



UNIVERSIDAD DE BUENOS AIRES

Facultad de Ciencias Exactas y Naturales

Departamento de Matemática

Tesis de Licenciatura

Dinámicas Estocásticas de Redes Neuronales

Pedro Maximiliano Altamirano Carlavan

Director: Matthieu Jonckheere

Marzo de 2016

# Agradecimientos

Son las 4:17 de la mañana. Estoy terminando de escribir la tesis después de 12 horas de trabajo y después de varios meses de corridas, frustraciones, resultados que no salían, otros que sí, etc. Estoy cansado...pero feliz. Y por eso, *necesito* agradecer. Agradecer a Dios por todo lo que me ha dado, que es mucho, y lo que no también. Agradecer a mis Papás, a los que les debo TODO, y entre otras cosas la perseverancia necesaria para mantener un ritmo de trabajo como el que tuve que mantener. En particular recuerdo con cariño esos fines de semana antes de los trimestrales en que mamá se sentaba a tomarme la lección de las 10 materias que tenía que rendir esa semana, con paciencia infinita. A mis hermanas, Guada y Agus, que se bancaron mis locuras matemáticas, el volver de la facultad y ponerme a contarles que había visto algo “rre-copado”, los Espacios Vectoriales, y acto seguido ponerme a definirles lo que era. A mis amigos: del CBC, de la facultad, del colegio, de la vida. A todos los quiero mucho, pero creo que en esta ocasión merecen mención especial aquéllos que compartieron estos años juntos: Desde mi primer año en física, con Esteban, Luciano, Mauro, Dani A., Rodrigo C., Guille; hasta el descubrimiento de mi “verdadera naturaleza” de matemático (¡Gracias Bonder y Laura!) con Yami, Dani, Dani G., Hernán, Javi G., Santi D., Santi V., Luz, Mer, y la lista sigue (¡perdón!). En especial quiero mencionar, por su relevancia en esta etapa que termina, a Yami y Dani, que fueron un gran soporte en todo sentido y una gran ayuda, explicándome mil cosas hasta el infinito, siempre con una sonrisa. A mis profesores, que fueron todos buenos, pacientes, y de todos me llevé buena matemática, y siempre algo más. Y al final de este tramo apareció como caído del cielo mi querido colega Fernando Dobarro, que me acompañó en las luchas encarnizadas contra las proposiciones que no salían, ¡gracias Fer! La lista seguiría y seguiría, pero me limito a que se sientan aludidos todos los que crean que merecen su saludo, y cierro mis agradecimientos con Matt, que aceptó emprender este corto pero intenso tramo final conmigo guiándome; que fue infinitamente paciente para repetirme mil veces conceptos que no entendía; del que aprendí un montón y del que tengo mucho que aprender todavía.



# Índice general

<b>1. Introducción</b>	<b>1</b>
<b>2. Procesos de Markov</b>	<b>9</b>
2.1. Espacio de Estados contable . . . . .	11
2.1.1. El Generador: . . . . .	13
2.1.2. Construcción de Procesos de Markov ( $S$ numerable) . . . . .	14
2.1.3. Medidas invariantes . . . . .	14
2.2. El espacio de estados Continuo: <b>PDMP</b> . . . . .	15
2.2.1. Conceptos generales . . . . .	15
2.2.2. Procesos de Markov Determinísticos a Trozos . . . . .	17
<b>3. Modelos de Redes Neuronales</b>	<b>19</b>
3.1. Contexto biológico . . . . .	19
3.2. Los modelos <i>integrate-and-fire</i> . . . . .	21
3.3. El modelo estocástico . . . . .	24
<b>4. Redes Neuronales Finitas</b>	<b>27</b>
4.1. Comportamiento asintótico en el tiempo. . . . .	27
4.2. ¿Cuándo muere la Red? . . . . .	32
4.3. El Proceso independiente del Estado . . . . .	36
<b>5. Redes Neuronales Infinitas: Resultados Asintóticos</b>	<b>43</b>
5.1. <b>Existencia y unicidad de Proceso límite:</b> . . . . .	43
5.2. <b>Convergencia al Proceso de McKean-Vlasov:</b> . . . . .	44
5.3. Análisis de las distribuciones invariantes . . . . .	45
5.3.1. Cantidad de Distribuciones Estacionarias . . . . .	47

