Estrategias de aprendizaje automático para discriminación de estados fotónicos.

Tomás Crosta

Director: Dr. Matías Bilkis

Asesor académico: Prof. Dr. Juan Mauricio Matera.

Mesa Jurada conformada por el Prof. Dr. Juan Mauricio Matera, Prof. Dr. Raul Rossignoli y el Prof. Dr. Carlos Lamas

Tesis de Licenciatura para optar el título de grado de Licenciado en Física

Departmento de Física de La Plata
Facultad de Ciencias Exactas,
Universidad Nacional de La Plata
La Plata, Pcia. de Buenos Aires, Argentina
7 de diciembre de 2023

Agradecimientos

En primer lugar, quiero agradecer a mi director, Matías Bilkis, por su valioso apoyo a lo largo de toda la tesis, su dedicación y ayuda cuando organizo proyectos futuros. También a Lorena Rebón por realizar la codirección de la tesis, contribuyendo en todas las reuniones del trabajo, y a mi asesor académico, Mauricio Matera, por sus comentarios y correcciones de este trabajo. Aparte de este trabajo, agradezco a Matias Tueros, por la buena onda y ayudarme a mejorar mis códigos, además de los test necesarios para validarlos.

Por sobre todo, quiero agradecer a mis papás por ayudarme, cuidarme y darme todo lo necesario para poder seguir lo que me gusta. Quiero reconocer también el gran trabajo que les costó no preguntarme sobre los exámenes finales, a pesar de las constantes fallas en esta área. También tengo que agradecer a mi hermana, con quien aunque no tenga muchos intereses en común, logró hacer todos los días mucho más ruidosos y entretenidos.

Al resto de mi familia, a mi abuela, Marito, Mirta y a mi abuelo que desafortunadamente no pudo llegar al último tramo de la carrera, gracias por estar ahí ayudando todos estos años. También quiero agradecer por estar siempre a las familias, Bernheim, Compalati y Penzi.

A todos los miembros de La Vagancia por estar desde el primer día de la carrera y ayudarme constantemente a entender los ejercicios y problemas que no salían. Al grupo de Quilmes, Martín y Redru, por una amistad de tantos años, incluso cuando desaparezco durante muchos meses. También quiero agradecer a Dido por bancarme desde el ingreso cuando no conocía a nadie y ahorrarme tantas horas de viaje, dándome un lugar para dormir durante las cursadas. A Jean por explicarme un montón de dudas, incluso en vacaciones, y a Tomi por acompañarme tanto en las últimas partes del trayecto y ayudando a armar diferentes proyectos.

Finalmente, a Celeste por darme un año tan lindo, ayudando a calmarme en momentos de estrés y escuchándome cuando me enojaba con alguna de todas las personas que les acabo de agradecer.

Índice general

Resumen vii

1 Introducción 1

2 Información cuántica & sistemas de variable continua 3
- 2.1 Operador de densidad . . . 3
  - 2.1.1 Sistemas Compuestos y Estados Reducidos . . . 4
  - 2.1.2 Evolución temporal . . . 4
  - 2.1.3 Observables y medidas . . . 4
- 2.2 Sistemas de variable continua . . . 6
  - 2.2.1 Sistemas Gaussianos . . . 6
  - 2.2.2 Transformaciones Gaussianas . . . 8
  - 2.2.3 Proyecciones sobre estados Gaussianos . . . 9

3 Discriminación de estados & modelado de canales reales. 11
- 3.1 Distinción de estados cuánticos . . . 11
  - 3.1.1 Límite de Helstrom . . . 12
- 3.2 Experimento de comunicación . . . 12
- 3.3 Modelos de canales ruidosos . . . 16
  - 3.3.1 Canales Gaussianos . . . 17
  - 3.3.2 Fondos no Gaussianos y conclusiones . . . 20

4 Inteligencias artificiales 23
- 4.1 Aprendizaje reforzado . . . 24
  - 4.1.1 Optimización de la política . . . 26
  - 4.1.2 Estimación de $V_\pi(s)$ . . . 27
- 4.2 Redes Neuronales . . . 27
  - 4.2.1 Inconvenientes del descenso por gradiente . . . 28

5 Implementación de inteligencias artificiales 31
- 5.1 Entorno y agentes . . . 32
  - 5.1.1 Agente: Sr. Magoo . . . 33
  - 5.1.2 Cambios en la energía de la señal . . . 34
  - 5.1.3 Una IA para gobernarlas a todas . . . 36

6 Conclusiones 39

7 Apéndice 41
- A Realización Teórica-Experimental . . . 41
- B Discriminaciones óptimas . . . 43

Bibliografía 47

Resumen

El primer objetivo de este trabajo será estudiar el uso de técnicas de aprendizaje automático, para la discriminación de estados coherentes en canales reales.

Para completar este objetivo, en el trabajo se repasarán los fundamentos de la mecánica cuántica, enfocándonos a dar una breve descripción de la computación cuántica en variables continuas. Una vez estudiado esto, describiremos un problema canónico de la literatura, donde las comunicaciones se hacen mediante la discriminación de estados con fases diferentes. De esta forma, se verán esquemas para diferenciarlos, como son las detecciones Homodinas, Kennedy, Kennedy optimizadas y Dolinar, comparando así las probabilidades de éxito de cada una en canales ideales. Posteriormente, se modelarán diferentes fuentes de fondo, característico en comunicaciones reales, observando así cambios en las probabilidades de éxito y variaciones en los parámetros óptimos del método de discriminación empleado. Esto motivará el uso de técnicas de aprendizaje automático que realizan optimizaciones a partir de medidas directas en vez de un modelo teórico, debido a que suele ser imposible una correcta caracterización de todos los fondos presentes en una comunicación. Para poder comenzar a implementar este tipo de optimizaciones, tendremos que estudiar la teoría necesaria para la implementación de los diferentes agentes optimizadores y como evolucionan las decisiones tomadas.

Luego de sentar estos conceptos, se verán 3 implementaciones de inteligencias artificiales para la calibración de un detector de tipo Kennedy en canales reales. Se discutirán las ventajas de las diferentes implementaciones en lo referente a tiempos de entrenamiento y formas de adaptarse a cambios en el entorno. En primer lugar, se seguirá a un agente ciego al entorno, el cual tendrá una rápida convergencia y con decisiones prácticamente iguales al límite teórico. Pero este agente no podrá adaptarse a ningún cambio en el sistema, teniendo que ser entrenado cada vez que este evoluciona. Luego se verá un agente que toma como observación al $|\alpha|$, pudiendo así adaptarse a variaciones en la intensidad del estado enviado. La última inteligencia será diseñada teniendo en cuenta que ninguno de los agentes descritos anteriormente se pueden adaptar a cambios en el fondo. Para mejorar este aspecto se permitió a un último agente utilizar como observación a todos los parámetros obtenidos de una serie de experimentos anteriores.

1 Introducción

Jamás hablamos del futuro
Solo el presente era seguro
El resto humo
Cavernicolas, eso fuimos.

Ricardo Arjona.

Durante los últimos 20 años, la facilidad de enviar información sobre largas distancias generó cambios muy importantes en la vida de las personas. Un ejemplo de esto es la posibilidad de comunicarnos con nuestro director de tesis al otro lado del mundo de forma casi instantánea. Así cambió la disponibilidad de información, la forma de socializar e incluso alteró la economía mundial [1].

Estos avances requirieron no solo una infraestructura tecnológica muy grande, sino que también fue sumamente importante desarrollar un formalismo matemático. Haciendo uso de estas herramientas fuimos capaces de entender los fundamentos teóricos de la información, permitiendo desarrollar nuevos y mejores protocolos para su transmisión y procesamiento [2].

La teoría de la información permite entonces determinar límites para los cuales la información puede ser transmitida a través de un canal, así como protocolos para alcanzar estas cotas [3]. Sin embargo, este paradigma está basado en la física clásica, la cual está limitada a la hora de describir determinados sistemas.

Separadamente, desde 1925 hasta la actualidad, se desarrolló una de las teorías con mayor poder predictivo del siglo XX: la mecánica cuántica. Esta teoría nos permite describir sistemas tan pequeños como un átomo y tan colosales como estrellas de neutrones [4]. Este mayor entendimiento de la materia fue fundamental para desarrollar tecnologías como la energía nuclear, medicamentos, materiales semiconductores, etc.

Teniendo en cuenta los avances que la mecánica cuántica supuso en muchas áreas, algunos investigadores, como von Neumann, notaron que el desarrollo matemático de la teoría de la información estaba basado en teorías clásicas. En este sentido, la teoría cuántica de la información surge como una reformulación (tanto de la física cuántica como de la teoría de la información). Entendiendo así a los sistemas físicos (incluidos los cuánticos) como información, y desarrollando una serie de métodos para describir la manipulación de sistemas cuánticos, como extensión de protocolos en teoría de la información. En este sentido, la explotación de recursos cuánticos supone una promesa para el desarrollo de nuevas tecnologías, que superan los límites de las actuales (por ejemplo, una computadora o un teléfono) a la hora de procesar información. De forma que existen tecnologías cuánticas como las computadoras y las comunicaciones cuánticas, que permiten (al menos teóricamente) implementar protocolos de una forma exponencialmente más eficiente que su análogo clásico. En esta tesis nos centraremos en estudiar un problema canónico dentro el campo de comunicaciones cuánticas.

En este sentido, la teoría cuántica de la información plantea nuevos retos, en particular a la hora de distinguir información. Uno de los resultados fundamentales de esta teoría es que dos estados cuánticos no ortogonales son, en general, imposibles de distinguir con probabilidad de éxito uno, incluso siendo ambos estados puros. Es por esto que uno de los problemas más importantes a la hora de transmitir información cuántica es el de discriminar estados con la mayor probabilidad de éxito posible. Esto a su vez maximiza las posibilidades de poder transmitir información entre una parte y otra de forma correcta. Un ejemplo de esto son las comunicaciones de larga distancia utilizando pulsos láser, *e.g.* comunicación satélite-Tierra. En este escenario, información clásica (por ejemplo, el valor de un bit) es *codificada* en el estado cuántico, que es enviado desde el satélite a la Tierra, a través de la atmósfera. La señal que llega a la Tierra debe ser *decodificada*; sin embargo, es tan débil que efectos cuánticos (como la imposibilidad de distinguir perfectamente una señal

de otra) se vuelven dominantes. Es por esto que es de suma importancia desarrollar métodos para poder incrementar la probabilidad de discriminación correcta entre señales, sean tanto referidos a la implementación experimental de medidas cuánticas óptimas, como a la calibración de instrumentos subóptimos en entornos experimentales adversos.

En este trabajo estudiaremos dichos sistemas haciendo medidas sobre observables de variable continua, y en particular nos concentraremos en el problema de la discriminación entre dos estados coherentes con diferente fase. Este es un problema canónico en comunicaciones cuánticas sobre largas distancias, que plantea desafíos experimentales tanto en la implementación de medidas óptimas, como en su calibración en entornos adversos. En este sentido, si bien el protocolo de discriminación entre dos estados coherentes es conocido desde los inicios de la información cuántica [5], su implementación experimental resulta menos trivial. Más allá de ciertas pruebas de concepto llevadas a cabo recientemente [6, 7], la transmisión de información se realiza mediante canales como la atmósfera, cuya caracterización resulta difícil de modelar. Por ejemplo, dependiendo de las condiciones climáticas, la medida óptima será potencialmente diferente, ya que los estados que arriben al instrumento detector tendrán otras características [8]. Más aún, diversas fuentes de ruido, como absorciones, cuentas oscuras en el instrumento de detección, pueden potencialmente degradar la calidad de la trasmisión.

Recientemente se han desarrollado métodos para calibrar de manera agnóstica el instrumento de medida [9], basados en técnicas de aprendizaje automático y por refuerzo. En este sentido, durante los últimos 10 años surgió una gran herramienta para resolver problemas donde no conocemos todas las variables del sistema. Mediante observaciones del entorno y sin un modelo (que potencialmente puede funcionar como un sesgo), este tipo de algoritmos permite encontrar configuraciones óptimas mediante prueba y error del experimento. Estos algoritmos lograron revolucionar todo tipo de áreas, desde ajedrez con AlphaZero hasta autos que se manejan solos [10, 11].

La propuesta principal de este trabajo será la de emplear diferentes algoritmos de inteligencia artificial, ampliando las capacidades de los propuestos en la referencia [9]. Las características principales de los mismos es que no están dotadas de un modelo sobre el problema, sino que el algoritmo de inteligencia artificial debe calibrar un instrumento de detección de señales mediante prueba y error. Esto lo hace intentando minimizar el número de repeticiones del experimento, pero maximizando los aciertos a la hora de distinguir las señales. En este sentido, nos enfocaremos en la calibración agnóstica del detector de Kennedy, que consiste en generar un desplazamiento de la señal, y medir el estado resultante mediante un fotodetector [12]. Finalmente, estudiaremos la eficacia de los métodos propuestos en canales ruidosos.

Para cumplir con estas metas, la tesis está estructurada de la siguiente manera. Al terminar este capítulo introductorio, repasaremos los fundamentos teóricos necesarios para estudiar el problema de discriminación de estados coherentes. En el siguiente capítulo haremos una breve introducción sobre discriminación de estados coherentes con fases diferentes y la optimización de parámetros en diferentes métodos de detección. Además, se discutirán algunos modelos de canales ruidosos que pueden estar presentes en comunicaciones reales. Esto motivará la necesidad de usar modelos de inteligencia artificial, capaces de encontrar parámetros óptimos de discriminación mediante experimentación y sin emplear un modelo teórico. Para poder desarrollar implementaciones de esto, en el cuarto capítulo veremos los fundamentos de inteligencia artificial, definiendo las diferentes recompensas y formas en que un agente puede mejorar sus decisiones. Sentadas las bases teóricas de inteligencias artificiales, en el último capítulo estudiaremos tres implementaciones diferentes para aplicarlas en el problema de discriminación de estados con fases diferentes. Comenzaremos trabajando con agentes ciegos al entorno, que requieren realizar un entrenamiento cada vez que el sistema cambia. Luego, se estudiarán agentes que puedan adaptarse a cambios del entorno mediante observaciones, sin necesidad de realizar nuevamente el entrenamiento. Finalmente, haremos una discusión sobre el trabajo realizado y una ruta sobre posibles caminos futuros.

2 Información cuántica & sistemas de variable continua

> Para poner a prueba la realidad, hemos de verla en la cuerda floja. Cuando las verdades se hacen acróbatas podemos juzgarlas.
>
> *El retrato de Dorian Gray.*

En este capítulo haremos un breve repaso del formalismo cuántico para luego discutir sistemas de variable continua. Este formalismo describe, entre otros, los sistemas fotónicos. Empezaremos por las definiciones básicas en el campo de mecánica cuántica, como los conceptos de estados cuánticos, evolución temporal, observables y medidas cuánticas. Luego estudiaremos sistemas de variable continua, donde la clase de estados Gaussianos toma un rol prominente a la hora de estudiar analíticamente ciertos sistemas que se consideran *lineales*, debido a su caracterización en el espacio de fases. Además, dichos sistemas son experimentalmente accesibles, conformando lo que en la literatura es conocido como el *lineal-optics toolkit*. De esta manera, nuestra discusión sobre sistemas de variable continua prosigue con la definición de estados Gaussianos y operaciones Gaussianas, para finalmente concluir con métodos de distinción de estos estados en el caso de comunicaciones.

2.1. Operador de densidad

Un sistema $\mathcal{S}$ es descrito en mecánica cuántica mediante los observables que tiene asociados. En este formalismo, la distribución de probabilidades de todo observable puede ser obtenida mediante un operador de densidad $\hat{\rho}$. Este operador es autoadjunto $\hat{\rho} \in B(\mathcal{H})$, semidefinido positivo ($\hat{\rho} \geq 0$) y de traza uno $\mathrm{Tr}[\hat{\rho}] = 1$. Estos operadores viven sobre un espacio de Hilbert d-dimensional $\mathcal{H}$; donde la dimensión d puede ser tanto finita como infinita.

Si nos restringimos a un sistema de dimensión finita, dada una base ortonormal $\{|i\rangle\}_{i=1}^{d}$, un estado cuántico puede ser escrito como

$$\hat{\rho} = \sum_{i,j} \rho_{ij} \, |i\rangle\langle j| \, , \tag{2.1}$$

donde ρ_{ij} son las componentes de la matriz $\hat{\rho}$. Como $\hat{\rho} \geq 0$, el operador es autoadjunto, admitiendo así una descomposición espectral en una autobase asociada $\{|k\rangle\}$, de forma que

$$\hat{\rho} = \sum_{k} p_k \, |k\rangle\langle k| \, , \quad 0 \leq p_k \leq 1, \quad \sum_{k} p_k = 1, \tag{2.2}$$

las condiciones sobre p_k surgen del hecho que $\hat{\rho}$ es semi definida, positiva y de traza 1.

La pureza de un estado se define según $\mathrm{Tr}[\hat{\rho}^2]$. Siguiendo esta definición, los *estados puros* están representados por operadores cuya pureza sea 1 y los *estados mixtos* son todos los que tienen una pureza estrictamente menor que 1.

2.1.1. Sistemas Compuestos y Estados Reducidos

Se puede notar que estos operadores pueden dar información no solo del sistema, sino que también sobre los subsistemas del mismo. Si denotamos a estos como A y B, se pueden escribir las matrices de densidad $\hat{\rho}_A$ y $\hat{\rho}_B$ tales que estas determinan los valores de expectación locales de A y B respectivamente. En particular, si el estado es separable, la matriz de densidad del sistema conjunto viene dada por una combinación convexa de estados producto:

$$\hat{\rho} = \sum_k p_k \, \hat{\rho}_A^{(k)} \otimes \hat{\rho}_B^{(k)} \, . \tag{2.3}$$

Sin embargo, existen estados cuánticos que no son separables, y que en general presentan correlaciones cuánticas. Por ejemplo, los estados de Bell

$$|\psi_0\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} \, ; |\psi_1\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}} \, ; |\psi_2\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}} \, ; |\psi_3\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}} \, , \tag{2.4}$$

no pueden ser escritos de la forma de la ecuación (2.3).

Para obtener información del estado de los subsistemas, se define la *traza parcial* tal que

$$\hat{\rho}_A := \mathrm{Tr}_B \, \hat{\rho}_{AB} \, . \tag{2.5}$$

De esta forma, la matriz de densidad $\hat{\rho}_A$ puede determinar todos los valores de expectación de los observables locales del subsistema.

Siguiendo esta definición, el operador de densidad $\hat{\rho}_A$ tiene que contener toda la información del estado del subsistema. Esto implica que $\langle i|\hat{\rho}_A|j\rangle = \sum_k \langle ik|\hat{\rho}_{AB}|jk\rangle$, donde $|ik\rangle \equiv |i\rangle \otimes |k\rangle$ son los estados de una base producto ortonormal de $H_A \otimes H_B$. Por esta razón, la traza parcial es determinada mediante la expresión

$$\mathrm{Tr}_A(\hat{\rho}_{AB}) = \sum_r \langle r| \, \hat{\rho}_{AB} \, |r\rangle \, . \tag{2.6}$$

Una vez conocido esto, estamos interesados en ver como evolucionan los operadores de densidad ante la evolución temporal.

