



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS FÍSICAS

**PARÁMETROS ALEATORIOS E INFERENCIA
BAYESIANA EN METROLOGÍA CUÁNTICA**

TESIS

**QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS (FÍSICA)**

PRESENTA:

ESTEBAN MARTÍNEZ VARGAS

TUTOR

DR. PABLO BARBERIS BLOSTEIN (IIMAS)

MIEMBROS DEL COMITÉ TUTOR

DR. CARLOS PINEDA (IF)

DR. FRANÇOIS LEYVRAZ (ICF)

CIUDAD DE MÉXICO, JUNIO 2016.



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

Resumen

Dada la importancia de la estimación de parámetros en la física y en la ciencia en general, es muy importante el estudio detallado del proceso y de las cotas mínimas de estimación de parámetros que nos son permitidas. Más hoy en día ya que la detección de efectos requiere a veces de una precisión impresionante. La colaboración LIGO reportó recientemente [1] que se tuvo que detectar un cambio de longitud de trayectoria óptica del orden de 10^{-18} metros para corroborar la existencia de ondas gravitacionales predichas por Einstein. Es por tanto de interés investigar los límites últimos de la precisión que puede tener un experimento.

Hasta donde sabemos, la mecánica cuántica es la teoría más fundamental con la que contamos. Es de gran interés investigar los límites de precisión que nos permite esta teoría. Investigamos en esta tesis las cotas de error en la estimación de parámetros. El proceso de estimación es el siguiente: partimos de un modelo teórico de un sistema, y en base de datos experimentales encontramos los parámetros que mejor describan el comportamiento de nuestro sistema. El área de metrología cuántica se centra en encontrar las cotas mínimas de error que son permitidas en un proceso de estimación.

Es muy útil entender la desigualdad de Cramér-Rao para el desarrollo de una teoría rigurosa de dichas cotas. Veremos que el caso cuántico conlleva maximizar sobre un conjunto de operadores para encontrar la cantidad llamada información cuántica de Fisher F_Q .

Esta tesis es una crítica a F_Q . Así también es una propuesta, ya que expresamos la necesidad de, en vez de usar la teoría usual, debemos movernos a una versión Bayesiana de la misma. Introducimos la cantidad Z y vemos cómo difiere de lo desarrollado anteriormente. Hacemos el cálculo explícito de Z para un sistema cuántico.

Proponemos también un método de Inferencia Bayesiana que mejora la cota de error con cada medición. Aplicamos dicho método al ejemplo del que hicimos el cálculo explícito. Vimos que usando Z y nuestro método propuesto el error mínimo es menor que el obtenido utilizando un método comparable

(al que llamaremos método de Fujiwara) que utiliza la cantidad F_Q .

Abstract

Given the importance of parameter estimation in physics and in science in general, it is very important the detailed study of the process and the allowed minimum bounds in parameter estimation. Even more nowadays where the detection of effects requires sometimes an impressive precision. The LIGO collaboration has recently reported [1] that a change of length of the order of 10^{-18} meters had to be detected to corroborate the existence of gravitational waves predicted by Einstein. Therefore, it is interesting to investigate the ultimate limits of the precision that an experiment can achieve.

As far as we know, Quantum Mechanics is the most fundamental theory that we have. It is very interesting to investigate the limits of precision allowed by this theory. We investigate in this thesis the error lower bounds in parameter estimation. The process of parameter estimation is the following: we start from a theoretical model of a system and based on experimental data we find the parameters that best describe our systems behaviour. The area of Quantum Metrology is focused on the lower bounds in the allowed error in a parameter estimation process.

It is very useful to understand the Cramér-Rao inequality for the development of a rigorous theory of such bounds. We will see that the quantum version leads to maximize over a set of operators to find the quantity that is called Quantum Fisher Information F_Q .

This thesis is a critique of F_Q . It is also a proposal because we explain the necessity of, instead of using the usual theory, we have to move to a Bayesian version of such theory. We introduce the quantity Z and we see how it differs from what was developed previously. We do the explicit calculation of Z for a quantum system.

We propose also a Bayesian inference method that improves the error bound with every measurement. We apply such method to the example that we calculated explicitly. We saw that using Z and our proposed method, that the obtained minimum error is lower than the obtained with a comparable method (which we will name Fujiwara Method) that uses F_Q .

Agradecimientos

Agradezco a mis padres y hermanas, cuyo soporte y aliento siempre me han impulsado a hacerle frente a todo.

Agradezco a Grecia su cariño y la inspiración que me brinda.

Agradezco a mis amigos su comprensión y cercanía, así como todos los momentos alegres a su lado.

En esta tesis colaboró de manera muy cercana todo el comité tutor. Quiero agradecer a Carlos Pineda ya que los métodos numéricos fueron idea suya. Así también utilicé programas que él me proporcionó para partes de los cálculos numéricos. Agradezco a François Leyvraz ya que ayudó a clarificar mucho los conceptos que teníamos en mente. Agradezco a Pablo Barberis su paciencia conmigo, además de que de él es la propuesta de la cantidad Z y el método de Inferencia Bayesiana.

Agradezco el apoyo financiero del proyecto PAPIIT-DGAPA número IN103714.

Índice general

1. Introducción	3
1.1. Preliminares	3
1.2. Proceso de estimación de Parámetros de sistemas cuánticos	4
1.3. Cotas de error	6
1.4. Objetivos y resultados	6
2. Algunas consideraciones	9
2.1. Variables aleatorias	9
2.2. Inferencia Bayesiana	10
2.3. POVMs	10
2.3.1. Mediciones proyectivas	11
2.3.2. POVMs	11
3. Desigualdades de Cramér-Rao	13
3.1. Información de Fisher	13
3.1.1. Estimadores y error cuadrado medio	14
3.1.2. Cota de Cramér-Rao	14
3.2. Información de Van Trees	16
3.3. Información Cuántica de Fisher	17
3.4. Crítica de F_Q	19
4. Información Cuántica de Van Trees	21
4.1. Definición Estándar	21
4.2. Crítica de la información cuántica de Fisher generalizada V_Q	22
4.3. Definición alternativa	23
5. Aplicaciones	25
5.1. Distribución Gaussiana con POVM de 2 opciones	25
5.2. Distribución uniforme y POVM continuo	30

6. Inferencia Bayesiana	35
6.1. Método de Fujiwara	35
6.2. Inferencia Bayesiana	36
6.3. Ejemplo POVM de 2 elementos con una Gaussiana	36
7. Conclusiones	41
Bibliografía	45

*All truths wait in all things,
They neither hasten their own delivery nor resist it,
They do not need the obstetric forceps of the surgeon,
The insignificant is as big to me as any,
(What is less or more than a touch?)*

-Walt Whitman, *Leaves of Grass*

Capítulo 1

Introducción

1.1. Preliminares

Todo procedimiento científico cuyo fin sea el de crear conocimiento involucra, en primera, o en última instancia, mediciones, desde la perspectiva de la ciencia moderna. Las mediciones, en principio, nos proporcionan los datos sobre el mundo, sobre los que se pueden realizar teorías de la realidad. También a la inversa, la única forma que se conoce de poner a prueba una teoría, es corroborando con mediciones. Las mediciones, para la ciencia, equivalen a las percepciones de los sentidos. El tema del conocimiento a partir de los sentidos es de suma importancia en la historia del pensamiento occidental. Es, por tanto, muy importante para la ciencia moderna, entender bien cómo se hacen dichas mediciones, y tener mediciones precisas y confiables. Más hoy en día, cuando las mediciones de efectos físicos, a veces requieren una gran precisión para medir efectos minúsculos. Por ejemplo, recientemente la colaboración LIGO [1], ha publicado la detección de ondas gravitacionales predichas por Einstein. Sin embargo, para detectar este efecto, se tuvo que detectar un cambio de longitud en el camino óptico de un láser del orden de 10^{-21} . Por esto, los brazos son de 4 km de longitud. A pesar de dicha extensión, los efectos son minúsculos. Se requirió una precisión impresionante para la detección de las ondas gravitacionales.

La metrología cuántica tiene como propósito estudiar las mediciones de sistemas cuánticos, para encontrar límites últimos de precisión que permite la teoría cuántica. Es decir, se considera el carácter probabilista de la teoría cuántica. Está fuera del enfoque de ésta área el investigar o decir algo al respecto del llamado “problema de la medición” en mecánica cuántica, no es la intención ir más allá de la mecánica cuántica. Tampoco se tiene la in-

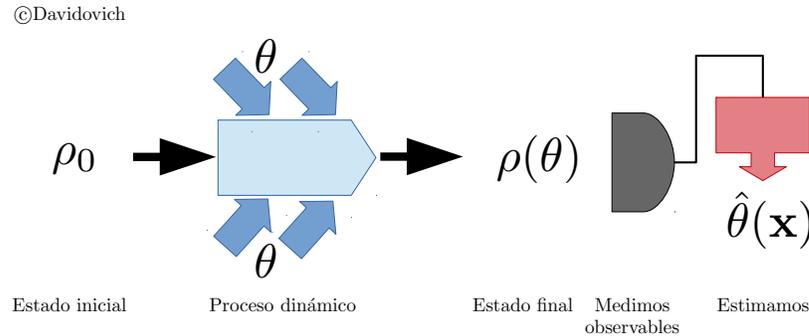


Figura 1.1: Proceso de estimación de parámetros de sistemas cuánticos.

tención el de hacer aparatos que midan mejor algún observable. Lo que se desea es desarrollar métodos de alta precisión para *estimar parámetros*. El nombre de metrología es un poco impreciso en este respecto.

Para hacer un análisis teórico de este tipo de problemas nos concentraremos en las cotas del error mínimo en la estimación de un parámetro. En específico, el error mínimo que permite la mecánica cuántica. Nuestro tratamiento será solamente del cálculo de estas cotas, sin meternos en el problema de estimación de lleno.

1.2. Proceso de estimación de Parámetros de sistemas cuánticos

Es necesario describir el proceso de estimación de parámetros. Partimos de un modelo teórico con algunos parámetros que no se puedan calcular desde la teoría misma. Entonces, usando éste modelo y los resultados de mediciones, buscamos el valor de los parámetros con el que mejor se ajuste a nuestro modelo y a los valores que se obtienen de las mediciones.

