en [79] se logró implementar el POVM (6.24) para un paquete de onda de un fotón utilizando retroalimentación cuántica, dejando aún abierto el problema de implementación para estados coherentes.

Por tal motivo, las estrategias Gaussianas, las cuales se pueden realizar de forma simple experimentalmente, se han convertido en las técnicas estándar para atacar el problema de estimación de fase en estados coherentes. El punto de partida natural dentro de estas estrategias es la detección heterodina, cuya varianza está acotada inferiormente por  $1/2|\alpha|^2$  [118]. Múltiples estrategias adaptativas Gaussianas han superado la cota inferior en la varianza de la detección heterodina [31, 118]. Sin embargo, estas estrategias no ha sido capaz de alcanzar el desempeño de la medición canónica [118].

Actualmente, la mejor estrategia Gaussiana adaptativa que se conoce es la llamada **Mark-II (MK II)**, la cual esta basada en un esquema adaptativo de detecciones homodinas [118]. La estrategia MK II tiene una varianza en el límite asintótico dada por

$$\text{Var} \left[ \hat{\phi}_{MKII} \right] \approx \frac{1}{4|\alpha|^2} + \frac{1}{8|\alpha|^3}, \quad (6.26)$$

la cual supera la de la detección heterodina pero no puede alcanzar el desempeño de la medición canónica [118].

$$\text{Var} \left[ \hat{\phi}_{CAN} \right] \approx \frac{1}{4|\alpha|^2} + \frac{5}{32|\alpha|^4}. \quad (6.27)$$

Por lo tanto, encontrar una estrategia de estimación que pueda ser implementada físicamente y que aproxime al límite fundamental de precisión es un problema abierto.

El objetivo de este capítulo es introducir una estrategia adaptativa no-Gaussiana para el problema de estimación de fase de disparo único en estados coherentes basada en conteo de fotones [31], y probar que esta estrategia de estimación es capaz, bajo ciertas configuraciones, de superar el desempeño de MK II con una mejor razón de convergencia que cualquier otra estrategia de estimación expuesta en la literatura. Con este objetivo en mente se continuará con una exposición de los conocimientos previos necesarios para proceder al resultado principal de este capítulo.

### 6.3. Diseño Bayesiano óptimo de experimentos

La estrategia de estimación adaptativa para la estimación de fase en estados coherentes basada en conteo de fotones se puede describir en el contexto de la teoría de diseño Bayesiano óptimo de experimentos. En esta sección se dará un breve resumen de esta teoría matemática.

El diseño óptimo de experimentos permite mejorar las inferencias estadísticas de cantidades de interés al escoger apropiadamente el valor de ciertas variables de control llamadas diseños [21, 105, 109]. Aquí, se hace el supuesto que los datos observados  $y \in \mathcal{Y} \subset \mathbb{R}$  son salidas de una variable aleatoria con una densidad o función de probabilidad  $p(y | d, \theta)$ , la cual pertenece una familia indexada por  $\theta \in \Theta$

$$P_{\Theta} = \{p(y | d, \theta), d \in \mathcal{D}, \theta \in \Theta\}, \quad (6.28)$$

donde  $d$  es un parámetro llamado diseño escogido de algún conjunto  $\mathcal{D}$  llamado espacio de diseños,  $\theta \in \Theta$  es un parámetro desconocido y los datos pertenecen a un espacio muestra  $\mathcal{Y} \subset \mathbb{R}$ . En

este contexto, los diseños  $d$  son interpretados como parámetros de control dentro del experimento que pueden ser ajustados antes de realizar cualquier medición sobre el sistema. Esto permite que la optimización de dicha variable de control pueda mejorar la estimación del parámetro desconocido  $\theta$ .

En el enfoque Bayesiano [21, 105], la falta de información acerca de  $\theta$  se modela con una distribución de probabilidad apriori  $p(\theta)$ . Entonces, cuando se mide sobre el sistema se puede reducir la incertidumbre acerca de  $\theta$  aplicando el teorema de Bayes sobre la distribución apriori:

$$p(\theta | d, y) \propto p(\theta)p(y | d, \theta). \quad (6.29)$$

Para el problema de estimación, los valores óptimos para  $d$  son aquellos que maximizan el valor esperado de una función de costo  $U(d, \theta, y)$  sobre el espacio de diseños  $\mathcal{D}$  condicionada a los posibles datos futuros  $y$  y al espacio paramétrico  $\Theta$ :

$$\begin{aligned} d^{opt} &= \arg \max_{d \in \mathcal{D}} \mathbb{E} [U(d, \theta, y)] \\ &= \arg \max_{d \in \mathcal{D}} \int_{\mathcal{Y}} \int_{\Phi} U(d, \theta, y) p(\theta, y | d) d\theta dy \\ &= \arg \max_{d \in \mathcal{D}} \int_{\mathcal{Y}} \int_{\Phi} U(d, \theta, y) p(\theta | d, y) d\theta p(y | d) dy. \end{aligned} \quad (6.30)$$

Un enfoque estándar es fijar  $d^{opt}$  al comienzo del experimento y entonces recolectar datos de forma independiente e idénticamente distribuidos con una distribución dada por  $p(y | d^{opt}, \theta)$ . Un segundo enfoque es tomar la distribución a posteriori (resultado del teorema de Bayes (6.29)) como la nueva prior en cada medición para ir actualizando  $d^{opt}$  de forma adaptativa, conforme más datos van siendo recolectados. La ventaja de este enfoque es que las estrategias adaptativas nunca son menos eficientes que las no adaptativas [95].

La teoría de diseño Bayesiano óptimo de experimentos tiene libertad en la elección de la función de costo, la cual depende del problema bajo estudio. Para el problema de estimación de parámetros, la función de costo más común en el problema de diseño Bayesiano óptimo de experimentos es la divergencia de Kullback-Liebert (KLD - acrónimo tomado del inglés) entre la distribución a priori y distribución a posteriori

$$\begin{aligned} U(d, y) &= D_{KL} [p(\theta | d, y) || p(\theta)] \\ &= \int_{\Theta} p(\theta | d, y) \log \left[ \frac{p(\theta | d, y)}{p(\theta)} \right] d\theta. \end{aligned} \quad (6.31)$$

La divergencia de Kullback-Liebert provee una distancia entre la distribución de probabilidad  $p(\theta | d, y)$  y la información acerca de  $\theta$ ,  $p(\theta)$  [24]. Si  $p(\theta | d, y)$  es igual a  $p(\theta)$ , entonces  $U(d, y) = 0$  y no se obtiene una ganancia de información acerca de  $\theta$  por haber medido  $y$  con el diseño  $d$ .

Otra función de costo comúnmente utilizada en el diseño óptimo de experimentos es el negativo de entropía condicional entre los valores plausibles de  $\theta$  y  $y$ ,

$$\begin{aligned} U(d) &= -H(\theta | Y) \\ &= - \sum_{y \in \mathcal{Y}} p(y) \int_{\Theta} p(\theta | d, y) \log [p(\theta | d, y)] d\theta. \end{aligned} \quad (6.32)$$

la cual es una medida de cuánta información se necesita para describir las salidas de la variable aleatoria  $\theta$  dado el valor de otra variable aleatoria  $Y$  conocida [24]. No obstante, cabe destacar que las funciones de costo (6.31) y (6.32) se relacionan con la información mutua [24]:

$$\begin{aligned} U(d) &= MI(\theta | Y) = \mathbb{E}_y [D_{KL} [p(\theta | d, y) || p(\theta)]] \\ &= H(p(\theta)) - H(\theta | Y). \end{aligned} \quad (6.33)$$

Este hecho tiene como resultado que maximizar cualquiera de las funciones de costo (6.31), (6.32) y (6.33) sobre el espacio de diseños sea equivalente [21, 105].

La aplicación de la teoría de diseño óptimo de experimentos al problema de estimación de fase en estados coherentes con conteo de fotones se resume en los siguientes pasos:

- i) Utilizar una función de costo específica e información a priori para escoger el diseño.
- ii) Hacer una medición.
- iii) Usar el resultado del teorema de Bayes para actualizar la información apriori.
- iv) Repetir los pasos i)-iii) de forma recursiva.

De este modo, se obtiene una estimación para  $\theta$  en cada paso adaptativo utilizando el estimador de **máxima a posteriori (MAP)**, el cual se define como

$$\hat{\theta} = \arg \max_{\theta \in \Theta} p(\theta | y, d). \quad (6.34)$$

Este enfoque requiere que el estimador sea consistente conforme el número de mediciones se incrementa.

En el enfoque adaptativo del diseño Bayesiano óptimo de experimentos, Paninski [95] probó que bajo ciertas condiciones de regularidad y al utilizar funciones de costo basadas en la información mutua, se puede escoger diseños que producen un **MAP** asintóticamente consistente y eficiente con varianza

$$\sigma_{INFO}^2 = \left( \arg \max_{C \in co(I(\theta; d))} |C| \right)^{-1}, \quad (6.35)$$

donde  $co(I(\theta; d))$  denota la cerradura convexa del conjunto de informaciones de Fisher como función de los diseños  $I(\theta; d)$ . Formalmente, las condiciones de regularidad pueden ser establecidas como:

1. El espacio paramétrico  $\Theta$  es un espacio métrico compacto.
2. La función de verosimilitud logarítmica  $\log(p(y | \theta, d))$  es uniformemente Lipschitz en  $\theta$  con respecto a alguna medida dominante en  $\mathcal{Y}$ .
3. La función de verosimilitud es identificable para  $\theta$ , esto es, que la función de verosimilitud tenga un único máximo global.
4. La distribución a priori asigna una probabilidad a cualquier vecindad del valor real de  $\theta$ .
5. Las funciones de información de Fisher  $I(\theta; d)$  estén bien definidas para todo  $d \in \mathcal{D}$ .

6.  $\max_{C \in \text{co}(I(\theta; d))} |C| > 0$ , donde  $\text{co}(I(\theta; d))$  denota la envolvente convexa del conjunto de informaciones de Fisher  $I(\theta; d)$ .

Adicionalmente, cuando el parámetro a estimar es un escalar, la maximización de la información mutua es equivalente a minimizar el error cuadrático medio para el caso asintótico [95].

Como la fase en (6.22) es una cantidad escalar, en el límite asintótico, tanto la información mutua como el error cuadrático medio son funciones de costo apropiadas para producir estrategias de estimación óptimas. Entonces, uno puede utilizar la función de costo que le sea más sencilla de calcular. En la siguiente sección la estrategia no Gaussiana basada en conteo de fotones y retroalimentación será descrita utilizando las herramientas matemáticas de esta sección.

## 6.4. Esquema adaptativo no-Gaussiano

La estrategia adaptativa de estimación no Gaussiana basada en conteo de fotones y retroalimentación [31] tiene el objetivo de estimar  $\theta \in \Theta = [0, 2\pi)$  utilizando un número finito de pasos adaptativos  $l = 1, 2, \dots, L$  y la información a priori  $p(\theta)$  acerca de  $\theta$  de un estado coherente (6.22) con un número medio de fotones  $|\alpha|^2$ .

La estrategia consiste en dividir el pulso de luz descrito por un estado coherente (6.22), el cual contiene la información del parámetro codificada en la fase  $\theta$ , en  $L$  pedazos. En cada paso adaptativo  $l = 1, 2, \dots, L$ , el pulso se mezcla con otro pulso de luz proveniente de un oscilador local, cuya amplitud y fase se pueden elegir por alguna regla que en general puede depender de las mediciones de los pasos previos, de modo que el número medio de fotones para cada paso adaptativo antes de mezclarse con el oscilador local es  $|\alpha|^2/L$ .

Matemáticamente, el esquema adaptativo para  $L \in \mathbb{N}$  pasos adaptativos se puede describir como sigue [31]: Dada cierta información a priori  $P(\theta)$  acerca de  $\theta$  con densidad  $p(\theta)$ , se comienza con un número aleatorio  $\beta_0 \in \mathbb{C}$  y se aplica el POVM

$$\left\{ D(\beta_0) |n\rangle \langle n| D^\dagger(\beta_0) \right\}_{n=0}^{m-1} \cup \left\{ \mathbb{I} - \sum_{n=0}^{m-1} D(\beta_0) |n\rangle \langle n| D^\dagger(\beta_0) \right\} \quad (6.36)$$

sobre el estado  $|e^{i\theta}\alpha/\sqrt{L}\rangle \langle e^{i\theta}\alpha/\sqrt{L}|$ , donde  $D$  es el operador de desplazamiento introducido en la ecuación (6.9)  $D(\beta)|\alpha\rangle = |\alpha + \beta\rangle$  y  $m \in \mathbb{N}$ . En el POVM (6.36) el operador de desplazamiento modela la mezcla del estado coherente (6.22) con el oscilador local, mientras que la suma sobre los estados de Fock  $|n\rangle \langle n|$  modela un detector de fotones con resolución  $m$  [6].

Suponiendo que tras el primer paso  $l = 1$ , el valor  $n_1$  es observado, la función de la distribución a posteriori resultante es

$$p(\theta | n_1; \beta_0) \propto \mathcal{L}(n_1 | \theta; \beta_0) p(\theta), \quad (6.37)$$

donde  $\mathcal{L}(\theta | n_1; \beta_0)$  es la función de verosimilitud dada por

$$\mathcal{L}(\theta | n_1; \beta_0) = p(n_1 | \theta; \beta_0) = \text{Tr} \left[ D(\beta_0) |n_1\rangle \langle n_1| D^\dagger(\beta_0) \rho_\theta \right]. \quad (6.38)$$

La estimación de fase para este paso es  $\theta_1 = \hat{\theta}_1(n_1)$  obtenida con el  $\arg \max_{\theta \in \Theta} p(\theta | n_1; \beta_0)$  (estimador MAP). Utilizando esta distribución a posteriori (6.37) como la nueva apriori para el siguiente paso  $l = 2$ , se optimiza una función de costo  $U(\beta)$  para obtener el siguiente valor de  $\beta$  e implementar el POVM (6.36) con este  $\beta$ . Para el paso  $l \geq 2$ , se implemente el POVM

$$\left\{ D(\beta_{l-1}) |n\rangle \langle n| D^\dagger(\beta_{l-1}) \right\}_{n=0}^{m-1} \cup \left\{ \mathbb{I} - \sum_{n=0}^{m-1} D(\beta_{l-1}) |n\rangle \langle n| D^\dagger(\beta_{l-1}) \right\} \quad (6.39)$$

sobre el estado  $|e^{i\phi}\alpha/\sqrt{L}\rangle \langle e^{i\phi}\alpha/\sqrt{L}|$ , obteniendo una distribución a posteriori con densidad

$$\begin{aligned} p(\theta | \vec{n}, \vec{\beta}) &= p(\theta | n_l, \dots, n_1, \beta_{l-1}, \dots, \beta_0) \\ &\propto \prod_{i=1}^l p(n_i | \theta, \beta_i - 1) p(\theta), \end{aligned} \quad (6.40)$$

con  $n_i$  el dato observado en el paso  $i$ , produciendo, de esta manera, el  $l$ -ésimo estimador  $\hat{\theta}_l$ . Repitiendo este proceso hasta los  $L$  pasos adaptativos en cuestión.

Esta estrategia de estimación es un caso particular del diseño óptimo Bayesiano de experimentos, donde los parámetros  $\beta \in \mathbb{C}$  son los diseños y con el objetivo de estimar la fase  $\theta \in \Theta = [0, 2\pi)$ . Como el valor óptimo para  $\beta$  en cada paso adaptativo depende de los datos previos, la función de costo a optimizar es una función de la distribución a posteriori. Entonces, utilizar una función de costo adecuada como la información mutua (6.33) o el ECM pueden producir una secuencia de estimadores  $\theta_n$  que aproxime al valor de  $\theta$  [95].

Aún más, como en este caso la fase  $\theta$  es un parámetro cíclico, la distribución a posteriori es una función  $2\pi$ -periódica, por lo que el estimador MAP  $\hat{\theta}$  es una variable aleatoria cíclica. Por tal motivo, es conveniente utilizar la varianza de Holevo (4.22) como medida de dispersión del estimador. Para estimadores insesgados, la varianza de Holevo se expresa como

$$\text{Var}_H [\hat{\theta}] = \frac{1}{|E [e^{i\hat{\theta}}]|^2} - 1 \quad (6.41)$$

y como el estimador MAP es insesgado en el límite asintótico [67], minimizar la incertidumbre acerca de  $\theta$  requiere la maximización de la cantidad  $E [e^{i\hat{\theta}}]$  en la ecuación (4.22), la cual es llamada “sharpness”. Por lo tanto, la función de costo adecuada para este protocolo de estimación es la sharpness promedio

$$\bar{S}(\beta, m) = \sum_{n=0}^m p(n) \left| \int_{\Theta} e^{i\theta} p(\theta | n, \beta) \right|^2. \quad (6.42)$$

Para garantizar la consistencia asintótica del estimador del esquema adaptativo, se deben satisfacer las condiciones de regularidad 6.3. Para la estimación de fase, las condición de regularidad 1 se satisface si  $\theta$  es un punto interior de  $[0, 2\pi)$ . Además, dado que la probabilidad

$$\begin{aligned} p(n | \phi; \beta) &= \text{Tr} \left[ \left| \frac{\alpha e^{i\phi}}{\sqrt{L}} \right\rangle \left\langle \frac{\alpha e^{i\phi}}{\sqrt{L}} \right| D(\beta) |n\rangle \langle n| D^\dagger(\beta) \right] \\ &= \frac{\exp \left( - \left| \frac{\alpha e^{i\phi}}{\sqrt{L}} - \beta \right|^2 \right) \left| \frac{\alpha e^{i\phi}}{\sqrt{L}} - \beta \right|^{2n}}{n!}, \quad \alpha \in \mathbb{R}_+, \beta \in \mathbb{C} \end{aligned} \quad (6.43)$$

está bien definida y es dos veces diferenciable, las condiciones 2, 4, 5 y 6 se satisfacen directamente. No obstante, la condición 3, en general no se cumple. En particular, si uno escoge  $\beta$  como el valor que maximiza el información mutua (6.33) o el sharpness promedio (6.42), la función de verosimilitud resultante puede tener en general dos máximos, esto es, una función de verosimilitud no identificable [60]. Entonces, no hay razón matemática para obtener un estimador consistente.