2.1.2. Evolución temporal

La evolución de un sistema cuántico cerrado está dado por una transformación unitaria $\hat{U}_t$, generada por un Hamiltoniano $\hat{H}$. Escribiendo la ecuación de Schödinger, para una matriz de densidad $\hat{\rho}(t)$ tenemos

$$\partial_t \hat{\rho}(t) = -i[\hat{H}, \hat{\rho}(t)], \tag{2.7}$$

donde la matriz de densidad en función del tiempo puede ser escrita como

$$\hat{\rho}(t) = \hat{U}_t \hat{\rho}(0) \hat{U}_t^\dagger . \tag{2.8}$$

Si el Hamiltoniano no depende explícitamente del tiempo entonces en particular $\hat{U}_t = e^{-i\hat{H}t}$. Esto muestra que el Hamiltoniano es el generador de evoluciones temporales, y su naturaleza cumplirá un rol fundamental a la hora de caracterizar sistemas de variable continua Gaussianos. Pero antes de ahondar en esto, repasaremos el concepto de medida en mecánica cuántica.

2.1.3. Observables y medidas

Todas las cantidades físicas, tales como la posición, el momento angular o la energía, son descritos por operadores autoadjuntos, conocidos como observables. Como tales, tienen autoestados y un observable asociado, correspondientes a los estados.

El espectro de un observable representa entonces los posibles resultados de una medida individual, mientras que la distribución de probabilidad de esos resultados se relacionan con el estado del sistema vía la relación $\langle A \rangle = \mathrm{Tr}(\rho A)$. Esta relación matemática funciona en dos sentidos: el estado inicial del sistema es inferido a partir de los valores medios observados, y los resultados de nuevas medidas son predichos en términos del estado conocido.

Estos observables son una forma concisa de describir un concepto más general, las *medidas proyectivas*. Una medida proyectiva representa una función lineal del estado, que devuelve una distribución de probabilidades. Esta función se puede expresar como combinación de proyectores E_k. De forma que si asignamos valores reales λ_k a cada uno de los resultados posibles k, entonces todos los observables pueden ser escritos de la forma $\hat{A}_k = \sum_k \lambda_k \hat{E}_k$. Estos definen una medida tal que se cumple la relación $\langle \hat{A} \rangle = \mathrm{Tr}(\hat{\rho}\hat{A})$.

En particular, las *Medidas proyectivas* ortogonales consisten en proyecciones $\{E_k\}$ tales que $E_k E_j = \delta_{kj} E_j$, con $\sum_k E_k = \mathbb{I}_d$ y $E_k = E_k^\dagger$. Los posibles resultados de la medida son $\{k\}_{k=0}^{d-1}$, y la probabilidad de medir el resultado k está dado por el estado del sistema $\hat{\rho}$, según

$$p(k) = \mathrm{Tr}[\hat{\rho} E_k]. \tag{2.9}$$

Además de esto, siguiendo el tercer postulado de la mecánica cuántica, el estado del sistema después de la medida estará dado por

$$\hat{\rho} \to \rho_{|k} = \frac{E_k \hat{\rho} E_k}{\mathrm{Tr}[\hat{\rho} E_k]}, \tag{2.10}$$

donde notamos que la normalización está dada por la probabilidad de la medida obtenida. Por el contrario, si la medida es realizada, pero el resultado no es conocido, el *estado no condicional* está dado por

$$\hat{\rho} \to \sum_k E_k \hat{\rho} E_k. \tag{2.11}$$

Por ejemplo, consideremos a un observable $\hat{O}$. Podemos inmediatamente construir una medida proyectiva mediante su descomposición espectral, $\hat{O} = \sum_k o_k \left|k\right\rangle\left\langle k\right|$, llevando a $E_k = \left|k\right\rangle\left\langle k\right|$ y

$$\langle O \rangle_{\hat{\rho}} = \mathrm{Tr}[\hat{O}\hat{\rho}] = \sum_k p_k o_k, \tag{2.12}$$

donde $p_k = \mathrm{Tr}[\hat{\rho} \left|k\right\rangle\left\langle k\right|]$ es la probabilidad de encontrar al estado $\hat{\rho}$ en el k^{th} estado de $\hat{O}$.

Así mismo, si consideramos un sistema compuesto ρ_{AB}, y un observable de la forma $\hat{O} = O_A \otimes \mathbf{I}_B$, los valores de expectación locales pueden ser calculados según

$$\langle \hat{O}_A \rangle = \mathrm{Tr}\, \hat{\rho}_{AB} \hat{O}_A = \mathrm{Tr}_A\, \hat{\rho}_A \hat{O}_A\,. \tag{2.13}$$

Sin embargo, las medidas proyectivas no son las más generales. En general, una medida cuántica está constituida por una POVM (Positive Operator-Valued Measure), la cual surge cuando una medida proyectiva es llevada sobre un auxiliar con el que el sistema está acoplado. En general, estas medidas cuánticas generalizadas están dadas por una serie de operadores lineales $\{E_k\}_{k=1}^n$, cada uno asociado a un resultado posible $k \in \{1, ..., n\}$. La probabilidad de que el resultado k suceda condicionada a que el estado cuántico se encuentre como $\hat{\rho}$, viene dado por la regla de Born

$$p(k|\hat{\rho}) = \mathrm{Tr}[\hat{\rho} E_k]. \tag{2.14}$$

Debido a que esa probabilidad requiere ser positiva para todo estado cuántico, $E_k \geq 0$, son autoadjuntos. Aún más, ya que $\sum_k p(k|\hat{\rho}) = 1$ para cualquier estado, entonces $\sum_k E_k = \mathbb{1}_d$. Esto constituye la *Positive Operator-Valued Measure (POVM).*

En lo siguiente estudiaremos sistemas de variable continua, cuyo espacio de Hilbert asociado tiene dimensión infinita. El formalismo que veremos es útil a la hora de describir sistemas fotónicos (ver Apéndice A), como los que estudiaremos en siguientes capítulos referidos a la discriminación de estados de luz.

2.2. Sistemas de variable continua

Un estado cuántico de variable continua es definido como un sistema cuyos n grados de libertad, son descritos con $2n$ operadores canónicos $\hat{\boldsymbol{r}} = \left(\hat{q}_1, \hat{p}_1, ..., \hat{q}_n, \hat{p}_n\right)^T$, estos obedecen las relaciones canónicas de conmutación

$$\left[\hat{q}_j, \hat{p}_k\right] = i\hbar\delta_{jk}, \quad j, k = 1, ..., n. \tag{2.15}$$

Es posible notar que también podemos definir operadores de creación y destrucción como $\hat{a}_j = \frac{\hat{q}_j + i\hat{p}_j}{\sqrt{2}}$, y $\hat{\boldsymbol{a}} = \left(\hat{a}_1, \hat{a}_1^\dagger ..., \hat{a}_n, \hat{a}_n^\dagger\right)^T$.

Cada grado de libertad canónico j define un modo, asociado a un espacio de Hilbert de dimensión infinita $\mathcal{H}_j$. De forma que un sistema de n-modos está conformado por un espacio de Hilbert $\otimes_{j=1}^n \mathcal{H}_j$.

Un ejemplo muy empleado a la hora de discutir sistemas de variable continua es el oscilador armónico cuántico uni-dimensional, cuyo Hamiltoniano es $H = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{q}^2 = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right)$. Los autoestados de dicho Hamiltoniano constituyen la base de Fock $\{|n\rangle\}_{n=0}^\infty$, y la acción de los operadores de aniquilación y creación sobre dichos estados viene dada por que

$$\begin{aligned} \hat{a}_l\,|n_l\rangle &= \sqrt{n_l}\,|n_l - 1\rangle \\ \hat{a}_l\,|0\rangle &= 0 \\ \hat{a}_l^\dagger\,|n_l\rangle &= \sqrt{n_l + 1}\,|n_l + 1\rangle\ . \end{aligned}$$

Así podemos notar que los estados $|n\rangle$ resultan autoestados de un operador $\hat{N} = \hat{a}^\dagger\hat{a}$, de forma que

$$\hat{N}\,|n\rangle = n\,|n\rangle\ . \tag{2.16}$$

Si hacemos este análisis sobre campos electromagnéticos, una cantidad que resulta de interés son las *cuadraturas de campo*. Estas pueden ser interpretadas como posición y momento del campo electromagnético, siendo así definidas por

$$\hat{q} = \frac{1}{\sqrt{2\hbar}}\left(\hat{a} + \hat{a}^\dagger\right) \qquad \hat{p} = \frac{i}{\sqrt{2\hbar}}\left(\hat{a} - \hat{a}^\dagger\right). \tag{2.17}$$

En el Apéndice A se puede ver la cuantización del campo electro magnético, donde es de notar que este Hamiltoniano resulta cuadrático en las cuadraturas del campo, y que el autoestado fundamental $|0\rangle$ (es decir, el vacío) tiene una función de onda $\psi(x) = \langle x|\psi\rangle$ cuyo perfil es Gaussiano. En efecto, dicho estado es un ejemplo de la clase de estados Gaussianos, generados a partir de Hamiltonianos cuadráticos que discutiremos a continuación.

2.2.1. Sistemas Gaussianos

Se pueden definir a los estados Gaussianos como todos los estados cuánticos térmicos de un Hamiltoniano cuadrático

$$\rho_G = \frac{e^{-\beta\hat{H}}}{\mathrm{Tr}[e^{-\beta\hat{H}}]}, \quad \beta \in \mathbb{R}^+, \tag{2.18}$$

$$\hat{H} = \frac{\hat{\boldsymbol{r}}^T H \hat{\boldsymbol{r}}}{2} + \hat{\boldsymbol{r}}^T \vec{\boldsymbol{r}}, \tag{2.19}$$

donde $\vec{r}$ es un vector constante en el espacio de fases y H es la *Matriz Hamiltoniana*, la cual es definida positiva y simétrica. Notar que esta definición incluye al límite $\beta \to \infty$, que corresponde al caso de estados puros. Tal construcción nos permite encontrar los llamados modos normales de ρ_G, en donde el sistema se comporta como una serie de n osciladores armónicos desacoplados.

En particular, cualquier sistema Gaussiano puede entenderse como un sistema de osciladores armónicos desacoplados, junto con una transformación Gaussiana que los acopla [13]. Dicha transformación Gaussiana está generada por Hamiltonianos cuadráticos en las cuadraturas del campo, y puede descomponerse en ciertas transformaciones fundamentales (desplazamientos, squeezing, cambios de fase y beam-splitters) que revisaremos a continuación. Sin embargo, antes de esto, es conveniente definir el primer momento $\langle r| = \mathrm{Tr}[\hat{\rho}\bar{r}]$, con $\bar{r} = (x_1, p_1, ..., x_n, p_n)$ y la matriz de covarianza $\sigma^{(N)}$ cuyas componentes están dadas por

$$\sigma_{kl}^{(N)} = \frac{1}{2} \langle \{\hat{r}_k, \hat{r}_l\} \rangle - \langle \hat{r}_k \rangle \langle \hat{r}_l \rangle \,. \tag{2.20}$$

Para un sistema Gaussiano, puede demostrarse que es suficiente conocer los dos primeros momentos para caracterizarlo completamente, ya que su función característica, es una distribución Gaussiana caracterizada por dichos momentos.

Por ejemplo, para el vacío sobre un modo $|0\rangle$, el primer momento es $(0, 0)$, mientras que la matriz de covarianza es la identidad en dimensión dos. Por otro lado, un estado que usaremos como portador de información en los próximos capítulos, son los *estado coherentes* $|\alpha\rangle$. Estos describen el estado cuántico de un láser (ver el Apéndice A para más detalles del funcionamiento de un láser) y se definen como el autoestado del operador aniquilación $\hat{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle$.

Su descomposición en la base de Fock viene dada por

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_k \frac{\alpha^k}{\sqrt{k!}} |n\rangle \;, \tag{2.21}$$

y pertenece a la clase de estados Gaussianos, donde su primer momento es $\bar{r} = (2\Re[\alpha], 2\Im[\alpha])$ y su matriz de covarianza también es la identidad [14].

Tanto en la transmisión como en el procesamiento de información es importante conocer cómo manipular a este tipo de estados. En particular, se define como operación Gaussiana a las transformaciones que preservan el carácter Gaussiano de todo estado (Gaussiano) sobre el cual actúe. De esta forma su acción está determinada por los cambios en el primer momento y la matriz de covarianza. Así, la acción de una transformación unitaria Gaussiana sobre un estado Gaussiano dará pie a otro estado Gaussiano, y la acción de una medida Gaussiana sobre un estado Gaussiano dará pie a una distribución de probabilidad Gaussiana para los resultados de la medida. De esta forma, cada operación Gaussiana tiene asociada una matriz F y un vector d que generan la transformación

$$\sigma \to F\sigma F^T \qquad\qquad \bar{r} = F\bar{r} + d \,, \tag{2.22}$$

donde d es un vector real de dimensión $2n$ y F cumple que

$$F\Omega F^T = \Omega \,. \tag{2.23}$$

Además, vale la afirmación inversa y por cada matriz F que cumple la condición (2.23) y vector $d \in \mathcal{R}^{2n}$ existe una operación Gaussiana asociada.

En lo siguiente daremos ejemplos de transformaciones Gaussianas fundamentales y sus proyecciones para obtener resultados de algunos observables.

2.2.2. Transformaciones Gaussianas

Los *operadores de desplazamiento* se caracterizan por un valor complejo $\beta \in \mathcal{C}$ y están definidos operacionalmente como

$$\hat{D}_k = e^{\beta \hat{a}_k^\dagger - \alpha^* \hat{a}_k} \ . \tag{2.24}$$

Como los términos en la exponencial son solo lineales, existe una transformación simpléctica para la matriz de covarianza y los valores de expectación. En este caso la transformación viene dada por la matriz y vector

$$F = I_{2n}, \qquad d = \sqrt{2} \begin{pmatrix} Re(\alpha_1) \\ Im(\alpha_1) \\ \vdots \\ Re(\alpha_n) \\ Im(\alpha_n) \end{pmatrix} . \tag{2.25}$$

Estos estados son particularmente importantes, ya que desplazan a los estados coherentes en la cantidad α. Por ejemplo, si tenemos un estado coherente $|\beta\rangle$, entonces

$$\hat{D}(\alpha) \, |\beta\rangle = |\alpha + \beta\rangle \ . \tag{2.26}$$

En este sentido, podemos entender a un estado coherente $|\alpha\rangle$ como el resultado de aplicar un desplazamiento $\hat{D}_\alpha$ al vacío $|0\rangle$.

Experimentalmente, para realizar un desplazamiento sobre la señal, se combina la señal con un estado coherente de referencia, haciendo uso de un separador de haces. Como esta es una compuerta de un solo modo, podemos generalizar al operador que actúa sobre n modos como

$$\hat{D}(\vec{\alpha}) = \hat{D}_1(\alpha_1) \otimes ... \otimes \hat{D}_n(\alpha_n) \ . \tag{2.27}$$

Los *operadores de cambio de fase* actúan sobre un único modo, cambian a los operadores de creación y destrucción como

$$\hat{a} \to e^{-i\phi} \hat{a} \qquad \hat{a}^\dagger \to e^{i\phi} \hat{a}^\dagger \ . \tag{2.28}$$

Nuevamente, como esta es una operación Gaussiana, la matriz y vector que transforma a la matriz de covarianza viene dada por

$$F = \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix} = R(\phi) \qquad d = 0 \tag{2.29}$$

Si aplicamos un cambio de fase a varios modos, entonces obtenemos

$$F = \bigoplus_{k=1}^{n} R(\phi_k) \tag{2.30}$$

y $d = 0$. Para producir un cambio de fase se hace incidir a los fotones sobre un medio con índice de refracción diferente de 1 o en general, controlando el camino óptico efectivo.

Asimismo, la transformación de Squeezing tiene un efecto de *ensanchar* y *comprimir* las cuadraturas del campo (vistas en términos del espacio de fase y una distribución de cuasi-probabilidad). Esta compuerta depende de un parámetro $z \in \mathcal{C}$, y su operador asociado es

$$\hat{S}(z) = e^{\frac{1}{2}(z^* \hat{a}^2 - z \hat{a}^{\dagger^2})} \ . \tag{2.31}$$

Si escribimos al número complejo como $z = re^{i\theta}$, entonces también podemos encontrar las matrices que definen la transformación simpléctica como

$$F(z) = \cosh(r)\mathbb{I} - \sinh(r)S(\theta) \quad d = \vec{0}\,, \tag{2.32}$$

donde

$$S(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}. \tag{2.33}$$

Si queremos realizar esta operación sobre n modos, nuevamente podemos encontrar la transformación mediante la suma directa

$$F_N(\vec{z}) = \bigoplus_{k=1}^{N} F(z_k)\,. \tag{2.34}$$

Estos operadores son realizados experimentalmente haciendo uso de medios ópticos no lineales. En estos, los fotones interactúan siguiendo el Hamiltoniano

$$H = \chi\left(\hat{a}^2 + \hat{a}^{\dagger 2}\right). \tag{2.35}$$

Si la interacción es lo suficientemente grande, esta interacción produce una compresión sobre los estados de la forma descrita anteriormente.

Finalmente, consideramos el separador de haces, o beam-splitter, que es una compuerta de dos modos muy utilizada en los laboratorios ópticos. El operador que define a esta transformación viene dado por

$$\hat{U} = e^{i\eta(\hat{a}_1^\dagger \hat{a}_2 + \hat{a}_2^\dagger \hat{a}_1)}\,, \tag{2.36}$$

y la transformación asociada que define a este operador está dada por

$$F = \begin{pmatrix} \sqrt{\eta}\,\mathbb{I} & \sqrt{1-\eta}\,\mathbb{I} \\ -\sqrt{1-\eta}\,\mathbb{I} & \sqrt{\eta}\,\mathbb{I} \end{pmatrix}, \qquad d = 0\,. \tag{2.37}$$

Para realizar un separador de haces, se emplea un cristal reflectante, que permite el paso de una parte de ambas señales.

2.2.3. Proyecciones sobre estados Gaussianos

Para poder realizar medidas sobre los diferentes estados Gaussianos y así enviar información, es importante saber como calcular proyecciones y medidas sobre estos de forma analítica. En general, si queremos realizar una proyección de un estado Gaussiano con operador de densidad $\hat{\rho}_1$ sobre otro estado Gaussiano con operador de densidad $\hat{\rho}_2$, podemos utilizar la siguiente fórmula

$$Tr(\hat{\rho}_1\hat{\rho}_2) = \frac{1}{\sqrt{\Sigma}} e^{-\frac{1}{2}(r_1-r_2)^T\Sigma(r_1-r_2)}\,, \tag{2.38}$$

donde $\Sigma = \sigma_1 + \sigma_2$ es la suma de ambas matrices de covarianza y r_1, r_2 los valores de expectación.

Si tenemos un sistema bipartito ρ_{AB}, podemos escribir a la matriz de covarianza σ, con valores de expectación $\bar{r}$ como

$$\sigma = \begin{pmatrix} A & C^T \\ C & B \end{pmatrix} \qquad \bar{r} = \begin{pmatrix} a \\ b \end{pmatrix}, \tag{2.39}$$

donde A y B son las matrices de covarianza de los sistemas ρ_A y ρ_B respectivamente. Se puede demostrar que si proyectamos al estado B sobre el vacío, entonces la matriz de covarianza y los valores de expectación de ρ_A se vuelven

$$A \to A - C\left(B + \frac{1}{2}\mathbb{I}\right)^{-1} C^T \tag{2.40}$$

$$a \to a - C\left(B + \frac{1}{2}\mathbb{I}\right)^{-1} b\,. \tag{2.41}$$

De manera que el estado descrito por el subsistema ρ_A resulta Gaussiano [15].

Habiendo establecido como operar sobre los diferentes estados y como proyectar sobre otros, estamos en condiciones de discutir algunos métodos para distinguir entre el par de estados $|\pm\alpha\rangle$.

3 Discriminación de estados & modelado de canales reales.

Tengamos un asesinato que haga sonreír a toda la gente.

Koro Sensei
Assassination of classroom.

En este capítulo, nos dedicaremos a formalizar algunos esquemas de comunicaciones cuánticas basados en el marco teórico, presentado en el capítulo previo.