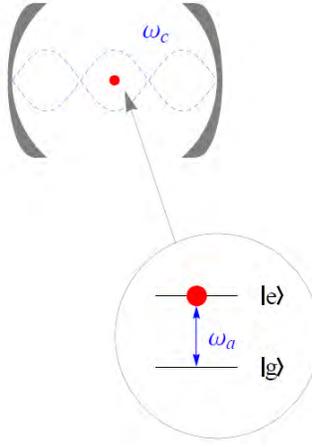
En la figura (1.1) se muestra un esquema del proceso de estimación de parámetros para sistemas cuánticos. Primeramente tenemos un estado inicial ρ_0 . Dicho estado es sometido a una evolución dinámica, lo que le imprime cierta dependencia en un parámetro θ , entonces, se termina con un estado $\rho(\theta)$ que depende de dicho parámetro. Se hacen mediciones, se miden observables del sistema, obteniendo un vector de resultados \mathbf{x} , y con ésta información, se estima el valor del parámetro θ . Éste esquema es bastante general. Note que no es imperativo conocer los detalles de la evolución del sistema para el pro-

1.2. PROCESO DE ESTIMACIÓN DE PARÁMETROS DE SISTEMAS CUÁNTICOS 5

$$\hat{H} = \hat{H}_{field} + \hat{H}_{atom} + \hat{H}_{int}$$

$$\hat{H}_{field} = \hbar\omega_c \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right)$$

$$\hat{H}_{atom} = \hbar\omega_a \frac{\hat{\sigma}_z}{2}$$

$$\hat{H}_{int} = \hbar g (\sigma_+ \hat{a} + \sigma_- \hat{a}^\dagger)$$


©Davidovich ©Wikipedia

Figura 1.2: Modelo de Jaynes-Cummings.

blema de estimación de parámetros. No es necesario saber un hamiltoniano específico, pues nos enfocamos en el problema de, una vez que se tiene $\rho(\theta)$, hacer una estimación de θ lo más precisa posible.

Para poner un ejemplo ilustrativo, considere el Modelo de Jaynes-Cummings [2]. Como puede verse en la figura (1.2) dicho modelo tiene varios parámetros, a saber, ω_c , ω_a y g . El problema de estimación de parámetros dentro de dicho modelo es el de, a partir de mediciones de estados que evolucionan con del hamiltoniano del sistema (\hat{H}), estimar el valor de alguno de estos parámetros. Esto se debe a que dichos estados dependerán al final de los parámetros ω_c , ω_a y g . En esta tesis nos centraremos en la estimación de un solo parámetro aunque la teoría se puede generalizar para considerar un número finito de parámetros.

Se ha demostrado que usando efectos puramente cuánticos es posible mejorar la precisión en la estimación de parámetros. Por ejemplo, en interferometría óptica, la determinación de una fase como función de N fotones escala como $1/\sqrt{N}$ en el límite estándar, mientras que estados de N fotones entrelazados conllevan un escalamiento proporcional a $1/N$, lo que es una mejor resolución con la misma cantidad de recursos [3, 4].

1.3. Cotas de error

Para hacer un estudio riguroso de los límites del error en la estimación de parámetros es muy útil usar un resultado muy elegante llamado la desigualdad de Cramér-Rao. Dicha desigualdad asigna una cota inferior del error en la estimación de un parámetro cuando se usa un estimador no sesgado (ver capítulo 3),

$$\zeta^2 \geq \frac{1}{F(\theta)}, \quad (1.1)$$

Donde ζ^2 es el error asociado a la estimación del parámetro θ y $F(\theta)$ se conoce como la información de Fisher. Como veremos en el capítulo 3, $F(\theta)$ es el error cuadrado mínimo que permite el uso de un estimador no sesgado, pues $F(\theta)$ es independiente de la estrategia de estimación que se adopte.

La información de Fisher está definida cuando se tiene la distribución de probabilidad de los observables \mathbf{x} , que dependen de un parámetro θ , es decir, siempre que exista $p(\mathbf{x}|\theta)$, una distribución de probabilidad. Es natural plantear el problema de las mediciones en sistemas cuánticos. Siguiendo la regla de Born, tenemos que para un estado $\rho(\theta)$ y un elemento de un POVM (vea capítulo 2) $M(x)$ (x es un número, no un vector) que,

$$p(x|\theta) = \text{Tr}[\rho(\theta)M(x)]. \quad (1.2)$$

Ahora bien, la probabilidad depende también del POVM que escojamos. Hay cierta arbitrariedad en el caso cuántico y hay que tomar en cuenta el POVM que se use para hacer una medición. Normalmente, se toma una cantidad que es un límite de $F(\theta)$, para sistemas cuánticos. Braunstein y Caves definieron [5] la cantidad

$$F_Q(\theta) = \text{Máx}_{POVM_s} F(\theta). \quad (1.3)$$

Que propusieron como la cota última de la desigualdad de Cramér-Rao (1.1) en el caso de sistemas cuánticos.

1.4. Objetivos y resultados

El propósito principal de ésta tesis es el de criticar la cantidad F_Q , y proponer una cantidad de información como el mínimo en el error obtenible para la estimación de parámetros en sistemas cuánticos. Argumentaré en esta tesis que la cantidad F_Q en general no tiene sentido físico. Así también, argumentaré que se debe adoptar otra perspectiva del problema. Se tiene

que considerar la información que se conoce del parámetro θ que deseamos estimar. Debemos rehacer la cota de Cramér-Rao en términos de probabilidad Bayesiana.

En este cambio de perspectiva, hay que asumir que el parámetro a estimar, θ , es una variable aleatoria, con una distribución de probabilidad $\lambda(\theta)$ que representa el conocimiento *a priori* del parámetro θ . Ahondaré en definir bien cada cantidad y ver la insuficiencia de F_Q en términos físicos. Presentaré una cota alternativa a F_Q que es de naturaleza Bayesiana a la que llamaré Z , y mostraré cómo nuestra propuesta difiere de lo que se ha hecho hasta ahora. Asimismo, presentaré cálculos explícitos y numéricos de un ejemplo concreto, que pretende ser uno sencillo.

Se verá que el cálculo explícito de las cotas de Cramér-Rao para sistemas cuánticos presenta varios retos. Tuve que utilizar un *ansatz* para obtener un resultado analítico en el primer ejemplo. Para corroborar mi *ansatz* se tuvieron que desarrollar métodos numéricos que se explican en la sección correspondiente. El método numérico que se desarrolló es una herramienta útil para encontrar dichas cotas de error, ya que se carece de una teoría matemática que nos indique como encontrarlas. En el segundo ejemplo usé otro *ansatz*, sin embargo, el método numérico indicó que no solucionaba el problema.

Finalmente, presentaré un método de inferencia Bayesiana en la que el error mínimo disminuye cada vez que se hace una medición y se cambia la información *a priori* del parámetro θ en cada paso. Éste método utiliza la cota Z que introdujimos. Hay un método que también es un proceso adaptativo de POVMs al que llamamos método de Fujiwara [6]. Dicho método utiliza la cantidad F_Q . Se obtuvo una gráfica al final que indica que el método que proponemos en esta tesis obtiene un error menor que el método de Fujiwara para el mismo ejemplo concreto.

En capítulo (2) presento algunos conceptos y terminología que usaré en el resto de la tesis, como POVMs, inferencia Bayesiana, etc. En el capítulo (3), presento una deducción de la desigualdad de Cramér-Rao, también, presento la desigualdad de Van Trees, que es la versión Bayesiana clásica. Así mismo presento la deducción de la información de Van Trees. Presento también la información de Fisher cuántica F_Q . En el capítulo (4) una versión Bayesiana de F_Q que existe actualmente en la literatura especializada. Sin embargo, dicha versión difiere de nuestra propuesta en que utiliza F_Q , mientras que nuestra propuesta no. Se hace énfasis en que el aspecto Bayesiano de nuestra propuesta viene de *considerar el conocimiento que se tiene* del parámetro que se desea estimar. En el capítulo (5) se hace un cálculo explícito usando dos ejemplos concretos. Se muestran los cálculos numéricos que se hicieron

así como analíticos. En el capítulo (6) se muestra el método de inferencia Bayesiana para uno de los ejemplos antes mencionados. En el capítulo (7) se presentan nuestras conclusiones.

Esta tesis es la continuación de mi trabajo de tesis de licenciatura [7]. En dicha tesis ya había introducido la cantidad Z pero la perspectiva de dicha tesis es muy distinta a la que tomo aquí. Además los cálculos analíticos como numéricos así como el método de inferencia Bayesiana es original de esta tesis.

Capítulo 2

Algunas consideraciones

En este capítulo revisaremos algunos conceptos que será útil tener en cuenta ya que se hará uso exhaustivo de ellos en el resto de la tesis.

2.1. Variables aleatorias

Una Variable Aleatoria es [8]

- Una función X que toma sus argumentos de un conjunto al que llamamos espacio de muestreo M , aunque a veces se llama conjunto de estados. Es decir,

$$X : M \rightarrow \mathbb{R}.$$

- Las salidas de X están moduladas por una distribución de probabilidad $p(x)$. $x \in \mathbb{R}$ es la salida de la variable aleatoria.

Definimos el valor de expectación de la variable aleatoria como

$$\langle X \rangle = \sum_i x_i p(x_i).$$

Para variables continuas unidimensionales $X : \mathbb{R} \rightarrow \mathbb{R}$ la distribución de probabilidad $p(x)$ es una función no negativa,

$$p(x) \geq 0,$$

y normalizada, en el sentido de

$$\int p(x) dx = 1.$$

La probabilidad de que $x \in (x, x + dx)$ esta dada por $p(x)dx$. Entonces, el valor de expectación es

$$\langle X \rangle = \int xp(x)dx.$$

2.2. Inferencia Bayesiana

El teorema de Bayes nos dice que una distribución de probabilidad conjunta se puede factorizar de manera conveniente [9],

$$P(x, \theta) = p(x|\theta)\lambda(\theta) = p(\theta|x)\gamma(x).$$

Sabiendo que,

$$\gamma(x) = \int P(x, \theta)d\theta = \int p(x|\theta)\lambda(\theta)d\theta.$$

Entonces, la consecuencia inmediata del teorema de Bayes es

$$p(\theta|x) = \frac{p(x|\theta)\lambda(\theta)}{\int p(x|\theta)\lambda(\theta)d\theta}. \quad (2.1)$$

Utilizando la ecuación (2.1) se puede hacer un proceso de inferencia. Sea Θ una variable aleatoria con salidas θ y con distribución $\lambda(\theta)$. Supongamos que se mide el valor x_1 , entonces, se tiene una distribución $p(\theta|x_1)$ usando la ecuación (2.1). Entonces, se tendrá una nueva distribución de probabilidad para la variable aleatoria Θ . Se ve entonces que el proceso se puede repetir de manera completamente análoga para cualquier número de pasos, con lo cual tendremos nuevas distribuciones de probabilidad. Con esto, se va cambiando el conocimiento que se tiene de la variable aleatoria Θ , en otras palabras, se va infiriendo la distribución de probabilidad de la variable aleatoria Θ con cada nueva medición.

2.3. POVMs

La formulación axiomática de la mecánica cuántica (vea la referencia [10] por ejemplo) nos indica que cualquier medición cuántica se describe por un conjunto $\{M_m\}$ de operadores, donde el sufijo m numera los resultados posibles de un experimento cuántico. Dichos operadores actúan sobre algún estado cuántico $|\psi\rangle$. El conjunto $\{M_m\}$ debe ser completo,

$$\sum_m M_m^\dagger M_m = \mathbb{1}.$$

La probabilidad de que el resultado m ocurra antes de hacer la medición está dada por

$$p(m) = \langle \psi | M_m^\dagger M_m | \psi \rangle.$$

El estado después de la medición es

$$\frac{M_m | \psi \rangle}{\langle \psi | M_m^\dagger M_m | \psi \rangle}.$$

2.3.1. Mediciones proyectivas

Muchas veces resulta conveniente no considerar el conjunto más amplio posible de mediciones que el axioma de mediciones cuánticas nos indica, sino restringirnos a uno específico que es el de mediciones proyectivas. En este caso los operadores M_m son proyectores ortogonales, esto es, son Hermitianos $M_m^\dagger = M_m$ y ortogonales, $M_{m'} M_m = \delta_{m'm} M_m$.