Para afrontar el reto de trabajar con funciones de verosimilitud no identificables, se pueden utilizar diseños con una relación fija entre la fase de  $\beta$  y la amplitud  $|\beta|$ , dada por  $\beta = f(\phi)e^{i\phi}$ , donde  $\phi = [0, 2\pi)$  y  $f(\theta)$  una función real. Entonces los diseños resultantes se obtienen de maximizar una función  $U$  que es función de  $\phi$ .

Para entender cómo esta estrategia resuelve el problema de no identificabilidad, reescribimos la ecuación (6.43) como

$$p(n | \theta; \beta = |\beta|e^{i\phi}) = \frac{e^{-\lambda}\lambda^n}{n!} \quad (6.44)$$

con  $\lambda = |\alpha|^2/L + |\beta|^2 - 2|\alpha||\beta| \cos(\theta - \phi) / \sqrt{L}$ . Para  $L$  pasos adaptativos con resultados  $n_i$ ,  $i = 1, \dots, L$ , la función de verosimilitud es el producto de  $L$  funciones de probabilidad dadas por la ecuación (6.44):

$$\mathcal{L}(\theta) = \prod_{i=1}^L p(n_i | \theta; \beta_i = |\beta|e^{i\phi_i}) = \prod_{i=1}^L \frac{e^{-\lambda_i}\lambda_i^{n_i}}{n_i!}. \quad (6.45)$$

Cada  $\phi_i$  depende de la función de costo y las salidas del POVM previo. La elección de los diseños experimentales con  $|\beta| = f(\phi)$  puede forzar a la estrategia a cambiar  $\phi_i$  en cada paso adaptativo. En este caso, la función de verosimilitud (6.45) se vuelve identificable ya que el producto de las funciones de probabilidad con diferentes  $\phi_i$  produce una función con un solo máximo global. Para observar este hecho notemos que si  $\phi_i = \phi$  es fijo, incluso si el parámetro  $|\beta|$  es diferente en cada paso adaptativo del protocolo, la función de verosimilitud producida por una secuencia de variables aleatorias independientes con función de probabilidad (6.44) tiene dos máximos sobre  $\Theta$ . Uno de estos máximos siempre estará alrededor de  $\theta$  y el otro dependerá del valor de  $\phi$ . La estrategia permite escoger  $\phi$  en cada paso adaptativo y hacer que el máximo alrededor de  $\theta$  supere cualquier otro máximo de la función, pues la verosimilitud será el producto de diferentes funciones de probabilidad como función de  $\theta$  con segundos máximos a diferentes posiciones  $\phi_i \neq \phi_j$ , ( $i \neq j$ ). Como resultado, tomando diseños experimentales  $\beta = f(\phi)e^{-i\phi}$ , la verosimilitud de la estrategia satisface todas las condiciones de regularidad 6.3.

#### 6.4.1. Correcta elección del diseño óptimo

Como el objetivo final de la estimación de parámetros es alcanzar la CICRC, la elección natural para  $|\beta|$  es aquella que maximice la información de Fisher. Entonces, de acuerdo a la ecuación (2.44), la información de Fisher para un diseño particular  $\beta = |\beta|e^{i\phi}$  está dada por

$$I(\theta; \beta) = \frac{4|\alpha|^2|\beta \sin^2(\theta - \phi)|^2}{|\alpha|^2 + L|\beta|^2 - 2|\alpha||\beta| \cos(\theta - \phi)\sqrt{L}}. \quad (6.46)$$

A consecuencia, al optimizar sobre  $\beta$  se encuentra que

$$I(\theta, \beta_{opt}) = 4|\alpha|^2/L, \quad (6.47)$$

donde

$$\beta_{opt} = \frac{|\alpha|}{\sqrt{L} \cos(\theta - \phi)} \quad (6.48)$$

es el valor de  $\beta$  que maximiza la información de Fisher.

De esta forma, al utilizar el método adaptativo en cuestión, con  $\beta$  como en la ecuación (6.48), se puede garantizar todas las condiciones de regularidad 6.3 y por lo tanto, el estimador de máxima a posteriori es un estimador asintóticamente consistente (con respecto al número de pasos adaptativos) con varianza dada por (6.35). Aún más, como

$$I_Q \geq pI(\theta, \beta_1) + (1 - p)I(\theta, \beta_2), \quad \beta_1, \beta_2 \in \mathcal{D} \text{ y } 0 \leq p \leq 1, \quad (6.49)$$

entonces la IFC es el supremo de la envolvente convexa del conjunto de informaciones de Fisher, por lo tanto se tiene que

$$\sigma_{INFO}^2 = \frac{1}{I_Q} = \frac{1}{4|\alpha|^2}. \quad (6.50)$$

Concluyendo que el estimador MAP puede saturar la desigualdad de cuántica de Cramér-Rao (6.23) cuando el número de pasos adaptativos tiende a infinito.

Desafortunadamente, como  $\beta_{opt}$  depende de  $\theta$  su implementación es inviable ya que se necesita a priori conocer el parámetro que se desea estimar. Para tratar este problema se puede aproximar a  $\beta_{opt}$  sustituyendo  $\theta$  por su estimador  $\hat{\theta}$

$$\hat{\beta}_{opt} = \frac{|\alpha|}{\sqrt{L} \cos(\hat{\theta} - \phi)}. \quad (6.51)$$

De esta forma, el diseño óptimo propuesto utiliza la fase  $\phi$  como parámetro de control para minimizar la incertidumbre y al optimizar la información mutua o el sharpness promedio se obtienen funciones de verosimilitud identificables.

Como en general las estimaciones de  $\hat{\theta}$  solo son aproximaciones al valor puntual de  $\theta$ , es de esperar que la información de Fisher de este diseño sea menor que la información de Fisher para  $\beta_{opt}$ . Definiendo  $\Delta = \theta - \phi$ , la información de Fisher para  $\hat{\beta}_{opt}$  se puede escribir como

$$I(\theta, \hat{\beta}_{opt}) = \frac{4 \sin^2 \Delta \alpha^2}{L(\cos^2(\delta + \Delta) - 2 \cos \Delta \cos(\delta + \Delta) + 1)}, \quad (6.52)$$

donde  $\delta = \hat{\theta} - \theta$ . Por tanto,  $I(\theta, \hat{\beta}_{opt})$  es una variable aleatoria que depende de  $\hat{\theta}$  a través de  $\hat{\beta}_{opt}$ .

Cuando  $\delta \neq 0$ , la información de Fisher depende fuertemente de  $\phi$ , por ejemplo, si  $\phi \rightarrow \theta$ ,  $I(\Delta = 0, \hat{\beta}_{opt}) \rightarrow 0$ , por lo que no se gana información cuando el sistema es medido. Para encontrar el valor óptimo de  $\phi$  se busca el valor de  $\Delta$  que maximiza el valor esperado de  $F(\Delta, \hat{\beta}_{opt})$ . Tomando en consideración que  $E[\delta^{2n+1}] = 0$ , con  $n \in \mathbb{N}$ , se encuentra que  $\Delta = \pi/2$ , por lo que una buena estrategia adaptativa debe tener un  $\Delta = \theta - \phi$  cercano a  $\pi/2$  en el  $L$ -ésimo paso adaptativo. Sin embargo, en el límite de  $\Delta \rightarrow \frac{\pi}{2}$  y  $\delta \rightarrow 0$ , se tiene que  $|\hat{\beta}_{opt}| \rightarrow \infty$ , haciendo a la medida inviable ya que cualquier detector de fotones realista no podrá detectar nada. De esta manera es razonable esperar una tendencia de  $\Delta$  a  $\pi/2$  conforme  $L$  tiende a infinito de modo que la intensidad del pulso de luz resultante obliga a aumentar la resolución en el número de fotones ([photon number resolution \(PNR\)](#)) del detector con el fin de poder capturar toda la información. Por lo cual, dado una resolución de fotones  $m < \infty$  existe un límite en el número de pasos adaptativos para el cual la medida adaptativa es útil.

Por otro lado, si se realizan  $L$  mediciones utilizando el diseño óptimo  $\beta_{opt}$ , la información de Fisher total será  $4|\alpha|$ , es decir, la información de Fisher cuántica. Sin embargo, al no conocer  $\beta_{opt}$



y utilizar  $\hat{\beta}_{opt}$  en su lugar, podemos ver qué tan diferente es su información de Fisher total para  $L$  pasos adaptativos desarrollando a segundo orden en  $\delta$  la ecuación (6.52)

$$I(\theta, \hat{\beta}_{opt}) \approx 4|\alpha|^2(1 - \delta^2), \quad (6.53)$$

observando que  $I(\theta, \hat{\beta}_{opt}) < I(\theta, \beta_{opt})$  e independiente de  $L$ . Por lo que, cualquier esquema adaptativo de estimación tendrá Información de Fisher menor a la información de Fisher cuántica, puesto que en cada paso adaptativo el error de la estimación de  $\theta$  se propaga al elegir el diseño óptimo y en consecuencia, el esquema adaptativo no-Gaussiano no satura la CICRC para el régimen de bajas energías  $|\alpha| \ll \infty$ .

Finalmente notemos que aunque  $\beta_{opt}$  tiene la ventaja de forzar a que la función de verosimilitud sea identificable en cada paso adaptativo, este diseño no necesariamente maximiza la función de costo  $U(\beta)$  en todo el plano complejo  $\mathbb{C}$ . No obstante, cuando el número de pasos adaptativos  $L$  tiende a infinito,  $\hat{\beta}_{opt}$  tiende al valor que maximiza a  $U$  obteniendo un diseño óptimo global [31]. En el resto del capítulo, se hará un análisis exhaustivo del método adaptativo con la finalidad de encontrar como actúa este en función de  $|\alpha|^2$  y se comparará con el desempeño de la medición canónica (6.25) y MK II (6.26).

## 6.5. Desempeño del esquema adaptativo no-Gaussiano

Con el objetivo de analizar la tendencia hacia el desempeño de la medida canónica de fase, numéricamente se estudió el esquema adaptativo no-Gaussiano para la estimación de fase como función del número de pasos adaptativos  $L$  y de la media de fotones  $|\alpha|^2$  en el estado coherente. Los resultados se comparan directamente con la mejor estrategia Gaussiana conocida MK II y el error producido por la medida canónica de fase. En concreto, como resultado original se resuelven las siguientes preguntas: ¿Es el esquema adaptativo no-Gaussiano mejor que MK II? y ¿Qué tan cercano es el desempeño de la estrategia adaptativa al error producido por la medida canónica de fase?. Para las simulaciones numéricas se utilizó la sharpness promedio como función de costo pues esta requiere para su cálculo un número menor de operaciones en comparación con las funciones de costo (6.31), (6.32) y (6.33), aunque la estrategia de estimación de fase se puede implementar de forma equivalente utilizando cualquiera de las funciones de costo (6.31), (6.32), (6.33).

La Fig. 6.1-A representa la varianza de Holevo para el método adaptativo como función del número de pasos adaptativos  $L$  para un estado coherente con media de fotones  $|\alpha|^2 = 1$  y un detector con diferentes valores en la resolución del número de fotones ( $m = 1, m = 3, m = 6$ ). Para  $m = 1$ , la estrategia adaptativa supera al desempeño de la estrategia MK II (línea discontinua roja) a partir de  $L \gtrsim 30$ . Además, la estrategia con  $m = 3$  y  $m = 6$  tienen una mejor razón de convergencia superando al desempeño de MK II para  $L \gtrsim 20$  y  $L \gtrsim 15$ , respectivamente, obteniendo una menor varianza.

Para estimar la varianza de Holevo de la estrategia de estimación adaptativa no-Gaussiana en el límite asintótico, se ajustaron los resultados de la simulación numérica a modelo exponencial

$$\hat{y} = Ae^{-BL} + C. \quad (6.54)$$

Del análisis de regresión se encontró para  $m = 1$  un coeficiente  $C = 0.751 \pm 0.002$ , para  $m = 3$  un coeficiente  $C = 0.719 \pm 0.004$  y para  $m = 6$  un coeficiente  $C = 0.714 \pm 0.003$ . Por siguiente, dado que el coeficiente  $C$  en (6.54) es mayor que la varianza de la medición canónica (0.6727) en



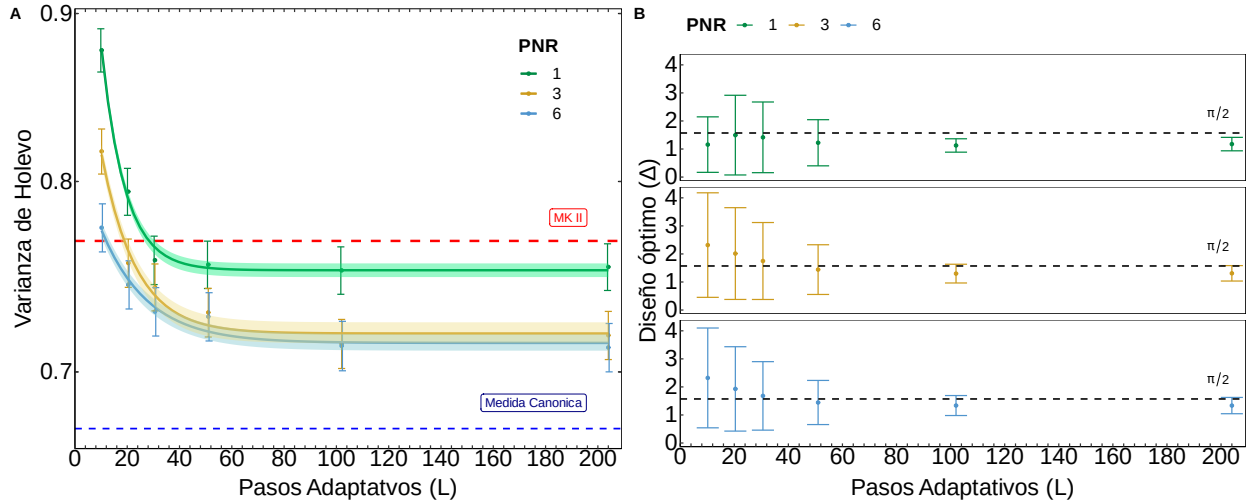


Figura 6.1: Panel A: Varianza de Holevo para  $\alpha = 1$  como función del número de pasos adaptativos  $L$  y distintos valores de resolución de fotones  $m$ . Conforme  $L$  aumenta, la precisión de las estimaciones mejora. La estrategia no-Gaussiana (curvas continuas) supera el desempeño de la estrategia MK II (línea discontinua roja a  $y = 0.767$ ) con  $L > 30$  en el caso de  $m = 1$  (puntos y curva de color verde), para  $L > 20$  en el caso de  $m = 3$  (puntos y curva de color amarillo) y  $L > 15$  en el caso de  $m = 6$  (puntos y curva de color azul). Las tres curvas se ajustaron a un modelo exponencial (6.54), la parte sombreada representa un desviación estándar. Todos los puntos de la simulación fueron calculados a partir de un bootstrap de  $10^4$  puntos. Panel B: Valor del diseño óptimo  $\Delta$  como función del número de pasos adaptativos  $L$  para  $\alpha = 1$  y diferentes valores de la resolución de fotones (PNR)  $m = 1, 3, 6$ . Si  $L$  aumenta su valor, el diseño óptimo se acerca  $\frac{\pi}{2}$ .

todos los casos, se concluye que incluso en el caso cuando el número de pasos tienda a infinito, el método no puede alcanzar el límite canónico cuando  $|\alpha| = 1$  y en consecuencia para cualesquiera valores mayores de energía. Cabe destacar que mientras el valor de  $C$  para el caso de  $m = 6$  es mayor que el caso de  $m = 3$ , debido a las incertidumbres de los parámetros no hay diferencia estadística entre estos. Este hecho permite extrapolar esta conclusión para valores mayores de  $m$ .

La Fig. 6.1-B muestra la fase óptima del oscilador local como función de  $L$ . Esta fase es el diseño óptimo  $\Delta$  introducido en la sección 6.4.1. De la Fig. 6.1-B se observa que al incrementar el número de pasos adaptativos, el diseño óptimo tiende a  $\frac{\pi}{2}$ . Este resultado es de esperarse de acuerdo a las predicciones de nuestro modelo teórico.

El estudio de la estrategia adaptativa no Gaussiana basada en conteo de fotones y retroalimentación se prosiguió con la exploración de su desempeño como función de  $|\alpha|^2$  y diferentes pasos adaptativos. Los resultados numéricos de este análisis se presentan en la Fig. (6.2). Cabe enfatizar que en el límite de altas energías  $|\alpha| \rightarrow \infty$  y con  $L = 100$ , la varianza para el esquema Bayesiano no-Gaussiano parece tener una tendencia hacia la cota cuántica de Crámer-Rao.