5.3.2. Estabilidad de la Solución Trivial . . . . .	50
<b>6. Conclusiones</b>	<b>55</b>
<b>7. Apéndice A</b>	<b>57</b>
7.1. Algunos resultados . . . . .	57
7.2. Histogramas de la Sección 4.2 . . . . .	59

# Capítulo 1

## Introducción

El cerebro humano es un órgano complejo. Muy complejo. Es probablemente la estructura biológica más compleja existente. En cerebros humanos se han gestado obras maestras como La Piedad o La Gioconda, invenciones tan espectaculares como la comunicación inalámbrica o el avión, construcciones imponentes como las Pirámides de Egipto o La Gran Muralla China; e incluso avances increíbles en el estudio de nuestro Universo, desde las teorías de la Mecánica Cuántica al mapeo del genoma humano. El cerebro es el centro de operaciones de básicamente todas las actividades de nuestro cuerpo, desde controlar el ritmo cardíaco hasta mantener el equilibrio o elaborar las teorías más complicadas.

La Neurociencia es rama de la ciencia que se especializa en el estudio del sistema nervioso del que el cerebro es protagonista; con el objetivo de comprender cada vez con mayor profundidad su funcionamiento, sus interacciones con el resto de los sistemas, etc. Este estudio se realiza de manera interdisciplinar. El cerebro es objeto de investigación de biólogos, neurólogos, químicos, físicos, matemáticos, filósofos, psiquiatras, entre otros; y dicha investigación se lleva a cabo en muchos niveles: desde el puramente químico, a nivel molecular, pasando por el nivel celular, hasta el nivel más conductual y de interacción con el entorno. Es dentro del nivel celular que podemos enmarcar el trabajo que nos ocupa en esta tesis: El estudio de las interacciones entre neuronas, las Redes Neuronales.

En las últimas décadas, la Neurociencia en general ha sido objeto de variadísimos y prolíficos avances, como el descubrimiento de las neuronas espejo, que podrían tener un papel importante en el aprendizaje por imitación; o la “plasticidad” del cerebro, la capacidad de generar nuevas conexiones o regenerarse en ciertas circunstancias.

El impacto que puede tener a mediano o largo plazo un mayor entendimiento del cerebro se extiende desde la educación, donde ya existen métodos que integran en sí los conocimientos aportados por la neurobiología; hasta la Medicina: La Organización Mundial de la Salud por ejemplo, ha dedicado muchos recursos para hacer frente a la presencia de los desórdenes mentales que afectan a cientos de millones de personas. [34]

El estudio de Redes Neuronales dentro del ámbito de las Matemáticas persigue, entre otros, el objetivo de describir la extrema complejidad de las interacciones entre las neuronas que conforman una red, o un cúmulo de redes. Para ello el desafío es plantear modelos que sean lo suficientemente complejos como para mantenerse fieles a la realidad que intentan describir,

pero de manera tal que sea posible manipular esos modelos para extraer información relevante a partir de ellos.

Dentro de este contexto abordamos nuestro trabajo, basándonos principalmente en el artículo “*On the dynamics of Random Neuronal Networks*” de Philippe Robert y Jonathan Touboul (trabajo en progreso) [43]. Presentaremos el modelo estocástico propuesto en este trabajo, y recorreremos sus resultados - agregando algunos - buscando dar una respuesta relevante a la tan importante cuestión de la *distribución de los tiempos de descarga* de una Red Neuronal. En ese sentido son de vital importancia los resultados relativos a la existencia de *medidas estacionarias*, pues éstas revelan cuál es el comportamiento hacia el que tiende un Proceso Estocástico en el tiempo; y se podrían interpretar desde el punto de vista biológico como el comportamiento “estable” de una Red Neuronal bajo ciertas circunstancias especificadas, es decir, lo esperable sería que en la experimentación se aprecien fluctuaciones *en torno* a este comportamiento estacionario. Como veremos, en el caso de una Red Neuronal Finita, en algunos casos este análisis nos llevará a concluir que la actividad de la red eventualmente “se extingue” (asociado a tener a la distribución trivial  $\delta_0$  como la única distribución invariante), o bien que existe una actividad auto-sostenida en el tiempo (cuando hay única distribución invariante y ésta es no trivial). En el primer caso, se podría quizás asociar a situaciones como la respuesta de una Red Neuronal en reposo, ante un estímulo específico; o la aislación de tejido cortical. Además, surgen naturalmente preguntas directamente relacionadas: ¿Cuánto tiempo tarda en llegar a ese estado de inactividad? ¿Cómo depende de la dimensión de la Red ese tiempo de extinción en esos casos?