Hay muchas analogías y parecidos clásicos que ayudan a generar intuición sobre la distinción de estados cuánticos, pero también existen diferencias que afectan al sentido físico de los diferentes fenómenos. Para notar diferencias no hay que ir demasiado lejos, al leer este trabajo de diploma necesitamos distinguir entre las diferentes letras. Pero supongamos que repentinamente el texto se vuelve más pequeño y cambia de formato, este efecto reduce la capacidad que tenemos de distinguir entre estados debido a ruidos que deforman las letras de forma menos reconocible. Esta es una particularidad de las señales clásicas, si la transmisión de información es libre de fondo, la probabilidad de distinguir entre señales clásicas es 1, viéndose reducida solamente cuando se tienen fuentes de ruido. Pero a diferencia de las señales clásicas, cuando se tiene un único estado cuántico, se puede observar que los estados no ortogonales no pueden ser distinguidos de forma perfecta, incluso en el caso libre de ruido.

3.1. Distinción de estados cuánticos

Conociendo esta diferencia fundamental, y en vistas a su aplicación en comunicaciones, en este trabajo de diploma nos concentraremos en el problema de discriminación de los estados cuánticos ρ_0 y ρ_1. Dada una única copia del estado ρ, el trabajo será determinar si ambas hipótesis $H_0 : \rho = \rho_0$ se mantienen, o la hipótesis $H_1 : \rho = \rho_1$ lo hace, cada una teniendo una probabilidad de suceder p_k, con $\sum_k p_k = 1$.

En este sentido, cabe destacar que el problema al cual nos enfrentamos en este trabajo es distinto al introducido en textos tales como el Nielsen [16], donde se estudia la discriminación inambigua entre estados cuánticos. En nuestro problema, el decodificador solamente puede predecir entre dos candidatos (no pudiendo calificar el experimento como *ambiguo*), lo que da pie a una probabilidad finita de error, que ha de ser minimizada sobre todas las posibles medidas cuánticas que B pueda implementar.

De esta manera, introducimos a la POVM de 2 salidas $\mathcal{M} = \{M_0, M_1\}$ tal que $\sum_{k=0}^{1} M_k = \mathbb{I}$ y $\mathcal{M}_k \geq 0$. Basado en que la medida tenga resultado k, la regla de decisión se lee como $\hat{k} = k$, mientras que la probabilidad del resultado está dado por la regla de Born, *e.g.* $p(k|\rho) = \mathrm{Tr}[M_k \rho]$. La probabilidad de éxito y de error del tal protocolo de discriminación está dado por

$$\begin{aligned} P_s(\mathcal{M}) &= \sum_{k=0,1} p_k \, p(\hat{k}|k) = \sum_{k=0,1} p_k \, \mathrm{Tr}[\rho_k M_k], \\ P_e(\mathcal{M}) &= \sum_{k=0,1} p_{\bar{k}} \, p(\hat{k}|\bar{k}) = \sum_{k=0,1} p_{\bar{k}} \, \mathrm{Tr}[\rho_{\bar{k}}] M_k = 1 - P_s(\mathcal{M}), \end{aligned} \tag{3.1}$$

donde definimos al complementario de la hipótesis $\bar{k} \equiv k + 1 \mod 2$.

3.1.1. Límite de Helstrom

Como la seguridad que tenemos de distinguir correctamente entre estados en una comunicación cuántica no puede ser absoluta, tiene que existir una probabilidad de éxito máxima. Por esta razón, a continuación discutiremos una cota superior a los valores que la probabilidad de éxito descrita en la ecuación 3.1 puede tomar, cuando optimizamos sobre todas las posibles medidas $\mathcal{M}$.

Supongamos entonces que luego de comenzar la comunicación, Beto ya recibió un estado y va a realizar operaciones sobre este para discriminarlos de forma correcta. Esto tiene asociada una probabilidad de éxito

$$P_{exito} = p_1 P_{Y|X}(0|0) + p_2 P_{Y|X}(1|1)\,. \tag{3.2}$$

Considerando que la POVM está descrita por dos operadores $\mathcal{M} = \{\Lambda_0, \Lambda_1\}$, podemos expresar la probabilidad de éxito según

$$P_{exito}(\Lambda) = p_1 \mathrm{Tr}(\Lambda_0 \rho_0) + p_2 \mathrm{Tr}(\Lambda_1 \rho_1)\,. \tag{3.3}$$

Dado que las medidas son completas, podemos usar que $\Lambda_0 + \Lambda_1 = \mathbb{I}$, llegando así a la expresión

$$P_{exito}(\Lambda) = p_1 \mathrm{Tr}(\Lambda_0 \rho_0) + p_2 \mathrm{Tr}((\mathbb{I} - \Lambda_0)\rho_1) = p_1 + \mathrm{Tr}(\Lambda_0 (p_0 \rho_0 - p_1 \rho_1))\,, \tag{3.4}$$

obteniendo así la probabilidad de éxito en función de cualquier POVM Λ. Si queremos saturar el límite superior tenemos que usar la POVM que maximiza esta probabilidad de éxito

$$P_{\text{éxito}}^{\text{máx}} = \max_{\Lambda} p_1 + \mathrm{Tr}(\Lambda_0 (p_0 \rho_0 - p_1 \rho_1))\,. \tag{3.5}$$

Siguiendo propiedades de la distancia traza [17], se puede demostrar que la discriminación máxima viene dada por la función

$$P_{exito} = \frac{1}{2}\left(1 + \frac{1}{2}||\rho_0 - \rho_1||_1\right)\,, \tag{3.6}$$

siendo este el límite teórico de discriminación entre 2 matrices de densidad, conocido como *Límite de Helstrom*.

Podemos notar que si los estados son ortogonales, entonces $||\rho_0 - \rho_1||_1 = 1$ y la discriminación es perfecta. En particular podemos calcular que el límite de Helstrom para 2 estados puros $|\psi\rangle$ y $|\psi'\rangle$ está dado por

$$P_{exito} = \frac{1}{2}\left(1 + \sqrt{1 - p_0 p_1 |c|^2}\right)\,, \tag{3.7}$$

donde $c = \langle\psi|\psi'\rangle$.

Finalmente, puede demostrarse [17] que la medida que satura el límite de Helstrom (es decir, la medida óptima para el problema de discriminación binaria) viene dada por una proyección sobre la parte positiva y negativa del operador $\rho_1 - \rho_0$. Para estados puros, esto se reduce a una proyección sobre superposiciones del tipo $|\psi\rangle + |\psi'\rangle$.

3.2. Experimento de comunicación

En particular, las señales que estudiaremos son de un problema canónico de discriminación de estados cuánticos, el cual está motivado por las comunicaciones sobre largas distancias. En este sentido, consideramos dos participantes: un emisor A que quiere enviar un bit clásico de información ($k \in \{0, 1\}$) codificándolo en la fase de un estado coherente ($|(-1)^k \alpha\rangle$), y un receptor B, quien recibirá la señal (enviada a través de un canal de transmisión), y que suyo objetivo es extraer la información original, a través de un *estimador*.

De esta forma, es imperioso que B maximice su probabilidad de éxito en la clasificación de las señales. En este contexto, el problema se reduce al de optimizar la probabilidad de éxito en la discriminación de dos posibles estados cuánticos que arriban al detector que utiliza B.

En lo siguiente, nos enfocaremos en el caso de un canal de transmisión ideal, por lo que las señales que arriban al detector vienen dadas por los dos posibles estados coherentes que A prepara con fases opuestas, es decir $\{|\alpha\rangle, |-\alpha\rangle\}$. De este modo, B debe medir el estado cuántico que le llega, y a raíz del resultado de la medida, decidir por uno de los dos posibles candidatos.

El problema de discriminación entre estados coherentes es un problema canónico en el campo de la información cuántica, tanto por su relevancia teórica [12, 18, 19, 20, 21] como experimental y tecnológica [6, 22, 23, 8].

Entendiendo el esquema descrito anteriormente en el contexto de comunicaciones clásicas a través de canales cuánticos, el problema de discriminación óptima de estados cuánticos cobra un rol fundamental a la hora de optimizar la *calidad* de la transmisión. En este sentido, podemos tomar como ejemplo las comunicaciones satelitales, donde un bit de información se codifica en la fase de un estado coherente (un pulso de láser); el mismo es enviado a través de la atmósfera a la Tierra. Modelando la atmósfera como un canal con perdidas, la señal que llega a la Tierra está descrita por un estado coherente, cuya intensidad se ve reducida según la distancia entre el satélite y la Tierra. En la práctica, la señal que llega a la Tierra es *casi* el vacío (en el sentido que es un estado coherente $|\alpha\rangle$ cuya amplitud $|\alpha| \ll 1$). Recordando que los estados coherentes no son ortogonales entre sí, es de esperar una probabilidad finita de error, que crecerá según al menos el solapamiento $|\langle\alpha| - \alpha\rangle|^2 = e^{-\alpha^2}$ entre los estados. Para poder distinguir entre este tipo de señales, se pueden seguir muchos tipos de medidas POVM, las cuales no necesariamente saturarán el límite de Helstrom, pero son experimentalmente accesibles.

En el caso particular de distinción de estados coherentes con fases opuestas, si se preserva la Gaussianidad del estado, las mediciones que maximizan la probabilidad de éxito, son conocidas como *medidas Homodinas* [24, 25, 26]. Estas medidas constan de proyectar al estado recibido sobre un estado coherente, midiendo así la cuadratura $\hat{q}$ y caracterizar al estado según cuál estado tiene mayores probabilidades de producir esa observación.

Para lograr hacer las proyecciones experimentalmente, un detector, como el esquematizado en la figura 3.1, compara el campo observado con uno de referencia, llamado "oscilador local", cuyo $|\alpha| \gg 1$ y obteniendo así información de la fase de forma indirecta. Para lograr esto, un fotodetector mide un modo electromagnético convirtiendo a fotones en electrones y por ende en una corriente eléctrica conocida como *foto corriente* $\hat{i}$. Siguiendo la idea de que los fotones se transforman mediante algún proceso en electrones, es coherente asumir que la corriente observada es proporcional $\hat{i} = q\hat{N} = q\hat{a}^\dagger\hat{a}$, siendo q una constante [27].

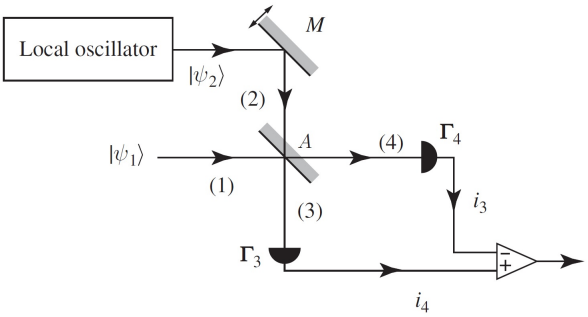

Figura 3.1: Diagrama de un detector que realiza medidas Homodinas.

De esta forma, si calculáramos como se verían las corrientes del diagrama, en función de los estados empleados, podríamos ver que la diferencia de señal eléctrica es

$$\bar{d} = -\alpha_2\varepsilon^{(1)}\Big(\cos\phi_2 \langle\psi_1(0)|\,\hat{E}_{Q1}\,|\psi_1(0)\rangle + \sin\phi_2 \langle\psi_1(0)|\,\hat{E}_{P1}\,|\psi_1(0)\rangle\Big)\,. \tag{3.8}$$

donde ψ_1 son los estados a determinar y ϕ_2 es la fase del oscilador local.

Esta expresión muestra que las fluctuaciones de cuadratura pueden ser medidas si el segundo término es despreciable (si $|\alpha_2|$ es lo suficientemente grande). Realizando estas medidas entonces podemos medir las cuadraturas de nuestros estados y basados en el resultado, podemos diferenciar con una cierta probabilidad cuál fue el estado recibido.

Cabe destacar que dentro de las proyecciones sobre estados Gaussianos, existe otra medida muy útil que no estudiaremos en esta tesis, pero de suma relevancia experimental, conocida como *medida heterodina*. La misma consiste en proyecciones sobre estados coherentes (es decir, proyecciones no ortogonales), y pueden ser realizadas experimentalmente con un esquema similar al descrito para la medida Homodina.

La probabilidad de éxito de este método puede ser determinada utilizando la ecuación (2.38) y realizando la proyección sobre los estados coherentes. Así se puede ver que la probabilidad de éxito de distinguir entre el par de estados $|\pm\alpha\rangle$ mediante una medición Homodina viene dado por

$$P_{\text{éxito}}^{HOM}(\alpha) = \frac{1}{2}\Big[1 + \text{Erf}\Big(\sqrt{2}\alpha\Big)\Big] \,. \tag{3.9}$$

Es importante remarcar que las transformaciones Gaussianas pueden generarse de forma relativamente fácil en los laboratorios ópticos, según hemos detallado para cada caso. Su acción en el espacio de fases es *lineal*, ya que están representadas por transformaciones afines parametrizadas por la matriz F y el vector d. Más aún, los estados cuánticos están unívocamente caracterizados por su función característica. Para los estados Gaussianos, dicha cantidad es una distribución Gaussiana, y por ende definida positiva en todo el espacio de fases. En este sentido, existen estados cuánticos para los cuales dicha cantidad puede tomar valores negativos, lo que es un efecto genuinamente cuántico [28]; dichos estados son por ende no-Gaussianos, y resultan ser un recurso para la computación cuántica en variable continua. A pesar de los problemas teóricos y experimentales que involucra emplear métodos no Gaussianos, la probabilidad de éxito obtenida anteriormente es siempre menor que el límite de Helstrom, viéndose así la necesidad de utilizar recursos no gaussianos para alcanzarlo.

Teniendo en cuenta la necesidad de *recursos No Gaussianos* para poder alcanzar el límite de Helstrom, describiremos el uso de detectores de fotones únicos.

Foto-contadores. Estos consisten en proyecciones sobre la base de Fock, donde idealmente es posible distinguir el número de excitaciones del campo (fotones), e implementan la medida $\{|n\rangle\langle n|\}_{n=0}^{\infty}$. Si bien su realización experimental es factible, en el presente son instrumentos delicados y costosos.

Fotodetectores Estos detectores tienen 2 salidas posibles, 0 cuando no se observa ningún fotón y 1 cuando se observa uno o más fotones, esto significa que las medidas del fotodetector son $\mathcal{M} = \{|0\rangle\langle 0| \,, \mathbb{I} - |0\rangle\langle 0|\}$. Esta medida es por ende de carácter no Gaussiano, y su implementación experimental resulta más sencilla que los foto-contadores (que resuelven número de partículas); resultan de suma importancia a la hora de considerar problemas de discriminación de estados en variable continua.

Haciendo uso de los recursos que estos detectores traen, las *Medidas de tipo Kennedy* consisten en generar un desplazamiento sobre los estados en una cantidad β fija, realizando la transformación $\{|-\alpha\rangle \,, |\alpha\rangle\} \to \{|-\alpha+\beta\rangle \,, |\alpha+\beta\rangle\}$, y luego realizar una medición con fotodetector. En general, para clasificar al estado recibido, se tendrá que hacer según la observación de nuestro experimento y el estado más probable en generar esa observación. En el caso de estados coherentes con diferente fase se puede escribir

$$P_s^{ken} = \sum_{o=0,1} \max_k p_k p(o| - \alpha^k) \,, \tag{3.10}$$

siendo $k = \{0, 1\}$ y o la observación del experimento. Si usamos este método para clasificar a los estados, se puede ver que para la distinción de $|\pm\alpha\rangle$ con probabilidades $\left(\frac{1}{2}, \frac{1}{2}\right)$, es óptimo clasificar al estado recibido

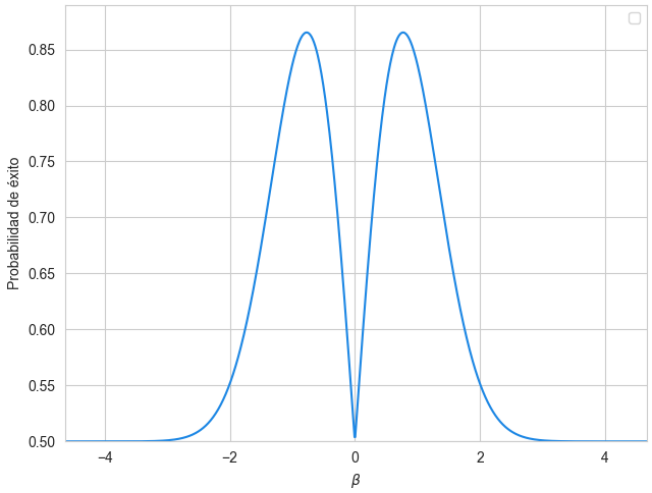

Figura 3.2: Probabilidad de éxito en medida tipo Kennedy variando el parámetro de desplazamiento β, teniendo en cuenta el criterio del estado más probable.

como $|-\alpha\rangle$ cuando no se observen fotones y en caso contrario clasificarlo como $|\alpha\rangle$, siempre que $\beta \geq 0$. Esto se debe a la paridad en la probabilidad de éxito según β, como se puede ver en la figura 3.2.

En general se llama discriminación *Kennedy* cuando se fija al parámetro de desplazamiento como $\beta = \alpha$. De forma que el par de estados resultará $\{|0\rangle, |2\alpha\rangle\}$.

En este caso, podemos discriminar unívocamente al estado $|\alpha\rangle$ si se observan fotones, mientras que en caso de no detectar ningún fotón, se clasificará al estado como $|-\alpha\rangle$ teniendo un error

$$P_{error} = P(0, 2\alpha) = e^{-|2\alpha|^2} ,$$

donde $P(0, 2\alpha)$ es la probabilidad de que el estado $|2\alpha\rangle$ produzca 0 fotones.

Este método tiene entonces una probabilidad de éxito asociada de

$$P_{exito} = 1 - e^{-|2\alpha|^2} , \tag{3.11}$$

por lo que se puede observar que la probabilidad de éxito siguiendo esta metodología es mayor que la detección Homodina solo si $|\alpha| \gtrapprox 0{,}4$.

Por otro lado, la discriminación *Kennedy optimizada* maximiza la probabilidad de éxito en función del parámetro β y del $|\alpha|$ recibido. En el caso de discriminación del par de estados coherentes $|\pm\alpha\rangle$, la probabilidad de éxito aplicando un desplazamiento en una cantidad β viene dada por

$$P_{\text{éxito}}(\beta) = \frac{1}{2}e^{-|\beta-\alpha|^2} + \frac{1}{2}\left(1 - e^{-|\beta+\alpha|^2}\right) , \tag{3.12}$$

siendo esta la función que queremos maximizar. Es importante notar que aunque la maximización en ausencia de ruido es sencilla, tratándose de un problema de optimización unidimensional, pero esta función no puede ser maximizada de forma analítica. Por esta razón, se necesitan utilizar métodos numéricos como el algoritmo Dual-Annealing, para encontrar los valores óptimos de β. Se puede observar en la figura (3.3), las diferentes probabilidades de éxito descritas.

A diferencia del detector Kennedy, este método no puede distinguir unívocamente a ninguno de los dos estados, pero a cambio otorga una mayor probabilidad de éxito, superando a la detección Homodina en todas las intensidades. A pesar de la mejora, esta probabilidad de éxito sigue siendo estrictamente menor que el límite de Helstrom.