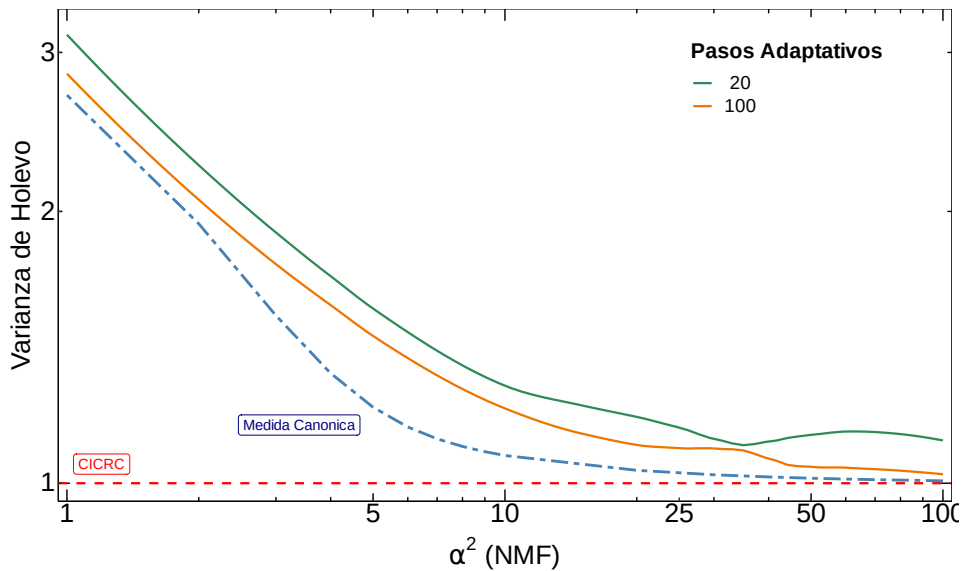


Figura 6.2: Varianza de Holevo como función de  $|\alpha|^2$  de la estrategia adaptativa no-Gaussiana para diferente número total de pasos adaptativos  $L = 20$  y  $L = 100$ . Las curvas sólidas muestran una tendencia descendente de la varianza cuando el número de pasos adaptativos se incrementa. En todos los casos se calcularon las curvas con un bootstrap de  $10^4$  puntos y con PNR  $m = 3$ . El desempeño de la medida canónica de fase (curva azul discontinua) se muestra como referencia.

### 6.5.1. Modelo para el rendimiento asintótico de la estrategia adaptativa

En el límite asintótico, el rendimiento de la estrategia adaptativa no-Gaussiana se puede describir razonablemente por tres potenciales modelos:

$$\text{Modelo 1: } \frac{A_1}{|\alpha|^2} + \frac{A_2}{|\alpha|^3} + \frac{A_3}{|\alpha|^4}, \quad (6.55a)$$

$$\text{Modelo 2: } \frac{A_1}{|\alpha|^2} + \frac{A_2}{|\alpha|^3}, \quad (6.55b)$$

$$\text{Modelo 3: } \frac{A_1}{|\alpha|^2} + \frac{A_3}{|\alpha|^4}, \quad (6.55c)$$

entonces, la cuestión es aplicar un criterio para decidir cual es el modelo que mejor describe el comportamiento de la varianza de la estrategia de estimación de fase. Para afrontar este problema se aplicó la técnica de “backward elimination”. Esta técnica consiste en los siguientes pasos [122]:

- i) Seleccionar un nivel de significancia  $\alpha$  (por ejemplo  $\alpha = 0.05$ ), el cual representará la probabilidad de tomar la decisión de rechazar la hipótesis nula cuando ésta es verdadera en el contraste de hipótesis del paso *iii*).
- ii) Ajustar los datos con el modelo con mayor cantidad de coeficientes.
- iii) Realizar una prueba de contraste de hipótesis entre la hipótesis nula  $H_0 : A_i = 0$  y la hipótesis alterna  $H_1 : A_i \neq 0$  para cada uno de los coeficientes  $A_i$  del modelo y registrar su valor  $p$  (la probabilidad de haber obtenido el resultado observado suponiendo que  $H_0$  es cierta).
- iv) Considerar el coeficiente con valor  $p$  más alto, si este es mayor al nivel de significancia  $\alpha$ , se prosigue al siguiente paso, de otro modo, se sigue con el paso *vii*).
- v) Se elimina ese coeficiente del modelo.
- vi) Ajustar el modelo sin ese coeficiente.
- vii) Se repite el proceso desde el paso *iii*) hasta llegar a un modelo donde el valor  $p$  de todos los coeficientes sea menor que el nivel de significancia  $\alpha$ .

Para esta técnica, el valor  $p$  de la prueba de hipótesis es el criterio para decidir qué coeficientes del modelo más general son los más adecuados para ajustar el comportamiento de los datos observados. El valor  $p$  se interpreta como la probabilidad de observar los datos obtenidos suponiendo que la hipótesis nula  $H_0$  es verdadera [29, 122]. Como el contraste de hipótesis es entre las premisas  $H_0 : A_i = 0$  y  $H_1 : A_i \neq 0$ , el hecho de que un coeficiente  $A_i$  tenga un valor  $p$  menor que el nivel de significancia  $\alpha$ , se entiende como el hecho que los datos observados proveen suficiente evidencia en contra de la hipótesis nula  $H_0$ , por lo que se rechaza esta hipótesis a favor de la hipótesis alterna  $H_1$ . Por el contrario, cuando el valor  $p$  es mayor que el nivel de significancia, el resultado se interpreta como el hecho que los datos proveen suficiente evidencia en contra de la hipótesis alterna  $H_1$ , por lo que se no se rechaza la hipótesis nula  $H_0$ . En otras palabras, si  $p < 0.05$  en algún contraste de hipótesis para algún coeficiente  $A_i$ , se puede concluir que los datos observados son evidencia estadísticamente significativa para suponer que  $A_i \neq 0$ .

De la tabla 6.1 se observa que  $p > 0.1$  para  $A_3$  y  $p < 0.001$  para los coeficientes  $A_1$  y  $A_2$ , para cualquier valor de  $L$ , por lo que se concluye que el modelo que mejor describe el comportamiento de los datos es  $A_1/|\alpha|^2 + A_3/|\alpha|^4$ . Además, se corroboró este resultado aplicando el criterio de

$L$	p-value de $A_1$	p-value de $A_2$	p-value de $A_3$
20	1.14e-21	0.41	4.10e-10
30	7.04e-21	0.65	3.84e-10
40	9.67e-20	0.62	7.17e-09
50	5.29e-21	0.53	9.55e-10
100	2.00e-16	0.11	1.78e-09

Cuadro 6.1: p-value para los diferentes coeficientes del ajuste del Modelo 2 (6.55) para diferentes valores de  $L$ .

información de Akaike, el cual es una métrica de la calidad relativa de un modelo de regresión, para un conjunto de datos [122]. La métrica de Akaike se aplicó a cada uno de los modelos expuestos en (6.55) obteniendo la misma conclusión que al utilizar “Backward elimination”. Con el resultado previo, se prosiguió a ajustar las curvas de la Fig. (6.2) al modelo 3 de (6.55) y se realizó un estudio de los coeficientes  $A_1$  y  $A_3$  como función del número de pasos adaptativos con el objetivo de obtener la tendencia del método de estimación en el límite asintótico ( $L \rightarrow \infty$ ). Este análisis se presenta en la Fig. 6.3.

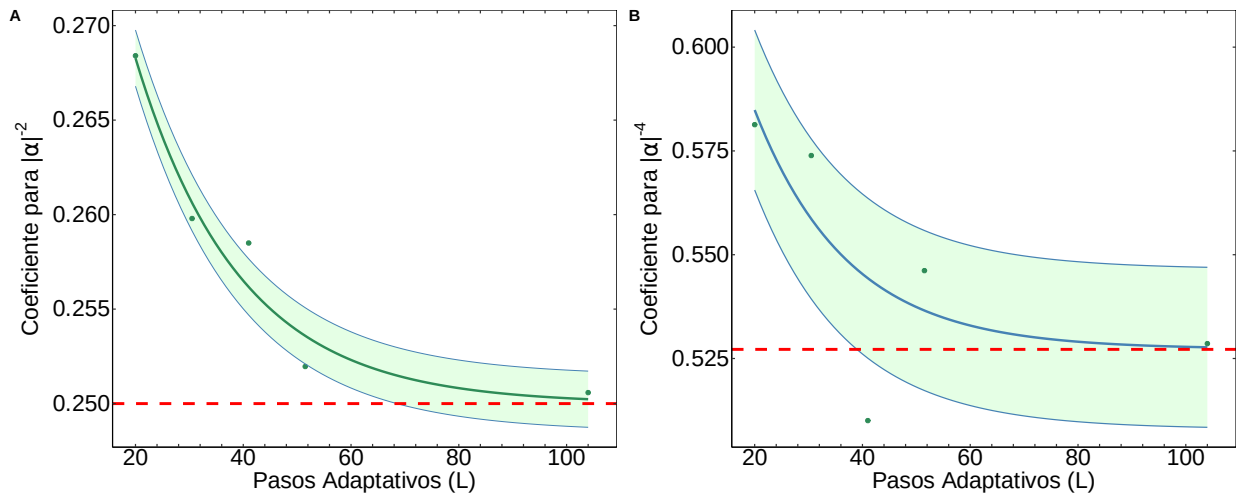


Figura 6.3: Coeficientes del ajuste modelo 3 (6.55) para la varianza del método como función del número de pasos adaptativos. En el panel izquierdo, se muestra el coeficiente  $A_1$  y en el panel derecho el coeficiente  $A_3$ . Las curvas muestran la tendencia  $A_1 \rightarrow 0.25$  y  $A_3 \rightarrow 0.53$  cuando el número de pasos adaptativos tiende a infinito. Ambas curvas se ajustaron a un modelo exponencial (6.54), con coeficientes  $(A, B, C, EER) = (0.0541 \pm 0.01975, 0.0542 \pm 0.02044, 0.2499 \pm 0.00270, 9.72 \times 10^{-5})$  para los datos del panel izquierdo, y  $(A, B, C, EER) = (0.0607 \pm 0.08782, 0.1942 \pm 0.30907, 0.5272 \pm 0.03063, 4.369 \times 10^{-3})$  para los datos del panel derecho. La región sombreada representa una desviación estándar.

Para extrapolar el comportamiento del escalamiento en la varianza cuando el número de pasos adaptativos tiende a infinito, se ajustaron los puntos de las curvas de la Fig. 6.3 a un modelo

exponencial (6.54). Para el ajuste de la curva del coeficiente  $A_1$  (panel izquierdo de Fig. 6.3) se encontró  $C = 0.2499 \pm 0.00270 \approx 1/4$  y para el ajuste de la curva del coeficiente  $A_3$  (panel derecho de Fig. 6.3) se encontró  $C = 0.5272 \pm 0.03063 \approx 0.53$ . Por tanto, la varianza (en el caso ideal) tiene un rendimiento que va como

$$\text{Var}_H [\hat{\phi}] \sim \frac{1}{4|\alpha|^2} + \frac{0.53}{|\alpha|^4}. \quad (6.56)$$

La ecuación (6.56) es uno de los resultados principales de este trabajo. El segundo término en la varianza es del orden  $\frac{1}{|\alpha|^4}$ , el mismo que el POVM canónico de fase (ver (6.27)), mientras que la estrategia MK II es del orden  $\frac{1}{|\alpha|^3}$  (ver (6.26)). La diferencia, para valores grandes de  $|\alpha|$ , entre el POVM canónico y la estrategia no-Gaussiana es el coeficiente en el segundo término, el cual es  $1/2$  en vez de  $5/32$ . Esto muestra que la estrategia Bayesiana adaptativa basada en conteo de fotones para estimar la fase produce un error de estimación más cercano a los límites físicos teóricos que la mejor estrategia Gaussiana conocida MK II. De igual manera, el número de pasos adaptativos  $L$  para alcanzar el comportamiento descrito por (6.56) es función de la resolución  $m$  del detector. De acuerdo con la discusión de la Fig. 6.1 es de esperar que a mayor valor de  $m$ , menor es el número de pasos adaptativos necesarios para llegar al desempeño asintótico descrito por el modelo (6.56). Para ejemplificar este hecho, en la Fig. 6.3) se presenta la varianza de Holevo para la estrategia no-Gaussiana a  $L = 100$ , como función de  $|\alpha|^2$  y diferentes valores de PNR,  $m = 1, 3, 6$ .

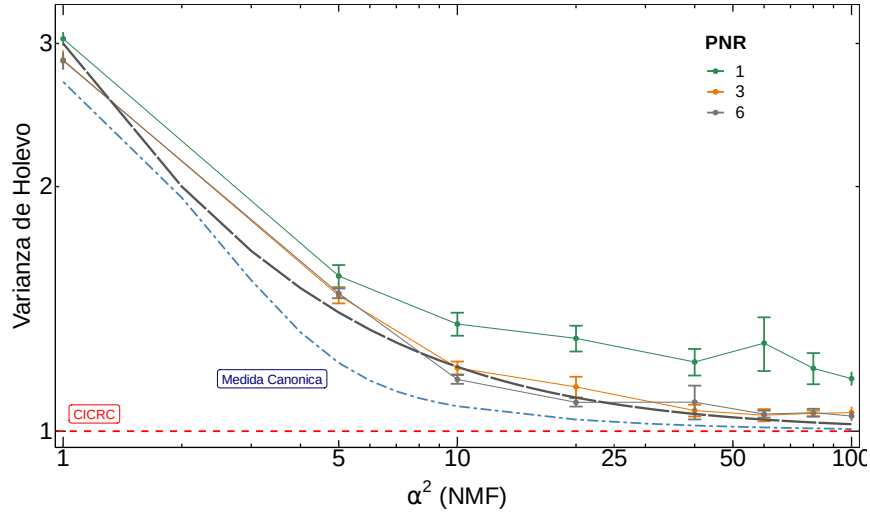


Figura 6.4: Varianza de Holevo como función de  $|\alpha|^2$  de la estrategia adaptativa no-Gaussiana para  $L = 100$  y diferente resolución del número de fotones  $m = 1, 3, 6$ . A mayor valor de  $m$  menor es el valor de la varianza de Holevo. En todos los casos se calcularon las curvas con un bootstrap de  $10^4$  puntos. El desempeño de la medida canónica de fase (curva azul discontinua) y el modelo ideal (6.56) se muestran como referencia.



## Capítulo 7

# Sensores con escalamiento Super-Heisenberg

El objetivo final de la teoría de estimación de parámetros en sistemas cuánticos es obtener la mayor precisión en las estimaciones a partir de mediciones del sistema. En estimación escalar, el límite fundamental en la precisión está dado por la CICRC (3.13). Como ya se discutió, alcanzar este límite es posible cuando se realiza un número que tiende a infinito de mediciones independientes con un POVM óptimo que produce un modelo estadístico regular. No obstante, un supuesto más realista es considerar el hecho que se tiene un número limitado de recursos tales como un número finito de copias del sistema, energía fija o tiempo escaso para realizar las mediciones experimentales. De esta manera, llegar a los límites asintóticos es imposible. Un ejemplo de esto fue el problema de estimación de fase de un solo disparo donde el error de los estimadores es función del número medio de fotones  $|\alpha|^2$  y sólo satura la CICRC para el caso cuando  $|\alpha| \rightarrow \infty$ .

Bajo la suposición de tener recursos limitados, el objetivo es encontrar estrategias de estimación que utilicen los recursos del sistema de la forma más eficiente posible. Con la finalidad de estudiar a las estrategias de estimación como función de los recursos del sistema, se han creado clasificaciones para las cotas de Crámer-Rao dependiendo de como estas escalan en función de los recursos del sistema. De este modo, dada una estrategia de estimación, el enfoque usual que se adopta es el de buscar condiciones del sistema donde se obtengan los mejores escalamientos de la CICRC. Como ejemplo, para el problema de estimación de fase en estados coherentes, el escalamiento de la CICRC va como el inverso del cuadrado de los recursos del sistema, este escalamiento se conoce como el límite cuántico estándar y para el problema de estimación de fase en estados comprimidos (6.10), la CICRC tiene un escalamiento conocido como límite de Heisenberg que va como el inverso de una potencia cuarta de los recursos del sistema. Este capítulo comienza con una introducción a estos resultados y después se procede a la descripción de un circuito cuántico capaz de producir una CICRC con escalamiento superior al de Heisenberg llamada Super-Heisenberg, el cual es un resultado original.

### 7.1. Límite cuántico estándar y límite de Heisenberg

Para una gran clase de experimentos, sobre todo en el área de metrología cuántica e interferometría, encontrar límites sobre la escala de la precisión de estrategias de medición como función de los recursos del sistema toma una gran importancia [15]. Dependiendo del contexto, se pueden en-

contrar diferentes límites, como por ejemplo el límite estándar cuántico (SQL del inglés [standard quantum limit](#)) y el límite de Heisenberg [15, 59].