El mismo ejercicio de asociación puede realizarse en el caso en que existe una actividad sostenida en el tiempo, expresada matemáticamente bajo la forma de una distribución estacionaria no trivial: ¿A qué fenómeno podemos asociar la actividad auto-sostenida descrita por el modelo? Quizás se podría asociar a la respuesta del cerebro ante un estímulo externo que se perpetúa en el tiempo, o con la actividad neuronal asociada a funciones biológicas que no se interrumpen en vida, como el latir del corazón.

Además, cuando la dimensión de la red neuronal tiende a infinito, veremos que con un escalamiento apropiado se obtiene un resultado de campo-medio (fundamental en muchas disciplinas para el tratamiento de sistemas infinitos de partículas similares entre sí), y en este caso, veremos que pueden aparecer incluso más de una distribución invariante no trivial; y esto incita a preguntarse por el fenómeno de **cuasi-estacionariedad**: en palabras simples, una situación de equilibrio local que finalmente desaparece. Esto se apoya en que hay situaciones en que sabemos que cierta red finita se extingue tarde o temprano, pero en el límite aparecen las mencionadas medidas no triviales. Desde el punto de vista interpretativo, este fenómeno es quizás el más interesante, porque a primera vista parecería el más factible de asociar a fenómenos reales como los que describimos párrafos más arriba.

De más está decir que, en el estado personal actual de experiencia, estas asociaciones se mueven en un ámbito más bien especulativo, pero con vistas a poder darle sustento teórico-experimental en el futuro próximo, mediante la profundización integral en el tema, y el aprovechamiento de la expertise de investigadores de relevancia, tanto en el ámbito nacional, como en el internacional. Sin embargo a lo largo de la tesis estas preguntas permanecen como telón de fondo y motivación trascendente a la matemática que se maneja, y guiarán la futura investigación en el tema.

El modelo en el que nos basamos para este trabajo consiste en un Proceso Estocástico De-

terminístico a Trozos (**PDMP** por sus siglas en inglés), cuyo estado representa el Potencial de las membranas de una red neuronal a tiempo  $t$ . El modelo combina las ideas fundamentales de los modelos de tipo *Leaky integrate-and-fire* (**LIF**) y supuestos del modelo *Linear non-linear Poisson Model*, explotando sus ventajas e intentando superar sus limitaciones (ver sección 3.2) para abordar una red neuronal *exitatoria*: es decir, una red en la que las interacciones - las *sinapsis*- producen un aumento de potencial en las neuronas que reciben dicha interacción. A esos fenómenos también los llamamos “*spikes*” o *descargas*. También corresponde mencionar que este modelo, a diferencia de otros ( ver por ejemplo [17]) se concentra en las sinapsis *químicas*, que son las descritas al principio del capítulo 3. Existen también sinapsis *eléctricas* o “brechas de salida” (“*gap junctions*” en inglés)<sup>1</sup>, pero son las primeras las que predominan, y por lo tanto las que se incluyen en el modelo.

En concreto la dinámica funciona así: Mientras no hay interacciones - sinapsis - entre las  $N$  neuronas, su potencial decae exponencialmente a su valor de reposo  $v_r$  (que tomamos como 0). Las interacciones ocurren cuando alguna de ellas emite un *spike*<sup>2</sup>, digamos la neurona  $i$  en el tiempo  $t$ . En ese instante, dicha neurona se resetea instantáneamente a su potencial de reposo, y las otras reciben un impulso exitatorio dado por las variables aleatorias ( $W_{ij}(t)$ ) que son no negativas, independientes e idénticamente distribuidas. Estos eventos se producen de acuerdo a  $N$  Procesos de Poisson no-homogéneos, cuya tasa de salto está dada por una función  $\lambda_i(t) = b(X_i(t))$  dependiente del estado de cada neurona respectivamente; con  $b : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  no decreciente. La ecuación, en su forma de Ecuación Diferencial Estocástica, es:

$$dX_i^N(t) = -X_i^N(t)dt + \sum_{j \neq i} W_{ji} \mathcal{N}_{b(X_j^N(t^-))}^j(dt) - X_i^N(t^-) \mathcal{N}_{b(X_i^N(t^-))}^i(dt) \quad (1.1)$$