2.3.2. POVMS

Si consideramos que $\{M_m\}$ es cualquier conjunto de operadores que cumpla con el axioma de medición, entonces tenemos que podemos hacer la siguiente identificación

$$M_m^\dagger M_m = E_m.$$

Al conjunto de operadores $\{E_m\}$ se le suele llamar un POVM [10] (*positive operator valued measure*). Vea que no se pide en este caso que los operadores sean proyectores. En general, el conjunto $\{E_m\}$ es una base sobre-completa de proyectores.

Capítulo 3

Desigualdades de Cramér-Rao

En este capítulo se hace la deducción de la desigualdad de Cramér-Rao así como la desigualdad de Van Trees. También se introduce la información cuántica de Fisher.

3.1. Información de Fisher

Un parámetro de un sistema físico es una idealización de la mente del que experimenta. Su realidad es algo *a priori* de la experiencia con el fin de entender algún fenómeno físico, es una herramienta del entendimiento. Parte del modelo idealizado que se tenga de algún sistema físico. Lo que irrevocablemente concierne a la experiencia son los observables. A partir de ellos se deduce, en base a el modelo físico que se tenga en mente, los parámetros que describen mejor la experiencia. Para estimar parámetros es crucial entonces hacer mediciones, pues es imposible obtener dichos valores del modelo teórico en mente, sin recurrir a otro modelo.

Supongamos que deseamos medir el valor de un parámetro θ de un sistema físico dado. Para esto hacemos observaciones de N , un número finito de observables del sistema en cuestión, obteniendo así un conjunto $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$ de resultados. Pues bien, dado este conjunto de resultados nos encontramos con un problema, a saber, el de deducir a partir de las observaciones \mathbf{x} un valor confiable sobre el parámetro θ . Este problema es bien conocido en estadística, le llamamos el problema de estimación de parámetros.

La manera más general de abordar este problema es introduciendo una función llamada estimador que denotaremos por $\hat{\theta}(\mathbf{x})$. Ahora bien, es natural pensar que, dado que el valor de θ está fijo, los valores de nuestro conjunto

de resultados \mathbf{x} dependa del valor de θ . Matemáticamente podemos expresar esta idea del modo siguiente: los elementos del conjunto \mathbf{x} siguen una distribución de probabilidad $p(\mathbf{x}|\theta)$.

3.1.1. Estimadores y error cuadrado medio

Queremos ver que tan bueno es el estimador $\hat{\theta}$ para estimar θ . Para esto es necesario considerar, en general, el error que se comete al usar $\hat{\theta}$. Una cantidad bastante usada es el error cuadrado medio $\zeta^2(\theta)$ que se tiene al usar $\hat{\theta}$ como estimador. El error cuadrado medio se define como,

$$\zeta^2 \equiv \int (\hat{\theta}(\mathbf{x}) - \theta)^2 p(\mathbf{x}|\theta) d\mathbf{x}.$$

Un estimador es mejor si ζ^2 es pequeño. Analicemos pues, esta cantidad. Vea que sumando y restando lo mismo tenemos que [9],

$$\begin{aligned} \zeta^2 &= \int ((\hat{\theta}(\mathbf{x}) - \langle \hat{\theta}(\mathbf{x}) \rangle) + (\langle \hat{\theta}(\mathbf{x}) \rangle - \theta))^2 p(\mathbf{x}|\theta) d\mathbf{x} \\ &= \langle (\hat{\theta}(\mathbf{x}) - \langle \hat{\theta}(\mathbf{x}) \rangle)^2 \rangle + (\langle \hat{\theta}(\mathbf{x}) \rangle - \theta)^2, \end{aligned}$$

donde

$$\langle \hat{\theta}(\mathbf{x}) \rangle = \int \hat{\theta}(\mathbf{x}) p(\mathbf{x}|\theta) d\mathbf{x}.$$

A $\langle (\hat{\theta}(\mathbf{x}) - \langle \hat{\theta}(\mathbf{x}) \rangle)^2 \rangle$ se le llama varianza y a $\langle (\langle \hat{\theta}(\mathbf{x}) \rangle - \theta)^2 \rangle$ se le llama sesgo. Un conjunto de estimadores de especial interés, son los estimadores no sesgados, que son estimadores que cumplen que $\langle \hat{\theta}(\mathbf{x}) \rangle = \theta$, es decir, cuyo sesgo es cero. A pesar de que los estimadores no sesgados en promedio estimen bien el parámetro, podemos investigar la varianza de este estimador, e investigar la condición de la varianza mínima.

3.1.2. Cota de Cramér-Rao

Deduciremos entonces la cota de Cramér-Rao a partir de la condición de no-sesgo. Para los estimadores no sesgados se sigue que,

$$\langle \hat{\theta}(\mathbf{x}) - \theta \rangle = \int (\hat{\theta}(\mathbf{x}) - \theta) p(\mathbf{x}|\theta) d\mathbf{x} = 0. \quad (3.1)$$

Derivando con respecto a θ tenemos que,

$$\int (\hat{\theta}(\mathbf{x}) - \theta) \frac{\partial p(\mathbf{x}|\theta)}{\partial \theta} d\mathbf{x} - \int p(\mathbf{x}|\theta) d\mathbf{x} = 0. \quad (3.2)$$

Observe que

$$\frac{\partial p}{\partial \theta} = p \frac{\partial \ln p}{\partial \theta}. \quad (3.3)$$

Entonces,

$$\int (\hat{\theta}(\mathbf{x}) - \theta) \frac{\partial \ln p(\mathbf{x}|\theta)}{\partial \theta} p(\mathbf{x}|\theta) d\mathbf{x} = 1. \quad (3.4)$$

Haciendo $p = \sqrt{p}\sqrt{p}$ y reagrupando, tenemos que

$$\int \left[\frac{\partial \ln p}{\partial \theta} \sqrt{p} \right] \left[(\hat{\theta}(\mathbf{x}) - \theta) \sqrt{p} \right] d\mathbf{x} = 1. \quad (3.5)$$

Utilizando la desigualdad de Schwarz, tenemos que

$$\left[\int \left(\frac{\partial \ln p}{\partial \theta} \right)^2 p d\mathbf{x} \right] \left[\int (\hat{\theta}(\mathbf{x}) - \theta)^2 p d\mathbf{x} \right] \geq 1. \quad (3.6)$$

La ecuación (3.6) se llama desigualdad de Cramér-Rao, en honor a los trabajos de Cramér y Rao. El término de más hasta la izquierda, es decir,

$$F(\theta) = \int \left(\frac{\partial \ln p}{\partial \theta} \right)^2 p d\mathbf{x}, \quad (3.7)$$

es llamado información de Fisher [3, 11, 12] en honor a Ronald Fisher, quien estudió esta función y la desigualdad de Cramér-Rao [13]. Reescribimos la ecuación (3.6) como

$$\varsigma^2 \geq \frac{1}{F(\theta)}. \quad (3.8)$$

El otro término es el error cuadrado medio el cual introdujimos más arriba. El error cuadrado medio es una medida general del error que se comete en una medición arbitraria. Es de notar que, mientras que el error cuadrado medio depende del estimador, la información de Fisher no. Esto quiere decir que la información de Fisher nos da el menor error cuadrado medio posible, que se logra cuando se satura la desigualdad de Cramér-Rao (3.8). La información de Fisher es una cota natural del error que existe al hacer una medición, pues solo depende de la distribución de probabilidad de los resultados de algún observable del sistema en cuestión. Esto es, la información de Fisher es una cota del error cuadrado medio que solamente depende del modelo del sistema en cuestión.

3.2. Información de Van Trees

Puede parecerse en un primer acercamiento a este esquema teórico que el hecho de que θ esté fijo es crucial para que sea posible dar una cota al error cuadrado medio asociado a una medición. Sin embargo, veremos en esta sección que aún si el parámetro θ es una variable aleatoria a la que se le asocia una distribución de probabilidad $\lambda(\theta)$, existe una desigualdad tipo Cramér-Rao. Primeramente, observe que tiene sentido pensar en este caso una distribución de probabilidad de un conjunto de resultados \mathbf{x} como en el caso de la información de Fisher, es decir, si θ es una variable aleatoria, podemos definir $p(\mathbf{x}|\theta)$. Si definimos la variable aleatoria $\theta \in [a, b]$, podemos, sin pérdida de generalidad suponer que la distribución de probabilidad asociada a esta variable, $\lambda(\theta)$ cumple que $\lambda \rightarrow 0$ cuando $\theta \rightarrow a, b$. Entonces, tenemos que

$$\int_a^b \hat{\theta}(\mathbf{x}) \frac{\partial(p(\mathbf{x}|\theta)\lambda(\theta))}{\partial\theta} d\theta = \hat{\theta}(\mathbf{x}) p(\mathbf{x}|\theta)\lambda(\theta) \Big|_a^b = 0, \quad (3.9)$$

porque λ se elimina en los extremos del intervalo. Note también que multiplicamos por una función que no depende de θ , así que sin ningún problema entra a la integral y no es afectada por el proceso de integración. Por otra parte, también podemos hacer la siguiente integral por partes,

$$\int_a^b \theta \frac{\partial(p(\mathbf{x}|\theta)\lambda(\theta))}{\partial\theta} d\theta = \theta p(\mathbf{x}|\theta)\lambda(\theta) \Big|_a^b - \int_a^b p(\mathbf{x}|\theta)\lambda(\theta) d\theta = - \int_a^b p(\mathbf{x}|\theta)\lambda(\theta) d\theta.$$

Utilizando las dos integrales anteriores e integrando en \mathbf{x} tenemos que,

$$\iint_a^b (\hat{\theta}(\mathbf{x}) - \theta) \frac{\partial(p(\mathbf{x}|\theta)\lambda(\theta))}{\partial\theta} d\theta d\mathbf{x} = \iint_a^b p(\mathbf{x}|\theta)\lambda(\theta) d\theta d\mathbf{x} = 1. \quad (3.10)$$

Podemos reescribir esta igualdad como sigue,

$$\iint_a^b (\hat{\theta}(\mathbf{x}) - \theta) \frac{\partial(\ln p(\mathbf{x}|\theta)\lambda(\theta))}{\partial\theta} p(\mathbf{x}|\theta)\lambda(\theta) d\theta d\mathbf{x} = 1. \quad (3.11)$$