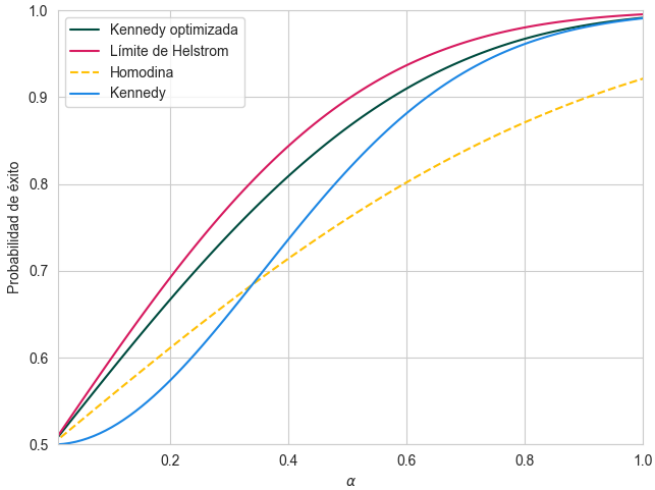

Figura 3.3: Comparación entre los métodos de discriminación Homodina, Kennedy, Kennedy optimizada y el límite de Helstrom.

Según lo descrito en la Sección anterior 3.1.1, la medida óptima para la discriminación entre dos estados coherentes vendrá dada por una proyección sobre un estado del tipo *gato*, es decir, una superposición de estados coherentes de la forma $|\alpha\rangle \pm |-\alpha\rangle$. Dichos estados son no-gaussianos, lo que conlleva una clara dificultad a la hora de analizar el problema. Notablemente, en este caso particular, una implementación experimental para la medida óptima (y de un gran interés teórico por su carácter secuencial) es conocida bajo el nombre de *detector tipo Dolinar* [18, 20], el cual es discutido en mayor profundidad en el Apéndice B. A pesar de su gran interés teórico y experimental, su implementación conlleva grandes dificultades, y solo ciertas pruebas de concepto han sido llevadas a la práctica [6]. Entre varios, uno de los problemas es cómo optimizar los desplazamientos para un esquema de detección finito (y no asintótico, donde técnicas de programación dinámica permitirían hacerlo [29]). Por otro lado, errores en los detectores impiden incluso batir el límite presentado por medidas gaussianas (a discutir a continuación). Finalmente, en condiciones realistas, resulta difícil modelar el canal de transmisión y el comportamiento de las señales y detectores. Esta es una de las motivaciones por las cuales trabajos recientes han propuesto la calibración de este tipo de medidas en forma automática [30, 31].

Pero en los métodos de discriminación vistos, la determinación de parámetros óptimos es dependiente de las probabilidades de éxito. De forma que si los estados que recibe B no son los estados que fueron enviados por A, los valores óptimos podrían ser sumamente diferentes a los esperados. Esto significa que en canales reales donde los estados recibidos no son iguales a los estados enviados (debido a interacciones con el medio), podríamos necesitar modelos teóricos robustos que permitan determinar esta probabilidad de éxito con precisión. Por esta razón, en la siguiente sección hablaremos de algunos modelados de cables reales y los problemas asociados a cada uno de ellos.

3.3. Modelos de canales ruidosos

Hasta ahora estudiamos comunicaciones ideales, donde los estados enviados son exactamente los deseados, los estados recibidos no sufrieron ninguna interacción con el medio y son detectados de forma ideal.

Pero en todo sistema real, los resultados teóricos sufren alteraciones debido a condiciones más allá de nuestro control. Para el caso de comunicaciones, estos factores externos son observados en forma de absorciones, fluctuaciones térmicas del canal, detectores imperfectos, etc. Por ello, para poder distinguir estados cuánticos de manera óptima, es importante tener modelos que describan de forma precisa los resultados de nuestros experimentos y los canales de transmisión.

En esta sección introduciremos algunos modelos de canales ruidosos frecuentemente utilizados, a la hora de implementar experimentalmente el problema de comunicación mediante estados coherentes introducido

anteriormente. De forma esperable, la calidad de la transmisión de los bits, cuantificada según la probabilidad de éxito definida en la ecuación (3.12), se verá degradada en la presencia de canales ruidosos; esto será ejemplificado a continuación, donde también estudiaremos el comportamiento de probabilidad de éxito de los detectores de tipo Kennedy Optimizados en función del parámetro de ruido del canal.

3.3.1. Canales Gaussianos

A continuación discutiremos ciertos modelos de ruido que resultan *accesibles* analíticamente, en el sentido que la Gaussianidad del sistema es preservado. Esto permite describir a nuestro canal mediante circuitos Gaussianos, pudiendo así realizar optimizaciones sobre estos de forma sencilla. Asimismo, también permiten obtener las probabilidades del resultado de una medida Gaussiana de forma analítica, según la ecuación 2.38.

Atenuación Gaussiana Para describir la atenuación que sufre un estado $|\pm\alpha\rangle$ podemos pensar a la transmisión como un sistema de 2 modos. Colocando un separador de haces y realizando medidas solo sobre el primer modo, el estado final es equivalente a atenuar el estado enviado. El esquema de este modelo puede ser visto en la figura 3.4.

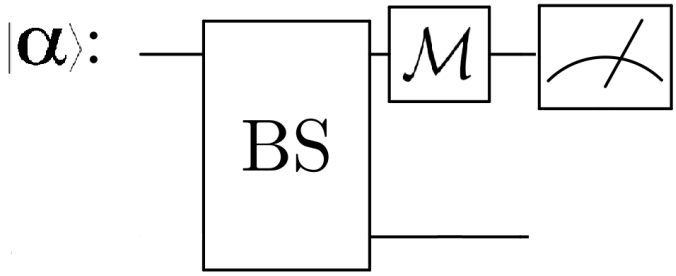

Figura 3.4: Esquema de discriminación con una atenuación Gaussiana realizando una medida generalizada $\mathcal{M}$.

Siguiendo este circuito, el estado final que recibe B es

$$|\alpha\rangle_{fin} = |\alpha\cos(\theta)\rangle \ . \tag{3.13}$$

Esto hace que la probabilidad de éxito, para la discriminación de tipo Kennedy, en función del parámetro β, venga dada por

$$P_{exito}(\beta) = \frac{1}{2}e^{-|-\alpha\cos(\theta)+\beta|} + \frac{1}{2}\left(1 - e^{|\alpha\cos(\theta)+\beta|}\right) . \tag{3.14}$$

Llevando a que una mayor atenuación aumenta el solapamiento entre los estados, y así la probabilidad de éxito se ve reducida para cualquier método de discriminación (ver por ejemplo ecuación (3.7).

Por ejemplo, se puede estudiar la distinción de estados coherentes para $\alpha = 0{,}5$, usando un detector de tipo Kennedy optimizado. Si realizamos incorrectamente la optimización, ignorando la atenuación sufrida por el estado, y variamos el parámetro θ del separador de haces, obtenemos la figura 3.5.

En este ejemplo podemos observar que la caída en la probabilidad de éxito es bastante abrupta, pero si tenemos en cuenta el modelo de atenuación al momento de optimizar el parámetro de desplazamiento, las probabilidades de éxito se ven considerablemente incrementadas. Esto se puede ver en la figura 3.6 donde se distinguió entre estados coherentes de mayor intensidad, fijando $\alpha = 1{,}5$.

La consideración de este modelo para la optimización de nuestros parámetros, afecta particularmente a la probabilidad de éxito cuando se distinguen estados de muy alta energía, pero es menos efectivo en estados de energía menor.

Cambio de fase. Un tipo de error muy común al enviar estados coherentes es que su fase cambie con respecto a la esperada. Si suponemos que todo el estado se *desfasa* manteniendo su Gaussianidad, lo podemos

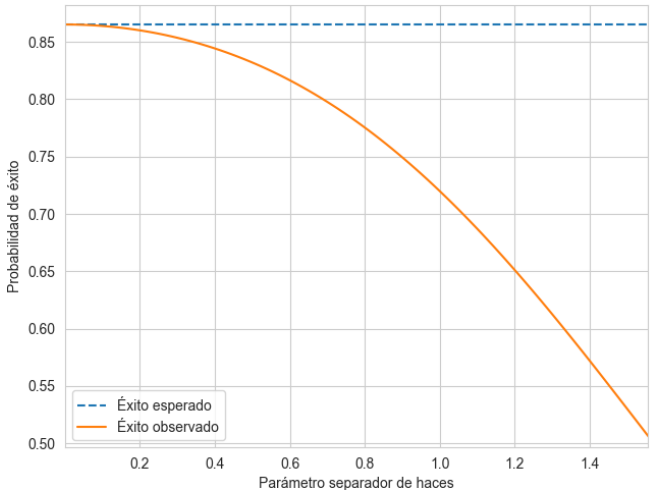

Figura 3.5: Probabilidad de éxito en una medición de tipo Kennedy entre estados $|\pm\alpha\rangle$ con $\alpha = 0{,}5$ en un canal con diferentes absorciones.

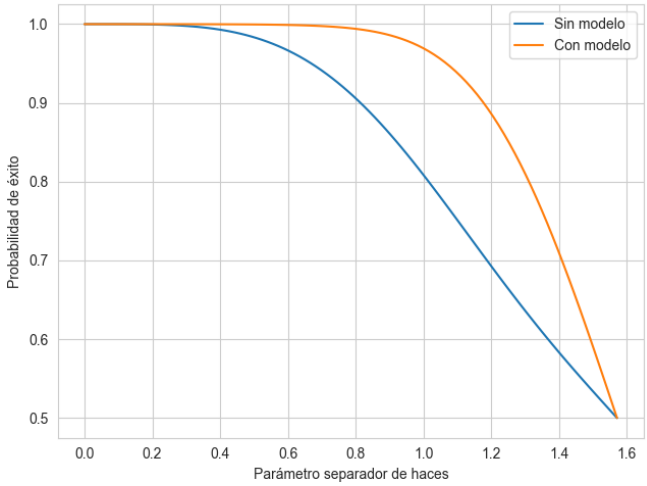

Figura 3.6: Comparación de la probabilidad de éxito para absorción Gaussiana en estados $|\pm 1{,}5\rangle$, eligiendo el parámetro β con y sin modelo teórico del fondo.

modelar aplicando la compuerta Gaussiana de cambio de fase al comienzo del circuito, transformando al par de estados en $|\pm(\alpha\cos(\phi) + i\sin(\phi))\rangle$, haciendo que las probabilidades de éxito resulten

$$P_{\text{éxito}} = \frac{1}{2}e^{-|-\cos(\theta)\alpha - i\sin(\theta)\alpha + \beta|} + \frac{1}{2}\left(1 - \frac{1}{2}e^{-|\cos(\theta)\alpha + i\sin(\theta)\alpha + \beta|}\right) \tag{3.15}$$

Este modelo solo emplea una compuerta unitaria, por lo que no altera las probabilidades de éxito y puede ser revertido. Pero este cambio hace los parámetros óptimos de β para la probabilidad de éxito cambien. Por ejemplo, en el caso de distinguir entre un par de estados con $\alpha \in \mathcal{R}$, la probabilidad de éxito se maximiza con un parámetro $\beta \in \mathcal{R}$. Pero al aplicar un cambio de fase, la probabilidad de éxito cambia y el parámetro óptimo se vuelve $\beta' \in \mathcal{C}$.

De esta forma, una mala determinación de la fase, junto a no determinar correctamente el estado más probable en generar el resultado observado, podría resultar en que las probabilidades de éxito sean aproximadamente 0 (menor que elegir estados de forma aleatoria), cuando la distancia entre estados se mantiene invariante, como se puede ver en la figura (3.7).

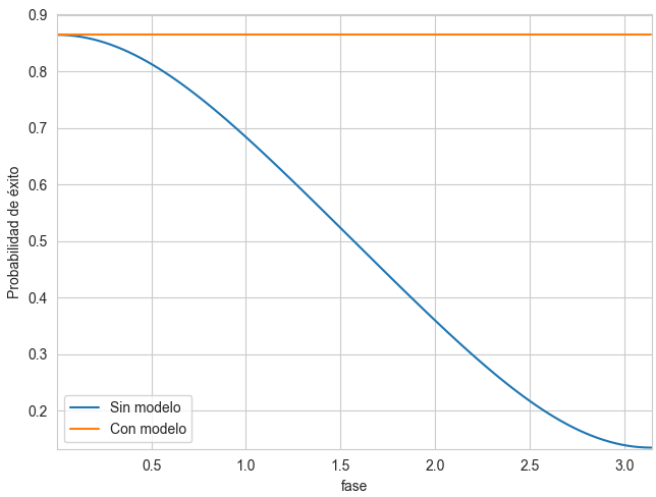

Figura 3.7: Comparación en la probabilidad de éxito, caracterizando correctamente e incorrectamente el cambio de fase Gaussiano en estados $|\pm\alpha\rangle$ con $\alpha = 0{,}5$.

Cuentas oscuras. Una fuente de ruido que afecta a la mayoría de las comunicaciones son las conocidas cuentas oscuras en los detectores. Estas suelen suceder por acumulaciones de carga en los materiales, existiendo así una probabilidad de que los detectores midan fotones, incluso en ausencia de señal externa.

Una distribución que se suele emplear para describir la cantidad de cuentas oscuras es utilizando una distribución de Poisson [32]. Teniendo esto en cuenta, si queremos distinguir entre el par de estados $|\pm\alpha\rangle$, con $\alpha \in \mathcal{R}$; proponemos un modelo empleando una compuerta de desplazamiento en una cantidad $\gamma \in \mathcal{I}$. De esta forma, si distinguimos entre estados, utilizando un detector de tipo Kennedy que tiene restringido su parámetro de desplazamiento a los valores reales $\beta \in \mathcal{R}$, este circuito siempre va a observar en media como mínimo $|\gamma|^2$ fotones, siguiendo una distribución de Poisson. Un diagrama del circuito propuesto puede ser observado en la figura 3.8.

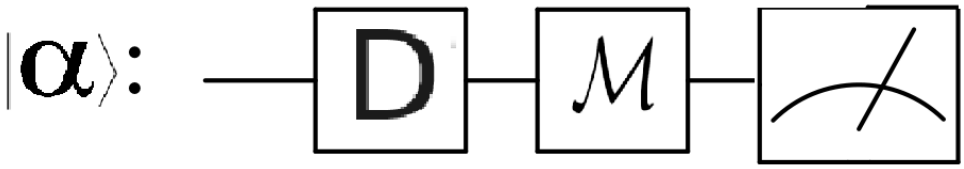

Figura 3.8: Propuesta de circuito para modelar las cuentas oscuras de un detector imperfecto.

Si fijamos la clasificación de estados como votar por $|-\alpha\rangle$ cuando no observemos fotones, la estrategia óptima consistiría en anular la señal externa lo más posible. Pero como limitamos a que el parámetro $\beta \in \mathcal{R}$, esto no es posible, generando una probabilidad de éxito

$$P_{\text{éxito}} = \frac{1}{2}e^{-|-\alpha+i|\gamma|+|\beta||^2} + \frac{1}{2}e^{-|\alpha+i|\gamma|+|\beta||^2} \,. \tag{3.16}$$

Aunque los parámetros óptimos son ligeramente diferentes al caso libre de fondo, a diferencia de los modelos que describimos anteriormente, estos no generan una mejora apreciable en la probabilidad de éxito (se producen diferencias en el orden de 10^{-12}), como puede ser visto en la figura 3.9 con la discriminación de estados $|\pm\alpha\rangle$ con $\alpha = 0{,}5$. Notar que este modelo solo fue analizado con detección Kennedy optimizada y que otros métodos de discriminación como Dolinar podrían tener la capacidad de mejorar la probabilidad de éxito con una calibración apropiada.

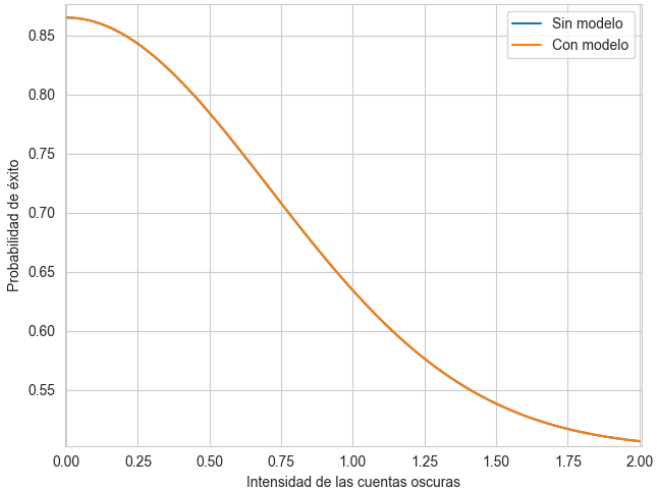

Figura 3.9: Distinción de estados $|\pm\alpha\rangle$ con $\alpha = 0{,}5$ en presencia de un fondo de cuentas oscuras variable.

Es importante también notar que este modelo de cuentas oscuras no representa perfectamente a un detector real, ya que no tiene en cuenta procesos como el tiempo muerto de procesamiento del detector [32].

3.3.2. Fondos no Gaussianos y conclusiones

Anteriormente, hemos visto diferentes modelos de canales ruidosos que preservan la Gaussianidad de los estados. Pero en comunicaciones reales esto no sucede necesariamente, volviendo importante el modelado no gaussiano de algunas interacciones con el entorno. Por esto mismo, en esta parte discutiremos un modelo de desfasaje puro y una amplitud amortiguada, el cual describe el comportamiento de fotones, por ejemplo, en fibras ópticas. Esto ilustrará la problemática de trabajar con este tipo de estados y sus optimizaciones.

Desfasaje y amplitud amortiguada: En esta sección discutiremos como evoluciona un modo bosónico que interacciona según una amortiguación de la amplitud y un cambio de fase. Para esto mencionaremos la ecuación de Lindblad $\dot{\rho} = \mathcal{L}\rho$, que nos servirá para describir las interacciones no gaussianas con el medio, pero las cuales superan los conceptos vistos en este trabajo [33].

Haciendo uso de esta ecuación se puede modelar a la amortiguación de amplitud como una evolución Gaussiana, que sigue la ecuación maestra

$$\mathcal{L}_{AD} := 2\hat{a}\hat{\rho}\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\hat{\rho} - \hat{\rho}\hat{a}^\dagger\hat{a} = \frac{d\hat{\rho}}{dt}\,, \tag{3.17}$$

y desfasaje puro según el modelo no gaussiano

$$\mathcal{L}_{PD}(\hat{\rho}) = \frac{d\hat{\rho}}{dt} = 2\hat{a}^\dagger\hat{a}\hat{\rho}\hat{a}^\dagger\hat{a} - \left(\hat{a}^\dagger\hat{a}\right)^2\hat{\rho} - \hat{\rho}\left(\hat{a}^\dagger\hat{a}\right)^2\,. \tag{3.18}$$

Para modelar una interpolación de ambos regímenes, podemos trabajar con la ecuación

$$\frac{d\hat{\rho}}{dt} = (1-\epsilon)\mathcal{L}_{AD}(\hat{\rho}) + \epsilon\mathcal{L}_{PD}(\hat{\rho})\,, \tag{3.19}$$

y viendo que en este caso los elementos conmutan, se puede despejar la solución como

$$\hat{\rho}(t) = e^{(1-\epsilon)tL_{AD}}e^{\epsilon t\mathcal{L}_{PD}}\rho(0)\,. \tag{3.20}$$

Operando esta solución se llega a que la matriz de densidad en la base de Fock, puede ser escrita como

$$\rho(t) = \sum_{m,n=0}^{\infty} C_{m,n}(t)\,|m\rangle\langle n|\,, \tag{3.21}$$

donde

$$C_{m,n} = e^{-Y_{m,n}(\epsilon)t} \sum_{l=0}^{\infty} C_{m+l,n+l}(0) \left[\binom{m+l}{l} \binom{n+l}{l} \right]^{1/2} f^l \tag{3.22}$$

con $Y_{m,n}(\epsilon) = 2(1-\epsilon)(m+n) + \epsilon(m-n)^2$ [34].