El límite cuántico estándar, aparece cuando dos variables satisfacen la relación de incertidumbre de Heisenberg. Suponiendo que el sistema se encuentra en un estado  $\rho_\theta$  que depende de un parámetro  $\theta \in \Theta$  y que este obtuvo su dependencia paramétrica por la evolución

$$\rho_\theta = e^{i\theta A} \rho e^{-i\theta A}, \quad (7.1)$$

donde  $\rho$  es un estado sonda inicial y  $A$  un operador auto adjunto en  $\mathcal{H}$  el espacio de Hilbert del sistema.  $A$  cumple la relación de incertidumbre de Heisenberg con otro observable  $M$  sobre  $\mathcal{H}$  si se satisface la siguiente desigualdad [57]

$$\text{Var}_\theta [M] \cdot \text{Var}_\theta [A] \geq \frac{1}{4} |\text{Tr} [\rho_\theta [A, M]]|^2. \quad (7.2)$$

En particular, para el estado (7.1), la desigualdad de Heisenberg se reescribe como

$$\text{Var}_\theta [M] \geq \frac{1}{4 \text{Var}_\theta [A]}, \quad (7.3)$$

donde  $\text{Var}_\rho [A] \equiv \text{Var}_\theta [A]$  ya que  $A$  conmuta con el grupo de transformaciones unitarias  $\{e^{i\theta A}\}_{\theta \in \Theta}$ . Tomando a  $M$  como una estrategia de estimación para  $\theta$ , se obtiene el SQL, una relación para las incertidumbres de  $A$  y  $M$ . Por otro lado, el límite de Heisenberg no es una relación de incertidumbre. Este límite es una relación entre la varianza de una estrategia de estimación  $M$  y el primer momento del generador de la dependencia paramétrica  $A$  [125],

$$\text{Var}_\theta [M] \geq \frac{1}{|\mathbb{E}_\theta [A]|^2}. \quad (7.4)$$

El límite de Heisenberg originalmente fue introducido en [59] en el contexto de estimación de fase con interferómetros. En este ámbito, producir muchas copias del sistema puede resultar bastante complicado, por lo que el problema de estimación de único disparo cobra mayor relevancia y tomar a la CICR como función de los recursos del sistema se vuelve un eje central a la hora de buscar una estrategia de estimación óptima.

Para ilustrar estos conceptos, consideremos nuevamente el problema de estimación de fase en estados coherentes, en este caso, se codifica una fase desconocida  $\theta$  con desfaseamiento *linear* vía el operador  $R(\theta) = \exp(-i\theta \hat{n})$ , donde  $\hat{n} = \hat{a}^\dagger \hat{a}$ . Es decir,  $A = \hat{n}$  en la ecuación (7.1). En interferometría, el valor esperado  $\mathbb{E}_\theta [\hat{n}] = |\alpha|^2 = \bar{n}$ , es identificado como el número promedio de fotones y considerado como el recurso del sistema [9, 41, 90]. De este modo,  $I_Q = 4\bar{n}$  y la precisión de las estrategias de estimación es acotada por el inverso de los recursos del sistema (SQL). En cambio, si el estado inicial es un estado comprimido del vacío  $\rho = |\phi_r\rangle \langle \phi_r|$  con  $|\phi_r\rangle = \hat{S}(r) |0\rangle$  y se continúa considerando a  $\bar{n}$  como el recurso del sistema, se obtiene que la información cuántica de Fisher es

$$I_Q[|\phi_r\rangle, \hat{n}] = 4(\text{Var}_{|\phi_r\rangle \langle \phi_r|} [\hat{n}]) = 8 \sinh(r)^2 \cosh(r)^2 = 8(\bar{n}^2 + \bar{n}), \quad (7.5)$$

obteniendo un mejor escalamiento como función de  $\bar{n}$ . En efecto, este escalamiento corresponde al límite de Heisenberg para la estimación de fase [93].

La definición del límite de Heisenberg ha causado a lo largo de la historia bastante controversia debido a que la definición de “recursos” tiene un grado subjetividad [78, 125]. Por ejemplo, en



[8, 70], se afirma que se puede superar el límite de Heisenberg para estados coherentes utilizando un generador *no lineal* de la forma  $\hat{n}^2$ . En esta situación, la fase  $\theta$  es codificada a partir de la transformación paramétrica  $\exp(-i\theta\hat{n}^2)$ , la cual produce la siguiente información de Fisher [102],

$$I_Q[|\alpha\rangle, \hat{n}^2] = 4(\text{Var}_\rho \hat{n}^2)|_\alpha^2 = 4(4\bar{n}^3 + 6\bar{n}^2 + \bar{n}), \quad (7.6)$$

observando que escala como  $\bar{n}^3$ , donde  $E_\theta[\hat{n}] = \bar{n}$ . Por consiguiente si se definen los recursos del sistema como  $\bar{n}$  se obtiene un escalamiento *super-Heisenberg* para la estimación de fase. Esta definición de escalamiento tipo Super-Heisenberg es habitualmente adoptada en trabajos de interferometría y detectores cuánticos y ha sido demostrada experimentalmente, usando interferómetros atómicos no lineales [107] y usando acoplamiento de muchos cuerpos en técnicas de NMR [87]. Aun así, esta forma de definir los recursos se contradice con la definición dada por [125], la cual toma los recursos como  $E_\theta[\hat{n}^2] = \bar{n}^2$ . De este modo, el límite de Heisenberg vendría dado por  $\frac{1}{\bar{n}^2}$ , por lo que desfases no lineales no sólo no superarían el límite de Heisenberg, sino que ni siquiera lo alcanzarían.

Fuera de esta controversia, en este capítulo adoptaremos el enfoque de los trabajos de interferometría, por lo que daremos el nombre de escalamiento de super-Heisenberg a escalas de la información de Fisher como función de  $\bar{n}$  con potencias mayores a 2. En el resto del capítulo se describirá una forma de alcanzar este tipo de escalamientos utilizando un circuito cuántico en el régimen óptico.

## 7.2. Protocolo para generar dinámica paramétrica no lineal

En los últimos años, se ha encontrado que los sistemas cuánticos no lineales pueden ser útiles para numerosas aplicaciones como el estudio de la decoherencia [124], mejoramiento de protocolos de metrología [5, 12, 36, 47, 70–73, 86, 102, 112, 120, 123], construcción de sensores cuánticos [7, 116] y encontrar ventajas para algoritmos de computación cuántica [2, 46, 84]. En particular, cuando se desea estimar un parámetro  $\theta$ , se ha probado que los sistemas no lineales pueden mejorar el escalamiento en la precisión de un estimador óptimo  $\hat{\theta}$ , superando (dependiendo de la definición de recursos del sistema) el llamado límite de Heisenberg. Sobre todo, los osciladores de Kerr, aquellos cuyo Hamiltoniano es proporcional al cuadrado del operador de número  $\hat{n}^2$ , han sido intensamente estudiados como generadores de dinámica no lineal sobre sistemas cuánticos. Sin embargo, generar este tipo de dinámica es en extremo complicado, ya que en la mayoría de los sistemas atómicos y ópticos, las no linealidades producidas conllevan demasiado ruido. Aunque recientemente, se han logrado generar estados cuánticos estables generados por osciladores tipo Kerr en el régimen de microondas [64], en el régimen óptico no se ha encontrado un mecanismo para obtener este tipo de estados de forma sencilla.

En este capítulo se mostrará cómo es posible generar dinámica no lineal sobre estados a partir de evoluciones tipo Kerr a partir de un estado óptico (el cual será llamado sonda) utilizando el resultado de medir un estado ancilla al cual se le aplicó una operación lineal. La medición sobre la ancilla se puede describir como un operador de Kraus sobre el sonda, la cual se puede descomponer en partes unitarias y no unitarias y la cual en cierto límite adecuado, aproxima a un generador de Kerr puro. Además se mostrará cómo este proceso puede ser usado para producir estimadores  $\hat{\theta}$  con  $\text{Var}(\hat{\theta}) \sim \frac{1}{\bar{n}^3}$ , donde  $\bar{n}$  es el número de fotones promedio del estado sonda, el cual es interpretado como el recurso del sistema.

El protocolo para generar dinámica no lineal para un estado sonda dentro del régimen óptico

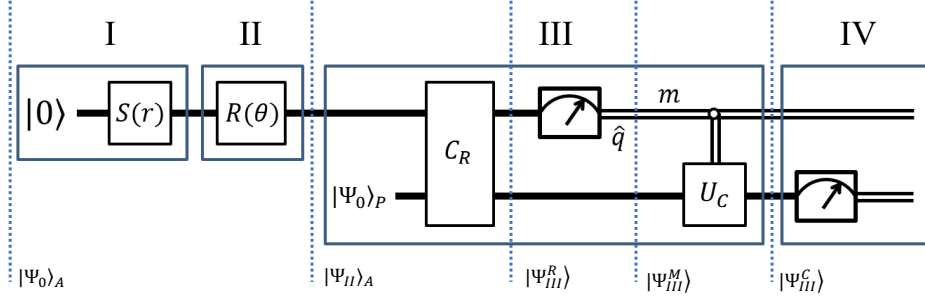


Figura 7.1: Protocolo: Ambas líneas del circuito representan modos bosónicos nombrados como ancilla (modo superior) y sonda (modo inferior). En la parte I se inicializa el modo ancilla en un estado vacío comprimido, donde  $S(r)$  es el operador de compresión. En la parte II se implementa una rotación de la ancilla por  $R(\theta)$ . En la parte III se implementa una “teleportación de Kerr” sobre el estado de entrada  $|\psi\rangle$  (el cual suponemos como un estado coherente), a través de una compuerta de rotación cruzada  $C_R$  (la cual rota el estado sonda una cantidad de momento dependiente de la ancilla, por ejemplo, generado por  $\hat{H}_R = \tilde{g} \hat{p}_A \otimes \hat{n}_P$ ), y una medición de la cuadratura de posición en el primer modo. Esto tiene como consecuencia la implementación de tanto una operación lineal  $\sim \hat{n}_P$  y una operación cuadrática  $\sim \chi \hat{n}_P^2$  sobre el modo sonda, donde la fuerza de Kerr  $\chi$  es función de  $r, \theta$  y  $g$ . Además  $U_C$  es un operador unitario dependiente de la salida de la medición del ancilla, realizada tras la conclusión del procedimiento de teleportación, lo cual permite *deshacer* las rotaciones lineales dejando solo la evolución tipo Kerr con un decaimiento. Finalmente, el estado sonda puede ser medido en la parte IV. Cada parte del circuito referida en el texto es separada en la figura por las líneas punteadas verticales.

está basado principalmente en la siguiente integral Gaussiana

$$\int_{-\infty}^{+\infty} dy e^{-ay^2+by} = \sqrt{\frac{\pi}{a}} e^{b^2/(4a)}, \quad (7.7)$$

la cual converge cuando  $\text{Re}(a) > 0$ . Esta integral tiene la propiedad de que el parámetro  $b$ , el cual inicialmente es lineal, termina siendo cuadrático en la exponencial del lado derecho de (7.7). De este modo, como los operadores unitarios pueden ser escritos como el operador exponencial de un generador Hermitiano, es posible usar esta propiedad para producir el cuadrado del generador. En particular, se puede hacer un *bootstrap* sobre el generador del oscilador armónico  $\hat{H}_{ho} \sim \hat{n}$  para convertirlo en un oscilador tipo Kerr  $\hat{H}_{ko} \sim \chi \hat{n}^2$ . Aunque este comportamiento solo podrá conseguirse de forma aproximada con recursos realistas [62].

El protocolo de *bootstrap* se puede describir por el circuito de la Fig. 7.1, el cual comienza con dos estados, un estado ancilla (superior) y otro llamado sonda (inferior). Dentro del circuito se realizan las siguientes operaciones

$$\hat{S}(r) = e^{-r(\hat{a}_A^2 - \hat{a}_A^{\dagger 2})/2}, \quad (7.8)$$

$$\hat{R}(\theta) = e^{i\theta \hat{n}_A}, \quad (7.9)$$

$$\hat{S}h(\beta) = e^{-i\beta \hat{p}_A^2}, \quad (7.10)$$

$$\hat{C}_R = e^{-ig \hat{p}_A \otimes \hat{n}_P}, \quad (7.11)$$

donde los subíndices  $A$  y  $P$ , denotan si los operadores están actuando sobre el modo ancilla o el modo sonda. El operador  $S(r)$  corresponde a un operador de compresión,  $R(\theta)$  a uno de rotación,  $Sh(\beta)$  a un cizallamiento [94, 114], y  $C_R$  a una compuerta de rotación cruzada.

De acuerdo con la Fig. 7.1, se comienza produciendo en el paso I y II el modo ancilla, al cual se le puede aplicar la operación de cizallamiento o una rotación. Por simplicidad en la descripción supondremos que se aplica  $\hat{S}h(\beta)$  en el paso II en vez que  $\hat{R}(\theta)$ , pero la idea general se puede extrapolar sin mayores complicaciones al caso de rotaciones, hecho que se discutirá posteriormente. De este modo, la ancilla tras finalizar el paso II se puede escribir como

$$\begin{aligned} |\psi_{II}\rangle_A &\equiv \hat{S}h(\beta) \hat{S}(r) |0\rangle_A \\ &= e^{-i\beta\hat{p}_A^2} \hat{S}(r) |0\rangle_A \\ &= \mathcal{N}_{II} \int dp e^{-\frac{p^2}{2\sigma^2}} e^{-i\beta p^2} |p\rangle_A \end{aligned} \quad (7.12)$$

donde  $|p\rangle_A$  es el eigenestado de  $\hat{p}_A$ , el operador de momento de la ancilla,  $\sigma^2 = e^{2r}$ , y  $\mathcal{N}_{II}$  es una constante de normalización. Posteriormente se introduce el estado sonda  $|\psi\rangle_P \equiv |\alpha\rangle_P$ , tomándolo como un estado coherente de magnitud  $\alpha$ , y se aplica el operador de rotación cruzada para obtener

$$|\Psi_{III}^R\rangle \equiv e^{-ig\hat{p}_A \otimes \hat{n}_P} |\psi_{II}\rangle_A \otimes |\alpha\rangle_P \quad (7.13)$$

$$= \mathcal{N}_{II} \int dp e^{-p^2 \left( \frac{1}{2\sigma^2} + i\beta \right)} e^{-igp\hat{n}_P} |p\rangle_A \otimes |\alpha\rangle_P. \quad (7.14)$$

En la parte III se aplica una medición  $\{|m\rangle \langle m|\}$ , donde  $|m\rangle$  es el eigenvector del operador hermitiano  $X \sim \hat{a}^\dagger \hat{a}$  sobre la ancilla, obteniendo un salida  $m$ . Al utilizar el hecho que  $\langle m|p\rangle_A = e^{imp}/\sqrt{2\pi}$ , se obtiene que el estado post-medición es

$$|\Psi_{III}^M\rangle = \mathcal{N}_{IIIM} \int dp e^{-p^2 \left( \frac{1}{2\sigma^2} + i\beta \right)} e^{ip(m-g\hat{n}_P)} |\alpha\rangle_P, \quad (7.15)$$

con  $\mathcal{N}_{IIIM} = 1/(\pi^{1/4} \sqrt{2\pi\sigma})$ . Ahora, integrando sobre la variable  $p$ , se llega al estado

$$|\Psi_{III}^M\rangle = \mathcal{N}'_{IIIM} \exp\left(-\frac{(g\hat{n} - m)^2}{4\left(\frac{1}{2\sigma^2} + i\beta\right)}\right) |\alpha\rangle_P \quad (7.16)$$

$$= \mathcal{N} \hat{U}(\beta) \hat{U}_c(\beta, m) \hat{K}(\beta) \hat{K}_c(\beta, m) |\alpha\rangle_P \quad (7.17)$$

donde  $\mathcal{N}$  es un factor de normalización.

Es importante notar que en la expresión (7.17) se descompuso al propagador  $\exp\left(-\frac{(g\hat{n}-m)^2}{4\left(\frac{1}{2\sigma^2} + i\beta\right)}\right)$  como el producto de cuatro operaciones. Estas incluyen dos operaciones unitarias  $\hat{U}$  y  $\hat{U}_c$  y dos no

unitarias  $\hat{K}$  y  $\hat{K}_c$ . Las cuales se pueden escribir explícitamente como

$$\hat{U}(\beta) = \exp\left(i\frac{\beta g^2}{\mu}\hat{n}^2\right) \quad (7.18)$$

$$\hat{K}(\beta) = \exp\left(-\frac{g^2}{2\sigma^2\mu}\hat{n}^2\right) \quad (7.19)$$

$$\hat{U}_c(\beta, m) = \exp\left(i\frac{2\beta mg}{\mu}\hat{n}\right) \quad (7.20)$$

$$\hat{K}_c(\beta, m) = \exp\left(-\frac{2mg}{2\sigma^2\mu}\hat{n}\right), \quad (7.21)$$

donde  $\mu = 4(\beta^2 + 1/(4\sigma^4))$ . De esta manera, la evolución del estado sonda consta de dos operaciones independientes del resultado de la medición sobre el ancilla,  $\hat{U}(\beta)$ , y  $\hat{K}(\beta)$ , y de dos operaciones dependientes del resultado de la medición  $m$  (las cuales llamaremos operaciones condicionadas),  $\hat{U}_c(\beta, m)$  y  $\hat{K}_c(\beta, m)$ .

Además, es importante resaltar que el operador unitario  $\hat{U}(\beta)$  es una evolución Kerr con fuerza  $\chi \equiv \beta g^2/\mu$ , la cual depende de  $\beta$  y  $\sigma$ , los parámetros de cizallamiento y compresión del ancilla. También es importante destacar que todos los operadores (7.18)–(7.21) conmutan independientemente del valor de los parámetros.

En el límite de compresión infinita  $\sigma \rightarrow \infty$ , se tiene que  $\mu \rightarrow 4\beta^2$ , por lo que los operadores no unitarios convergen al operador identidad,  $\chi \rightarrow g^2/4\beta$ , y  $\hat{U}_c$  pasa a ser una rotación dependiente de  $m$  con una fase de  $\phi = mg/2\beta$ . Entonces, en este límite, la evolución consiste solamente de una rotación no lineal independiente de  $m$  y una rotación lineal condicionada a  $m$ . En la sección final de la parte III de la Fig. 7.1 se puede, además, realizar una corrección unitaria sobre el estado sonda para lidiar con la dependencia de  $m$ .