Con  $d\mathcal{N}_z^i(t)$  un proceso de Poisson de tasa  $z$ , y ( $W_{ij}(t)$ ) variables aleatorias independientes e idénticamente distribuidas (i.i.d.), no-negativas. Es un proceso de Markov a valores en  $\mathbb{R}_+$ .

Siguiendo en gran medida el camino recorrido por Robert y Touboul en su artículo, en una primera parte comenzaremos por dar resultados para una red *finita* de neuronas. En concreto, mostraremos que si  $b(0) = 0$  (que se puede interpretar como ausencia de ruido externo) entonces la red eventualmente cesa su actividad, es decir, deja de emitir spikes; y la distribución trivial  $\delta_0$  (con masa 1 en  $0 \in \mathbb{R}_+$ ) es la única distribución invariante. Surge entonces la pregunta por el tiempo de extinción, y a modo de producción propia damos una respuesta explícita para un ejemplo bien sencillo de 2 neuronas; y para un caso con menos restricciones, de 2 neuronas también, damos una cota. Entre nuestros objetivos a futuro está el continuar esta línea de dar buenas cotas para los tiempos de extinción en casos más generales, y también está dispuesta en forma seminal la estrategia para analizar cómo depende de la dimensión el tiempo de extinción. En cambio si  $b(0) > 0$  (si hay ruido externo), se

<sup>1</sup>Las brechas de salida son uniones entre neuronas vecinas, que se encuentran a una distancia de 3-4 nm (mucho mayor que la distancia normal entre neuronas linkeadas en una sinapsis química). Estas sinapsis transmiten impulsos que siempre son menores o iguales al potencial de la neurona pre-sináptica, y que son cualitativamente más rápidos que las sinapsis químicas.

<sup>2</sup>También diremos que emite una descarga o que “se dispara”.



demostrará que el Proceso de Markov es Harris-recurrente y que tiene una única distribución no-trivial. En particular, esto indica que la actividad neuronal se perpetúa en el tiempo. Para un caso particular, en que la tasa de descargas es constante ( $(b(x) \equiv b > 0)$ ), damos una escritura explícita de un vector aleatorio con la ley de la distribución no-trivial del Proceso. Finalmente se muestra un primer resultado de campo-medio<sup>3</sup>: la independencia asintótica y la convergencia en distribución del proceso para la tasa constante, donde los  $W_{ij}$  se toman escalados como  $V_{ij}/N$ .

La segunda parte de la tesis aborda la cuestión de describir el sistema cuando el tamaño de la red se va a infinito. Para ello se utiliza la aproximación de campo-medio, ya clásica en la literatura para este contexto. Se lleva a cabo el mismo escalamiento de orden  $1/N$  que en la primer parte - en 4.7 - con lo que la ecuación del Proceso queda escrita como:

$$dX_i^N(t) = -X_i^N(t)dt + \frac{1}{N} \sum_{j \neq i} V_{ji} \mathcal{N}_{b(X_j^N(t^-))}^j(dt) - X_i^N(t^-) \mathcal{N}_{b(X_i^N(t^-))}^i(dt) \quad (1.2)$$

[43] y de esta manera en nuestro caso la aproximación de Campo-medio consiste en:

$$\frac{1}{N} \sum_{j \neq i} V_{ji} \mathcal{N}_{b(X_j^N(t^-))}^j(dt) \approx \mathbb{E}(V) \mathbb{E}(X_i^N(t))dt \quad (1.3)$$

con  $V \sim V_{ij}$ . Imponiendo algunas condiciones sobre  $b(x)$ , Robert y Touboul llegan a un resultado de Campo-Medio: la secuencia de procesos converge en distribución a un proceso de McKean-Vlasov. Nosotros simplemente enunciamos y damos un esquema sintético de la demostración.