Como se puede ver, colocando como condición inicial que el estado sea un estado coherente, es posible obtener la evolución del estado sobre el sistema. Pero el estado final no puede ser perfectamente descrito mediante su matriz de covarianza, ya que resulta ser no gaussiano. Esto hace que, a pesar de tener la matriz de densidad final, obtener analíticamente la interacción de este estado con diferentes compuertas, sin usar aproximaciones y calcular los valores de expectación, sea una tarea no trivial. Por estas complicaciones, en este trabajo no estudiamos el valor óptimo de β y los posibles aumentos en la probabilidad de éxito.

Hay que destacar que no todo modelo No Gaussiano va a llevar a una representación analítica de cada estado, dificultando incluso más la tarea de encontrar la información necesaria para la optimización de ciertos canales. Por otro lado, en una comunicación de larga distancia puede resultar muy difícil determinar un modelo de todas las interacciones y fallas en detectores, además de las intensidades en que cada fondo afecta a cada estado y como cambian en el tiempo. Esto hace que una correcta caracterización del fondo sea casi imposible en la mayoría de las comunicaciones reales.

Ante la falta de un modelo teórico que definan a las interacciones, necesitaremos encontrar formas de obtener los parámetros que maximizan la probabilidad de éxito, sin conocer modelos teóricos del sistema de estudio. Afortunadamente, en los últimos años, diferentes modelos de inteligencias artificiales revolucionaron áreas donde es imposible realizar modelos teóricos que permitan generar resultados óptimos. Áreas que fueron revolucionadas incluyen el manejo automático, juegos como el ajedrez y están comenzando a generar aplicaciones en áreas de la física. Por esta razón, será importante definir las bases del funcionamiento de estos algoritmos.

4 Inteligencias artificiales

Tres mil millones de vidas humanas se apagaron el 29 de agosto de 1997. Los supervivientes del fuego nuclear llamaron a aquella guerra el día del juicio final. Solo vivieron para hacer frente a otra pesadilla: la guerra contra las máquinas...

Sarah Connor, Terminator.

En el capítulo anterior hemos discutido un esquema fundamental de comunicación clásica a través de canales cuánticos. El mismo consiste en la utilización de estados coherentes, donde la información clásica (por ejemplo, un bit) es codificada en la fase del estado, el cual es enviado a través de cierto canal de transmisión, y decodificado por el recibidor. Hemos estudiado en detalle este problema, y caracterizado la figura de mérito para distintos esquemas de detección; según vimos, la probabilidad de éxito no solamente depende del aparato de medida, sino también del medio por el que es transmitido.

En esta línea, hemos discutido diferentes canales cuánticos de carácter Gaussiano y modelado un esquema no Gaussiano, con el objetivo de estudiar numéricamente cómo la calidad de la transmisión se ve afectada en términos de la intensidad del ruido. Bajo este paradigma — es decir, disponiendo un modelo del canal de transmisión ruidoso — es posible optimizar el detector de Kennedy mediante técnicas de optimización basados en modelo, por ejemplo el *Dual Annealing*, para obtener controles óptimos del detector (valores de desplazamiento).

Sin embargo, en la práctica resulta no trivial caracterizar correctamente dicho medio de transmisión (o incluso el aparato de detección). De este modo, falta de un modelo preciso que describa el medio, conlleva una dificultad a la hora de utilizar modelos de optimización que requieren total conocimiento del sistema. Para poner un ejemplo, la atmósfera funciona como un canal atenuador de la señal, y las condiciones climáticas al momento de la transmisión afectan al parámetro de atenuación [8].

Esto motiva la utilización de métodos que puedan *adaptarse* a las condiciones experimentales al momento de realizar la transmisión, incluso cuando no se tiene un modelo teórico. Para solucionar este tipo de problemas, en las últimas décadas se propuso un destacable cambio de paradigma tanto en el quehacer cotidiano como en la investigación científica [35] gracias al notorio avance en el campo de la inteligencia artificial y el aprendizaje automático. En particular, nos interesa discutir un campo del aprendizaje automático llamado *aprendizaje por reforzamiento*, el cual funciona bajo la premisa de encontrar *acciones* (o controles) óptimos para una tarea en concreto, sin necesidad de contar de antemano con un modelo teórico del sistema a estudiar. Estas técnicas están inspiradas en esquemas de aprendizaje mediante premios, como puede ser entrenar a un perro para que se siente; a partir de ejecutar ciertas acciones (como sentarse), el perro maximiza la cantidad de premios que recibe.

De esta forma, los modelos que estudiaremos a continuación son capaces de obtener información adaptando la estrategia de un *agente* a partir de su interacción con su *entorno*, recibiendo una recompensa según la calidad de los resultados obtenidos.

En este trabajo de diploma, implementaremos diferentes esquemas de aprendizaje automatizado y por refuerzo para el problema de la calibración de parámetros en el escenario de comunicación clásica con estados coherentes introducido anteriormente. Antes de discutir esta implementación, dejaremos sentadas las bases teóricas del funcionamiento de estas técnicas. Por esta razón, el objetivo de este capítulo será introducir algunos conceptos básicos de inteligencia artificial.

4.1. Aprendizaje reforzado

Al comenzar un experimento, todos los sucesos serán descritos en un entorno, este será el encargado de generar una observación s en el que describe la situación actual de forma aproximada, esta información será la que utilizará un *agente*, el cual es el encargado de realizar las diferentes acciones a, según la observación obtenida. Esto hará que el estado inicial cambie a un estado s' donde en general $s' \neq s$.

Para que el agente pueda mejorar sus futuras decisiones, necesitamos además de comunicar la observación, cuantificar que tan efectiva fue la acción a tomada anteriormente. Por esta razón, luego de aplicar una acción a, el entorno va a también devolver una recompensa $R(s') = R(a|s)$.

En general, estas recompensas van a ser una función de s y a, y van a ser determinadas por la persona que quiera entrenar al agente. Estas recompensas pueden estar basadas en un modelo teórico, pero también pueden ser determinadas a partir de cantidades que nos interesan maximizar. Por ejemplo, en un juego de *snake*, una recompensa podría ser $1 - d(s, c)$ siendo $d(s, c)$ la distancia entre la serpiente y la comida. Este tipo de interacción entre entorno y agente son conocidas como procesos de decisión de Markov y un diagrama de este puede ser visto en la figura 4.1. De esta forma, la interacción entre el entorno y el agente estarán da-

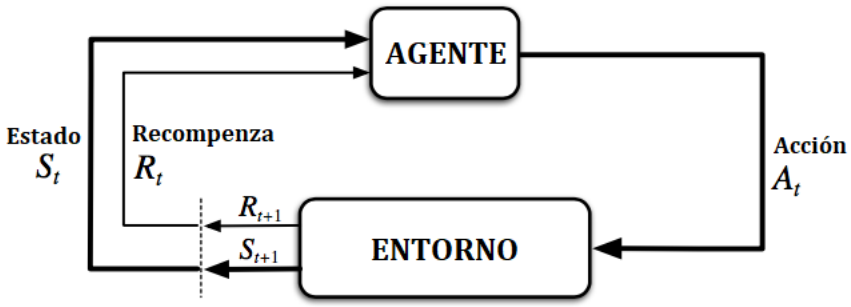

Figura 4.1: Diagrama del proceso de decisión de Márkov.

dos por pasos discretos en tiempo $t = 1, 2, ..., T$, definiendo así los*episodios* como la secuencia de acciones, estados y recompensas $S_1, A_1, R_1, ..., S_T$.

Ya teniendo estudiado como el entorno se ve afectado por la interacción de un agente, tenemos que ver como se produce una acción a dado un estado s. Para esto, el agente va a seguir una serie de reglas internas conocidas como *política* $\pi(s)$, la cual determina la acción a a tomar, cada vez que se observa un estado s. Estas políticas pueden ser deterministas o estocásticas, y pueden ser determinadas mediante algo tan sencillo como una serie de reglas pre-programadas, hasta modelos más complejos que hacen uso de redes neuronales (posteriormente se hará una introducción del tema).

Se define entonces el retorno medio como

$$G_t = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \,, \tag{4.1}$$

donde $\gamma \in (0, 1)$ cuantifica la incerteza que tenemos sobre decisiones muy futuras; en este sentido notemos que si bien la suma va hasta infinito, dicha suma puede ser truncada definiendo el valor de $R_{k'} = 0$ para valores $k' > T$. Un ejemplo típico de esta incerteza es en ajedrez, donde si jugamos correctamente podemos predecir que tan buena va a ser la posición dentro de dos o tres movimientos, pero es muy difícil predecir cómo va a seguir el juego dentro de setenta movimientos.

Como se puede intuir, el retorno es una cantidad que depende de la política que el agente siga, y del estado específico que se encuentre en el tiempo t. El objetivo principal del agente es entonces maximizar el retorno medio que se obtiene para un dado episodio, y para esto debe optimizar su política. De esta manera, se define la función *valor de un estado* según

$$V_\pi(s) = \mathbb{E}_\pi[G_t | S_t = s] \,, \tag{4.2}$$

que cuantifica cuán valioso es el estado s a futuro y en promedio, siguiendo la política π. Este promedio tiene un doble significado: por un lado, la manera de seleccionar acciones utilizando π puede ser estocástica (y de esta forma es un valor medio sobre esta distribución de probabilidad, dado s). Por otro lado, el estado del entorno observado por el agente, S_{t+1}, puede hacer la transición a partir de (S_t, A_t) de una forma estocástica. De esta manera, el valor medio de la ecuación de arriba también considera dicha estocasticidad.

Esta función cuantifica no solo que tan *valioso* es estar en el estado s, sino que determina que tan buena es la trayectoria completa seguida por el agente, al elegir π como estrategia para la toma de decisiones (acciones). Esto es especialmente importante para entornos donde una decisión presente puede afectar mucho a resultados futuros, por ejemplo en un juego de ajedrez.

Por otro lado, también se define la función *valor del estado-acción* según

$$Q_\pi(s, a) = \mathbb{E}_\pi[G|S_t = s, A_t = a] \,. \tag{4.3}$$

Es posible ver que la función valor de un estado s y su función valor estado-acción están relacionadas según

$$V_\pi(s) = \sum_{a \in A} Q_\pi(s, a)\pi(a|s) \,, \tag{4.4}$$

En general, el aprendizaje reforzado se enfoca en encontrar las políticas π que llevan al agente por un camino que maximiza a la función $V_\pi(s)$. Es decir, el objetivo es encontrar una política π que pueda guiar al agente hacia estados cuya función valor sea la más alta. La dificultad aquí radica en que dicha función valor depende no solo del estado, sino también de la política. Por esto mismo, el campo del aprendizaje por refuerzo tiene mucha similitud con la toma de decisiones, y los algoritmos que se utilizan para resolver este problema son muy variados [36]. Entre las muchas ideas, algoritmos y estrategias, podemos distinguir entre dos clases de *reglas internas*, utilizadas por el agente para resolver el problema de la toma de decisiones: *(i)* basadas en un modelo, donde el agente utiliza cierta representación del entorno y de la función objetivo para optimizar su estrategia, y *(ii)* libres de modelo, donde el agente aprende *a ciegas*, y necesita estructurar su búsqueda a partir de las diferentes interacciones con el entorno en pos de optimizar la cantidad de premios.

En este trabajo nos enfocaremos en la segunda línea, es decir, en los esquemas de aprendizaje que funcionan sin un modelo previo. Si bien la falta de un modelo (como por ejemplo disponer de la función probabilidad de éxito en función de los desplazamientos) supone una desventaja a la hora de optimizar una estrategia (ya que el agente debe, inevitablemente, interactuar con el entorno), esto también presenta ciertas ventajas. En primer lugar, el agente no estará sesgado frente al modelo del entorno que le sea dado; esto ha resultado ser un punto positivo a la hora de optimizar estrategias para juegos de mesa tales como el Go o el ajedrez, descubriendo estrategias nunca antes observadas a partir de una fase considerable de entrenamiento (notar que la cantidad de recursos computacionales utilizados para este entrenamiento alcanza del orden de millones de dólares). En segundo lugar, si consideramos el problema de comunicación con estados coherentes, la posibilidad de adaptarse frente a cambios en el entorno, tanto por una nueva circunstancia experimental como por una falla en el detector que no puede ser modelada, constituye una ventaja a favor de los algoritmos que funcionan sin un modelo. De este modo, el agente es más *versátil* a la hora de calibrar el detector. Pero también son más propensas a caer en mínimos locales, produciendo efectos como el *overfiting*, teniendo buenos resultados en algunas regiones y muy malos en otras. Estos agentes también suelen necesitar una mayor cantidad de datos en su entrenamiento, debido a que carecen de modelos que los ayuden al tomar decisiones.

En un esquema libre de modelo, la política a seguir por el agente comienza siendo aproximadamente aleatoria. Esto corresponde a una fase exploratoria del agente, que debe estudiar cuán efectivas son las acciones que dispone respecto a la recompensa observada. Según las observaciones y recompensas observadas durante el entrenamiento, el agente dispone de ciertas reglas, dadas por el algoritmo específico de aprendizaje automático, para modificar su política. No obstante, para poder realizar este cambio, no solo se necesita la

recompensa actual, sino que también es importante determinar que tan buenas son las recompensas futuras de un estado s. Esto se debe a que el agente busca maximizar la recompensa observada *en promedio* para cada estado.

4.1.1. Optimización de la política

Siendo el objetivo final el encontrar una política óptima, es decir, que maximiza $V_\pi(s)$ sobre todos los estados s, existen una gran variedad de estrategias. En este trabajo vamos a enfocarnos en estrategias que parametrizan a la política a partir de parámetros θ, de forma denotaremos a la política actual como $\pi_\theta(s) \equiv \pi(s)$. Si pensamos a los θ_i como direcciones, podemos buscar máximos siguiendo la dirección del gradiente de una función $\mathcal{J}(\theta)$, tal que π_θ al final de su entrenamiento va a maximizar siempre la función $V_\pi(s)$.

Supongamos entonces que tenemos un estado inicial S_1, luego podemos definir una función que cuantifique cuán buena es nuestra política según

$$\mathcal{J}(\theta) = V_{\pi_\theta}(S_1) \,. \tag{4.5}$$

Esta definición produce un valor numérico igual a la recompensa obtenida por una política π_θ que empieza en un estado S_1. Como la dirección $\vec{\nabla}\mathcal{J}(\theta)$ maximiza a la función $V_\pi(s)$, vamos a evolucionar a nuestros parámetros θ de forma

$$\theta \to \theta + \alpha\vec{\nabla}\mathcal{J}(\theta) \,, \tag{4.6}$$

donde α es un parámetro externo, que podrá servir para observar cambios en la función $\mathcal{J}(\theta)$. Este parámetro es sumamente útil para evitar mínimos locales en la función, y en la jerga del aprendizaje automático es calificado como un *hiper-parámetro* definido por el usuario.

A pesar de la técnica descrita, un agente libre de modelo no tiene acceso a la función $\mathcal{J}(\theta)$, ya que desconoce las funciones valor del estado $v_\pi(s)$. Esto hace en principio que sea imposible su determinación exacta, y haya que estimar dichas cantidades.

La forma más fácil de determinar un gradiente es mediante *pequeñas variaciones*. En este algoritmo se elijen muchos puntos al azar muy cercanos y se mueve a la política del agente en la dirección que generó la mayor recompensa $V(s)$. Si se tiene una suficiente cantidad de puntos, entonces la dirección elegida va a ser similar a la dirección del gradiente, pero a un costo computacional muy alto.

Para optimizar este algoritmo se suele aprovechar el hecho de que si α es pequeño, entonces en general se cumple que $\nabla\mathcal{J}(\theta_1) \approx \nabla\mathcal{J}(\theta_1 + \alpha\nabla\mathcal{J}(\theta_1))$. De esta forma, si se conoce una vez el gradiente en un punto θ_1, entonces bajo esta aproximación, la dirección del gradiente en un punto $\theta_2 \approx \theta_1$ tiene que ser similar al gradiente anterior. Así, eligiendo puntos en una dirección similar a la del gradiente en θ_1 anteriormente calculado, podemos estimar el gradiente en θ_2, sin necesidad de inspeccionar todo el espacio. La ventaja de este algoritmo es que es muy fácil de implementar, aunque sigue siendo muy ineficiente y no es capaz de ver cambios en la dirección del gradiente cuando α no es lo suficientemente pequeño [37].

Para reducir los tiempos de cómputo y adaptarse a cambios en el gradiente, podemos recurrir al *teorema de la política primaria* [37], según

$$\vec{\nabla}\mathcal{J}(\theta) = \mathbb{E}_{\pi_\theta}[\vec{\nabla}\ln\pi(a|s,\theta)Q_\pi(s,a)] \,. \tag{4.7}$$

La importancia de este teorema recae en que en general todos los términos de la derecha son conocidos o es posible estimarlos a partir de repetidas interacciones con el entorno, consiguiendo así una aproximación del gradiente buena y rápida. El pseudocódigo de la implementación puede ser visto en el Algoritmo 1.

Algoritmo 1 Teorema de la Política Primaria

```
for i = 1, 2, ... do
    for k = 1,2,..., N do
        ∇J(θ)[k] += (θ[k] - θ_i[k]) * R(θ_i) * P(θ_i[k])/ N
        θ[k] += α∇J(θ)[k]
    end for
end for
```

En este algoritmo $\theta[k]$ son las componentes de los parámetros de la política π_θ. Luego $P(\theta_i[k])$ es la probabilidad de que con una distribución Gaussiana, centrada en θ_{actual}, se obtenga el parámetro θ_i y $R(\theta_i)$ es la recompensa asociada a una política con θ_i.

4.1.2. Estimación de $V_\pi(s)$

Como vimos anteriormente, una función importante a conocer es $V_\pi(s)$, la cual en general resulta imposible de determinar con precisión. Esto nos lleva a determinar la evolución de nuestro agente mediante la función $Q_\pi(s, a)$ como estimador, aunque esta no siempre es sencilla de conseguir, existen muchas aproximaciones para realizar encontrarla, por esto en esta sección nos centraremos en las dos que vamos a utilizar.

El primer método es conocido como *Método Monte-Carlo* en el cual no se intentar modelar $V_\pi(s)$ de ninguna forma. Se parte de un estado s_0 y se permite que el agente interaccione con el sistema durante T episodios. De esta forma, el retorno medio de un estado s, con una política π_θ, está dado por

$$V_{\pi_\theta}(s) = \frac{\sum_{t=1}^{T} \mathcal{I}[S_t = s] G_t}{\sum_{t=1}^{T} \mathcal{I}[S_t = s]} , \tag{4.8}$$

donde $\mathcal{I}[S_t = s]$ es una función binaria que evita que se sumen recompensas de estados repetidos. Este método implica que la interacción entre agente y entorno se repite a través de varios episodios, en pos de estimar el valor de la función $V(s)$ para cada estado, bajo una política π.

Por otro lado, un agente puede también usar estimativos de la función mediante el algoritmo *Actor-Crítico*. En este caso la política $\pi_\theta(s)$ y la recompensa $V_w(s)$ estimadas, van a ser funciones con parámetros θ y w que serán ajustados durante el entrenamiento. Así este algoritmo comienza explorando zonas de forma aleatoria y a mientras las políticas y recompensas se hacen más precisas, se obtienen parametrizaciones óptimas de ambas funciones. Este algoritmo entonces utiliza dos modelos:

- Crítico: Actualiza los parámetros w viendo una serie de estados S y otorgándoles una recompensa R.
- Actor: El actor es el que cambia las decisiones de la política, en general va a ver estados diferentes S' y va a obtener una recompensa otorgada por el crítico $V_w(s)$.