Cabe destacar que aunque se utilizó el operador de cizallamiento  $\hat{S}h(\beta)$ , en vez del operador de rotación  $\hat{R}(\theta)$ , la descripción previa refleja la idea principal de cómo funciona el circuito. Sin embargo, si se desea obtener la descripción para una rotación, se debe reparametrizar la descripción anterior. Para encontrar esta nueva parametrización se utiliza la representación en el espacio Siegel para estados Gaussianos puros [81, 91]. En esta representación, los estados gaussianos son escritos como un número complejo  $z = v + iu$ , donde  $u > 0$  [81]. Entonces, como las matrices simplécticas correspondientes a la compresión y rotación pueden ser escritas como

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} e^{-r} & 0 \\ 0 & e^r \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (7.22)$$

se obtiene que

$$\frac{c + id}{a + ib} = z = v + iu, \quad (7.23)$$

donde  $u = 1/\Xi$ ,  $v = -\sin 2\theta \sinh 2r/\Xi$ , y  $\Xi = e^{-2r} \sin^2 \theta + e^{2r} \cos^2 \theta$ . De esta forma, uno puede expresar el estado que resulta de la parte II como

$$|\psi_{II}\rangle_A = \hat{R}(\theta)\hat{S}(r)|0\rangle \quad (7.24)$$

$$= \left(\frac{u}{\pi}\right)^{1/4} \int dp e^{-\frac{1}{2}p^2(u-iv)} |p\rangle_A, \quad (7.25)$$

por lo que repitiendo los pasos a partir de (7.12) – (7.17), se llega a

$$\frac{1}{2\sigma^2} = u, \quad (7.26)$$

$$\beta = -v. \quad (7.27)$$

Por lo tanto, al usar (7.26) y (7.27) uno puede obtener una expresión para los parámetros  $(\sigma, \beta) \rightarrow (r, \theta)$ . En la siguiente sección se explorará cómo utilizar este circuito para mejorar protocolos de estimación de parámetros.

### 7.3. Mejoría en la estimación de parámetros

Para el problema de estimación de un parámetro  $\theta \in \Theta \subseteq \mathbb{R}$ , el circuito 7.1 puede ser utilizado para producir un estimador  $\hat{\theta}$  cuya varianza escale como  $\frac{1}{\bar{n}_P^3}$ , donde  $\bar{n} = \mathbb{E}_\theta[\hat{n}]$  es el valor esperado de fotones calculado respecto al estado (sonda). Este escalamiento es mejor que el típico límite de Heisenberg, en el cual la varianza del estimador escala como  $\frac{1}{\bar{n}^2}$ . Cabe aclarar que estamos suponiendo que la ancilla tiene como un recurso fijo, el nivel de compresión y el recurso variable dentro del circuito es  $\bar{n}_P$ . En particular, este circuito puede ser usado para estimar el parámetro de la rotación (una fase) o el parámetro del operador de cizallamiento (un desplazamiento).

Suponiendo que se conocen los parámetros de compresión  $(r, g, m, \alpha_P)$  correspondientes a la fuerza de la compuerta de rotación cruzada  $C_R$ , salida de la medición  $m$ , y el parámetro  $\alpha_P$  que describe el estado sonda, el protocolo de estimación para  $\theta$  consiste en obtener un  $\rho_\theta(m) = |\Psi_{III}^M\rangle\langle\Psi_{III}^M|$  utilizando el circuito de la Fig. 7.1, aplicar un POVM óptimo sobre este estado y aplicar sobre los datos un estimador adecuado.

En especial, se puede sacar provecho de esta idea si además, se supone que el parámetro  $\theta$  que se desea estimar es un pequeño incremento de un valor base  $\theta_0$ , es decir,  $\theta = \delta\theta$  donde  $\theta_0 + \delta\theta \in \Theta$ . Ya que esta suposición permite producir con cierta facilidad modelos estadísticos regulares y abre la posibilidad de aplicar una rotación de corrección  $\hat{U}_c(f(r, g, \theta), m)$  al final del paso III del circuito 7.1, donde  $f$  es una función independiente de  $m$ . Esta corrección reduce el efecto estocástico sobre el estado sonda. Esta suposición sobre  $\theta$  aparece frecuentemente en trabajos de sensores cuánticos [39, 115].

El problema que conlleva este protocolo de estimación es que cada vez que uno repite el proceso, al ser  $m$  la salida de una variable aleatoria, se obtiene un estado  $\rho_\theta(m)$ , por lo que no solo encontramos información acerca de  $\theta$  en las medidas que se realizan al final sobre  $\rho_\theta(m)$ , sino también en los resultados  $m$  de la medición sobre el ancilla. Por tal motivo es necesario extender el concepto de la información de Fisher cuántica a uno que englobe el caso de tener estados condicionados a las salidas de otra variable aleatoria que también depende del parámetro a estimar, que en nuestro caso son conjunto de salidas producidas por la medida en la ancilla y las mediciones sobre el estado sonda al final del circuito 7.1 utilizadas para estimar  $\theta$ .

### 7.4. Información cuántica de Fisher para un ensamble post-medición

En los ejemplos de la subsección previa se consideraron las informaciones de Fisher cuánticas acerca de un parámetro  $\theta$ , el cual se codificó a un estado vía una evolución unitaria. No obstante, cada vez que se realiza el esquema descrito en la Fig. 7.1, se obtienen como resultados:  $(m(\theta), \rho_\theta(m))$ ,

donde  $m$  es la salida de la medición al ancilla y  $\rho_\theta(m)$  es el estado correspondiente al estado sonda condicionado al resultado de la medición. Ahora, si se desea estimar  $\theta$  a partir de un conjunto  $N$  de mediciones, se deberá aplicar el circuito  $N$  veces, pero como el resultado de la medición  $m$  es aleatorio, se obtendrá un ensamble de  $N$  diferentes estados  $\{\rho_\theta(m_i)\}_{i=1}^N$ , con  $m_i$  uno de los posibles resultados de la medición de la ancilla, de tal manera que la evolución promedio sobre el estado sonda será no unitaria. Por tal motivo, para calcular la desigualdad de Crámer-Rao de este caso, se debe generalizar el concepto de información de Fisher cuántica. Con este fin, se aplicará el enfoque descrito en [76].

En [76] se considera una extensión del sistema que consiste en el sistema de interés (el estado sonda) y un entorno al que el sistema se acopla de forma unitaria (en nuestro caso sería el ancilla). Este sistema extendido se describe por una matriz unitaria  $\rho_\theta^{\text{ext}} = U_{\text{ext}}(\theta)(|E_0\rangle\langle E_0| \otimes \rho_0)U_{\text{ext}}^\dagger(\theta)$ , donde  $|E_0\rangle\langle E_0|$  es el estado inicial del entorno, el cual es independiente de  $\theta$  y  $\rho_0$  es el estado inicial del sistema. Entonces, calculando la traza del entorno con respecto a una base independiente de  $\theta$ ,  $\{|E_l\rangle\}$ , se obtiene la matriz de densidad reducida del sistema, la cual puede ser escrita como

$$\rho_\theta^{\text{sys}} = \text{Tr}_E \rho_\theta^{\text{ext}} = \sum_l \hat{\Pi}_l(\theta) \rho_0 \hat{\Pi}_l^\dagger(\theta) = \sum_l \tilde{\rho}_\theta(l) \quad (7.28)$$

donde  $\Pi_l(\theta) = \langle E_l|U_{\text{ext}}(\theta)|E_0\rangle$ , son operadores de Kraus actuando sobre el sistema. Esencialmente  $\tilde{\rho}_\theta(l)$  son un conjunto de trayectorias cuánticas que ocurren con probabilidad  $P(l|\theta) = \text{Tr}[\tilde{\rho}_\theta(l)]$ , y en nuestro caso serán estados puros. Dentro de este esquema, el POVM óptimo para estimar  $\theta$  es aquel que llega a la información de Fisher generalizada

$$\mathbb{F}(\theta) = I_{\{P_l\}} + \sum_l P(l|\theta) I_Q[\bar{\rho}(l)], \quad (7.29)$$

donde  $I_{\{P_l\}}$  es la información clásica de Fisher de la distribución de las salidas de mediciones dada por  $I_{\{P_l\}} = \sum_l (\partial_\theta P(l|\theta))^2 / P(l|\theta)$ ,  $\bar{\rho}(l) = \tilde{\rho}_\theta(l) / P(l|\theta)$  es la matriz densidad reducida del sistema condicionado al resultado de la medición  $l$ , e  $I_Q$  es la información de Fisher cuántica obtenida a partir de (3.14).

Usando esta nueva medida de información, se puede generalizar la desigualdad de Crámer-Rao a

$$\overline{\text{Var}_\theta[\hat{\theta}]} = \sum_l P(l|\theta) \text{Var}_\theta[\hat{\theta}] \geq \frac{1}{N \mathbb{F}} \quad (7.30)$$

con  $N$  el número de repeticiones del experimento.

En nuestro protocolo, el estado condicionado del sistema,  $\bar{\rho}(l)$ , corresponde a la matriz densidad del estado sonda  $\bar{\rho}_\theta(m)$ , la resultante de la parte III del circuito mostrado en la Fig. 7.1, escalada por  $P(m|\theta)$ , la probabilidad de que la medición sobre la ancilla de el valor  $m$ . Es decir,  $\bar{\rho}_\theta(m) = \tilde{\rho}_\theta(m) / P(m|\theta)$ , donde  $\tilde{\rho}_\theta(m)$  es el estado final sin normalizar post-medición obtenido tras completar la parte III de la Fig. 7.1. Por lo que la información de Fisher generalizada en nuestro caso es

$$\mathbb{F} = \int_{-\infty}^{\infty} \frac{(\partial_\theta P(m|\theta))^2}{P(m|\theta)} dm + \int_{-\infty}^{\infty} P(m|\theta) I_Q[\bar{\rho}_\theta(m)] dm \quad (7.31)$$

donde el primer término es la información de Fisher asociada a la distribución de probabilidad de las salidas de la medición, la cual se denotará como  $\mathbb{I}_C$ , y el segundo término es la información

de Fisher cuántica promedio, la cual se denotará por  $\mathbb{I}_Q$ . Para calcular  $I_Q[\bar{\rho}_\theta(m)]$  se puede utilizar (3.19), pues el estado normalizado post-medición, es un estado puro (ver (7.17)).

Finalmente, cabe destacar que el último operador unitario que se aplica en la parte III del circuito mostrado en la Fig. 7.1 es una corrección unitaria dependiente del resultado de la medición. Esta corrección introduce un operador adicional  $U_l^C$  en (7.28) obteniendo  $\rho_\theta^{\text{sys}} = \text{Tr}_E \rho_\theta^{\text{ext}} = \sum_l U_l^C \Pi_l(\theta) \rho_0 \Pi_l^\dagger(\theta) U_l^{C\dagger} = \sum_l \tilde{\rho}_\theta(l)$ . Esta operación no altera  $P(m | \theta)$ , y salvo por el cambio para  $\bar{\rho}_\theta(l) \rightarrow \bar{\rho}_\theta^C(l) \equiv U_l^C \bar{\rho}_\theta(l) U_l^{C\dagger}$ , la derivación de la información de Fisher cuántica generalizada puede aplicarse de forma directa evaluando (7.31) en el estado  $\bar{\rho}_\theta(l)^C$ .

## 7.5. Sensores con escalamiento Super-Heisenberg

Es bien sabido que en el contexto de estimación de parámetros, los estados obtenidos a partir de una evolución tipo Kerr  $\rho_\theta[\hat{n}^2]$  pueden producir estimadores con una escala en su precisión mejor que estados producidos por una evolución lineal  $\rho_\theta[\hat{n}]$ . Si los recursos del sistema son definidos como  $\bar{n} = E_\theta[\hat{n}]$ , entonces los estados  $\rho_\theta[\hat{n}^2]$  producen una cota de Crámer-Rao que supera el límite de Heisenberg. A este tipo de escalamientos se les llama *super-Heisenberg* [70].

En la sección 7.2, se discutió que el estado post-medición,  $\rho_\theta(m)$ , es similar a un estado derivado de una evolución tipo Kerr de un estado inicial  $\rho$ . Es decir, a un estado  $|\psi\rangle_{\text{kerr}} \langle\psi| \sim \exp(i\chi\hat{n}^2)|\alpha\rangle \langle\alpha| \exp(-i\chi\hat{n}^2)$ . Por lo que se puede utilizar el circuito mostrado en la Fig. 7.1 para producir un protocolo de estimación con escalamiento super-Heisenberg para el ángulo de rotación  $\theta$  que aparece en la parte II del circuito. En particular, supondremos que el parámetro que se quiere estimar es una pequeña desviación  $\delta\theta$  de un valor base  $\theta_0 \in \Theta$  conocido. Es decir,  $\theta \equiv \delta\theta$ .

En el circuito hay una medición construida con los eigenestados del operador de posición que regresa valores  $m$ , los cuales siguen una distribución de probabilidad  $P(m | \theta)$ . Debido a esto, con el objetivo de conocer que tipo de escalamiento se obtiene en la cota de Crámer-Rao para los estimadores, se debe estudiar cómo la información cuántica de Fisher (7.31) escala con los recursos  $\bar{n}$  del estado sonda. Pues se puede obtener, ya sea el límite cuántico estándar, correspondiente a  $\mathbb{F} \sim \bar{n}$ , el escalamiento de Heisenberg  $\mathbb{F} \sim \bar{n}^2$ , o un escalamiento tipo *super-Heisenberg* correspondiente a  $\mathbb{F} \sim \bar{n}^\nu$  con  $\nu > 2$ .

Para comenzar este análisis, tomemos el ket del estado inicial (sonda)  $|\Psi_0\rangle_P$ , que de acuerdo con (7.16) y utilizando el homeomorfismo  $(\beta, \sigma) \leftrightarrow (\phi, r)$  con  $\phi$  el parámetro de la rotación  $R(\phi) = \exp[i\phi\hat{n}]$  aplicado en la sección II del circuito, se llega a que

$$|\Psi_{III}^M\rangle = N'_{III} \exp\left[-\frac{\sigma^2(g\hat{n} - m)^2}{2(1 + 2i\beta\sigma^2)}\right] |\Psi_0\rangle_P = N'_{III} \exp\left[-f(m, \phi) \left(1 - \frac{g}{m} \hat{n}\right)^2\right] |\Psi_0\rangle_P, \quad (7.32)$$

donde

$$f(m, \phi) = \frac{m^2}{4} \frac{1 - i e^{2r} \cot(\phi)}{e^{2r} - i \cot(\phi)}. \quad (7.33)$$

Si ahora se considera el límite de alta compresión  $r \gg 1$ , se obtiene que  $f(m, \phi) \sim -i m^2/4 \cot(\phi)$ ,



y el efecto del circuito sobre el estado sonda es puramente unitario,

$$\begin{aligned} |\Psi_{III}^M\rangle &= \exp\left[\frac{i}{4}\cot(\phi)(m-g\hat{n})^2\right]|\Psi_0\rangle_P \\ &= \exp\left[-\frac{i}{2}\cot(\phi)mg\hat{n}\right]\exp\left[\frac{i}{4}\cot(\phi)m^2\right]\exp\left[\frac{i}{4}\cot(\phi)g^2\hat{n}^2\right]|\Psi_0\rangle_P. \end{aligned} \quad (7.34)$$

Aunque el término  $\exp\left[\frac{i}{4}\cot(\phi)m^2\right]$  es una fase global sin influencia en la información de Fisher generalizada, la fase en la rotación  $\exp\left[-\frac{i}{2}\cot(\phi)mg\hat{n}\right]$  es una función de  $(\phi, g, m)$ . Hecho que justamente evita obtener un escalamiento óptimo para la información de Fisher cuántica generalizada, pues esencialmente se pierde información acerca del parámetro en cada corrida del circuito. Pero es justamente en este límite, donde uno puede aplicar una corrección unitaria  $U_C = \exp[igm\cot(\phi)\hat{n}/2]$  al estado sonda para remover este efecto aleatorio no deseado. No obstante, para su implementación física es necesario conocer el valor de  $\phi$ .