Finalmente seguimos el análisis que se propone en [43] de las distribuciones invariantes del proceso límite, con resultados sobre la existencia y cantidad de soluciones invariantes, y su estabilidad, según la relación entre  $b(x)$  y  $\mathbb{E}(V)$ . Al respecto es importante observar que en el caso de un Proceso de Markov **homogéneo** la pregunta por la cantidad de soluciones no tiene demasiado sentido, puesto que sólo hay 2 opciones: o tiene única distribución invariante, o no tiene distribuciones invariantes.

Los resultados que mostramos en esta sección son los siguientes: si  $b(0) = 0$ ,  $\delta_0$  sigue siendo siempre una posible distribución invariante (y se da un resultado sobre la estabilidad de la solución); pero existen casos en que pueden aparecer más distribuciones, no-triviales. Teniendo en cuenta que en el capítulo 4, demostramos que eventualmente la red *muere*, la existencia de distribuciones estacionarias no-triviales sugiere una *cuasi-estacionariedad*, que será objeto de nuestro estudio en futuros trabajos. La cuasi-estacionariedad ocurre en un Proceso de Markov en el cual hay un estado absorbente, por lo que no tiene sentido el estudio de la medida estacionaria, pues ésta consiste en el proceso anclado en el estado absorbente. En cambio toma relevancia estudiar la distribución del proceso condicionado a no haber sido absorbido, y se habla de una cuasi-estacionariedad cuando existe “estabilidad” bajo esa distribución condicionada.

---

<sup>3</sup>ver 1

**Teoría de Campo-Medio** La teoría de Campo-Medio merece una mención aparte. También conocida como Teoría de Campo Auto-Consistente, consta del estudio del comportamiento de modelos complejos de gran tamaño -en general estocásticos - formados por una gran cantidad de pequeños componentes similares que interactúan entre sí. La estrategia para abarcar su estudio es aproximar el efecto que producen todos los demás individuos sobre uno de ellos por un sólo efecto ponderado: el Campo Medio, que se aproxima arbitrariamente al modelo original a medida que la dimensión del sistema se vuelve infinita.

Originalmente, el concepto surge de la física, a partir de los trabajos de Weiss y Pierre Curie sobre la transición de fase [50] [25]; esto es, cambios cualitativos bruscos que experimentan ciertas cantidades termodinámicas al variar alguna otra cantidad (como la temperatura) en un sistema de una gran cantidad de componentes estadísticamente idénticas.

Esta estrategia probó su efectividad para describir estos sistemas [23], y su idea fundamental se ha importado desde entonces a otros sistemas no necesariamente termodinámicos, en campos como modelos epidémicos [27], teoría de colas [5] y modelos de redes estocásticas en general [7], entre las que se encuentran las Redes Neuronales Estocásticas.

En este sentido hay una amplia literatura dedicada a atacar el problema de describir la dinámica de una red neuronal haciendo uso de la aproximación de campo-medio; por ejemplo los trabajos de Cáceres [12] [11], o de Pakdaman et al. [36] [37] [38]. En este contexto particular de Redes Neuronales, estos trabajos - y en general en la literatura sobre campo-medio sobre este tema - suelen aproximar el componente estocástico generado por la interacción entre las neuronas por un efecto dado por un Movimiento Browniano, de tal manera que la dinámica toma la forma de un proceso de Ornstein-Uhlenbeck.

Así es como, en una dinámica del tipo *Leaky integrate-and-fire* para  $N$  neuronas en una red excitatoria que está dada por una ecuación de la forma:

$$\frac{dv}{dt} = -(v - v_r) + W_N \sum_{i=1}^N \sum_j \delta(t - t_j^i) \quad (1.4)$$

(con  $\delta$  la Delta de Dirac en 0,  $W_N$  la *eficacia sináptica*, y para cada  $i$  ( $t_j^i$ ) representa un Proceso de Poisson de tasa  $\lambda$ ,  $t_j^i$  representando el  $j$ -ésimo spike de la  $i$ -ésima neurona), se aproxima el término de interacciones (que llamamos  $I(t)$ ) así:

$$I(t) dt \approx \mu dt + \sigma dB(t)$$

donde  $B(t)$  es el Movimiento Browniano Estándar, y  $\mu$  juega el rol del valor esperado del potencial de interacción, en este caso  $\mu = NW_N \lambda$ . Ver [12].