De forma que si el Crítico obtiene recompensas similares a las reales sin hacer el experimento, el Actor puede determinar de forma exacta los máximos locales, sin necesidad de realizar todos los experimentos. Basados en esta idea aparecen muchos algoritmos como *DQN* o *PPO* [38, 37]. Este último va a ser utilizado más adelante y en breve será introducida la idea del algoritmo.

4.2. Redes Neuronales

Un método para modelar a $\pi_\theta(s)$ y a $V_w(s)$ que es muy utilizado hoy en día, es mediante la implementación de redes neuronales.

Las redes neuronales LTU (linear threshold unit) están organizadas como una serie de nodos conectados entre sí. Formando varias capas, como se puede observar en la figura 4.2. La primera capa es conocida como "Capa de entrada", la última es conocida como "Capa de salida" y las intermedias "Capas ocultas". La capa

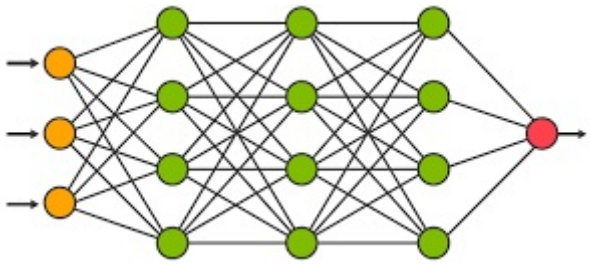

Figura 4.2: Diagrama de la estructura de una red neuronal.

de entrada es la que va a recibir las observaciones. Cada nodo de entrada toma un valor real x_i y va a tener conexiones a las cuales se les asigna un peso w_i.

Para calcular el valor de los siguientes nodos van a tomar entonces los valores $f(x_1w_1+x_2w_2+...+x_nw_n)$ siendo x_i los valores de todos los nodos que están conectados a este, w_i los pesos asignados a cada uno y $f(x)$ la llamada "función de activación". Esto es repetido en todos los nodos de la red hasta llegar a la capa de salida. Un ejemplo de función de activación muy utilizado es la función sigmoide

$$f(o) = A\frac{1}{1+e^{-o/\eta}} \tag{4.9}$$

con η un valor a elegir. Esta función es útil, ya que tiene un crecimiento continuo y va a tomar valores en un intervalo determinado $(0, A)$.

Para optimizar los pesos w_i de la red se puede utilizar el método del gradiente y métodos basados en diferenciación automática [37]. A continuación discutiremos ciertos problemas comunes que se presentan a la hora de recurrir a estos métodos.

4.2.1. Inconvenientes del descenso por gradiente

Este método de optimización de todos modos no es óptimo, ya que en general el descenso por gradiente lleva al máximo local más cercano. Esto hace que si la función $\mathcal{J}$ tiene muchos máximos locales (como en una red neuronal con cientos de parámetros), pueda resultar imposible converger a máximos globales. Para solucionar este problema, existen muchos métodos para intentar evitar caer en los mismos, en esta subsección nombraremos otros 3 comunes y simples que podrían resultar de interés

- **Muchos puntos:** Una de las estrategias más simples es en vez de entrenar una inteligencia artificial, podemos entrenar muchas, que partan todas de diferentes puntos y quedarnos con la que generó el mejor resultado.

- **Optimización por momento:** Otra forma es cambiar el parámetro α de la ecuación (4.6) según una función
$$\alpha \to \beta\alpha + \eta\vec{\nabla}\mathcal{J}(s)\,, \tag{4.10}$$
donde β y η son nuevos parámetros que cuantizan cuanto momento mantiene el sistema.

  De esta forma, si el cambio es muy grande, entonces los parámetros θ van a cambiar mucho, haciendo que los máximos locales más pequeños puedan ser evitados.

- **Optimización térmica:** En este algoritmo se asigna una temperatura inicial T_0, al avanzar en el entrenamiento, esta temperatura se va a ir viendo reducida. Luego se define una función $f(T)$, tal que si $T \gg 1$ entonces $f(T)$ resulta en un comportamiento aleatorio, pero si $T \ll 1$ entonces $f(T) \approx$

$\vec{\nabla}\mathcal{J}(\theta)$. Este movimiento semi-aleatorio permite evitar converger a mínimos locales y al seguir en parte la dirección del gradiente, el algoritmo tiende a converger a un máximo, luego de evitar la mayoría de los mínimos locales.

El optimizador más empleado hoy en día es un algoritmo llamado Adam [37], el cual implementa varias de estas ideas, incluido la optimización por momento.

Existen muchas formas de utilizar las redes neuronales para generar agentes más versátiles capaces de adaptarse a diferentes observaciones y evitar la mayoría de máximos locales. Por ejemplo, en los últimos años surgió un algoritmo que tomó gran protagonismo llamado *Primary Policy Optimization* (PPO). Este algoritmo fue desarrollado por OpenAI y fue capaz de aprender a jugar el juego *DOTA 2* y ganarle al equipo humano campeón del mundo de ese año [38, 39]. Aunque la implementación exacta de este algoritmo supera los contenidos de este trabajo (para más detalles consultar el apéndice B), en resumen, este algoritmo consta de 2 redes neuronales, una que intenta predecir las futuras recompensas y otra que aprende a elegir los caminos óptimos siguiendo estas recompensas. Este algoritmo es sumamente útil y en el futuro será utilizado para permitir que nuestro agente se adapte a diferentes entornos.[38]

En resumen, en este capítulo hemos revisado los conceptos básicos de aprendizaje por refuerzo, y enumerado diferentes estrategias a la hora de estimar y optimizar las políticas de un agente que se enfrenta a un entorno desconocido. Con esto, estamos en posición de discutir las diferentes implementaciones que sean útiles a la hora de distinguir estados en canales reales.

5 Implementación de inteligencias artificiales

Digamos que hay un pedazo de carne. Los piratas tendrían un banquete y se lo comerían, pero los héroes lo compartirían con otras personas. ¡Yo quiero toda la carne!

Monkey D. Luffy

En el tercer capítulo hemos discutido un esquema de comunicación clásica a través de canales cuánticos, el cual captura un problema fundamental proveniente de la teoría de la información cuántica. El mismo consiste en codificar el valor de un bit clásico $k \in \{0, 1\}$ en la fase de un estado coherente mediante estados $|\pm\alpha\rangle$. Este experimento, por ejemplo, describe a un pulso láser que utiliza como canal de transmisión a la atmósfera. De esta forma, el receptor realiza una medida cuántica sobre el estado, y a través del resultado de dicha medida provee un estimador $\hat{k}$ del valor del bit transmitido.

De esta manera, el receptor tiene una libertad muy grande a la hora de elegir la medida cuántica que realizar sobre el estado coherente que contiene la información a decodificar. Particularmente, los detectores tipo Kennedy optimizado, consisten en aplicar un desplazamiento al estado, de forma que se maximice la probabilidad de éxito al realizar una medida con un fotodetector.

Si bien la probabilidad de éxito del detector tipo Kennedy puede ser escrita según la ecuación (3.12), y los desplazamientos óptimos pueden ser encontrados mediante técnicas de optimización como el *dual-annealing*. En este trabajo estamos interesados en el estudio de técnicas que no utilizan un modelo teórico, es decir, que no tienen acceso directo a la probabilidad de éxito. Como mencionamos anteriormente, esto tiene diferentes justificaciones, que van desde las dificultades que se encuentran a la hora de modelar situaciones experimentales y de caracterizar los tipos de fondos, hasta problemas como la imposibilidad experimental de acceder a los observables de forma perfecta debido a errores en los detectores.

A continuación pondremos foco en un esquema totalmente centrado en el experimento de comunicación con detección de tipo Kennedy optimizado, donde el agente necesita implementar un desplazamiento, observar el resultado de la medida y clasificar el estado recibido. Basado en si el agente determina correctamente o no el estado original, se le otorgará un *premio*, cuantificando el desempeño del agente en dicho experimento.

En el esquema más básico, la referencia [9], utiliza un premio de carácter binario, siendo 1 si el agente adivina correctamente, y 0 caso contrario; tanto intuitiva como analíticamente puede verse que, en promedio, el valor del premio para una configuración fija de desplazamientos estará dado así por la probabilidad de éxito asociada. De esta manera, en la referencia [9] se propusieron y estudiaron diferentes esquemas de aprendizaje por refuerzo para la calibración de detectores de estados coherentes bajo el paradigma descrito, haciendo foco en un algoritmo llamado *Q-learning*. Sin embargo, la implementación de dicho algoritmo llevada a cabo en [9] tiene ciertas limitaciones. Por ejemplo, no tienen en cuenta nociones de aproximación e interpolación, ya que no pueden utilizar información acumulada durante interacciones previas con el entorno (es decir, experimentos anteriores durante el entrenamiento). Por otro lado, si bien la definición de premio descrita en el trabajo original converge a la probabilidad de éxito, no queda descartado que existan otras funciones de premio que *aceleren* el esquema de aprendizaje en términos de repeticiones del experimento. Además, este método solo encuentra desplazamientos óptimos haciendo uso de comportamientos en media, por lo que resulta muy complicada la implementación de agentes que puedan cambiar sus decisiones teniendo en consideración comportamientos observados a corto plazo.

Basados en las observaciones del párrafo anterior, en este capítulo propondremos y estudiaremos tres modelos de aprendizaje por refuerzo diferentes, que permiten a un agente encontrar configuraciones óptimas — o cercanas a las óptimas — de un detector tipo Kennedy en la presencia de fuentes de ruido, de forma completamente agnóstica. Estas implementaciones deben entenderse como una alternativa a aquellas de la referencia [9], donde aquí el objetivo es avanzar sobre los problemas enumerados en el párrafo anterior.

A continuación se presentarán métodos basados en diferentes técnicas que realizan una optimización de la política, maximizando una recompensa. En primer lugar, estudiamos un algoritmo libre de observaciones que sigue el descenso por gradiente haciendo uso de un método Monte-Carlo. Luego propondremos una modificación que utiliza como observación $|\alpha|$ y sigue la misma metodología para maximizar la recompensa. Por último, estudiamos un algoritmo que utilizando las observaciones de varios experimentos, maximiza la probabilidad de éxito siguiendo una optimización de tipo PPO.

5.1. Entorno y agentes

Para implementar diferentes inteligencias artificiales haremos uso del mismo sistema de comunicación que en los capítulos anteriores, un emisor A envía uno de los estados $\{|-\alpha\rangle\ ,\ |\alpha\rangle\}$ a un receptor B, el cual intentará distinguir entre ellos realizando una medida de tipo Kennedy.

Para obtener el valor óptimo de desplazamiento en este tipo de comunicaciones, el agente será entrenado en un *periodo de calibración*. Donde, el entorno y las interacciones con el agente van a seguir los siguientes pasos:

1. A enviará 20 veces el estado $|-\alpha\rangle$. La cantidad de estados que envía A no fue determinada de forma rigurosa, sino que se eligió por observaciones del rendimiento de muchas inteligencias artificiales diferentes.

2. A enviará el estado $|\alpha\rangle$ otras 20 veces.

3. B le aplicará a los estados recibidos un desplazamiento β_i con $\beta \in \mathbb{R}$ y realizará una medida de tipo Kennedy.

4. B clasifica al estado recibido como $|-\alpha\rangle$ si no obtiene fotones y votará por el estado $|\alpha\rangle$ en caso contrario. De esta forma, luego de seguir un criterio de estado más probable, comparará con las clasificaciones correctas y así estimará la probabilidad de éxito del parámetro β_i propuesto.

5. Una vez estimada la probabilidad de éxito, el entorno devolverá una recompensa, utilizando la función

$$R = 2\Big(P_{\text{extio}}^{\text{exper}}(\beta) - 0{,}5\Big)\ .$$

6. Se repite el proceso desde el primer paso y se vuelven a calcular los estimadores de la probabilidad de éxito, para el desplazamiento propuesto por B.

Notar que si la cantidad de estados enviados por A fuera N con $N \to \infty$, entonces la probabilidad de éxito podría ser estimada exactamente. De todos modos, esto resulta inviable en un experimento real tomar una gran cantidad de datos, ya que requeriría de mucho tiempo de señal y no es posible enviar información durante el periodo de calibración. En general, tampoco se puede implementar computacionalmente N muy grande, porque realizar las simulaciones consume mucho tiempo y recursos.

Por otro lado, si el valor de N es demasiado chico, entonces la estimación de la probabilidad de éxito para un dado β tiene fluctuaciones demasiado grandes, de forma que solo se pueden estudiar comportamientos en media a largo plazo. Esto hace que determinar la cantidad óptima de estados enviados por A no sea trivial. Para determinar $N = 20$ en este trabajo, se realizaron múltiples experimentos con un mismo ordenador, se compararon los resultados llegados por los diferentes agentes y sin mayor criterio se determinó este valor.

Teniendo en cuenta la descripción que hicimos del entorno y sus recompensas, en la siguiente sección describiremos un esquema muy básico que puede ser utilizado para determinar parámetros óptimos de desplazamiento, en una comunicación real con bajas fluctuaciones en el fondo.

5.1.1. Agente: Sr. Magoo

La primera implementación que haremos está basada en un agente ciego al entorno, que no posee observaciones del mismo, pero encuentra los parámetros óptimos realizando una maximización de las recompensas. Podemos ver el pseudocódigo que ilustra el comportamiento del algoritmo 2, el cual está basado en el descenso por gradiente. Pero la implementación exacta del código puede ser vista en detalle en el repositorio [40].

Algoritmo 2 Sr.Magoo

```
gamma = 1
Recompensas = []
Acciones = []
for i = 1, 2, ..., pasos do
    while not done do
        β = Desplazamiento + RandomGaussian(0, gamma)
        Observación, Recompensa, done = entorno.paso(β)
        Recompensas.append(Recompensa)
        Acciones.append(Acción)
    end while
    Desplazamiento = agente.Actualizar_Politica(recompensas, acciones, gamma)
    gamma *= 0.995
end for
```

Donde la función Actualizar_Política() actualiza el valor del desplazamiento siguiendo la dirección del gradiente, determinado por el teorema de la política primaria. Además, entorno.paso(Acción) es una función que simula el experimento utilizando el desplazamiento β y devuelve las observaciones y recompensas.

Como la recompensa otorgada lleva en media al valor óptimo de β, este agente debería obtener resultados similares a los de un modelo de optimización por gradiente, que tenga acceso a la probabilidad de éxito. Entonces, para tener una comparación entre la optimización Dual-annealing utilizando la probabilidad de éxito y nuestro método que no utiliza un modelo teórico, se estudió una comunicación en un entorno libre de fondo en el intervalo de intensidades $|\alpha| \in [0, 1]$, donde la probabilidad de éxito es menor al 99 %.

De esta forma se entrenó 10 veces al agente, con los estados con $|\alpha| = \{0{,}1; 0{,}2; ...; 1\}$, y se compararon las probabilidades de éxito obtenidas en el entrenamiento, con las determinadas por el algoritmo Dual-Annealing usando la probabilidad de éxito teórica. Como se puede apreciar en la figura 5.1, ambas probabilidades de éxito resultan prácticamente idénticas, mostrando así la utilidad de este algoritmo para encontrar soluciones óptimas, cuando no se tiene un modelo teórico de la probabilidad de éxito.

Para entender como el algoritmo encontró los parámetros óptimos, podemos observar la figura 5.2, que muestra las decisiones propuestas por el agente en diferentes pasos de su entrenamiento. El cambio en las decisiones comienza siendo rápido y va reduciendo su velocidad cuando tiene una estadística lo suficientemente grande. De forma que se evitan la mayoría de mínimos locales debido a la observación de un rango más amplio de valores posibles y sus respectivas recompensas.

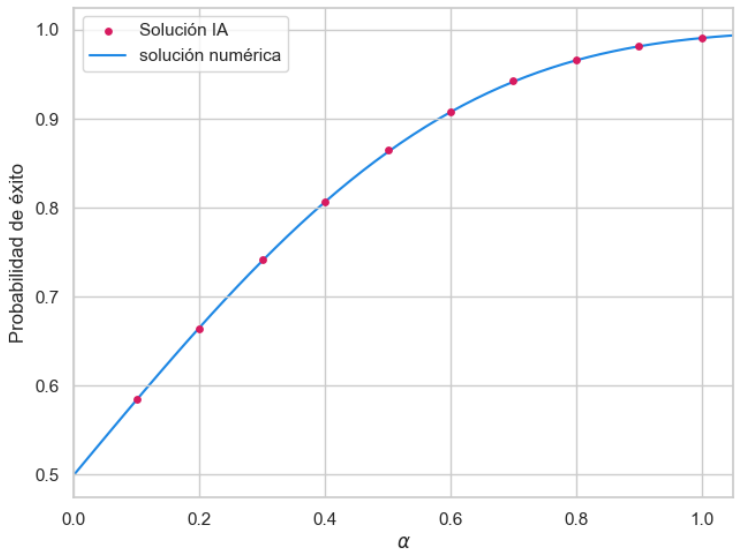

Figura 5.1: Rendimiento de inteligencia artificial comparado con rendimiento del detector Kennedy optimizado.

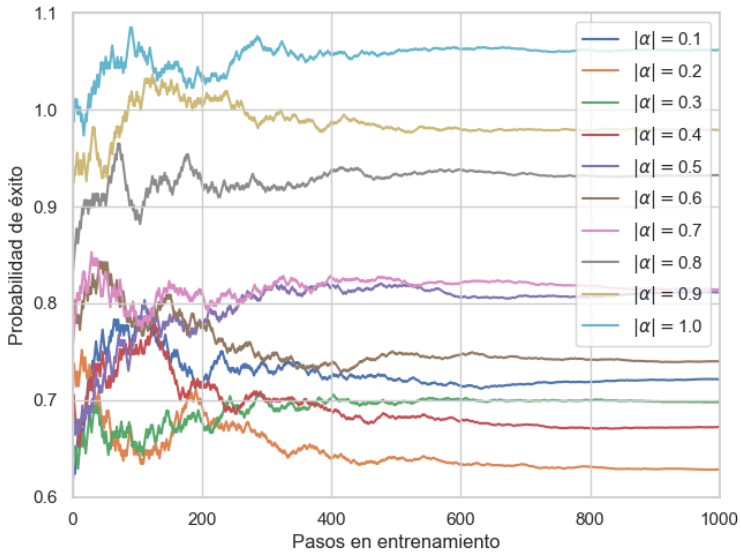

Figura 5.2: Evolución de la decisión tomada por el agente.

Aunque este algoritmo en general converge relativamente rápido y de forma similar a otros métodos de optimización teóricos, el agente no tiene observaciones del sistema. Esto hace que sea incapaz de adaptarse sobre el mismo, teniendo así que repetir el entrenamiento del agente cada vez que el sistema cambia. Esto motiva la idea de utilizar agentes que puedan usar observaciones del entorno para adaptarse mejor a los cambios del mismo.

5.1.2. Cambios en la energía de la señal

En particular, supongamos que tenemos una comunicación cuántica y una clásica. De forma que A va a comunicar clásicamente el valor $|\alpha|$ con una cierta precisión y luego envía a uno de los estados $\{|-\alpha\rangle, |\alpha\rangle\}$. Teniendo esto en cuenta, un agente puede tomar el valor de $|\alpha|$ como observación, para así adaptarse a cambios en el mismo, sin necesidad de rehacer toda la calibración. De esta forma, para usar esta información, el algoritmo que utilizaremos es similar al visto en la subsección anterior, donde la principal diferencia entre los algoritmos 3 y 2, es que existen múltiples parámetros que ajustan a la política del nuevo agente y la decisión tomada depende de $|\alpha|$.