Justamente, la suposición de que el parámetro  $\theta$  a estimar es un desplazamiento pequeño  $\delta\theta$  para un valor base  $\theta_0$  conocido, nos permitirá implementar el operador  $U_c = \exp[igm\cot(\theta_0)]$ , el cual nos produce una corrección de al menos un primer orden en  $\delta\theta$ . Dado que en este caso, se tiene que  $\phi = \theta_0 + \delta\theta$ ,

$$\cot(\phi) = \cot(\theta_0) - (1 + \cot^2\theta_0)\delta\theta + \dots, \quad (7.35)$$

y

$$|\Psi_{III}^M\rangle = \exp\left[\frac{i}{4}(\cot(\theta_0) - (1 + \cot^2\theta_0)\delta\theta)(m-g\hat{n})^2\right]|\Psi_0\rangle_P. \quad (7.36)$$

Notemos que si aplicamos una corrección  $\exp[igm\cot(\theta_0 + \delta\theta)\hat{n}/2]$  al caso general,  $r$  finito, el estado final que se obtiene es  $\rho_{\delta\theta}^C = |\Psi_{III}^C\rangle\langle\Psi_{III}^C|$ , donde

$$|\Psi_{III}^C\rangle = N(m, r, \delta\theta)\exp[f(\delta\theta, \hat{n})]|\Psi_0\rangle_P \quad (7.37)$$

con

$$f(\delta\theta, \hat{n}) = \frac{i}{2}gm\hat{n}(\cot\theta_0 - (1 + \cot^2\theta_0)\delta\theta) - \frac{m^2}{4}\frac{1 - ie^{2r}\cot(\theta_0 + \delta\theta)}{e^{2r} - i\cot(\theta_0 + \delta\theta)}\left(1 - \frac{g}{m}\hat{n}\right)^2, \quad (7.38)$$

y  $N(m, r, \delta\theta)$  una constante de normalización que puede ser tomada real. Por consiguiente, para calcular la información de Fisher cuántica generalizada  $\mathbb{F}$ , es necesario evaluar

$$\langle d\Psi_{III}^C | d\Psi_{III}^C \rangle = \left|\frac{dN}{N}\right|^2 + \left(\frac{dN}{N}\right)^* \mathbb{E}_{\delta\theta}[df] + \left(\frac{dN}{N}\right) \mathbb{E}_{\delta\theta}[df]^* + \mathbb{E}_{\delta\theta}[df^\dagger df] \quad (7.39)$$

$$|\langle \Psi_{III}^C | d\Psi_{III}^C \rangle|^2 = \left|\frac{dN}{N}\right|^2 + \left(\frac{dN}{N}\right)^* \mathbb{E}_{\delta\theta}[df] + \left(\frac{dN}{N}\right) \mathbb{E}_{\delta\theta}[df]^* + |\mathbb{E}_{\delta\theta}[df]|^2 \quad (7.40)$$

donde  $d \equiv d/d(\delta\theta)$  y  $f$  es la función (7.38). De aquí, la información cuántica de Fisher está dada por

$$I_Q[|\Psi_{III}^C\rangle\langle\Psi_{III}^C|] = 4\left(\langle d\Psi_{III}^C | d\Psi_{III}^C \rangle - |\langle \Psi_{III}^C | d\Psi_{III}^C \rangle|^2\right), \quad (7.41)$$

donde los valores esperados son tomados con respecto al estado normalizado  $\rho_{\delta\theta}^C$  con  $\theta = 0$ . Esta información de Fisher  $I_Q$  es función de  $(r, \theta_0, g, m)$ , y entonces, para obtener el término derecho de



la información de Fisher generalizada (7.31), se debe promediar  $I_Q$  sobre todas las posibles salidas  $m$ , es decir,

$$\mathbb{I}_Q(r, \theta_0, g) = \int_{-\infty}^{\infty} P(m | \theta) I_Q(r, \theta_0, g, m) dm, \quad (7.42)$$

donde  $P(m | \theta) \equiv \langle \Psi_{III}^M | \Psi_{III}^M \rangle$ .

Luego, pensando que el estado sonda  $|\Psi_0\rangle_P$  es un estado coherente, es decir,  $|\Psi_0\rangle_P = |\alpha\rangle$ , con valor esperado de fotones  $\bar{n}_P$ , las expresiones de (7.41), y  $P(m | \theta)$ , se pueden calcular de forma analítica. Sin embargo, (7.42) no puede ser evaluada de forma analítica, excepto para los casos más simples, por lo que es necesario hacer un análisis numérico, teniendo en cuenta que es necesario revisar la convergencia de las expresiones numéricas con respecto al número de Fock que se utilice para truncarlas, y tener cuidado en el hecho que (7.42) puede oscilar mucho si se trabaja con un dominio no compacto para  $m$ .

Como el estado sonda es un estado coherente, con  $\text{Tr}[\hat{n}|\alpha\rangle\langle\alpha|] = \bar{n}_P = E[\hat{n}]$ , y suponiendo que es posible hacer una cancelación a primer orden de la fase dependiente de  $m$  como en (7.38), se puede estudiar el escalamiento de la información de Fisher generalizada  $\mathbb{F}$  (7.31) acerca de  $\theta$  usando nuestro protocolo como función del número  $\bar{n}_P$  del estado sonda. La Fig. 7.2 muestra el comportamiento de  $\mathbb{F}$ , para varios valores de  $\theta_0$ , acoplamientos  $g$ , y valores de compresión  $r$ . También se muestra la gráfica del exponente  $\eta$ , dado por  $\mathbb{F} \sim \bar{n}_P^\eta$ , definido como  $\eta = d(\ln \mathbb{F})/d(\ln \bar{n}_P)$ , recordando que el límite cuántico estándar, escalamiento de Heisenberg y escalamiento super-Heisenberg corresponden a  $\eta = 1.0, 2.0, > 2.0$ , respectivamente.

De este análisis numérico también se encontró que  $\mathbb{I}_C \ll \mathbb{I}_Q$ , por lo que  $\mathbb{F} \sim \mathbb{I}_Q$  (7.31). Además, de la Fig. 7.2, se observan dos resultados claros: Primero,  $\mathbb{F}$  es mucho más grande que la información de Fisher cuántica de  $\theta$  correspondiente a codificar el parámetro en el ancilla, es decir, en un estado comprimido con el mismo valor  $\bar{n}_P$ . Segundo, el escalamiento con respecto a  $\bar{n}_P$  es mejor que el encontrado al usar un estado comprimido en protocolos estándar cuya evolución paramétrica es lineal, y por tanto, se obtiene un escalamiento tipo super-Heisenberg. Utilizando un estado comprimido como fuente de recursos, el escalamiento asintótico es aproximadamente  $\mathbb{F} \sim \bar{n}_P^2$  para valores grandes de  $\bar{n}_P$ , mientras que para nuestro esquema, se aproxima a  $\mathbb{F} \sim \bar{n}_P^3$ .

También notamos el rol que juega el valor  $\theta_0$ , ya que habrá escenarios donde sea posible escoger el valor de  $\theta_0$  que maximice la información de Fisher. Resultados para  $\theta_0 = (0.01, 0.1, 1.0)$  rad, son mostrados en la Figura 7.2. De este análisis numérico se observa que cuando  $\theta_0 \rightarrow 0$ , se obtiene un valor mayor de la información de Fisher generalizada  $\mathbb{F}$ , pero el escalamiento de  $\mathbb{F}$  con respecto a  $\bar{n}_P$  se debilita y varía dentro del rango de  $\bar{n}_P$ . También, se estudió cómo la información de Fisher generalizada depende del valor de  $r$  en el ancilla. Esto es mostrado en la Fig. 7.3. Conforme el valor de compresión  $r$  es incrementado, se obtiene un valor mayor para la información de Fisher y un mejor escalamiento que en el caso de un protocolo estándar de estimación usando un estado comprimido con recursos  $\bar{n}_P$ .

Una pregunta natural es ¿qué tan bien funciona la corrección lineal a primer orden?, es decir, ¿qué tan bueno es el resultado de la información de Fisher generalizada utilizando (7.38) comparado con el caso donde no se aplica la corrección unitaria  $U_C$ , (tomando el estado (7.32))? Esta situación se estudió en la Fig. 7.4. Al igual que en la gráfica anterior, se observó que a menor valor de  $\theta_0$  se obtiene un mayor valor de  $\mathbb{F}$  pero peor escalamiento. También se observó que sin ninguna corrección, en ningún caso el protocolo logró superar el límite estándar cuántico. No obstante, con la corrección se alcanza el escalamiento super-Heisenberg y se obtiene una información de Fisher mayor que al usar un esquema de estimación estándar con un estado comprimido con el mismo

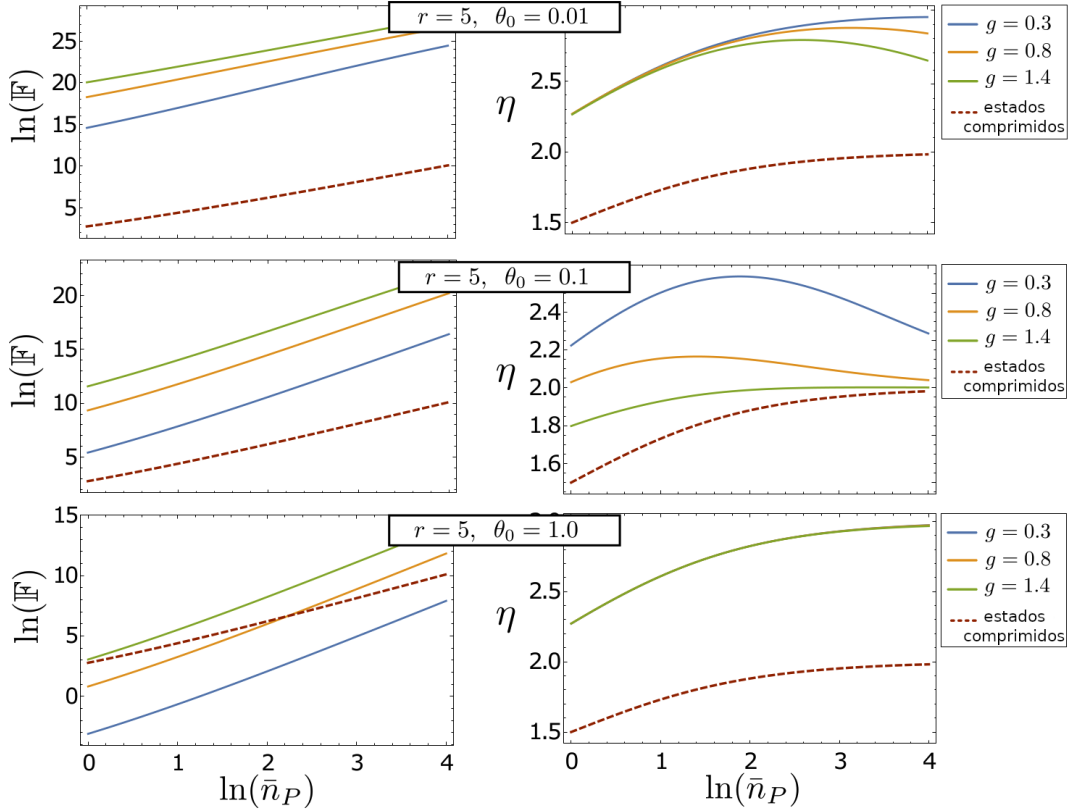


Figura 7.2: Gráfica de la información de Fisher generalizada para la estimación de fase  $\mathbb{F}$  (columna izquierda) y el exponente de escalamiento  $\eta$ , de  $\mathbb{F} \sim \bar{n}_P^\eta$ , (columna derecha) como una función del número promedio de fotones del estado sonda  $\bar{n}_P$ , y para el escalamiento de un protocolo de estimación estándar donde se utiliza un estado comprimido con la misma cantidad de recursos  $\bar{n}_P$  como condición inicial. Se consideraron varios valores de acoplamiento en el operador de rotación-cruzada  $g$ , valores de compresión sobre el ancilla  $r$ , y valores iniciales  $\theta_0$ . Las curvas mostradas en la columna derecha son las pendientes de las curvas mostradas en los paneles izquierdos. Se tomó el eje horizontal como el logaritmo natural de  $\bar{n}_P$ , el valor esperado en el número de fotones del estado sonda, y en el caso de un protocolo estándar con un estado comprimido con el mismo valor de recursos. Las curvas de color (azul, naranja, verde), corresponden a  $g = (0.3, 0.8, 1.4)$ , respectivamente, mientras que las curvas discontinuas de color rojo corresponden a un protocolo estándar con estados comprimidos. Cabe resaltar que el escalamiento Super-Heisenberg corresponde a  $\eta > 2$ , y en los paneles de la derecha se observan varias situaciones donde se alcanza este comportamiento utilizando nuestro protocolo. Para estos cálculos numéricos, se truncó la base de Fock en  $N_{trunc} = 260$ , y el dominio de  $m$  fue restringido al conjunto  $m \in [-900, 900]$ .

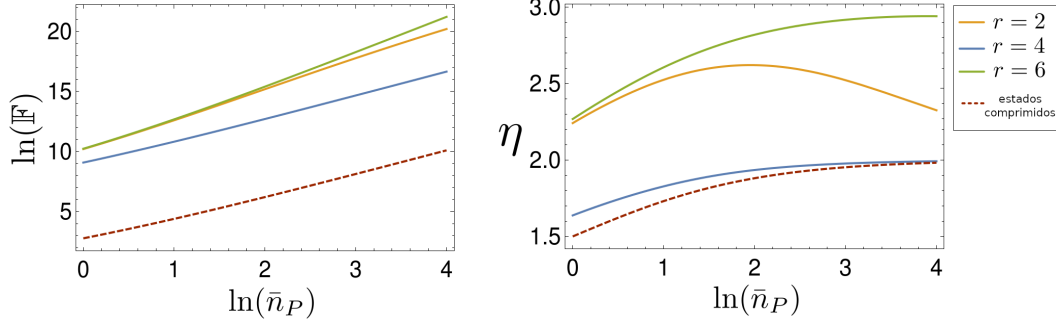


Figura 7.3: Gráfica para la información de Fisher generalizada como función de  $\bar{n}_P$  de nuestro esquema para varios valores de  $r$  en el ancilla y un protocolo estándar donde un estado comprimido con los mismos recursos es utilizado. Las curvas (naranja, azul y verde) corresponden a  $\mathbf{r} = (2, 4, 6)$ , respectivamente, con  $\theta_0 = 0.1, g = 1.0$ . La curva roja corresponde al protocolo estándar con un estado comprimido.

valor de  $\bar{n}_P$ .

En resumen, a partir del análisis numérico de la información de Fisher generalizada para el esquema con una corrección lineal y en el límite de  $r \gg 1$  para el ancilla, se encontró que para un rango de acoplamientos  $g$ , y ángulos iniciales  $\theta_0$ , el protocolo exhibe un escalamiento de  $\mathbb{F} \sim \bar{n}_P^\eta$ , donde  $\eta > 2$ , es decir, un escalamiento tipo *super-Heisenberg*. La causa central detrás de esto es que la integral (7.7) produce el cuadrado de  $b$  en el argumento de la exponencial, por lo que al aplicar la rotación cruzada  $\exp(-ig \hat{p}_A \otimes \hat{n}_P)$  en la parte III, se produce una evolución tipo Kerr en el estado sonda  $\sim \exp(ig^2/4 \cot(\theta) \hat{n}^2)$ . Concretamente, lo que hace el circuito es aplicar un operador  $\exp(ig^2/4 \cot(\theta) \hat{A}^2)$  sobre el estado sonda, en vez de aplicar una evolución  $\exp(-ig \hat{p} \otimes \hat{A}_P)$ , obteniendo las ventajas de una evolución con generador no lineal.

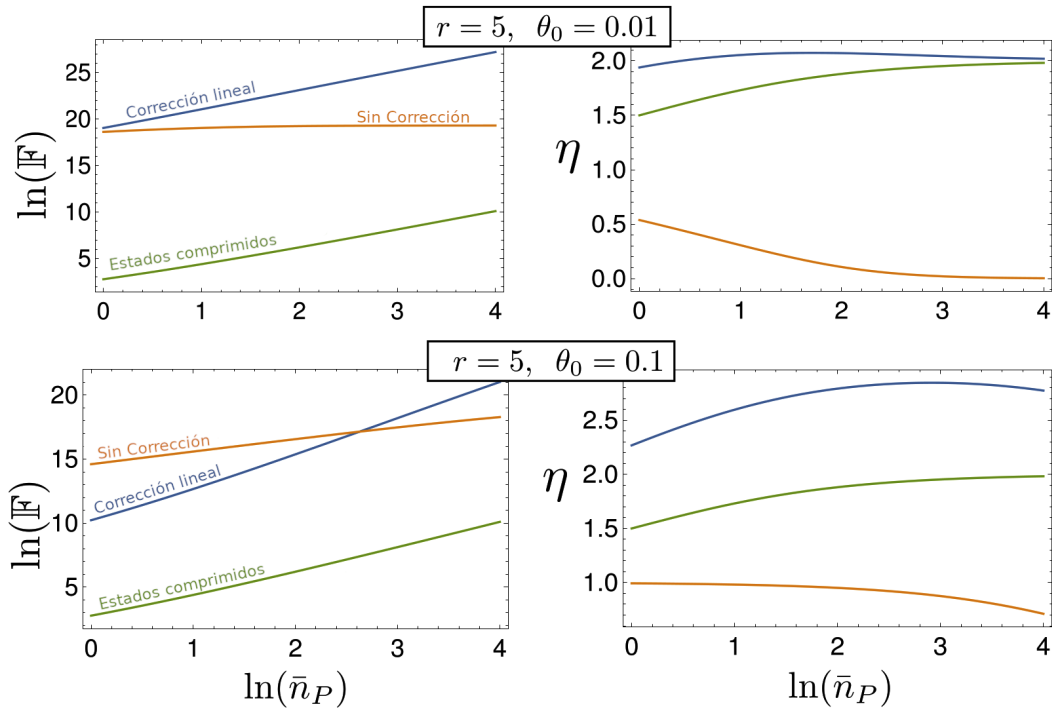


Figura 7.4: Comparación del esquema con corrección lineal y el esquema sin corrección. Las gráficas muestran la información de Fisher generalizada  $\mathbb{F}$  (columna izquierda) y el exponente de escalamiento  $\eta$  (columna derecha). Las curvas (naranja, verde, azul) corresponden a los esquemas (sin corrección, estándar con estados comprimidos, y protocolo con corrección lineal), con  $g = 1.0$ .

## Capítulo 8

# Conclusiones

El objetivo de este trabajo fue dar una descripción sistemática acerca del problema de estimación de parámetros en sistemas cuánticos haciendo especial énfasis en el análisis clásico de datos que muchas veces se omite en los trabajos del área. Este problema tiene un amplio espectro de aplicaciones como metrología cuántica, sensores cuánticos, comunicación cuántica, computación cuántica, entre muchos otros, por lo que existe una gran motivación por encontrar protocolos de estimación que puedan llegar a los límites físicos de precisión. Teóricamente, el límite físico en la precisión de un estimador para un parámetro desconocido es la llamada [cota inferior de Crámer-Rao Cuántica \(CICRC\)](#). Para obtener una estrategia óptima que alcance este límite de precisión es necesario optimizar sobre el espacio de mediciones cuánticas y buscar un estimador clásico eficiente.