Esta estrategia de concentrar la aleatoriedad en un Proceso Gaussiano, además de ser una aproximación aceptable desde el punto de vista empírico, se apoya en que tiene un resultado de campo-medio cuando se escala de manera correcta la eficacia sináptica; y los resultados estándar del cálculo estocástico (es decir, el cálculo de Itô) son aplicables de manera casi directa.

El enfoque que adoptan Robert y Touboul en el que nosotros nos basamos es esencialmente distinto, ya que describimos los tiempos de disparo a partir de Procesos de Poisson no-homogéneos. Es un enfoque que parece no estar tan presente en la literatura como el que

describimos recién; y esto, según argumentan Robert y Touboul [43][Sección 1.2], es debido probablemente a que el cálculo estocástico asociado resulta más extraño de manejar.

Sin embargo, hay otros trabajos recientes que abordan el “problema” de campo-medio desde esta perspectiva: recientemente De Masi et al. [17], por ejemplo, prueban un resultado de esta índole para un modelo Markoviano determinístico a trozos, con la tasa de descargas dependiendo del potencial de las neuronas, aunque en lugar de considerar un término de pérdida (un “*leakage*”), tienen en cuenta las sinapsis eléctricas, y la eficacia sináptica es considerada un valor fijo escalado correspondientemente (en concreto,  $W_N = 1/N$ ). Con este modelo, y mediante un argumento de acoplamiento con un proceso - discreto en el tiempo - que lo aproxima, prueban un resultado de campo-medio cuando las condiciones iniciales ( $X^N(0)$ ) tienen soporte compacto.

En esta misma línea, Fournier y Löcherbach toman este mismo modelo y avanzan en los resultados, relajando las hipótesis sobre la condición inicial y demostrando que el proceso tiene una tasa de convergencia de orden al menos  $1/\sqrt{N}$ . Un problema que ambos trabajos enfrentan es demostrar la *propagación del caos* (es decir, la independencia asintótica de los elementos del sistema) cuando la función  $b$  que determina la tasa de descarga de las neuronas en función de su potencial no es globalmente Lipschitz. En un estudio más detenido de las técnicas utilizadas en estos trabajos esperamos encontrar una fuente de inspiración para abordar estas técnicas de campo-medio para el modelo que tratamos en esta tesis, para poder avanzar con respecto a los resultados ya obtenidos; de manera análoga a como se logró para el modelo propuesto por De Masi et al.

Para finalizar los comentarios la teoría de campo-medio, es necesario mencionar que existe otro paradigma de modelos, los “*conductance-based models*”, de entre los cuales se destaca el clásico Modelo de Hodgkin-Huxley; y que abordan el problema de describir el comportamiento de una Red Neuronal desde un punto de vista más bio-fisiológico, modelando con cierta precisión los intercambios químicos que se observan en las sinapsis químicas reales. Ha habido trabajos recientes sobre resultados de campo-medio para modelos de esta clase, como los de Touboul [49], [48] o Baladron, Touboul et al. [6]. En general estos modelos llevan a ecuaciones extremadamente complejas, que sin embargo pueden simplificarse un poco en algunos casos, permitiendo resultados relevantes [47].

**Estructura de la Tesis** De esta manera, la estructura de la tesis está dada así:

- En el Capítulo 2, desarrollamos la teoría matemática básica para entender el modelo que presentamos. En concreto, definiremos los *Procesos de Markov* y enunciaremos algunas de sus propiedades. En especial, nos interesa presentar los *Procesos de Markov Determinísticos a trozos (PDMP)*, ya que el modelo que queremos estudiar es simple describir y entender en ese marco teórico.
- En el Capítulo 3, hablaremos sobre el contexto biológico, aportando algunas nociones básicas de dan sustento al modelo; luego presentaremos un recuento histórico de modelos clásicos de Redes Neuronales; y finalmente motivaremos el modelo de Robert y Touboul que vamos a estudiar, que presentaremos a través de una Ecuación Diferencial Estocástica, y también bajo la forma de un **PDMP**.