Algoritmo 3 Cambios de intensidad.

```
i = 1
gamma = 1
Recompensas = []
Acciones = []
for i < pasos do
    while not done do
        Acción = agente.Acción(|α|) + RandomGaussian(0, gamma)
        Observación, Recompensa, done = entorno.Paso(Acción, |α|)
        Recompensas.append(Recompensa)
        Acciones.append([Acción, |α|])
    end while
    agente.Parámetros = agente.Actualizar_Política(recompensas, acciones, gamma)
    gamma *= 0.995
    i++
end for
```

A pesar del esquema definido anteriormente, todavía se necesita definir una política que genere los valores óptimos de β para una intensidad dada. Para encontrarla, podemos considerar que si la curva que describe a los parámetros óptimos es suave y no varía demasiado con $|\alpha|$, entonces la curva óptima puede ser aproximada con un polinomio. Así proponemos definir a la política en función de $|\alpha|$ como

$$\pi_\theta(|\alpha|) = \sum_{i=0}^{3} \theta_i |\alpha|^i \,. \tag{5.1}$$

Notar que esta política no es ideal y en caso de querer realizar un experimento real, sería prudente basarla parcialmente en algún modelo teórico. A pesar de este comentario, podemos comparar la probabilidad de éxito obtenida bajo esta aproximación, con los resultados del algoritmo Dual-Annealing haciendo uso de la probabilidad de éxito en un canal libre de fondo. Si graficamos las probabilidades de éxito en el caso libre de fondo, obtenemos la figura 5.3. Donde se puede observar que la probabilidad de éxito de este método difiere en menos de un 2 % con respecto al límite teórico.

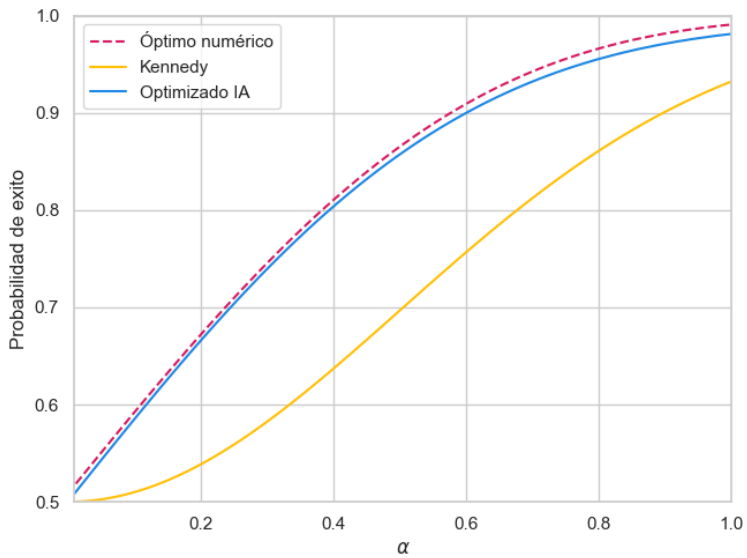

Figura 5.3: Comparación entre detección Kennedy, Kennedy Optimizada y predicción de la IA.

Para estudiar cómo cambian las decisiones tomadas por el agente al discriminar entre estados $|\pm\alpha\rangle$, se puede fijar $|\alpha| = 0{,}5$ y graficar la decisión tomada en diferentes instancias del entrenamiento. Esto da forma al gráfico de la figura 5.4, donde se nota que el agente es más versátil, pero a cambio requiere cinco veces más datos para su calibración.

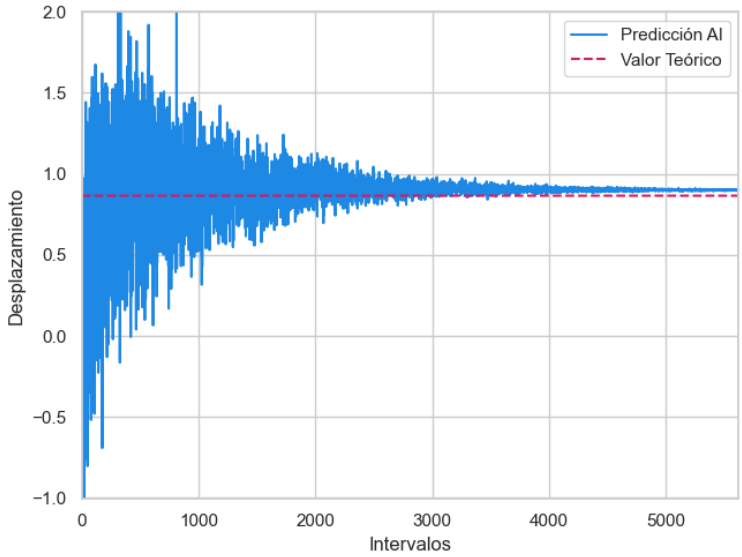

Figura 5.4: Desfasajes propuestos por la inteligencia, para discriminación de estados con $|\alpha| = 0{,}5$, en función del capítulo del entrenamiento.

Finalmente, es importante ver que esta inteligencia solo utiliza como observación $|\alpha|$. Esto significa que no tiene ninguna información sobre la interacción del canal con los estados, haciendo así que agente no pueda adaptarse a ningún cambio en el fondo, ya que no es capaz de verlo. Por esta razón, queda motivado el último agente que estudiaremos. Este será basado en el algoritmo PPO y usará como observación información tomada de los últimos 10 experimentos realizados, teniendo así la posibilidad de ver como diferentes desplazamientos afectan al estado recibido.

5.1.3. Una IA para gobernarlas a todas

Como vimos a lo largo de este capítulo, se pueden diseñar agentes muy diferentes, según las observaciones devueltas por el entorno. Esto también mostró que los resultados y los tiempos de entrenamiento son muy dependientes de estas observaciones.

Hasta ahora estudiamos 2 agentes, uno ciego al entorno y otro capaz de adaptarse a cambios en $|\alpha|$. Pero estos modelos no tienen la cantidad de observaciones necesarias para adaptarse a cambios en el fondo. Esto resulta un inconveniente en situación donde el fondo cambian relativamente rápido, como podría ser una comunicación satelital en un día nublado. El objetivo de esta sección será entonces implementar una inteligencia artificial capaz de adaptarse a cualquier cambio en el sistema.

Para esto vamos a comenzar teniendo en consideración los 3 valores: $|\alpha|$, un desplazamiento $\beta_i \in \mathcal{C}$ y la probabilidad de éxito estimada, asociada al parámetro β_i. Para que el agente obtenga el parámetro óptimo, este seguirá los siguientes pasos:

1. El entorno da como observación el valor $|\alpha|$ con el que se realizarán los experimentos.

2. El agente propondrá un desplazamiento β_i.

3. El entorno realizará el experimento descrito en la sección 5.1 devolviendo una recompensa y las observaciones β_i y $P_{\text{exito}}^{\text{estim}}(\beta_i)$.

4. El agente realizará nuevamente el experimento, pero ahora como observación agrega los valores β_i y $P_{\text{exito}}^{\text{estim}}(\beta_i)$.

5. Repite este proceso k veces, de forma que en cada propuesta de β añade más observaciones, hasta encontrar el valor óptimo de β.

A diferencia de las implementaciones anteriores, en vez de utilizar al estimativo de la probabilidad de éxito como recompensa, vamos a utilizar a la función

$$R = P_{\text{exito}}^{\text{estim}} \gamma^{N-n} \,, \tag{5.2}$$

donde n es la cantidad de experimentos realizados hasta el momento, N es la cantidad total de experimentos y γ es un parámetro que representa cuando es importante que el agente realice las decisiones más cercanas a la óptima, en nuestro caso fijamos a $|\gamma| = 0{,}5$. Al definir así a las recompensas, se permite que el agente no sea penalizado por sus primeras decisiones, consiguiendo así más información del entorno al principio y convergiendo al valor óptimo sobre el final, donde las recompensas están basadas en llegar al desplazamiento que maximiza la probabilidad de éxito.

Para ver la capacidad que tiene este agente de producir mejoras en la probabilidad de éxito, cuando no se tiene un modelo del fondo, podemos usar para entrenar y comprar a un canal de comunicaciones tomando el circuito de la figura 5.5, el cual está inspirado en los modelos de fondo anteriormente estudiados. En particular, para realizar la comparación, podemos fijar a los parámetros como

$$\begin{aligned}
\text{Estado de Fock} &\rightarrow |n\rangle = |1\rangle \\
\text{Squeezing de canal} &\rightarrow z = \frac{i}{N} \\
\text{Separador de Haces} &\rightarrow \theta = \frac{\pi}{8} \\
\text{Desplazamiento} &\rightarrow \Delta = \frac{i\alpha}{2N} \\
\text{Fase} &\rightarrow \phi = \frac{i\pi}{2N} \,,
\end{aligned}$$

donde i es un entero $\{0, 1, ..., k\}$. De esta forma, si graficamos la probabilidad de éxito estimada del agente

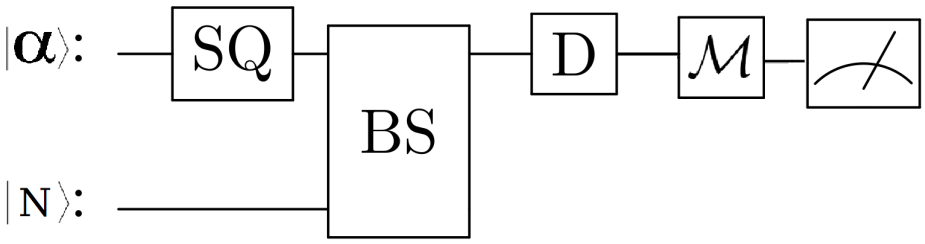

Figura 5.5: Modelo de fondo sobre el que trabajará nuestra inteligencia artificial.

y la comparamos con la estimada para el desplazamiento obtenido de optimizar la probabilidad de éxito sin fondo, obtenemos la figura 5.6. Donde se puede ver que los desplazamientos predichos por el agente producen probabilidades de éxito ligeramente superior en media al modelo teórico que ignorar al fondo.

Hay que destacar que para lograr estos resultados, el agente tuvo que entrenar con varios millones de estados, esto hace que los entrenamientos tengan que constar de días de señal o hechos en parte mediante simulaciones (aunque en principio se harían una única vez). A pesar de que los resultados obtenidos no son completamente satisfactorios, hay que notar varios factores. El primero es que al no tener un modelo teórico, no somos capaces de determinar la capacidad máxima de mejora para este canal particular. Lo segundo y más importante es que estas decisiones fueron hechas empleando tan solo 200 estados, de forma que una implementación diferente del agente y cambios en las recompensas, podrían generar resultados ampliamente superiores a los vistos en este trabajo.

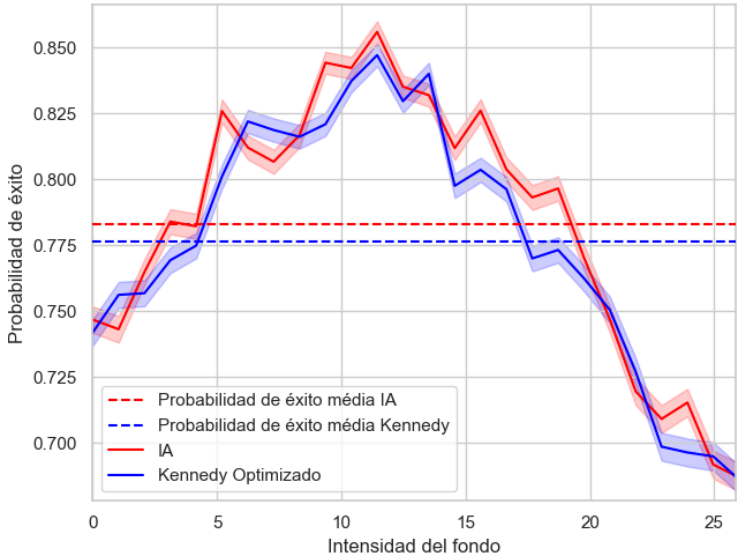

Figura 5.6: Probabilidad de éxito estimada por el agente y el desplazamiento β (obtenido de optimizar la probabilidad de éxito sin fondo), variando la intensidad de fondo.

6 Conclusiones

El camino fue un infierno, pero me trajo al paraíso

Taylor Swift

En este trabajo estudiamos el uso de inteligencias artificiales para la discriminación de estados coherentes en canales reales.

Para esto primero se estudiaron los fundamentos de la mecánica cuántica, definiendo operadores de densidad, viendo como obtener valores de expectación y como estudiar los observables locales de subsistemas. Luego pasamos a estudiar sistemas de variable continua donde vimos como calcular la interacción entre estados Gaussianos y compuertas que preservan esta propiedad. Esto se hizo con el fin de trabajar en un ejercicio canónico de comunicación cuántica, en el que se trata de distinguir entre estados coherentes con fases opuestas. De esta forma vimos varios métodos de discriminación en canales ideales, como la discriminación Homodina, Heterodina y de tipo Kennedy, mencionando las de tipo Dolinar y comparando así con el límite de *Helstrom*. Una vez estudiados estos métodos vimos como las comunicaciones reales se caracterizan por diferentes fuentes de ruido y estudiamos algunos modelos, junto a como cambian parámetros óptimos en la discriminación de estados.

Esto nos llevó a notar que en una comunicación real puede ser sumamente difícil determinar todas las interacciones del fondo con nuestros estados, surgiendo así la necesidad de buscar métodos de optimización libre de modelo para encontrar los parámetros óptimos, en particular de una medida Kennedy optimizada. Proponiendo así el uso de inteligencias artificiales, libres de modelos, que realizando experimentos sobre el sistema puedan encontrar los parámetros óptimos.

Una vez sentadas las bases teóricas necesarias para implementar inteligencias artificiales al problema de discriminación de estados, estudiamos la capacidad de un agente ciego al entorno para encontrar parámetros óptimos. Para poder ver que la capacidad del agente de encontrar parámetros óptimos, se compararon las probabilidades de éxito, en un sistema libre de fondo, con los resultados del modelo teórico utilizando un algoritmo Dual-Annealing. Viendo así que ambas probabilidades de éxito eran iguales.

De todos modos, el principal inconveniente de este agente es que al no tener observaciones, es incapaz de adaptarse a cambios en el fondo. Esto hace que se tenga que realizar un nuevo entrenamiento, cada vez que el sistema cambia. Teniendo en cuenta este inconveniente, la siguiente idea fue proponer un agente que pueda tomar observaciones del experimento y así adaptarse al mismo.

Así diseñamos una inteligencia capaz de adaptarse a cambios en la intensidad $|\alpha|$ usándola como observación. Este agente utilizó una política cúbica y realizaba una suerte de ajuste a los valores óptimos, haciendo uso únicamente de las recompensas que generaban los diferentes experimentos. Haciendo esto, las probabilidades de éxito difirieron siempre en menos de un 2 % al límite teórico en un canal libre de fondo, pero necesitó aproximadamente 5 veces más experimentos para converger.

A pesar de ser un poco más versátil, ninguna de las inteligencias empleadas hasta este punto eran capaces de adaptarse a cambios en el entorno. Por esta razón, el último agente propuesto consta de una red neuronal basada en el algoritmo PPO, con el objetivo de conseguir adaptaciones más rápidas a cualquier cambio en el fondo del sistema. Para esto, el entorno envía como observaciones a la estimación de la probabilidad de éxito, para un desplazamiento β_i, el desplazamiento β_i y el valor de $|\alpha|$. De esta forma, el agente proponía un desplazamiento y guardaba los 3 nuevos números en forma de observaciones, para tenerlos en consideración en futuras decisiones. Este esquema seguido por el agente de realizar experimentos y agregar observaciones fue repetido un total de 10 veces, luego de esto el agente propone un parámetro óptimo. Para determinar la

calidad de los desplazamientos propuestos por este agente en situaciones ruidosas, se compararon las probabilidades de éxito de los desplazamientos obtenidos por el agente y la de los desplazamientos teóricos para un canal libre de fondo.

Realizando esta comparación en un modelo de canal real, el algoritmo mostró una mejora del 1 % en media frente a ignorar el fondo. Aun cuando el resultado no es completamente satisfactorio, creemos que una mejor implementación de este algoritmo podría haber generado resultados ampliamente superiores.

En futuros trabajos se podrían implementar estas inteligencias artificiales en comunicaciones reales, para ver si se observan las mismas mejoras y la viabilidad de cada una en diferentes entornos.

7 Apéndice

-He estado corriendo durante 3 años, 2 meses, 14 días y 16 horas.-Para de correr
-Callaos, va a decir algo.
-Estoy muy cansado, ahora me voy a casa.

Forest Gump

En este apéndice haremos un poco de énfasis en los temas que no resultan centrales para el problema de discriminación estudiados con inteligencias artificiales, pero que son de interés para un entendimiento más completo del mismo. Para esto dividiremos el apéndice en 2 secciones. Comenzaremos la siguiente sección estudiando la realización experimental de un láser y terminando como cuantizar los campos electromagnéticos.

A. Realización Teórica-Experimental

Fundamentos del funcionamiento de un láser

Hoy en día existen varias formas de realizar experimentalmente estados coherentes según las aplicaciones necesarias. El modelo más sencillo de un láser es mediante la emisión de un fotón por el decaimiento de un átomo. De esta forma, los láseres funcionan excitando y des excitando un sistema de 4 niveles. La figura 1 muestra un diagrama ilustrativo de los niveles de un láser y sus partes.

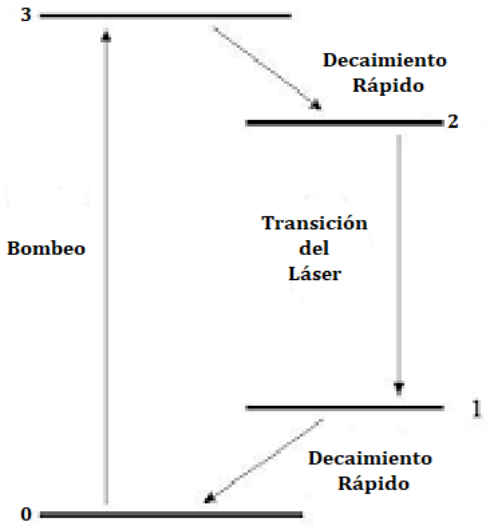

Figura 1: Esquema de los niveles de energía de un láser y sus partes.

Naturalmente, los electrones de niveles excitados tienden a decaer al nivel fundamental, generando en el proceso un fotón con energía igual a la diferencia. Pero si se induce un campo electromagnético, entonces este proceso puede generar otro fotón, además del producido por la transición de niveles, amplificando así el efecto de emisión. De esta forma, necesitamos comenzar con un elemento que tenga la mayoría de sus electrones en niveles excitados. Para poder realizar este efecto primero se tiene que ”invertir la población”,

haciendo que existan más átomos en el nivel excitado frente a la cantidad de átomos en el estado fundamental. En materiales semiconductores esto se logra inyectando una corriente. [41]

Como dijimos, los láseres suelen funcionar en sistemas de 4 niveles, esto se debe a que la transición entre niveles de energía suele ser muy lenta, de forma que si queremos vaciar rápido los niveles electrónicos, necesitamos niveles intermedios que puedan almacenar una cierta cantidad de electrones, para generar intercambios rápidos. Finalmente, para filtrar el pulso, se introduce a la señal del láser dentro de una cavidad resonante que mantiene únicamente a la frecuencia deseada. [41]

Clásicamente, esto genera por definición a un estado coherente, ya que obtendríamos una serie de fotones, todos en fase y con la misma frecuencia. Cuánticamente, se puede modelar el decaimiento de las diferentes frecuencias en una cavidad resonante. De esta forma se puede observar que los láseres, al menos a tiempos cortos, producen una matriz de densidad equivalente a la de los estados coherentes, según el modelo de Scully-Lamb.