Es común ver en la literatura de estimación de parámetros cuánticos la optimización sobre el espacio de mediciones cuánticas y no siempre el análisis clásico. Sin embargo, el análisis clásico sirve para mostrar que incluso al encontrar una medida óptima, es posible no saturar la CICRC ya que estas medidas pueden producir modelos estadísticos no regulares. En particular, cuando las medidas óptimas producen funciones de probabilidad periódicas, las funciones de verosimilitud involucradas son no identificables. Esta condición es recurrente en problemas de estimación de fase o estimación de frecuencias. En esta tesis se estudian diferentes estrategias de estimación, resaltando las condiciones que debe satisfacer el modelo estadístico para producir estimadores que saturen la CICRC. En particular, se estudió el problema de estimación de fase y frecuencia en qubits, el cual es un ejemplo claro de esta disyuntiva. Posteriormente se estudió el problema de estimación en estados coherentes, el cual ejemplifica el acercamiento de la teoría cuando la dimensión del espacio de Hilbert del sistema es infinita.

La investigación original de este trabajo comienza en el capítulo 4, donde se estudiaron diversas estrategias de estimación para el problema de estimación de fase en qubits. Concluyendo que el factor común que produce que estas no sean óptimas, en general, es que sus funciones de verosimilitud son no identificables. En este trabajo se introduce un nuevo método de estimación, el cual utiliza estimaciones covariantes para construir un intervalo de confianza capaz de corregir el problema de no identificabilidad. En este sentido, nuestro método proporciona una estrategia de estimación más robusta y eficiente capaz de alcanzar la CICRC para cualquier condición inicial.

Como una extensión de la estrategia de estimación basada en intervalos de confianza, en el capítulo 5 se trató en detalle el problema de estimación de frecuencia en qubits. El objetivo ideal para este problema es incrementar el tiempo de medición para obtener un error de estimación

menor. Sin embargo, conforme se aumenta el tiempo las funciones de verosimilitud tienden a volverse no identificables, por lo que estrategias estándar de estimación no son óptimas a menos que se trabaje con un intervalo de tiempo restringido, el cual es dado por el orden de magnitud de la frecuencia, hecho que limita el error acerca de las estimaciones. En este caso extendimos el método de estimación basado en intervalos de confianza para poder incrementar el tiempo de medición y obtener una precisión, en los estimadores de frecuencia, mucho menor. Además, se aplicó esta propuesta al problema de estimación sobre qutrits, mejorando aún más el error de los estimadores y teniendo una generalización natural de la propuesta de estimación a sistemas cuánticos de mayor dimensión.

Estos resultados requieren que se tenga acceso experimental a  $N$  copias idénticas del estado del cual se desea estimar el parámetro. No obstante, para una gran cantidad de aplicaciones no es posible contar con esto, por lo que se debe trabajar en el llamado problema de estimación de un solo disparo, donde las estimaciones se obtienen a partir de una única copia del sistema. En este contexto, los objetivos principales son buscar estrategias de estimación con la mínima varianza, optimizar la cota cuántica de Crámer-Rao como función de los recursos del sistema y buscar condiciones donde existan estrategias de estimación que puedan saturar esta cota inferior.

En el capítulo 6 estudiamos el problema de estimación de fase en estados coherentes. Este es el problema típico de estimación de un solo disparo. Aquí, aún cuando por construcción la medida canónica de fase es la de mínima varianza, no se ha encontrado una propuesta experimental para implementarla. Por tal motivo, las mediciones Gaussianas han servido como métodos estándar para estimar la fase de estos estados aunque con un error mayor que la medida canónica. En este capítulo se describió una estrategia de estimación no-Gaussiana basada en conteo de fotones y retroalimentación cuyo desempeño supera al de la mejor estrategia Gaussiana. Este método adaptativo utiliza conteo de fotones como mediciones sobre un estado coherente desplazado. Los desplazamientos se pueden ver como diseños óptimos del esquema, los cuales se eligen maximizando una función de costo. Pese a que las mediciones de conteo de fotones pueden producir funciones de verosimilitud no identificables, para este método es posible explotar el hecho de optimizar sobre el espacio de diseños para garantizar la identificabilidad de las funciones de verosimilitud y así obtener estimadores consistentes, obteniendo una estrategia de estimación que puede aproximar al desempeño de la medición canónica. Además como resultado original se prueba que el rendimiento de esta estrategia de estimación escala de mejor manera que la mejor estrategia de estimación basada en mediciones Gaussianas conocida hasta la fecha.

Finalmente el análisis del capítulo 7 cierra la tesis. Este capítulo se dedicó al problema de la optimización de la desigualdad de Crámer-Rao cuántica como función de los recursos en el problema de un único disparo. En este caso se suele fijar a la energía como recurso y buscar diferentes estados que produzcan mayor información cuántica de Fisher como función del recurso. Históricamente se ha trabajado con dos escalamientos principales: el error cuántico estándar y el escalamiento de Heisenberg. El primero corresponde cuando la información de Fisher es una potencia cuadrática de la energía del sistema y el segundo cuando se tiene una potencia a la cuarta. Pese a que el escalamiento de Heisenberg es absoluto, en la literatura se suele utilizar el término “escalamiento súper-Heisenberg” a aquellos escalamientos donde se obtienen mayores potencias de la información de Fisher. Físicamente, estos escalamientos no toman en cuenta el total de los recursos del sistema y utilizan energía adicional para obtener estos escalamientos sobre un estado objetivo. Este enfoque es el que se adopta en el resto del capítulo 7 donde se presenta un circuito cuántico original para producir un escalamiento tipo súper-Heisenberg para la energía de un

estado coherente con la ayuda de un estado comprimido ancilla.

El circuito cuántico consiste en una serie de compuertas cuánticas sobre el estado coherente y el ancilla para producir al final una transformación no lineal sobre el estado coherente y así una información de Fisher con potencias superiores a la cuarta. Como los estados coherentes modelan rayos láser, este resultado toma gran importancia ante la falta de propuestas para implementar este tipo de transformaciones no lineales en el régimen óptico. Sin embargo, físicamente solo se puede obtener esta cota de forma aproximada (aunque manteniendo la escala super-Heisenberg) ya que dentro del circuito cuántico se realiza una medición sobre la ancilla perdiendo un poco de información. Para cuantificar esta pérdida de información se trabaja con la información generalizada de Fisher. Analizando sus potencias para diferentes condiciones del sistema se concluyó que es posible obtener el escalamiento tipo súper-Heisenberg.

Este hecho es de particular utilidad cuando se desea estimar pequeños desplazamientos o fases sobre el estado coherente, ya que esta suposición produce modelos estadísticos regulares y la implementación de técnicas de estimación estándar se puede ejecutar sin ningún problema extra. En efecto, esta suposición se encuentra con regularidad en el estudio de sensores cuánticos. Por tal motivo, este resultado presenta una opción para mejorar la precisión de estas tecnologías en el espectro óptico.

Los resultados presentados en esta tesis ciertamente muestran que el análisis de problemas de estimación en sistemas cuánticos también requiere de un estudio clásico de los modelos estadísticos producidos por las mediciones sobre el sistema. Esto implica que no basta con encontrar una medida cuántica que produzca una CICRC teóricamente, el objetivo real es construir una medición cuántica que en conjunto con el estado bajo estudio produzcan modelos estadísticos regulares donde los estimadores a utilizar puedan saturar la CICRC, y así extraer toda la información del sistema físico de la que se pueda disponer.





# Bibliografía

- [1] J. Abadie, B. P. Abbott, R. Abbott, T. D. Abbott, M. Abernathy, C. Adams, R. Adhikari, C. Afeldt, B. Allen, and G. S. Allen. A gravitational wave observatory operating beyond the quantum shot-noise limit. *Nature Phys.*, 7:962–965, 2011.
- [2] V. V. Albert, S. O. Mundhada, A. Grimm, S. Touzard, M. H. Devoret, and L. Jiang. Pair-cat codes: autonomous error-correction with low-order nonlinearity. *Quantum Science and Technology*, 4(3):035007, 6 2019. ISSN 2058-9565. doi: 10.1088/2058-9565/ab1e69.
- [3] M. Avesani, D. G. Marangon, G. Vallone, and P. Villoresi. Source-device-independent heterodyne-based quantum random number generator at 17 gbps. *Nature Communications*, 9(1):5365, Dec 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-07585-0.
- [4] O. E. Barndorff-Nielsen and R. D. Gill. Fisher information in quantum statistics. *Journal of Physics A: Mathematical and General*, 33(24):4481–4490, 2000. ISSN 03054470. doi: 10.1088/0305-4470/33/24/306.
- [5] M. Beau and A. del Campo. Nonlinear Quantum Metrology of Many-Body Open Systems. *Physical Review Letters*, 119(1):010403, 7 2017. ISSN 0031-9007. doi: 10.1103/PhysRevLett.119.010403.
- [6] F. E. Becerra, J. Fan, and A. Migdall. Photon number resolution enables quantum receiver for realistic coherent optical communications. *Nat. Photonics*, 9, 2015.
- [7] J. Beltrán and A. Luis. Breaking the Heisenberg limit with inefficient detectors. *Physical Review A*, 72(4):045801, 10 2005. ISSN 1050-2947. doi: 10.1103/PhysRevA.72.045801.
- [8] J. Beltrán and A. Luis. Breaking the heisenberg limit with inefficient detectors. *Phys. Rev. A*, 72:045801, Oct 2005. doi: 10.1103/PhysRevA.72.045801.
- [9] D. W. Berry and H. M. Wiseman. Adaptive measurements and optimal states for quantum interferometry. *Technical Digest - Summaries of Papers Presented at the Quantum Electronics and Laser Science Conference, QELS 2001*, (5):60–61, 2001. doi: 10.1109/QELS.2001.961853.
- [10] D. W. Berry, B. L. Higgins, S. D. Bartlett, M. W. Mitchell, G. J. Pryde, and H. M. Wiseman. How to perform the most accurate possible phase measurements. *Physical Review A - Atomic, Molecular, and Optical Physics*, 80(5):1–22, 2009. ISSN 10502947. doi: 10.1103/PhysRevA.80.052114.

- [11] S. Boixo and R. D. Somma. Parameter estimation with mixed-state quantum computation. *Phys. Rev. A*, 77:052320, May 2008. doi: 10.1103/PhysRevA.77.052320.
- [12] S. Boixo, S. T. Flammia, C. M. Caves, and J. M. Geremia. Generalized Limits for Single-Parameter Quantum Estimation. *Physical Review Letters*, 98(9):090401, 2 2007. ISSN 0031-9007. doi: 10.1103/PhysRevLett.98.090401.
- [13] J. J. Bollinger, Wayne M. Itano, D. J. Wineland, and D. J. Heinzen. Optimal frequency measurements with maximally correlated states. *Phys. Rev. A*, 54:R4649–R4652, Dec 1996. doi: 10.1103/PhysRevA.54.R4649.
- [14] M. Born. Zur quantenmechanik der stoßvorgänge. *Zeitschrift für Physik*, 37(12):863–867, Dec 1926. ISSN 0044-3328. doi: 10.1007/BF01397477.
- [15] V. B. Braginsky, F. Y. Khalili, and K. S. Thorne. *Quantum Measurement*. Cambridge University Press, 1992. doi: 10.1017/CBO9780511622748.
- [16] S. L. Braunstein and C. M. Caves. Statistical distance and the geometry of quantum states. *Physical Review Letters*, 72(22):3439–3443, 1994. ISSN 00319007. doi: 10.1103/PhysRevLett.72.3439.
- [17] S. L. Braunstein, C. M. Caves, and G. J. Milburn. Generalized uncertainty relations: Theory, examples, and Lorentz invariance. *Annals of Physics*, 1996. ISSN 00034916. doi: 10.1006/aphy.1996.0040.
- [18] D. Bruss and G. Leuchs. *Quantum Information, 2 Volume Set: From Foundations to Quantum Technology Applications*. Wiley, 2019. ISBN 9783527805778.
- [19] G. Casella and R.L. Berger. *Statistical Inference*. Duxbury advanced series in statistics and decision sciences. Thomson Learning, 2002. ISBN 9780534243128.
- [20] C. M. Caves and G. J. Milburn. Qutrit entanglement. *Optics Communications*, 179(1):439 – 446, 2000. ISSN 0030-4018. doi: 10.1016/S0030-4018(99)00693-8.
- [21] K. Chaloner and I. Verdinelli. Bayesian Experimental Design: A Review. *Statistical Science*, 10(3):273 – 304, 1995. doi: 10.1214/ss/1177009939.
- [22] F. Chapeau-Blondeau. Optimizing qubit phase estimation. *Physical Review A*, 94(2):1–14, 2016. ISSN 24699934. doi: 10.1103/PhysRevA.94.022334.
- [23] D. Cohen, T. Gefen, L. Ortiz, and A. Retzker. Achieving the ultimate precision limit with a weakly interacting quantum probe. *npj Quantum Information*, 6(1):83, Sep 2020. ISSN 2056-6387. doi: 10.1038/s41534-020-00313-x.
- [24] T. M. Cover and J. A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, USA, 2006. ISBN 0471241954.
- [25] D. A. R. Dalvit, R. L. De Matos Filho, and F. Toscano. Quantum metrology at the Heisenberg limit with ion trap motional compass states. *New Journal of Physics*, 8, 2006. ISSN 13672630. doi: 10.1088/1367-2630/8/11/276.

- [26] E. B. Davies and J. T. Lewis. An operational approach to quantum probability. *Communications in Mathematical Physics*, 1970. ISSN 00103616. doi: 10.1007/BF01647093.
- [27] T. Decker, D. Janzing, and M. Roetteler. Implementation of group-covariant povms by orthogonal measurements. *Journal of Mathematical Physics*, 46, July 2004.
- [28] Dick P. Dee and Arlindo M. da Silva. Maximum-likelihood estimation of forecast and observation error covariance parameters. part i: Methodology. *Monthly Weather Review*, 127(8): 1822 – 1834, 1999. doi: 10.1175/1520-0493(1999)127<1822:MLEOFA>2.0.CO;2.
- [29] M.H. DeGroot and M.J. Schervish. *Probability and Statistics*. Addison-Wesley, 2012. ISBN 9780321500465.
- [30] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă. The elusive heisenberg limit in quantum-enhanced metrology. *Nature Communications*, 3:1063, 2012.
- [31] M. T. DiMario and F. E. Becerra. Single-shot non-gaussian measurements for optical phase estimation. *Phys. Rev. Lett.*, 125:120505, Sep 2020. doi: 10.1103/PhysRevLett.125.120505.
- [32] S. J. Dolinar. *A Class of Optical Receivers Using Optical Feedback*. PhD thesis, Research Laboratory of Electronics, Massachusetts Institute of Technology, Massachusetts, 1976.
- [33] B. Efron. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1):1 – 26, 1979. doi: 10.1214/aos/1176344552.
- [34] M. Eie and S.T. Chang. *A Course on Abstract Algebra*. World Scientific, 2010. ISBN 9789814271905.
- [35] R. A. Fisher. Theory of Statistical Estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 22(5):700–725, 1925. ISSN 14698064. doi: 10.1017/S0305004100009580.
- [36] J. Flórez, E. Giese, D. Curic, L. Giner, R. W. Boyd, and J. S. Lundeen. The phase sensitivity of a fully quantum three-mode nonlinear interferometer. *New Journal of Physics*, 20(12):123022, 12 2018. ISSN 1367-2630. doi: 10.1088/1367-2630/aaf3d2.
- [37] F. Fröwis, M. Skotiniotis, B. Kraus, and W. Dür. Optimal quantum states for frequency estimation. *New Journal of Physics*, 2014. ISSN 13672630. doi: 10.1088/1367-2630/16/8/083010.
- [38] A. Fujiwara. Strong consistency and asymptotic efficiency for adaptive quantum estimation problems. *Journal of Physics A: Mathematical and Theoretical*, 44(7), 2011. ISSN 17518113. doi: 10.1088/1751-8113/44/7/079501.
- [39] J. Gao, C. Picciotto, and W. Jackson. Displacement sensing and estimation theory and applications. *Applied Physics A*, 80(6):1265–1278, Mar 2005. ISSN 1432-0630. doi: 10.1007/s00339-004-3153-5.
- [40] J. P. Gazeau. *Coherent states in Quantum Optics: An oriented overview*. 10 2018. doi: 10.1007/978-3-030-20087-9.