Entonces, haciendo el truco de antes ($\lambda p = \sqrt{\lambda p} \sqrt{\lambda p}$) y aplicando la desigualdad de Schwarz tenemos que,

$$\left[\iint_a^b \left(\frac{\partial(\ln p(\mathbf{x}|\theta)\lambda(\theta))}{\partial\theta} \right)^2 p(\mathbf{x}|\theta)\lambda(\theta) d\theta d\mathbf{x} \right] \left[\iint_a^b (\hat{\theta}(\mathbf{x}) - \theta)^2 p(\mathbf{x}|\theta)\lambda(\theta) d\theta d\mathbf{x} \right] \geq 1. \quad (3.12)$$

Al segundo término del lado izquierdo de la desigualdad se le llama error cuadrado medio Bayesiano, que no es más que el error cuadrado medio del caso de Fisher promediado con la distribución $\lambda(\theta)$, i.e.

$$\zeta^2 = \iint_a^b \left(\hat{\theta}(\mathbf{x}) - \theta \right)^2 p(\mathbf{x}|\theta) \lambda(\theta) d\theta d\mathbf{x}.$$

Al primer término, es decir,

$$V = \iint_a^b \left(\frac{\partial(\ln p(\mathbf{x}|\theta) \lambda(\theta))}{\partial \theta} \right)^2 p(\mathbf{x}|\theta) \lambda(\theta) d\theta d\mathbf{x}.$$

le llamamos información de Van Trees, en honor a Harry Van Trees quien estudió por primera vez la desigualdad de Cramér-Rao expuesta en esta sección [14]. La información de Van Trees es el análogo a la información de Fisher cuando el parámetro que estamos estimando es una variable aleatoria. Reescribimos la ecuación (3.12) como

$$\zeta^2 \geq \frac{1}{V}. \quad (3.13)$$

3.3. Información Cuántica de Fisher

Hasta ahora la información de Fisher y de Van Trees han surgido por falta de control y conocimiento de los resultados de un experimento. Es decir, hemos usado estas medidas en el marco de incertidumbre clásica, proveniente de desconocimiento por parte del experimentador. Es bien sabido que las mediciones en mecánica cuántica tienen asociada una probabilidad intrínsecamente, es decir, parte esencial de la teoría es la imposibilidad de predecir el resultado antes de hacer una medición, lo que si se sabe son los resultados posibles y la probabilidad asociada a cada uno. Entonces, dado que lo único que la información de Fisher utiliza son distribuciones de probabilidad que vienen de los resultados de una medición, podemos usar las distribuciones de probabilidad que nos da la mecánica cuántica para calcular una información de Fisher sin ningún problema. Esta información de Fisher nos dice entonces, hasta donde nos permite la teoría cuántica disminuir el error cuadrado medio por su carácter intrínsecamente probabilista.

Primeramente, encontraremos la información de Helstrom H [15]. Supongamos que tenemos un conjunto de matrices de densidad que están parametrizadas por el parámetro que deseamos estimar $\rho(\theta)$. Para definir H necesitamos definir el operador L_θ llamado Derivada Logarítmica Simétrica (DLS) [11].

Sea L_θ el operador autoadjunto que soluciona la ecuación

$$\frac{\partial \rho}{\partial \theta} = \frac{1}{2}(L_\theta \rho(\theta) + \rho(\theta) L_\theta).$$

Para clarificar las cosas, consideremos el caso más sencillo de un solo observable x en lugar del vector \mathbf{x} .

Ahora bien, uno de los postulados de la mecánica cuántica [10] es que para un POVM $\{E(x)\}_x$ que define cierta medición, podemos definir la probabilidad de obtener el resultado y al medir en el sistema $\rho(\theta)$ como $p(y|\theta) = \text{tr}[E(y)\rho(\theta)]$ (ver capítulo (2)).

Entonces,

$$\frac{\partial p(x|\theta)}{\partial \theta} = \text{tr} \left[E(x) \frac{\partial \rho(\theta)}{\partial \theta} \right] = \text{Re}(\text{tr}[E(x)\rho(\theta)L_\theta]). \quad (3.14)$$

Entonces, usando la regla de Born podemos escribir una información de Fisher en este caso como sigue

$$F(\theta) = \int \frac{(\text{Re}(\text{tr}[E(x)\rho(\theta)L_\theta]))^2}{\text{tr}[E(x)\rho(\theta)]} dx. \quad (3.15)$$

Tomemos en cuenta la desigualdad de Cauchy-Schwarz:

$$|\text{tr}[A^\dagger B]|^2 \leq \text{tr}[A^\dagger A] \text{tr}[B^\dagger B].$$

Así también observe que si $A^\dagger = A$, se sigue que $\sqrt{A} = \sqrt{A^\dagger}$. Entonces, tenemos que, siguiendo a [11]

$$\begin{aligned} F(\theta) &\leq \int \left| \frac{\text{tr}[E(x)\rho(\theta)L_\theta]}{\sqrt{\text{tr}[E(x)\rho(\theta)]}} \right|^2 dx, \\ &= \int \left| \text{tr} \left[\frac{\sqrt{E(x)}\sqrt{\rho(\theta)}\sqrt{\rho(\theta)}L_\theta\sqrt{E(x)}}{\sqrt{\text{tr}[E(x)\rho(\theta)]}} \right] \right|^2 dx, \\ &\leq \int \text{tr}[\rho(\theta)L_\theta E(x)L_\theta] dx, \\ &= \text{tr}[\rho(\theta)L_\theta^2]. \end{aligned}$$

Para pasar del segundo al tercer renglón se utilizó la desigualdad de Schwarz. Esto demuestra que la información de Fisher está acotada por arriba por una función que llamaremos información Helstrom, es decir,

$$F(\theta) \leq H(\theta) \equiv \text{tr}[\rho(\theta)L_\theta^2]. \quad (3.16)$$

Braunstein y Caves definieron una cota para la desigualdad de Cramér-Rao [5] para sistemas cuánticos como sigue

$$F_Q(\theta) = \underset{E(x)}{\text{Máx}}\{F(\theta)\}. \quad (3.17)$$

Vea que cada POVM define una $F(\theta)$. Llamamos POVM *óptimo* al POVM para el cual se alcanza F_Q .

Braunstein y Caves afirmaron que $F_Q(\theta) = H(\theta)$. Sin embargo, se ha demostrado que F_Q puede ser distinta de H [16]. La definición (3.17) es muy clara en términos de una implementación en un laboratorio, hay que encontrar un POVM tal que cualquier otro de menor o igual información de Fisher. H no depende de encontrar un POVM.

En casos en donde $H > F_Q$, H no tiene una interpretación directa ya que es una cota inalcanzable por medio de mediciones. La debilidad de H estriba en que depende únicamente en la estructura del modelo físico a considerar y que no toma en cuenta ninguna estrategia de medición.

3.4. Crítica de F_Q

Entonces, es preferible quedarnos con F_Q que H . Sin embargo, F_Q no está bien definida siempre. Vea que el POVM óptimo puede depender del valor del parámetro θ . Se ha demostrado que hay sistemas en los que esto sucede [16]. Entonces, para encontrar F_Q hay que saber el valor de θ . Sin embargo, deseamos estimar el valor del parámetro θ , es el punto de todo el problema, no lo podemos presuponer. Por lo tanto, no podemos calcular F_Q . Hay casos en los que se puede calcular F_Q , sin embargo, son casos especiales, cuando el POVM con el cual se alcanza F_Q no depende del valor de θ . En general no se puede calcular.

Capítulo 4

Información Cuántica de Van Trees

4.1. Definición Estándar

Primeramente, observamos que la información de Van Trees se puede reescribir de manera más entendible.

$$\begin{aligned} V &= \iint_a^b \left(\frac{\partial \ln p(\mathbf{x}|\theta)\lambda(\theta)}{\partial \theta} \right)^2 p(\mathbf{x}|\theta)\lambda(\theta) d\theta d\mathbf{x} \\ &= \iint_a^b \left(\frac{\partial \ln p(\mathbf{x}|\theta)}{\partial \theta} + \frac{\partial \ln \lambda(\theta)}{\partial \theta} \right)^2 p(\mathbf{x}|\theta)\lambda(\theta) d\theta d\mathbf{x} \\ &= \iint_a^b \left(\frac{\partial \ln p(\mathbf{x}|\theta)}{\partial \theta} \right)^2 p(\mathbf{x}|\theta)\lambda(\theta) d\theta d\mathbf{x} + 2 \iint_a^b \frac{\partial \ln p(\mathbf{x}|\theta)}{\partial \theta} \frac{\partial \ln \lambda(\theta)}{\partial \theta} p(\mathbf{x}|\theta)\lambda(\theta) d\theta d\mathbf{x} \\ &\quad + \iint_a^b \left(\frac{\partial \ln \lambda(\theta)}{\partial \theta} \right)^2 p(\mathbf{x}|\theta)\lambda(\theta) d\theta d\mathbf{x}. \end{aligned}$$

Vea que

$$2 \iint_a^b \frac{\partial \ln p(\mathbf{x}|\theta)}{\partial \theta} \frac{\partial \ln \lambda(\theta)}{\partial \theta} p(\mathbf{x}|\theta)\lambda(\theta) d\theta d\mathbf{x} = 2 \int_a^b \frac{\partial \lambda}{\partial \theta} d\theta \int \frac{\partial p(\mathbf{x}|\theta)}{\partial \theta} d\mathbf{x} = 0. \quad (4.1)$$

Sea

$$\varphi[\lambda(\theta)] \equiv \int_a^b \left(\frac{\partial \ln \lambda(\theta)}{\partial \theta} \right)^2 \lambda(\theta) d\theta.$$

Puede entenderse a $\varphi[\lambda]$ como la información de Fisher de la distribución de conocimiento *a priori* $\lambda(\theta)$. Entonces V se puede escribir como

$$V = \int F(\theta)\lambda(\theta)d\theta + \varphi[\lambda]. \quad (4.2)$$

Omitiremos por simplicidad evaluar la integral en los límites a y b de aquí en adelante.

En la literatura [17, 11, 18] se hace el cambio $F \rightarrow F_Q$ (donde F_Q es la información cuántica de Fisher) en la ecuación (4.2), con lo que se obtiene la siguiente cota para la desigualdad de Cramér-Rao,

$$V_Q = \int F_Q(\theta)\lambda(\theta)d\theta + \varphi[\lambda(\theta)]. \quad (4.3)$$

A V_Q suele llamarse información cuántica de Fisher generalizada [11].

4.2. Crítica de la información cuántica de Fisher generalizada V_Q

La información de Van Trees se utiliza cuando el parámetro a estimar es una variable aleatoria. Esto tiene algunas interpretaciones. Puede ser que la distribución *a priori* $\lambda(\theta)$ represente el conocimiento que se tiene del parámetro θ , a lo que puede llamarse una interpretación Bayesiana. También puede ser que el parámetro θ cambie aleatoriamente de medición a medición. Sea la razón que fuere, θ es un parámetro aleatorio.