- En el Capítulo 4, comenzaremos el cuerpo principal de la tesis, enunciando y demostrando propiedades y resultados referidos a una Red *Finita* de Neuronas. Mostraremos que existe un cubo en  $\mathbb{R}_+^N$  - dependiente de  $N$  y de la eficacia sináptica  $W$  - tal que para un sistema con cualquier condición inicial  $x \in \mathbb{R}_+^N$ , se tiene esperanza finita (acotada por la norma 1 de  $x$ ) de entrar en ese conjunto de estados. Esto da cierto control (aunque no muy “fuerte”, como veremos) sobre el potencial del sistema, suficientemente útil para dar un resultado sobre la existencia y unicidad de medidas estacionarias para el proceso, en función del comportamiento de la tasa de descargas de la red neuronal: En el caso donde no hay ruido externo, veremos que con probabilidad 1 la red “muere”, en particular la única distribución invariante es la trivial; mientras que si existe ruido externo, se tiene actividad sostenida, reflejada en la existencia de una medida invariante no trivial. Para esta última demostración, ampliamos y corregimos lo realizado en [43]. Luego, a modo de producción propia, realizamos el cálculo explícito del tiempo de extinción para un modelo simplificado de 2 neuronas y condición inicial particular, y damos una cota para un caso un poco más general, también de 2 neuronas. Algunas simulaciones en Octave permiten verificar estos resultados, y se comentan otros algoritmos para próximas estimaciones. En este sentido, dentro de las perspectivas a futuro, una cuestión interesante es la de poder dar cotas eficientes para el tiempo de extinción en el caso de  $N$  neuronas, y un posible camino es el de construir, a partir de  $(X(t))$ , un **PDMP**  $(\hat{X}(t))$  eligiendo los tiempos de salto **y** la medida de transición inteligentemente tal que  $\tau_{ext} \leq \hat{\tau}_{ext}$ .

En otra sección, analizamos el caso en que se tiene tasa de descargas constante ( $b(x) = b$ ), caso en el que damos una formulación explícita para la distribución invariante, y su transformada de Laplace. Bajo la misma hipótesis, damos un primer resultado de tipo campo-medio, en el que ampliamos la demostración esbozada en [43].

- En el Capítulo 5, analizamos el comportamiento de la red cuando el número de neuronas tiende a infinito. En el artículo de Robert y Touboul se demuestra existencia y unicidad de un Proceso Estocástico de McKean-Vlasov al que la red, según también se demuestra, converge cuando  $N$  tiende a infinito. Nosotros nos limitaremos a enunciar esos resultados y dar un esquema sintético de las ideas de esa demostración. Demostramos algunos resultados en relación a las distribuciones invariantes de este proceso límite, en específico damos una fórmula explícita para su función de densidad en los casos en que existe; bajo una formulación que permite obtener los resultados que damos luego sobre la cantidad de soluciones no triviales posibles, en función de la tasa de descargas y la eficacia sináptica. Veremos también que estos parámetros determinan la “estabilidad” de la solución no trivial; y finalmente especificamos condiciones para la cantidad de posibles soluciones no estacionarias para el caso particular en que  $b(x)$  es lineal. Estos resultados de estacionariedad - que incluyen la existencia de soluciones estacionarias no triviales - abren una discusión sobre un fenómeno de cuasi-estacionariedad, al compararlo con los resultados del capítulo anterior, que permiten vislumbrar perspectivas muy interesantes de investigación futura.



## Capítulo 2

# Procesos de Markov

Comenzaremos con un capítulo que da un marco teórico a los desarrollos que realizaremos durante la tesis. Principalmente, vamos a recorrer de manera sintética las definiciones y propiedades fundamentales en relación a los Procesos de Markov, ya que el modelo que presentamos entra dentro de esta clase de objetos matemáticos.

Comencemos por dar una noción de Proceso Estocástico. Un *Proceso Estocástico* es esencialmente una colección de variables aleatorias  $(X_t)_{t \in T}$  definidas en el mismo espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ , a valores en un espacio medible  $(S, \mathcal{S})$ , denominado espacio de estados. Este objeto suele representar el comportamiento de cierta magnitud  $X$  a lo largo del tiempo, en el marco del estudio de algún fenómeno complejo que requiere la introducción de aleatoriedad en el modelo. De esta manera,  $X_t$  sería el estado o valor de  $X$  en el momento  $t \in T$  ( $T$  sería el lapso de tiempo a considerar; usualmente  $T$  es  $\mathbb{N}_0$  o  $\mathbb{R}_+$ ). Dicho proceso puede contar adicionalmente con cierta estructura adicional, distinta según el caso, que típicamente da información sobre la *evolución* del mismo.