- [41] V. Giovannetti, S. Lloyd, and L. Maccone. Quantum-Enhanced Measurements: Beating the Standard Quantum Limit. *Science*, 306(1330), 2004.
- [42] V. Giovannetti, S. Lloyd, and L. Maccone. Quantum metrology. *Phys. Rev. Lett.*, 96:010401, 2006.
- [43] V. Giovannetti, S. Lloyd, and L. Maccone. Advances in quantum metrology. *Nature Photon.*, 5:222–229, 2013.
- [44] Vittorio Giovannetti, Seth Lloyd, and Lorenzo MacCone. Advances in quantum metrology. *Nature Photonics*, 5(4):222–229, 2011. ISSN 17494885. doi: 10.1038/nphoton.2011.35.
- [45] Z. Griliches, R. Engle, M.D. Intriligator, D. McFadden, J.J. Heckman, and E.E. Leamer. *Handbook of Econometrics*. Handbook of Econometrics. Elsevier Science, 1983. ISBN 9780444887665.
- [46] A. L. Grimsmo, J. Combes, and B. Q. Baragiola. Quantum Computing with Rotation-Symmetric Bosonic Codes. *Physical Review X*, 10(1):11058, 2020. ISSN 21603308. doi: 10.1103/PhysRevX.10.011058.
- [47] C. Gross, T. Zibold, E. Nicklas, J. Estève, and M. K. Oberthaler. Nonlinear atom interferometer surpasses classical precision limit. *Nature*, 464(7292):1165–1169, 2010. ISSN 00280836. doi: 10.1038/nature08919.
- [48] B.C. Hall. *Quantum Theory for Mathematicians*. Graduate Texts in Mathematics. Springer New York, 2013. ISBN 9781461471165.
- [49] B. Harris. *Theory of Probability*. Addison-Wesley series in statistics. Addison-Wesley, 1966. ISBN 9780201028003.
- [50] M. Hayashi. *Asymptotic Theory of Quantum Statistical Inference: Selected Papers*. World Scientific, 2005. ISBN 9789812560155.
- [51] M. Hayashi. *Quantum Information Theory: Mathematical Foundation*. Graduate Texts in Physics. Springer Berlin Heidelberg, 2016. ISBN 9783662497258.
- [52] Masahito Hayashi. *A Group Theoretic Approach to Quantum Information*. 2017. ISBN 9783319452395. doi: 10.1007/978-3-319-45241-8.
- [53] C. W. Helstrom. Minimum mean-squared error of estimates in quantum statistics. *Physics Letters A*, 25(2):101–102, 1967. ISSN 03759601. doi: 10.1016/0375-9601(67)90366-0.
- [54] C. W. Helstrom. Quantum detection and estimation theory, 1969. ISSN 00224715.
- [55] A. S. Holevo. Statistical decision theory for quantum systems. *Journal of Multivariate Analysis*, 1973. ISSN 10957243. doi: 10.1016/0047-259X(73)90028-6.
- [56] A. S. Holevo. Covariant measurements and uncertainty relations. *Reports on Mathematical Physics*, 1979. ISSN 00344877. doi: 10.1016/0034-4877(79)90072-7.

- [57] A. S. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory*. 1982. doi: 10.1007/978-88-7642-378-9.
- [58] A. S. Holevo. Asymptotic estimation of a shift parameter of a quantum state. *Theory of Probability and its Applications*, 49(2):207–220, 2005. ISSN 0040585X. doi: 10.1137/S0040585X97981044.
- [59] M. J. Holland and K. Burnett. Interferometric detection of optical phase shifts at the heisenberg limit. *Phys. Rev. Lett.*, 71:1355–1358, Aug 1993. doi: 10.1103/PhysRevLett.71.1355.
- [60] Z. Huang, K. R. Motes, P. M. Anisimov, J. P. Dowling, and D. W. Berry. Adaptive phase estimation with two-mode squeezed vacuum and parity measurement. *Physical Review A*, 2017. ISSN 24699934. doi: 10.1103/PhysRevA.95.053837.
- [61] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac. Improvement of frequency standards with quantum entanglement. *Phys. Rev. Lett.*, 79:3865–3868, Nov 1997. doi: 10.1103/PhysRevLett.79.3865.
- [62] M. T. Johnsson, P. M. Poggi, M. A. Rodriguez, R. N. Alexander, and J. Twamley. Generating nonlinearities from conditional linear operations, squeezing, and measurement for quantum computation and super-heisenberg sensing. *Phys. Rev. Research*, 3:023222, Jun 2021. doi: 10.1103/PhysRevResearch.3.023222.
- [63] P.D. Jonathan. Quantum optical metrology – the lowdown on high-noon states. *Contemporary Physics*, 49(2):125–143, 2008. doi: 10.1080/00107510802091298.
- [64] G. Kirchmair, B. Vlastakis, Z. Leghtas, S. E. Nigg, H. Paik, E. Ginossar, M. Mirrahimi, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf. Observation of quantum state collapse and revival due to the single-photon Kerr effect. *Nature*, 495(7440):205–209, 3 2013. ISSN 0028-0836. doi: 10.1038/nature11902.
- [65] A. B. Klimov, L. L. Sánchez-Soto, H. de Guise, and G. Björk. Quantum phases of a qutrit. *Journal of Physics A: Mathematical and General*, 37(13):4097–4106, mar 2004. doi: 10.1088/0305-4470/37/13/012.
- [66] K. Landsman. *Foundations of Quantum Theory: From Classical Concepts to Operator Algebras*. Fundamental Theories of Physics. Springer International Publishing, 2018. ISBN 9783319847382.
- [67] E.L. Lehmann and G. Casella. *Theory of Point Estimation*. Springer Texts in Statistics. Springer New York, 2011. ISBN 9781441931306.
- [68] H. Liero and S. Zwanzig. *Introduction to the Theory of Statistical Inference*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press, 2016. ISBN 9781466503205.
- [69] P. Liuzzo-Scorpo, L. A. Correa, F. A. Pollock, A. Górecka, K. Modi, and G. Adesso. Energy-efficient quantum frequency estimation. *New Journal of Physics*, 20(6):063009, jun 2018. doi: 10.1088/1367-2630/aac5b6.

- [70] A. Luis. Nonlinear transformations and the Heisenberg limit. *Physics Letters A*, 329(1-2): 8–13, 8 2004. ISSN 03759601. doi: 10.1016/j.physleta.2004.06.080.
- [71] A. Luis and A. Rivas. Nonlinear Michelson interferometer for improved quantum metrology. *Physical Review A*, 92(2):022104, 8 2015. ISSN 1050-2947. doi: 10.1103/PhysRevA.92.022104.
- [72] A. Luis and L. L. Sánchez-Soto. Breaking the standard quantum limit for interferometric measurements. *Optics Communications*, 89(2-4):140–144, 1992. ISSN 00304018. doi: 10.1016/0030-4018(92)90147-J.
- [73] A. Luis, I. Morales, and A. Rivas. Nonlinear fiber gyroscope for quantum metrology. *Physical Review A*, 94(1):013830, 7 2016. ISSN 2469-9926. doi: 10.1103/PhysRevA.94.013830.
- [74] S. Luo. Quantum fisher information and uncertainty relations. *Letters in Mathematical Physics*, 53(3):243–251, Aug 2000. ISSN 1573-0530. doi: 10.1023/A:1011080128419.
- [75] González-Barríos J. M. *Monografía: Lecture notes on probability theory*. IIMAS-UNAM, 2011. ISBN 978-607-02-2817-9.
- [76] Y. Ma, M. Pang, L. Chen, and W. Yang. Improving quantum parameter estimation by monitoring quantum trajectories. *Physical Review A*, 99(3):032347, 3 2019. ISSN 2469-9926. doi: 10.1103/PhysRevA.99.032347.
- [77] M. Manisera and P. Zuccolotto. Identifiability of a model for discrete frequency distributions with a multidimensional parameter space. *Journal of Multivariate Analysis*, 140:302–316, 2015. ISSN 10957243. doi: 10.1016/j.jmva.2015.05.011.
- [78] J. Marcin and R. Demkowicz-Dobrzański. True precision limits in quantum metrology. *New Journal of Physics*, 17(1):013010, jan 2015. doi: 10.1088/1367-2630/17/1/013010.
- [79] L. S. Martin, W. P. Livingston, S. Hacoheh-Gourgy, H. M. Wiseman, and I. Siddiqi. Implementation of a canonical phase measurement with quantum feedback. *Nature Physics*, Jul 2020. ISSN 1745-2481. doi: 10.1038/s41567-020-0939-0.
- [80] Esteban Martínez-Vargas, Carlos Pineda, Fran çois Leyvraz, and Pablo Barberis-Bloststein. Quantum estimation of unknown parameters. *Phys. Rev. A*, 95:012136, Jan 2017. doi: 10.1103/PhysRevA.95.012136.
- [81] N. C. Menicucci, S. T. Flammia, and P. van Loock. Graphical calculus for Gaussian pure states. *Physical Review A*, 83(4):042335, 4 2011. ISSN 1050-2947. doi: 10.1103/PhysRevA.83.042335.
- [82] E. Merzbacher. *Quantum Mechanics*. Wiley, 1998. ISBN 9780471887027.
- [83] A. Monras. Optimal phase measurements with pure gaussian states. *Phys. Rev. A*, 73:033821, Mar 2006. doi: 10.1103/PhysRevA.73.033821.
- [84] W. J. Munro, K. Nemoto, and T. P. Spiller. Weak nonlinearities: a new route to optical quantum computation. *New Journal of Physics*, 7:137–137, 5 2005. ISSN 1367-2630. doi: 10.1088/1367-2630/7/1/137.

- [85] H. Nagaoka. *in Asymptotic Theory of Quantum Statistical Inference: Selected Papers (World Scientific, Singapore, 2005), Chap. 10.*
- [86] M. Napolitano and M. W. Mitchell. Nonlinear metrology with a quantum interface. *New Journal of Physics*, 12(9):093016, 9 2010. ISSN 1367-2630. doi: 10.1088/1367-2630/12/9/093016.
- [87] X. Nie, J. Huang, Z. Li, W. Zheng, C. Lee, X. Peng, and J. Du. Experimental demonstration of nonlinear quantum metrology with optimal quantum state. *Science Bulletin*, 63(8):469–476, 4 2018. ISSN 20959273. doi: 10.1016/j.scib.2018.03.007.
- [88] M.A. Nielsen and I.L. Chuang. *Quantum Computation and Quantum Information*. Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 2000. ISBN 9780521635035.
- [89] Changhun O. and Wonmin S. Sub shot-noise frequency estimation with bounded a priori knowledge. *Journal of Physics A: Mathematical and Theoretical*, 48(4):045304, dec 2014. doi: 10.1088/1751-8113/48/4/045304.
- [90] Lee C. Rockstuhl C. et al. Oh, C. Optimal gaussian measurements for phase estimation in single-mode gaussian metrology. *npj Quantum Inf*, 5, 2019. doi: 10.1038/s41534-019-0124-4.
- [91] T. Ohsawa and C. Tronci. Geometry and dynamics of gaussian wave packets and their wigner transforms. *Journal of Mathematical Physics*, 58(9):092105, 2017. doi: 10.1063/1.4995233.
- [92] R. Okamoto, M. Iefuji, S. Oyama, K. Yamagata, H. Imai, A. Fujiwara, and S. Takeuchi. Experimental demonstration of adaptive quantum state estimation. *Phys. Rev. Lett.*, 109:130404, Sep 2012. doi: 10.1103/PhysRevLett.109.130404.
- [93] S. Ole and S. Stefan. Approaching the heisenberg limit with two-mode squeezed states. *Journal of Optics B: Quantum and Semiclassical Optics*, 6(3):S66–S70, mar 2004. doi: 10.1088/1464-4266/6/3/011.
- [94] M. Oliva and O. Steuernagel. Quantum kerr oscillators’ evolution in phase space: Wigner current, symmetries, shear suppression, and special states. *Phys. Rev. A*, 99:032104, Mar 2019. doi: 10.1103/PhysRevA.99.032104.
- [95] L. Paninski. Asymptotic theory of information-theoretic experimental design. *Neural Comput.*, 17(7):1480–1507, July 2005. ISSN 0899-7667. doi: 10.1162/0899766053723032.
- [96] Matteo G.A. Paris. Quantum estimation for quantum technology. *International Journal of Quantum Information*, 7:125–137, 2009. ISSN 02197499. doi: 10.1142/S0219749909004839.
- [97] J. Pellonpää and J. Schultz. Measuring the canonical phase with phase-space measurements. *Phys. Rev. A*, 88:012121, Jul 2013. doi: 10.1103/PhysRevA.88.012121.
- [98] L. Pezzè and A. Smerzi. Quantum theory of phase estimation. In *Atom Interferometry*. 2014. ISBN 9781614994480. doi: 10.3254/978-1-61499-488-0-691.
- [99] M. E. Plenio. Logarithmic negativity: A full entanglement monotone that is not convex. 95: 090503, 2005.

- [100] E. Polino, M. Valeri, N. Spagnolo, and F. Sciarrino. Photonic quantum metrology. *AVS Quantum Science*, 2(2):024703, 2020. doi: 10.1116/5.0007577.
- [101] M. Reed and B. Simon. *I: Functional Analysis. Methods of Modern Mathematical Physics*. Elsevier Science, 1981. ISBN 9780080570488.
- [102] A. Rivas and A. Luis. Precision Quantum Metrology and Nonclassicality in Linear and Nonlinear Detection Schemes. *Physical Review Letters*, 105(1):010403, 6 2010. ISSN 0031-9007. doi: 10.1103/PhysRevLett.105.010403.
- [103] M. A. Rodríguez-García, I. P. Castillo, and P. Barberis-Blostein. Efficient qubit phase estimation using adaptive measurements. *Quantum*, 5:467, June 2021. ISSN 2521-327X. doi: 10.22331/q-2021-06-04-467.
- [104] S.M. Ross. *Introduction to Probability Models*. Elsevier Science, 2006. ISBN 9780123756879.
- [105] E. Ryan, C. Drovandi, J. McGree, and T. Pettitt. A review of modern computational algorithms for bayesian optimal design. *International Statistical Review*, 84(1):128–154, 2016. doi: 10.1111/insr.12107.
- [106] Rosen S. Finite transformations in various representations of  $su(3)$ . *Journal of Mathematical Physics*, 673(12), 1971. doi: 10.1063/1.1665634.
- [107] R. J. Sewell, M. Napolitano, N. Behbood, G. Colangelo, F. Martin Ciurana, and M. W. Mitchell. Ultrasensitive Atomic Spin Measurements with a Nonlinear Interferometer. *Physical Review X*, 4(2):021045, 6 2014. ISSN 2160-3308. doi: 10.1103/PhysRevX.4.021045.
- [108] T. Sugiyama, P. S. Turner, and M. Muraio. Adaptive experimental design for one-qubit state estimation with finite data based on a statistical update criterion. *Phys. Rev. A*, 85:052107, May 2012. doi: 10.1103/PhysRevA.85.052107.
- [109] J. Suzuki. Quantum-state estimation problem via optimal design of experiments, 2020.
- [110] M. Taboga. *Lectures on Probability Theory and Mathematical Statistics - 2nd Edition*. CreateSpace Independent Publishing Platform, 2012. ISBN 9781480215238.
- [111] F. Toscano, W. P. Bastos, and Ruynet L. de Matos F. Attainability of the quantum information bound in pure-state models. *Phys. Rev. A*, 95:042125, Apr 2017. doi: 10.1103/PhysRevA.95.042125.
- [112] D. V. Tsarev, T. V. Ngo, Ray-Kuang Lee, and A. P. Alodjants. Nonlinear quantum metrology with moving matter-wave solitons. *New Journal of Physics*, 21(8):083041, 8 2019. ISSN 1367-2630. doi: 10.1088/1367-2630/ab398e.
- [113] S. Virmani and M. B. Plenio. Construction of extremal local positive-operator-valued measures under symmetry. *Phys. Rev. A*, 67:062308, Jun 2003. doi: 10.1103/PhysRevA.67.062308.
- [114] P. Václav and S. M. Barnett. On the exponential form of the displacement operator for different systems. *Physica Scripta*, 90:065208, 6 2015. ISSN 0031-8949. doi: 10.1088/0031-8949/90/6/065208.



- [115] J. Wang, L. Davidovich, and G. S. Agarwal. Quantum sensing of open systems: Estimation of damping constants and temperature. *Phys. Rev. Research*, 2:033389, Sep 2020. doi: 10.1103/PhysRevResearch.2.033389.
- [116] C. Wei and Z. Zhang. Improving the phase sensitivity of a Mach–Zehnder interferometer via a nonlinear phase shifter. *Journal of Modern Optics*, 64(7):743–749, 4 2017. ISSN 0950-0340. doi: 10.1080/09500340.2016.1260777.
- [117] E.P. Wigner. *Symmetries and Reflections: Scientific Essays of Eugene P. Wigner*. Ox Bow Press, 1979. ISBN 9780918024169.
- [118] H. M. Wiseman. Adaptive phase measurements of optical modes: Going beyond the marginal  $q$  distribution. *Phys. Rev. Lett.*, 75:4587–4590, Dec 1995. doi: 10.1103/PhysRevLett.75.4587.
- [119] H. M. Wiseman, D. W. Berry, S. D. Bartlett, B. L. Higgins, and G. J. Pryde. Adaptive measurements in the optical quantum information laboratory. *IEEE Journal of Selected Topics in Quantum Electronics*, 15(6):1661–1672, Nov 2009. doi: 10.1109/JSTQE.2009.2020810.
- [120] M. J. Woolley, G. J. Milburn, and C. M. Caves. Nonlinear quantum metrology using coupled nanomechanical resonators. *New Journal of Physics*, 10(12):125018, 12 2008. ISSN 1367-2630. doi: 10.1088/1367-2630/10/12/125018.
- [121] K. Yamagata. *Asymptotic quantum statistical estimation*. PhD thesis, Osaka University, 2013.
- [122] D.S. Young. *Handbook of Regression Methods*. A Chapman & Hall Book. CRC Press, 2017. ISBN 9781498775298.
- [123] J. Zhang, Z. Zhang, L. Cen, J. Hu, and Y. Zhao. Nonlinear phase estimation: Parity measurement approaches the quantum Cramér-Rao bound for coherent states. *Physical Review A*, 99(2):022106, 2 2019. ISSN 2469-9926. doi: 10.1103/PhysRevA.99.022106.
- [124] W. H. Zurek. Decoherence, einselection, and the quantum origins of the classical. *Reviews of Modern Physics*, 75(3):715–775, 5 2003. ISSN 0034-6861. doi: 10.1103/RevModPhys.75.715.
- [125] M. Zwiernik, C. A. Pérez-Delgado, and P. Kok. General optimality of the heisenberg limit for quantum metrology. *Phys. Rev. Lett.*, 105:180402, Oct 2010. doi: 10.1103/PhysRevLett.105.180402.