Ahora bien, para calcular F_Q es necesario conocer un POVM (vea la ecuación (3.17)). Como sabemos, dicho POVM puede depender del valor del parámetro que se quiere estimar. Sin embargo, en el contexto de la información de Van Trees el parámetro es una variable aleatoria. Antes de una medición, carece de sentido preguntarse por el valor específico del parámetro en este caso. Entonces, si θ es aleatorio, el POVM que se requiere para calcular F_Q no está definido previo a una medición, solamente estará definido para los casos en que el POVM que alcance F_Q no dependa del valor del parámetro a estimar.

Observe también que V_Q es calculable si F_Q es calculable. Como vimos al final del capítulo (3) en general no se puede calcular.

4.3. Definición alternativa

A partir de la información de Van Trees V , proponemos la siguiente cantidad,

$$Z = \underset{POVM_s}{\text{Máx}} \left\{ \int F(\theta)\lambda(\theta)d\theta \right\} + \varphi[\lambda(\theta)]. \quad (4.4)$$

Note que $\varphi[\lambda(\theta)]$ no depende de un POVM. Z es una cota para la desigualdad de Van Trees (3.13) que se alcanza en sistemas cuánticos. Llamamos a Z información cuántica de Van Trees.

Z difiere de V_Q pues solamente es necesario un POVM para calcularse, a diferencia de un continuo de POVMs necesario para calcular V_Q . Ciertamente en general $Z \leq V_Q$, sin embargo, Z es implementable en un esquema de medición general mientras que, como vimos, V_Q no tiene sentido físico. Z toma en cuenta el modelo teórico del sistema cuántico, así como las estrategias de medición implementables. Esto es crucial pues las estrategias de medición es a lo que se tendrá acceso en un proceso de estimación paramétrica, pues, como dijimos al principio del capítulo 3 para estimar parámetros es crucial hacer mediciones sobre observables.

Observe que Z tiene una interpretación Bayesiana clara. $\lambda(\theta)$ representa el conocimiento *a priori* que se tiene del parámetro θ . Si aceptamos esta interpretación, Z puede usarse como herramienta de estudio para la metrología cuántica, es decir, en lugar de utilizar F_Q . La ventaja de hacer esto es que Z no tiene los problemas de generalidad para calcularse como lo tiene F_Q . Esto se debe a que Z no presupone que el valor del parámetro a estimar está fijo, entonces es una cota que no depende del valor del parámetro θ que se desconoce en principio. Sin embargo Z es una cota que depende de lo que *ya se conoce* del parámetro a estimar (lo cual está codificado en $\lambda(\theta)$). Veremos en el capítulo 6 que esto conlleva naturalmente un proceso iterativo de inferencia Bayesiana. Entonces, lo que se conoce del parámetro a estimar puede ir mejorando con mediciones sucesivas.

Capítulo 5

Aplicaciones

En este capítulo aplicaremos la definición de la información de Van Trees cuántica (4.4) dada en el capítulo anterior. Compararemos los resultados de usar Z con lo que se obtendría usando V_Q .

5.1. Distribución Gaussiana con POVM de 2 opciones

Tenemos como estado inicial un estado coherente $|\alpha\rangle$ y le aplicamos el operador $e^{i\hat{n}\theta}$, esto resulta en un estado que depende del parámetro θ . El estado coherente $|\alpha\rangle$ se define como

$$\hat{a}|\alpha\rangle \equiv \alpha|\alpha\rangle,$$

donde \hat{a} es el operador de aniquilación del oscilador armónico. Como mencioné en la introducción, no es necesario saber los detalles de cómo se da el cambio de $|\alpha\rangle$ a $|e^{i\theta}\alpha\rangle$ para el problema de estimación del parámetro θ , solamente necesitamos el estado final sobre el que se harán mediciones.

Lo que sucede en el cambio $|\alpha\rangle \rightarrow |e^{i\theta}\alpha\rangle$ se observa más claramente en las correspondientes funciones de Wigner de estos estados. En la figura (5.1) se muestra el cambio en la función de Wigner. En dicha imagen se ve que θ solamente mueve el centro de la gaussiana un ángulo θ sin cambiar su distancia con respecto al origen.

Tenemos entonces que,

$$|\psi(\theta)\rangle = e^{i\hat{n}\theta}|\alpha\rangle = |e^{i\theta}\alpha\rangle. \quad (5.1)$$

θ entonces representa un cambio de fase en un oscilador armónico. Estudiamos este sistema debido a que está muy estudiada la información de Fisher

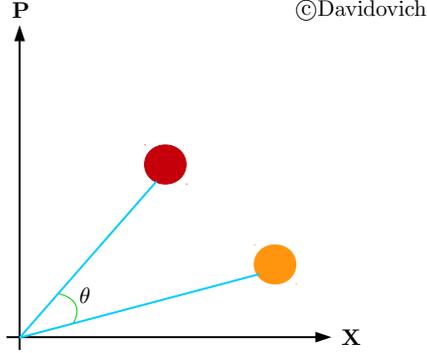


Figura 5.1: Evolución $|\alpha\rangle \rightarrow |e^{i\theta}\alpha\rangle$ en términos de sus funciones de Wigner, en el espacio $P \times X$.

cuántica F_Q correspondiente en la literatura [3].

A diferencia de [3] supondremos que θ es una variable aleatoria cuya distribución de probabilidad es una Gaussiana centrada en cero. Es decir, sea

$$\lambda(\theta, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\theta^2}{2\sigma^2}}, \quad (5.2)$$

la distribución de probabilidad asociada a la variable aleatoria θ .

Definamos ahora la familia de POVMs $\{E_0, E_1\}_\xi$ que nos será muy útil,

$$\begin{aligned} E_0(\xi) &= |\psi(\xi)\rangle\langle\psi(\xi)|, \\ E_1(\xi) &= \mathbf{1} - |\psi(\xi)\rangle\langle\psi(\xi)|, \end{aligned}$$

donde $|\psi(\xi)\rangle$ es una copia de $|\psi(\theta)\rangle$ pero con el valor ξ en lugar de θ . ξ es un parámetro que determinaremos posteriormente. Vea que tenemos una familia de POVM's, uno para cada valor ξ .

Para calcular Z usando la definición dada en (4.4) sería necesario encontrar el máximo sobre todos los POVM's posibles. Sin embargo, esto no es fácil ya que no hay una forma trivial de investigar el espacio de los POVM's posibles. Usaremos la familia de POVM's $\{E_0, E_1\}_\xi$ y veremos que, numéricamente podemos mostrar que esta familia *nos basta* para encontrar el máximo de Z sobre los POVM's, es decir, el máximo de V sobre todos los POVM's se encuentra en esta familia. En este caso, al tener dos posibles resultados de

la medición, a saber $x = 0$ ó $x = 1$, tenemos que

$$p(x|\theta) = \begin{cases} p(\theta, \xi) & \text{si } x = 0 \\ 1 - p(\theta, \xi) & \text{si } x = 1. \end{cases}$$

Donde $p(\theta, \xi) = \langle \psi(\theta) | E_0(\xi) | \psi(\theta) \rangle$.

Al ser un POVM discreto y finito, la información de Fisher es una suma en vez de una integral,

$$F(\theta, \xi) = \frac{1}{p(\theta, \xi)(1 - p(\theta, \xi))} \left[\frac{\partial p(\theta, \xi)}{\partial \theta} \right]^2.$$

Vea que $p(\theta, \xi) = |\langle \psi(\xi) | \psi(\theta) \rangle|^2 = |\langle \alpha | e^{i\hat{n}(\theta-\xi)} | \alpha \rangle|^2$, escribiendo $|\alpha\rangle$ en la representación del espacio de Fock y haciendo las cuentas,

$$\begin{aligned} \langle \alpha | e^{i\hat{n}(\theta-\xi)} | \alpha \rangle &= e^{-|\alpha|^2} \sum_{n,m=0}^{\infty} \frac{\alpha^n (\alpha^*)^m}{\sqrt{n!} \sqrt{m!}} e^{in(\theta-\xi)} \langle m | n \rangle, \\ &= e^{-|\alpha|^2} \sum_{n=0}^{\infty} \frac{|\alpha|^{2n}}{n!} e^{in(\theta-\xi)}, \\ &= e^{|\alpha|^2(e^{i(\theta-\xi)} - 1)}. \end{aligned}$$

Entonces,

$$p(\theta, \xi) = |e^{|\alpha|^2(\cos(\theta-\xi) - 1 + i \sin(\theta-\xi))}|^2 = e^{2|\alpha|^2(\cos(\theta-\xi) - 1)}.$$

Vea que en este caso

$$\begin{aligned} \varphi[\lambda] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int \left(\frac{\partial e^{-\frac{\theta^2}{2\sigma^2}}}{\partial \theta} \right)^2 e^{\frac{\theta^2}{2\sigma^2}} d\theta, \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int \left(\frac{\theta}{\sigma^2} \right)^2 e^{-\frac{\theta^2}{\sigma^2} + \frac{\theta^2}{2\sigma^2}} d\theta, \\ &= \frac{1}{\sqrt{2\pi\sigma^5}} \int \theta^2 e^{-\frac{\theta^2}{2\sigma^2}} d\theta, \\ &= \frac{1}{\sqrt{2\pi\sigma^5}} \sqrt{2\pi\sigma^3}, \\ &= \frac{1}{\sigma^2}. \end{aligned}$$

Como dijimos antes, en este ejemplo vamos a maximizar sobre una familia de POVM's, no sobre todos los POVM's posibles, posteriormente veremos que esto basta. Entonces para calcular Z hacemos el siguiente cambio,

$$\underset{E(\xi)}{\text{Máx}} \rightarrow \underset{\xi}{\text{Máx}}. \quad (5.3)$$

Entonces, ya tenemos todos los ingredientes para calcular Z . Tenemos que

$$Z = \underset{\xi}{\text{Máx}} \left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \int \frac{e^{-\frac{\theta^2}{2\sigma^2}} 4|\alpha|^4 \sin^2(\theta - \xi) e^{2|\alpha|^2(\cos(\theta - \xi) - 1)}}{1 - e^{2|\alpha|^2(\cos(\theta - \xi) - 1)}} d\theta \right\} + \frac{1}{\sigma^2}. \quad (5.4)$$

Esta integral es complicada, sin embargo podemos hacer una simplificación. Suponiendo que $|\alpha|^2$ tiene un valor pequeño, haciendo

$$e^{2|\alpha|^2(\cos(\theta - \xi) - 1)} \rightarrow (1 + 2|\alpha|^2(\cos(\theta - \xi) - 1))$$

tenemos que,

$$\begin{aligned} Z - \frac{1}{\sigma^2} &\approx \underset{\xi}{\text{Máx}} \left\{ \frac{2}{\sqrt{2\pi\sigma^2}} \int e^{-\frac{\theta^2}{2\sigma^2}} |\alpha|^2 (\cos(\theta - \xi) + 1) d\theta \right\} \\ &= \underset{\xi}{\text{Máx}} \left\{ \frac{2}{\sqrt{2\pi\sigma^2}} \int e^{-\frac{\theta^2}{2\sigma^2}} |\alpha|^2 \left(\frac{e^{i(\theta - \xi)} + e^{-i(\theta - \xi)}}{2} \right) d\theta \right\} + 2|\alpha|^2 \\ &= \frac{2}{\sqrt{2\pi\sigma^2}} \underset{\xi}{\text{Máx}} \left\{ e^{i\xi} \int e^{-\frac{\theta^2}{2\sigma^2} - i\theta} \frac{|\alpha|^2}{2} d\theta + e^{-i\xi} \int e^{-\frac{\theta^2}{2\sigma^2} + i\theta} \frac{|\alpha|^2}{2} d\theta \right\} + 2|\alpha|^2 \\ &= \frac{2}{\sqrt{2\pi\sigma^2}} \underset{\xi}{\text{Máx}} \left\{ \left(\frac{e^{i\xi} + e^{-i\xi}}{2} \right) |\alpha|^2 \sqrt{2\pi\sigma^2} e^{-\frac{\sigma^2}{2}} \right\} + 2|\alpha|^2 \\ &= 2|\alpha|^2 (e^{-\frac{\sigma^2}{2}} + 1). \end{aligned}$$