De esta forma vemos que para producir un estado coherente, necesitamos de un material cuyos electrones no se encuentren en el nivel fundamental, un campo electromagnético externo y una cavidad resonante, pero para poder estudiar en más detalle la teoría de estos estados, es prudente determinar como se cuantiza al Hamiltoniano de las ondas electromagnéticas.

Cuantización del campo

Para obtener la versión cuantizada del campo electromagnético, se comienza trabajando en el régimen clásico. Si escribimos a los campos eléctricos y magnéticos en el Gauge de Coulomb ($\vec{\nabla} \cdot \vec{A} = 0$), entonces los campos resultan

$$\vec{E} = -\frac{\partial \vec{A}}{\partial t} \qquad \vec{B} = \vec{\nabla} \times \vec{A}\,, \tag{.1}$$

donde $\vec{A}$ queda definido mediante la ecuación diferencial

$$\nabla^2 \vec{A} - \frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} = 0\,. \tag{.2}$$

Se puede ver que una solución de esta ecuación diferencial son las ondas planas, por lo que podemos expandir a cualquier potencial vector $\vec{A}$, en la base

$$\vec{A} = \sum_k \left(\vec{\alpha}_k e^{i\vec{k}\vec{r}} + \vec{\alpha}_k^* e^{-i\vec{k}\vec{r}}\right)\,, \tag{.3}$$

con $\alpha_k \propto e^{-i\omega_k t}$, $\omega_k = |\vec{k}|$ y donde $\vec{\alpha}_k \cdot \vec{k} = 0$.

Para escribir el Hamiltoniano en cantidades canónicas conjugadas se realiza la transformación

$$\vec{Q}_k = \frac{1}{\sqrt{4\pi}}(\vec{\alpha}_k + \vec{\alpha}_k^*) = \frac{1}{\omega_k^2}\frac{\partial \vec{P}_k}{\partial t} \tag{.4}$$

$$\vec{P}_k = \frac{-i\omega_k}{\sqrt{4\pi}}(\vec{\alpha}_k + \vec{\alpha}_k^*) = \frac{\partial \vec{Q}_k}{\partial t}\,. \tag{.5}$$

Los operadores $\vec{Q}$ y $\vec{P}$ son llamados *cuadraturas de campo*. Usando esta transformación, el potencial vector puede ser escrito como

$$\vec{A} = \sqrt{4\pi}\sum_k \left(\vec{Q}_k \cos(\vec{k}\vec{r}) - \frac{1}{\omega_k}\vec{P}_k \sin(\vec{k}\vec{r})\right)\,. \tag{.6}$$

Así la densidad de energía y por ende al Hamiltoniano pueden ser escritas según las *cuadraturas de campo* como

$$\mathcal{H} = \frac{1}{2}\sum_k (\vec{P}_k^2 + \omega_k^2 \vec{Q}_k^2)\,. \tag{.7}$$

Una vez obtenido el Hamiltoniano clásico de la radiación, estamos en posición de cuantizarlo. En general, para cuantizar un campo tenemos que seguir 3 pasos

- asociar las cantidades canónicas con operadores.
- imponer la conmutación canónica entre los operadores.
- ”conmutadores Cero” corresponden a operadores de otros modos, ya que estos modos están desacoplados.

Para este punto, ya tenemos a los operadores canónicos $\hat{Q}$ y $\hat{P}$, con los cuales encontramos el Hamiltoniano clásico. Una vez identificados, imponemos las condiciones canónicas entre operadores y la conmutación nula para operadores que actúan en modos diferentes

$$\left[\hat{Q}_l, \hat{P}_l\right] = i\hbar\delta_{l,l'} \tag{.8}$$

$$\left[\hat{Q}_l, \hat{Q}_{l'}\right] = \left[\hat{P}_l, \hat{P}_{l'}\right] = 0\,. \tag{.9}$$

A estos operadores los podemos asociar a un modo normal de amplitud compleja α_l y α^*_{-l} asociados a operadores $\hat{a}_l$ y $\hat{a}_l^\dagger$

$$\hat{a}_l := \frac{1}{\sqrt{2\hbar}}\left(\hat{Q}_l + i\hat{P}_l\right)\,, \tag{.10}$$

$$\hat{a}_l^\dagger := \frac{1}{\sqrt{2\hbar}}\left(\hat{Q}_l - i\hat{P}_l\right)\,. \tag{.11}$$

De la misma forma resulta que la conmutación de estos operadores es

$$\left[\hat{a}_l, \hat{a}_{l'}^\dagger\right] = \delta_{l,l'} \qquad\qquad [\hat{a}_l, \hat{a}_{l'}] = 0\,, \tag{.12}$$

y siguiendo el Hamiltoniano clásico de la radiación, podemos cuantizar a este operador mediante

$$\hat{H}_R = \sum_l \hbar\omega_l|\alpha_l|^2 \to \hat{H}_R = \sum_l \frac{\omega_l}{2}\left(\hat{Q}_l^2 + \hat{P}_l^2\right)\,. \tag{.13}$$

Si queremos obtener la expresión en función de operadores de creación y destrucción $\hat{a}$ y $\hat{a}^\dagger$ obtenemos la expresión

$$\hat{H}_R = \sum_l \frac{\hbar\omega_l}{2}(\hat{a}_l\hat{a}_l^\dagger + \hat{a}_l^\dagger\hat{a}_l) = \sum_l \frac{\hbar\omega_l}{2}\left(\hat{a}_l^\dagger\hat{a}_l + \frac{1}{2}\right)\,. \tag{.14}$$

Esto nos da el Hamiltoniano de la radiación libre.

B. Discriminaciones óptimas

Durante el trabajo estudiamos varios métodos de discriminación de estados coherentes y calculamos un límite teórico de que tan grande puede ser la probabilidad de éxito. Pero ninguno de los métodos estudiados eran capaces de alcanzarlo, por esta razón en la próxima subsección del apéndice se realizará una introducción de lo que es conocido como *Detector Dolinar*, el cual haciendo uso de $N \to \infty$ modos es capaz de alcanzar el límite de Helstrom.

Detector Dolinar

El detector tipo Dolinar consiste en un esquema sofisticado de control vía *feedback*, el cual comentaremos a continuación. La señal es dividida mediante un *separador de haces* donde la parte trasmitida es desplazada y luego medida mediante un fotodetector. El resultado de dicha medida, condiciona al próximo desplazamiento, a efectuarse sobre la parte reflejada de la señal. Este procedimiento es repetido en forma asintótica, sobre la señal recibida (separando el haz de luz en infinitas partes), y una *decisión* sobre el bit inicial es efectuada de acuerdo al criterio del estado *más probable*. Un diagrama del circuito puede ser observado en la figura 2.

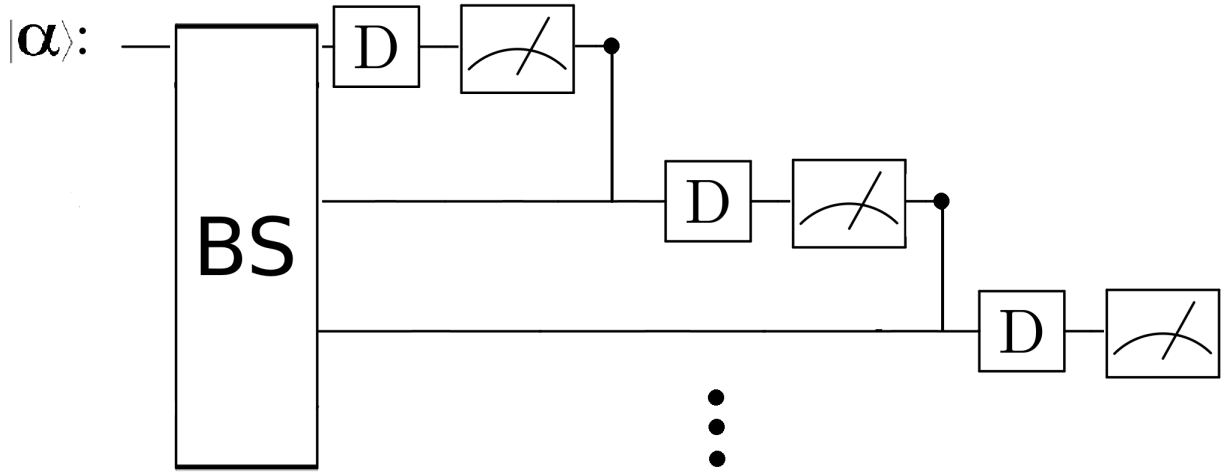

Figura 2: Circuito que describe la discriminación Dolinar.

Pero en estos métodos de discriminación óptimos siguen existiendo fuentes de ruido externas, que pueden afectar a la eficacia de los mismos. Este protocolo es sumamente complejo de maximizar, ya que en general se trata de un problema de complejidad exponencial. Por esta razón, protocolos de optimización como PPO podrían ser de gran ayuda al momento de buscar parámetros óptimos en canales con ruido.

PPO

Para generar las decisiones, el algoritmo usa una política π_θ, la cual en general suele ser una red neuronal. Luego utiliza un método Monte-Carlo, haciendo N iteraciones de un experimento con esta política, obteniendo así los *valores de la acción* Q_t con $t = \{0, 1, ..., N\}$.

Una vez calculados los valores de la acción, otra red neuronal que solo ve los estados S_t intenta predecir los valores Q_t, ya que en caso de que las predicciones sean *buenas*, podríamos determinar valores de expectación sin necesidad de hacer el experimento muchas veces. De esta forma el pseudocódigo del funcionamiento puede ser visto en el algoritmo 4.

Algoritmo 4 PPO. Estilo Actor-Critico

```
for i = 1, 2, ... do
    for k = 1,2,..., N do
        correr política π_θviejo en el entorno por T pasos.
        Computar la ventaja estimada Q_1, Q_2, ..., Q_T.
    end for
    Optimizar L con θ, cada K bloques de datos.
    θ_viejo = θ
end for
```

Además de esto, a diferencia de los algoritmos explicados durante el trabajo, que siguen la dirección del gradiente según el teorema de la política primaria, de la ecuación (4.7), el algoritmo PPO usa como dirección a seguir, a la que minimiza la función

$$L_t^{PPO}(\theta) = \hat{E}_t[L_t^{CLIP}(\theta) - c_1 L_t^{VF}(\theta) + c_2 S[\pi_\theta](s_t)] \,. \tag{.15}$$

Para entender el objetivo de seguir esta dirección (en vez de la dirección del gradiente), tendremos que estudiar por secciones que significa cada término $L_t(\theta)$. El primero expandido se escribe como

$$L_t^{CLIP}(\theta) = \hat{E}_t\Big[\text{mín}(r_t(\theta)\hat{Q}_\pi, CLIP(r_t(\theta), 1-\epsilon, 1+\epsilon))\Big] \,, \tag{.16}$$

donde $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}$ determina la probabilidad de que con los nuevos parámetros se vuelva a llegar a este estado y $\epsilon \approx 0{,}2$ es un hiperparámetro.

Para entenderla, nuevamente tenemos que ver esta función por partes, el primer término dentro del mínimo es $r_t(\theta)\hat{Q}_\pi$, la cual es una aproximación de la política primaria. Finalmente, el término $CLIP(r_t(\theta), 1-\epsilon, 1+\epsilon))$ limita qué tanto nos podemos mover del valor actual y nos permite volver al punto anterior del gradiente si es que vemos que la decisión tomada anteriormente fue mala.

Si volvemos a la ecuación (.15) y expandimos el segundo término tenemos

$$L_t^{VF}(\theta) = (V_\theta(s_t) - V_t^{\text{target}})^2 \,, \tag{.17}$$

siendo este el error cuadrático entre la recompensa vista y la recompensa predicha por la red neuronal. Finalmente, el término $S[\pi_\theta](s_t)$ se denota como el *bonus entrópico*, el cual permite que el agente vea una mayor cantidad de estados antes de converger.

Bibliografía

1. MaryamFarhadi, RahmahIsmail y MasoodFooladi. «Information and Communication Technology Use and Economic Growth», 2012.
2. C.E.Shannon. «A Mathematical Theory of Communication», 1948.
3. V. Vedral. «The role of relative entropy in quantum information theory», 2002.
4. J.J. Sakurai. «Modern Quantum Mechanics», 2011.
5. Carl W Helstrom. *Quantum Detection and Estimation Theory*. 1976.
6. Robert L. Cook et al. «Optical coherent state discrimination using a closed-loop quantum measurement». *Nature* 446:7137, 2007, págs. 774-777. ISSN: 14764687. DOI: 10.1038/nature05655. URL: https://link.aps.org/doi/10.1103/PhysRevA.79.022314%20http://www.nature.com/doifinder/10.1038/nature05655%20http://www.nature.com/articles/nature05655.
7. M. T. DiMario y F. E. Becerra. «Robust Measurement for the Discrimination of Binary Coherent States». *Phys. Rev. Lett.* 121:2, 2018, pág. 023603. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.121.023603. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.023603.
8. Stefano Pirandola. «Satellite quantum communications: Fundamental bounds and practical security». *Phys. Rev. Res.* 3:2, 2021, pág. 023130. ISSN: 2643-1564. DOI: 10.1103/PhysRevResearch.3.023130. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.3.023130.
9. M.Bilkis et al. «Real-time calibration of coherent-state receivers: Learning by trial and error», 2020.
10. D. Silver et al. «Mastering chess and shogi by self-play with a general reinforcement learning algorithm.», 2017.
11. Scott McLachlan et al. «The Self-Driving Car: Crossroads at the Bleeding Edge of Artificial Intelligence and Law», 2022.
12. Robert S. Kennedy. «Near-Optimum Receiver for the Binary Coherent State Quantum Channel». *MIT Res. Lab. Electron. Q. Prog. Rep.* 108, 1973, pág. 219. URL: https://dspace.mit.edu/handle/1721.1/56346.
13. Alessio Serafini. *Quantum Continuous Variables: A Primer of Theoretical Methods*. CRC Press, 2017.
14. Michael A. Nielsen e Isaac L. Chuang. «Quantum Computation and Quantum Information», 2010.
15. Jonatan Bohr Brask. «Gaussian states and operations – a quick reference», 2022.
16. Michael A. Nielsen e Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2010. ISBN: 9780511976667. DOI: 10.1017/CBO9780511976667. arXiv: arXiv:1011.1669v3. URL: http://ebooks.cambridge.org/ref/id/CBO9780511976667.
17. Mark M. Wilde. «From Classical to Quantum Shannon Theory», 2019.
18. S. J. Dolinar. «Communication and sciences engineering». *Q. Prog. Rep. (Research Lab. Electron.* 111, 1973, pág. 115. URL: https://dspace.mit.edu/handle/1721.1/56414.
19. Masao Osaki, Masashi Ban y Osamu Hirota. «Derivation and physical interpretation of the optimum detection operators for coherent-state signals». *Phys. Rev. A* 54:2, 1996, págs. 1691-1701. ISSN: 1050-2947. DOI: 10.1103/PhysRevA.54.1691. URL: http://link.aps.org/doi/10.1103/PhysRevA.54.1691.

20. Antonio Assalini, Nicola Dalla Pozza y Gianfranco Pierobon. «Revisiting the Dolinar receiver through multiple-copy state discrimination theory». *Phys. Rev. A* 84, 2 2011, pág. 022342. DOI: 10.1103/PhysRevA.84.022342. URL: https://link.aps.org/doi/10.1103/PhysRevA.84.022342.
21. Masahiro Takeoka et al. «Implementation of projective measurements with linear optics and continuous photon counting». *Phys. Rev. A - At. Mol. Opt. Phys.* 71:2, 2005, págs. 1-10. ISSN: 10502947. DOI: 10.1103/PhysRevA.71.022318. arXiv: 0410133 [quant-ph].
22. F E Becerra et al. «Experimental demonstration of a receiver beating the standard quantum limit for multiple nonorthogonal state discrimination». *Nat. Photonics* 7:2, 2013, págs. 147-152. DOI: 10.1038/nphoton.2012.316. URL: http://www.nature.com/doifinder/10.1038/nphoton.2012.316.
23. M. T. Dimario y F. E. Becerra. «Robust Measurement for the Discrimination of Binary Coherent States». *Phys. Rev. Lett.* 121:2, 2018, pág. 023603. ISSN: 10797114. DOI: 10.1103/PhysRevLett.121.023603. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.023603.
24. David E. Roberson et al. «Limit of Gaussian operations and measurements for Gaussian state discrimination and its application to state comparison». *Phys. Rev. A* 103, 2 2021, pág. 022423. DOI: 10.1103/PhysRevA.103.022423. URL: https://link.aps.org/doi/10.1103/PhysRevA.103.022423.
25. Masahiro Takeoka y Masahide Sasaki. «Discrimination of the binary coherent signal: Gaussian-operation limit and simple non-Gaussian near-optimal receivers». *Phys. Rev. A* 78:2, 2008, pág. 022320. ISSN: 1050-2947. DOI: 10.1103/PhysRevA.78.022320. URL: http://link.aps.org/doi/10.1103/PhysRevA.78.022320.
26. Krishna Kumar Sabapathy y Andreas Winter. *Bosonic data hiding: power of linear vs non-linear optics*. 2021. eprint: arXiv:2102.01622.
27. Samuel L. Braunstein y Peter van Loock. «Quantum information with continuous variables», 2004.
28. Mattia Walschaers. «Non-Gaussian Quantum States and Where to Find Them». *PRX Quantum* 2, 3 2021, pág. 030204. DOI: 10.1103/PRXQuantum.2.030204. URL: https://link.aps.org/doi/10.1103/PRXQuantum.2.030204.
29. JM Geremia. «Distinguishing between optical coherent states with imperfect detection». *Phys. Rev. A* 70:6, 2004, pág. 062303. ISSN: 1050-2947. DOI: 10.1103/PhysRevA.70.062303. arXiv: 0407205 [quant-ph]. URL: https://link.aps.org/doi/10.1103/PhysRevA.70.062303.
30. M. Bilkis et al. «Real-time calibration of coherent-state receivers: Learning by trial and error». *Phys. Rev. Res.* 2:3, 2020, pág. 033295. ISSN: 2643-1564. DOI: 10.1103/PhysRevResearch.2.033295. arXiv: 2001.10283. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.2.033295.
31. M. Bilkis, Matteo Rosati y John Calsamiglia. «Reinforcement-learning calibration of coherent-state receivers on variable-loss optical channels». En: *2021 IEEE Information Theory Workshop (ITW)*. 2021, págs. 1-6. DOI: 10.1109/ITW48936.2021.9611396.
32. Nicole Menkart et al. «Dark Current and single photon Detection by 1550 nm Avalanche Photodiodes: Dead Time Corrected Probability Distributions and Entropy rates», 2022.
33. H. P. Breuer y F. Petruccione. «The theory of open quantum systems.», 2002.
34. Laleh Memarzadeh y Stefano Mancini. «Minimumoutputentropyofanon-Gaussianquantumchannel», 2018.
35. Anna Dawid et al. *Modern applications of machine learning in quantum sciences*. 2022. eprint: arXiv:2204.04198.
36. Richard Sutton y Andrew G. Barto. *Reinforcement Learning Sutton*. MIT Press, 2018. ISBN: 9780262039246.
37. Aurélien Géron. «Hans-On Machine Learning with Scikit-Learn and TensorFlow», 2017.
38. John Schulman et al. «Proximal Policy Optimization Algorithms», 2017.

39. OpenAI. «Dota 2 with Large Scale Deep Reinforcement Learning», 2019.
40. Tomás Crosta. «Repositorio de código subido a GITHUB», 2023.
41. Benjamín Alonso Fernández, Rocío Borrego Varillas y Carlos Hernández García.et al. «El láser, la luz de nuestro tiempo.», 2010.