Si hacemos este cálculo usando la información cuántica de Fisher, es decir, para el mismo estado dependiente de θ , $|\psi(\theta)\rangle$, tenemos, siguiendo a [3], que la información de Fisher cuántica es en este caso,

$$F_Q(\theta) = 4|\alpha|^2.$$

Como F_Q es una constante en este caso, vemos que

$$V_Q = 4|\alpha|^2 + \frac{1}{\sigma^2}. \quad (5.5)$$

Vemos que entonces, hay una relación interesante entre Z y V_Q , ya que como

$$Z = 2|\alpha|^2 (e^{-\frac{\sigma^2}{2}} + 1) + \frac{1}{\sigma^2}, \quad (5.6)$$

vemos que si σ es muy pequeño, la primera parte es similar a la información de Fisher cuántica, sin embargo, la parte de la información *a priori* es muy grande. Si σ , por el contrario, es muy grande, la información *a priori* es muy pequeña, casi cero, sin embargo, el primer sumando es la mitad de lo que resulta usando la información cuántica de Fisher.

Ahora bien, es importante justificar el uso de la familia particular de POVM's. Hicimos un cálculo numérico para observar que es una buena familia para maximizar. Tomamos un Hamiltoniano \hat{A} generado al azar por una matriz aleatoria perteneciente al ensamble GUE usando la medida de Haar, lo cual garantiza que todos los miembros del GUE son igualmente probables [19]. Después, tomamos la base ortonormal de los eigenvectores de \hat{A} y con esto formamos un conjunto completo de proyectores. Entonces, por cada Hamiltoniano aleatorio \hat{A} tenemos un POVM aleatorio $\{\hat{E}(\xi)\}$. Sabemos que estos no son todos los POVM's ya que solo tenemos medidas de Von Neumann, esto también se tomó en cuenta como veremos posteriormente. Entonces, con cada POVM aleatorio $\{\hat{E}(\xi)\}$ calculamos una $p(\xi|\theta) = \langle \psi(\theta) | \hat{E}(\xi) | \psi(\theta) \rangle$. Dada una distribución de probabilidad, hacemos el cálculo

$$F_A(\theta) = \sum_{\xi=1}^N \left(\frac{\partial \ln p(\xi|\theta)}{\partial \theta} \right)^2 p(\xi|\theta),$$

$$V_A = \int F_A(\theta) \lambda(\theta) d\theta + \frac{16}{\pi^2},$$

donde el subíndice A viene de que es una información de Fisher aleatoria, y N es la dimensión del Hamiltoniano aleatorio \hat{A} , que es por tanto igual a la dimensión del POVM aleatorio. Note aquí algo importante, como N es finita, y, la expresión que usamos del estado con el que hacemos el cálculo, es decir, la expresión de $|\psi(\theta)\rangle$ es infinita en principio, al imponer N estamos haciendo una aproximación finita de $|\psi(\theta)\rangle$. Entonces, si tomamos 100,000 valores aleatorios de V_A , para un valor suficientemente grande de N , el máximo del conjunto resultante debería aproximarse a la Z , que en principio no tiene porqué coincidir con el cálculo analítico que hicimos previamente. Lo que hicimos fue, dado un valor de $|\alpha|^2$, calcular V_A para 100,000 POVMs generados aleatoriamente y escoger el máximo. Si hacemos esto para distintos valores de $|\alpha|^2$ tenemos entonces que podemos hacer una identificación de un valor de Z^{Aprox} con cada valor de $|\alpha|^2$. En la figura 5.2 vemos los puntos negros, calculados numéricamente, una curva continua y una hecha de rayas pequeñas rojas. La curva continua es la gráfica numérica de la ecuación (5.4), recuerde que la distribución *a priori* era una Gaussiana centrada en π y con $\sigma = \pi/4$. La curva de rayas rojas es la parábola que se obtiene al graficar la

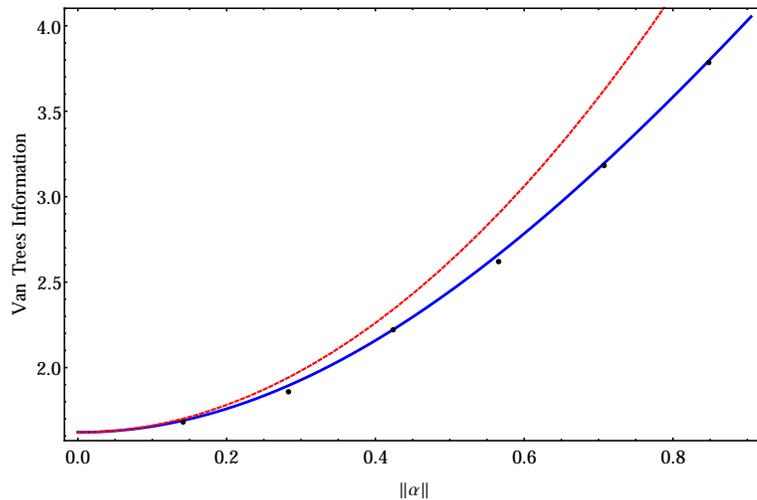


Figura 5.2: Puntos negros calculados numéricamente, las curvas azul continua y roja rayada corresponden a las ecuaciones (5.4) y (5.5) respectivamente.

ecuación (5.5).

Vemos una coincidencia muy buena de los puntos numéricos con la curva calculada analíticamente usando la familia de POVM's mencionada anteriormente. Corroboramos también que usando un POVM que no fuese una medición de Von Neumann, es decir, usando una base sobre-completa de proyectores, los puntos no subieran más que usando las bases completas. Vimos que usando bases sobre-completas los puntos subían menos con más iteraciones. Entonces, la evidencia numérica indica que *basta restringirnos* a esta familia de POVM's.

Se hizo otra prueba con variaciones pequeñas sobre el POVM obtenido con el muestreo aleatorio. Se avanzaba en todas las direcciones posibles del POVM para hasta que no subiera el valor de V . Usando este método los puntos no variaban perceptiblemente, por lo que asumimos que el muestreo bastaba.

5.2. Distribución uniforme y POVM continuo

Presentaré este ejemplo ya que fue interesante de realizar, sin embargo, no se obtuvo al final una correspondencia buena entre lo calculado numéricamente y lo hecho analíticamente. Supongamos que lo único que sabemos del parámetro que deseamos estimar, es que se encuentra en el intervalo $[0, 2\pi)$. Entonces, sin saber nada más, nuestra distribución de probabilidad *a priori*

es una distribución uniforme del intervalo $[0, 2\pi)$. Definiremos nuestro estado inicial en la base de Fock como sigue:

$$|\beta\rangle = \frac{1}{\sqrt{I_0(2|\beta|)}} \sum_{n=0}^{\infty} \frac{\beta^n}{n!} |n\rangle,$$

donde I_0 es la función modificada de Bessel de primera especie de orden 0. Se puede verificar fácilmente que este es un estado normalizado. Escogemos este estado porque simplifica las cuentas enormemente. Llamaremos al estado $|\beta\rangle$, estado exponencial β^1 . Vea que al igual que para los estados coherentes, hay uno de estos estados por cada número complejo. Entonces, el estado dependiente de θ en este caso será

$$|\psi(\theta)\rangle = e^{i\hat{n}\theta} |\beta\rangle.$$

Ahora bien, tratamos de Maximizar la información de Van Trees sobre todos los POVM's posibles, pero esto resulta difícil, ya que no conocemos manera de inspeccionar fácilmente el espacio de los POVMs. Tendremos que basarnos en la intuición y las simetrías del problema para encontrar este POVM. Observe que, dado que la distribución *a priori* es plana, le da la misma probabilidad a cualquier valor de θ , esto sugiere que el mejor POVM sea un conjunto de operadores de medición que incluya con el mismo cada posible valor de θ . Es decir, un POVM de proyectores que tome a todos los valores del parámetro θ como iguales y posibles. Dicho POVM es conocido, para presentarlo definiremos primeramente los estados de fase [20],

$$|\phi\rangle = \sum_{n=0}^{\infty} \frac{e^{in\phi}}{\sqrt{2\pi}} |n\rangle.$$

Entonces,

$$\begin{aligned} \int_0^{2\pi} |\phi\rangle\langle\phi| d\phi &= \int_0^{2\pi} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{e^{i\phi(n-m)}}{2\pi} |n\rangle\langle m| d\phi, \\ &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \delta_{n,m} |n\rangle\langle m|, \\ &= \sum_{n=0}^{\infty} |n\rangle\langle n|, \\ &= \mathbb{1}. \end{aligned}$$

¹ Ya existía en la literatura el estado exponencial $|\beta\rangle$ (con otro nombre), vea [20]. Sin embargo, esto yo lo supe *a posteriori*. La única razón de hacer el cálculo con dicho estado inicial es porque lo simplificaba.

Por lo que el conjunto $\{|\phi\rangle\langle\phi| \mid \forall \phi \in [0, 2\pi)\}$ es un POVM. Vea que este POVM tiene un proyector para cada valor de ϕ , y no prefiere a ninguno sobre otro (es posible demostrar que es el único con estas cualidades), así pues, es un POVM acorde al argumento anterior que alude a la intuición y la simetría. Entonces, ya que este parece ser un buen candidato para ser el POVM que maximiza V , supongamos que este es dicho POVM, el cálculo es completamente clásico, pues tenemos una sola distribución de probabilidad, a saber, $p(\phi|\theta) = |\langle\psi(\theta)|\phi\rangle|^2$. Llamaremos V^{tent} (V tentativa) a la información de Van Trees calculada con este POVM. Tenemos entonces que

$$p(\phi|\theta) = \frac{e^{2\operatorname{Re}\{e^{i(\theta-\phi)}\beta\}}}{2\pi I_0(2|\beta|)}.$$

Sacamos el logaritmo natural y derivamos,

$$\begin{aligned} \ln p(\phi|\theta) &= 2\operatorname{Re}\{e^{i(\theta-\phi)}\beta\} - \ln(2\pi I_0(2|\beta|)), \\ \frac{\partial \ln p(\phi|\theta)}{\partial \theta} &= ie^{i(\theta-\phi)}\beta - ie^{-i(\theta-\phi)}\beta^*, \\ &= -2\operatorname{Im}\{e^{i(\theta-\phi)}\beta\}, \\ \therefore \left(\frac{\partial \ln p(\phi|\theta)}{\partial \theta}\right)^2 p(\phi|\theta) &= \frac{e^{2\operatorname{Re}\{e^{i(\theta-\phi)}\beta\}}}{2\pi I_0(2|\beta|)} (2\operatorname{Im}\{e^{i(\theta-\phi)}\beta\})^2. \end{aligned}$$

Vea por un lado que ya que en este caso

$$\lambda(\theta) = \begin{cases} \frac{1}{2\pi} & \text{si } \theta \in [0, 2\pi) \\ 0 & \text{si } \theta \notin [0, 2\pi), \end{cases}$$

entonces

$$\varphi[\lambda(\theta)] = 0.$$

Podemos calcular V^{tent} en este caso como sigue,

$$V^{tent} = \int_0^{2\pi} \int_0^{2\pi} \frac{e^{2\operatorname{Re}\{e^{i(\theta-\phi)}\beta\}}}{2\pi I_0(2|\beta|)} (2\operatorname{Im}\{e^{i(\theta-\phi)}\beta\})^2 \frac{d\theta d\phi}{2\pi}.$$

Entonces, para este POVM, escribimos tentativamente,

$$V^{tent} = 2|\beta| \frac{I_1(2|\beta|)}{I_0(2|\beta|)}.$$

Es necesario verificar si el POVM que escogimos es en efecto el que maximiza V . Esto se hizo análogamente al caso anterior, es decir, numéricamente. En

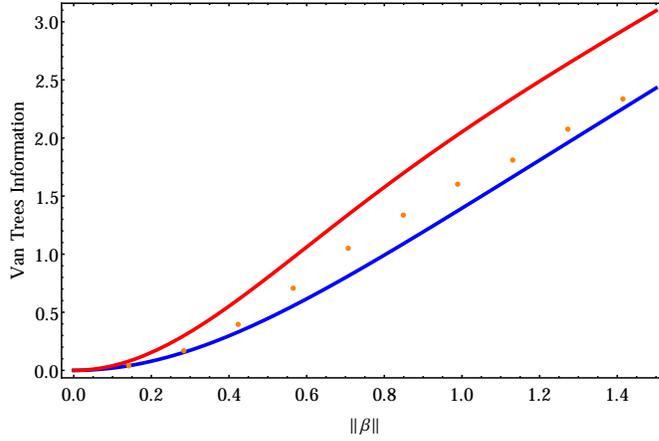


Figura 5.3: Puntos anaranjados calculados numéricamente, la curva azul es lo obtenido con V^{tent} , la curva roja es lo obtenido usando la información de Fisher para este problema.

este caso, observe que

$$F_A(\theta) = \sum_{\xi=1}^N \left(\frac{\partial \ln p(\xi|\theta)}{\partial \theta} \right)^2 p(\xi|\theta),$$

$$V_A = \int_0^{2\pi} \frac{F_A(\theta)}{2\pi} d\theta.$$

Entonces, hacemos muchas pruebas para muchos POVMs aleatorios y escogemos el valor máximo para la salida V_A . Esto se hace para varios valores de $|\beta|$. Graficamos el valor de este máximo como función de $|\beta|$. Vemos en la figura 5.3 que escogiendo POVMs al azar, hay algunas bases que alcanzan valores más altos para V_A . Entonces, a pesar de que V^{tent} parecía como un buen candidato para Z , vemos que hay POVMs que dan un valor mayor para diversos puntos, existe pues, otro POVM que nos da una información de Van Trees mayor.

Capítulo 6

Inferencia Bayesiana

En este capítulo mostraremos una consecuencia inmediata del hecho de que la definición propuesta en esta tesis (ecuación (4.4)) sea de naturaleza Bayesiana. En el contexto de la nueva definición es natural definir un proceso de inferencia Bayesiana en el que se cambia la creencia sobre la mejor estimación de un parámetro.

Primeramente presentamos un método de inferencia que se basa en la información de Fisher cuántica propuesta por Nagaoka, aunque Fujiwara demostró su convergencia asintótica [6]. Dicho método evita la contradicción intrínseca en el hecho de que para encontrar la información cuántica de Fisher sea necesario conocer *a priori* el parámetro a estimar y asintóticamente encuentra el valor del parámetro a estimar. Posteriormente presentamos un método alternativo que se basa en la inferencia Bayesiana y en la ecuación (4.4). Finalmente aplicamos ambos métodos al ejemplo de estimación de la fase en un oscilador armónico con una Gaussiana como distribución *a priori*, que presentamos en el capítulo anterior.

6.1. Método de Fujiwara

En [6] se presenta un método iterativo para estimar parámetros en los casos en que el POVM que maximiza la información de Fisher depende del valor del parámetro a estimar, es decir, θ . Para que dicho método funcione, es necesario saber el POVM que maximiza la información de Fisher para cada valor posible de θ . Entonces, de manera completamente arbitraria, se da una primera adivinanza, θ_0 , como primera estimación de θ . Se calcula la información de Fisher cuántica, ya que dado θ_0 , se sabe cual es el POVM que maximiza $F(\theta_0)$. Se tiene entonces, una distribución de probabilidad

asociada a este POVM. Supongamos ahora que se hace una medición cuyo resultado es x_0 . Entonces la función de verosimilitud (*likelihood function* en inglés) se define como

$$L_1(\theta) = p(x_0|\theta; E_{\theta_0}),$$

donde E_{θ_0} es el POVM que maximiza la información de Fisher para θ_0 . El máximo de esta función lo llamaremos θ_1 y será nuestra siguiente adivinanza. Entonces, sabiendo E_{θ_1} , se hace una medición con resultado x_1 , lo que nos da la siguiente función de verosimilitud

$$L_2(\theta) = L_1(\theta)p(x_1|\theta; E_{\theta_1}).$$

Entonces, para el paso n ,

$$L_n(\theta) = \prod_{i=0}^{n-1} p(x_i|\theta; E_{\theta_i}).$$

En [6] se demuestra que el máximo de L_n converge a el valor de θ cuando $n \rightarrow \infty$.

6.2. Inferencia Bayesiana

Se propone un método alternativo al de [6]. Primeramente, se tiene una distribución de probabilidad $\lambda(\theta)$ que codifica nuestro conocimiento del parámetro. Se calcula Z con un estado inicial dado. Se obtiene así el POVM que maximiza V , lo que nos define una distribución de probabilidad $p(\xi|\theta)$. Se puede realizar una medición con dicho POVM con resultado x_0 . Entonces, se aplica la regla de Bayes para obtener una nueva distribución de probabilidad *a priori*

$$\lambda_1(\theta|x_0) = \frac{p(x_0|\theta)\lambda(\theta)}{\int p(x_0|\theta)\lambda(\theta)d\theta}.$$

Se usa $\lambda_1(\theta)$ en lugar de $\lambda(\theta)$ para calcular Z . Se obtiene un nuevo valor para Z , lo que nos da un nuevo POVM. Entonces, se obtiene otra distribución de probabilidad y se puede hacer el proceso nuevamente. Se propone hacer este proceso n veces.

6.3. Ejemplo POVM de 2 elementos con una Gaussiana

Supusimos que en ambos casos se tiene un conocimiento *a priori* de un parámetro θ a estimar. Suponemos que θ es una fase, por lo que está en el

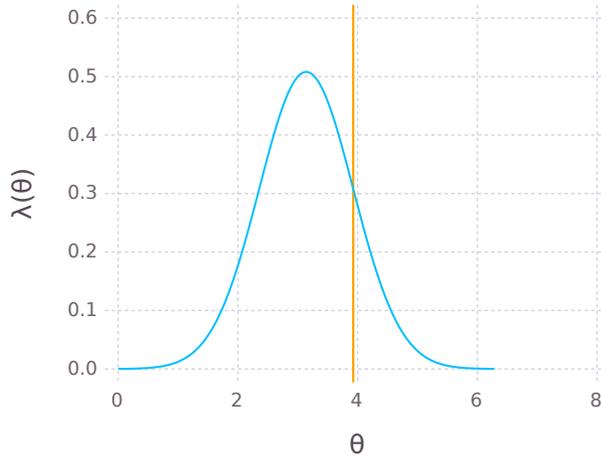


Figura 6.1: Planteamiento inicial del ejemplo.

intervalo $[0, 2\pi]$. Dicho conocimiento es que sabemos que hay una distribución de probabilidad λ que es una Gaussiana, centrada en π y con desviación estándar $\pi/4$. Cortamos la Gaussiana en 0 y en 2π y la normalizamos. La situación de nuestro ejemplo es la que se ilustra en la figura (6.1).

Suponemos que el valor en la naturaleza del parámetro a estimar es $5\pi/4$, por lo que, aunque λ nos da una buena aproximación al valor del parámetro, es posible mejorar el conocimiento que se tiene sobre el valor del parámetro θ .

Para el método de Fujiwara se tomó la función de Likelihood en cada paso y se normalizó. Observe que la función de Likelihood depende de todos los experimentos efectuados, es decir para el paso m , la función de Likelihood está dada en base a los resultados de los primeros $m - 1$ experimentos. Entonces, para calcular la información cuántica de Fisher del experimento en el paso m se multiplicó la función de Likelihood L_m a la distribución de probabilidad para considerar la información de los resultados anteriores, es decir, se calculó la información cuántica de Fisher en el paso m usando la distribución

$$\tilde{p}_m(\xi|\theta) = L_m(\theta)p(\xi|\theta).$$

El método de Van Trees se aplicó como se expuso en la sección previa. Vea que, a diferencia del método de Fujiwara, no es necesario agregar nada más, debido a que ya considera la información obtenida de mediciones previas.

Como expusimos anteriormente, en ambos métodos cada paso depende del valor medido en un paso para calcular la distribución de probabilidad que se usará en el paso siguiente. Sin embargo, las salidas son aleatorias, por lo que se hizo un análisis estadístico del problema. Entonces, calculamos, en ambos métodos todas las medidas posibles y les asignamos a cada una un peso estadístico. Dicho peso se calcula de la siguiente manera. Ya que en cada experimento hay dos posibles salidas, a las que designaré 0 y 1. Entonces, a cada paso el número de posibles salidas se duplica con respecto al anterior, puesto que cada resultado de medición dictamina el próximo POVM. Supongamos que los resultados de n experimentos sucesivos es $\{x_1, x_2, \dots, x_n\}$ donde $x_i \in \{0, 1\}$. Experimentos sucesivos en este contexto quiere decir que el POVM va cambiando usando los métodos antes expuestos, cada experimento depende del resultado de todos los anteriores. Entonces el peso estadístico que se le asigna a los resultados $\{x_1, x_2, \dots, x_n\}$ está dada por:

$$p(x_1|5\pi/4; E_{\theta_1})p(x_2|5\pi/4; E_{\theta_2}) \dots p(x_n|5\pi/4; E_{\theta_n}).$$

Es fácil ver que sumando todos los términos posibles hasta n nos da 1. A cada valor de Z se le pesa de esta forma y se suman. A cada función $F_Q(\theta)$ obtenida se le pone el peso correspondiente, y se encuentra su máximo en $[0, 2\pi)$. En la figura (6.2) vemos la comparación entre ambos métodos.

Vemos que método Bayesiano crece más para las primeras iteraciones. Además, los puntos del método de Fujiwara son del máximo de la información de Fisher, pues en este caso, la información de Fisher depende del valor del parámetro θ . En cambio, en el método Bayesiano, solamente hay un número asociado a esta medición.

Si el número de iteraciones tiende a infinito, se espera que tanto la información de Fisher cuántica F_Q , como la información de Van Trees Cuántica Z , tiendan a infinito de la misma manera, ya que la distribución *a priori* usada en el método Bayesiano tenderá a una delta de Dirac.

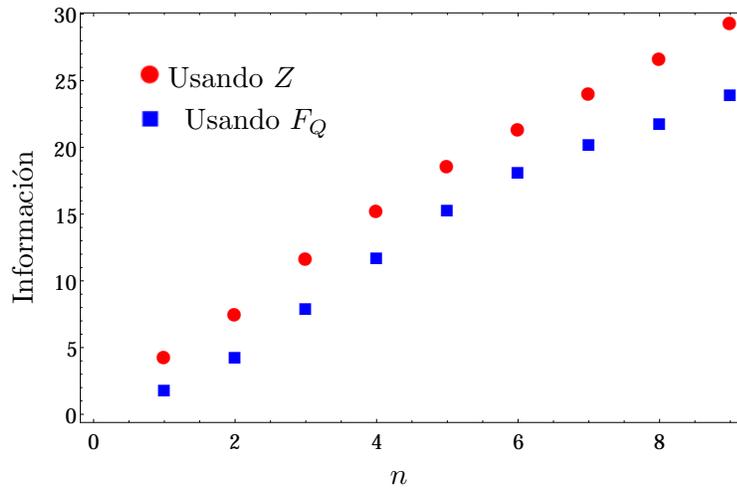


Figura 6.2: Los círculos rojos corresponden al método de inferencia Bayesiana mientras que los cuadros azules corresponden al método de Fujiwara.

Capítulo 7

Conclusiones

En esta tesis se analizó de manera general el problema de encontrar cotas mínimas para el error cuadrado medio en la estimación de parámetros de sistemas cuánticos. Expusimos la utilidad de usar la cota de Cramér-Rao. Revisamos cómo, en la literatura, se ha traducido a sistemas cuánticos en términos de la información de Helstrom H . Sin embargo, vimos la precariedad de H y la conveniencia de utilizar otra cota. Dicha cota involucra un caso límite de la información de Fisher F , a la que llamamos F_Q . Vimos que F_Q puede definirse matemáticamente de manera consistente, sin embargo, argumentamos que en general no se puede aplicar en un experimento concreto careciendo así de significado físico. Esto nos fuerza a buscar alternativas si es que deseamos encontrar cotas mínimas para el error en la estimación de parámetros de sistemas cuánticos.

Proponemos que debe usarse una versión Bayesiana de F_Q . Sin embargo, no podemos usar la información cuántica de Fisher generalizada V_Q (ya existente en la literatura) puesto que su cálculo depende de que podamos calcular F_Q . Además, aunque lo pudieramos calcular, esta cota no nos dice nada útil pues necesitamos un continuo de POVMs a la mano para su implementación en el laboratorio.

Proponemos una cota en términos del conocimiento que se tiene del parámetro a estimar θ a la que llamamos Z . Dicha cota, aunque depende del conocimiento *a priori* del experimentador, no presenta los problemas de F_Q ni de V_Q , se puede calcular en general.

Para ilustrar el cálculo de Z propusimos un ejemplo del cálculo de la fase de un estado coherente. Supusimos que conocíamos que dicha fase tiene una distribución gaussiana. Utilizamos un *ansatz*, una familia específica de POVMs que funcionaba para el cálculo de F_Q . Hicimos una prueba numérica para

saber si esta familia de POVMs contenía el POVM óptimo. Haciendo un muestreo aleatorio de POVMs encontramos evidencia de que nuestro *ansatz* es cierto en nuestro ejemplo.

Se hizo un cálculo explícito de otro ejemplo, en el que en vez de una gaussiana se tenía una distribución uniforme. Para este ejemplo se utilizó un POVM específico formado por los proyectores de los estados de fase para calcular una cantidad V^{tent} . La idea de este cálculo es que es un candidato para el valor de Z . Se hizo el cálculo analítico para un estado inicial. Haciendo el muestreo aleatorio se observó que el POVM específico no funcionaba, seguramente debe haber otro que resuelva el problema, sin embargo, no lo encontré.

Se propuso un método de inferencia Bayesiana, en el que, para cada medición el error disminuye, pues mejora nuestro conocimiento sobre el valor del parámetro a estimar θ . Este método viene de considerar Z , es decir, pensar en la cantidad de información que conocemos. Si pensamos en V_Q no es tan clara la interpretación que tendría un método de inferencia Bayesiana. Se comparó el método propuesto con un método alternativo que utiliza F_Q al que se llamó método de Fujiwara. Se observó que usando el método de inferencia Bayesiana se le gana por una medición en la cota del error al método de Fujiwara. Sin embargo yo tomé el valor más alto de F_Q en este caso. Recuerde que $F_Q = F_Q(\theta)$ entonces, como $\theta \in [0, 2\pi)$, nada nos asegura que el punto más alto sea el valor de θ . Estrictamente, tendría que haber tomado una especie de promedio, lo que reduciría el valor encontrado en los puntos para el método de Fujiwara.

Quedan muchas cosas por investigar. Por ejemplo, el problema de encontrar el POVM que alcanza Z para la distribución uniforme. Sospecho que el estado exponencial $|\beta\rangle$ será de alguna utilidad en la resolución de dicho problema. Un problema que se me hace muy interesante pero que es de aspecto puramente matemático es el de la maximización sobre todos los POVMs. No he encontrado ninguna teoría matemática que diga cómo buscar sistemáticamente en el espacio de POVMs. Una teoría sobre optimización en el espacio de POVMs sería de suma utilidad.

Una perspectiva a futuro es calcular Z así como el método de inferencia Bayesiana para una distribución que no sea tan uniforme como una Gaussiana, es decir, una para la que el máximo y el promedio de la distribución no coincidan. Puede que en estos casos se vea más claro la conveniencia de Z sobre F_Q . Podría usarse este tipo de métodos para calcular cotas de error al estimar cantidades complicadas en sistemas cuánticos como la intensidad de un campo magnético o eléctrico, o por ejemplo algún parámetro de un haz de

luz estructurado, etc. También es posible aplicar todo esto en el problema de estimación de parámetros un poco más abstractos, por ejemplo, la fidelidad o el entrelazamiento.

Claramente hay dos posibles vías de generalización de los métodos introducidos en esta tesis. Una, para el caso multiparamétrico finito. Como se ve en la literatura, la información de Fisher para varios parámetros es una matriz en vez de una función escalar [21]. Otra vía de generalización es elaborar una teoría para sistemas abiertos. Al principio presenté la regla de Born para matrices de densidad, sin embargo, me enfoqué en estados puros en esta tesis. Esta generalización está descrita en [3] usando F_Q .

En lo personal, una perspectiva a futuro que me llama la atención es investigar el área de Machine Learning en sistemas cuánticos. Machine Learning se basa en la estimación de parámetros a partir de un conjunto de datos. Es decir, hacer un modelo matemático a partir de datos. La desigualdad de Cramér-Rao así como los métodos Bayesianos son una parte esencial de Machine Learning. A mi modo de ver, es una línea que representa una continuación natural de lo que se hizo en esta investigación.

Bibliografía

- [1] B. P. Abbott *et al.* Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116:061102, Feb 2016.
- [2] S. Hacyan, R. Jauregui-Renaud, R. Jáuregui, and R. Lopez-Pena. *Latin-American School of Physics XXXI ELAF: New Perspectives on Quantum Mechanics*. AIP Conference Proceedings. American Inst. of Physics, 1999.
- [3] B. M. Escher, R. L. Matos Filho, and L. Davidovich. Quantum Metrology for Noisy Systems. *Brazilian Journal of Physics*, 41(4-6):229–247, September 2011.
- [4] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum-enhanced measurements: Beating the standard quantum limit. *Science*, 306(5700):1330–1336, 2004.
- [5] Samuel Braunstein and Carlton Caves. Statistical distance and the geometry of quantum states. *Physical Review Letters*, 72(22):3439–3443, May 1994.
- [6] Akio Fujiwara. Strong consistency and asymptotic efficiency for adaptive quantum estimation problems. *Journal of Physics A: Mathematical and General*, 39(40):12489, October 2006.
- [7] Esteban Martínez Vargas. *Información de Van Trees en la Metrología Cuántica*. Universidad Nacional Autónoma de México, 2014.
- [8] N.G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland Personal Library. Elsevier Science, 2011.
- [9] Sergios Theodoridis. *Machine Learning: A Bayesian and Optimization Perspective*. .Net Developers. Academic Press, 1 edition, 2015.

- [10] M.A. Nielsen and I.L. Chuang. *Quantum Computation and Quantum Information*. Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 2000.
- [11] MATTEO G. A. PARIS. QUANTUM ESTIMATION FOR QUANTUM TECHNOLOGY. *International Journal of Quantum Information*, 07(supp01):125–137, January 2009.
- [12] B.R. Frieden. *Science from Fisher Information: A Unification*. Cambridge University Press, 2004.
- [13] R. A. Fisher. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 222(594-604):309–368, 1922.
- [14] Harry L. Van Trees. *Detection, Estimation, and Modulation Theory*. John Wiley & Sons, April 2004.
- [15] C.W. Helstrom. Minimum mean-squared error of estimates in quantum statistics. *Physics Letters A*, 25(2):101–102, July 1967.
- [16] O. E. Barndorff-Nielsen and R. D. Gill. Fisher information in quantum statistics. *Journal of Physics A: Mathematical and General*, 33(24):4481, June 2000.
- [17] Richard D. Gill and Boris Y. Levit. Applications of the van trees inequality: a bayesian cramér-rao bound. *Bernoulli*, 1(1-2):59–79, March 1995.
- [18] S.D. Personick. Application of quantum estimation theory to analog communication over quantum channels. *IEEE Transactions on Information Theory*, 17(3):240–246, May 1971.
- [19] Christian P. Robert (auth.). *The Bayesian Choice: A Decision-Theoretic Motivation*. Springer Texts in Statistics. Springer New York, 1994.
- [20] F. Soto-Eguibar, B. M. Rodríguez-Lara, and H. M. Moya-Cessa. Phase state and related nonlinear coherent states. *J. Opt. Soc. Am. B*, 31(6):1335–1338, Jun 2014.
- [21] Alexander S. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory*. Springer Science & Business Media, May 2011.