Por ejemplo, un proceso simpático e ilustrativo en varios niveles, es el del mono tipeando el Quijote: Supongamos que colocamos a un mono a tipear en una máquina de escribir, 1 letra por unidad de tiempo, de manera independiente y aleatoria. De esta manera podríamos modelar el problema como una secuencia de variables aleatorias *i.i.d.* (independientes e idénticamente distribuidas)  $(X_n)_{n \in \mathbb{N}}$ , donde  $X_n$  representa el carácter que el mono ha tipeado en el  $n$ -ésimo paso; es decir:

$$X_n(\omega) \in S = \{\text{caracteres de una máquina de escribir}\} \quad \forall n \in \mathbb{N}, \omega \in \Omega$$

y tal que, para cada carácter  $x \in S$

$$\mathbb{P}(X_n = x) = p_x > 0 \quad (\text{asumimos})$$

Esta construcción, - aunque sea de juguete - resulta ser un proceso estocástico. La razón por la que vemos una colección de variables aleatorias como un todo, se apoya en que queremos responder preguntas sobre la evolución del fenómeno. En este ejemplo, podríamos preguntarnos cuál es la probabilidad que el mono tipeador, bajo las hipótesis planteadas, logre *eventualmente* tipear todo El Quijote. Y la respuesta (que se puede obtener rigurosamente con pocas herramientas de Probabilidad, como el Lema de Borel-Cantelli) es que la probabilidad de que

el mono escriba, no una, sino infinitas (!) veces la obra completa de Cervantes es 1. Es decir, bajo estas hipótesis, el mono *eventualmente* terminará escribiendo El Quijote tantas veces como quiera (si tengo suficiente paciencia, y suponemos el mono poco menos que inmortal e incansable).

De entre los numerosos Procesos Estocásticos que son objeto de Estudio en Teoría de Probabilidades, el *Proceso de Markov* tiene un lugar realmente destacado: Es por un lado el proceso más sencillo y elegante en sus resultados y propiedades, lo cual permite dar mucha y muy detallada información acerca de su comportamiento, como por ejemplo el cálculo explícito de muchas de sus cantidades, e incluso existen abundantes métodos computacionales para simularlo. Y mientras este hecho podría llevar a pensar que dicha sencillez se logra en detrimento de su aplicabilidad, la realidad es que el Proceso de Markov resulta de una utilidad amplísima en infinidad de aplicaciones, tanto computacionales (Cadena de Markov Monte-Carlo) como en variadísimos ámbitos del estudio de fenómenos: Crecimiento Poblacional, Genética (como el modelo de Wright-Fisher), Teoría de Potenciales (que incluye teoría sobre Gravedad, Electroestática, difusión del Calor, etc), entre otras. Incluso su aplicación se extiende a otros procesos y objetos de estudio de Teoría de Probabilidades, como las Martingalas, los Paseos al Azar o el Movimiento Browniano; Teoría de Colas y de Redes, entre las principales. En relación a esta última afirmación, puede aseverarse que, en algún sentido u otro, prácticamente todo otro proceso estocástico en alguna manera representa una generalización del proceso de Markov. En palabras de James Norris en su libro “Markov Chains” [32], los procesos de Markov son la *metáfora que guía* el desarrollo de los demás procesos.

## El Proceso de Markov

Para esta sección, nuestras referencias principales han sido los siguientes libros: *Markov Chains*, de J.R. Norris [32]; *Continuous Markov Chains: An applications-oriented approach*, de W.J. Anderson [2]; *Continuous Time Markov Processes*, de T.M. Liggett [30] y *Applied Probability and Queues*, de S. Asmussen [3].

La idea fundamental detrás de un Proceso de Markov es que *no tiene memoria de lo que ocurrió en el pasado*. En consecuencia, **lo que ocurre en el futuro sólo depende del estado actual del proceso**.

Al Proceso de Markov a tiempo continuo (es decir  $T$  no numerable, esencialmente  $\mathbb{R}_+$ ) lo llamaremos simplemente *Proceso de Markov* mientras que en el caso en que  $T$  es contable - finito o numerable (principalmente  $\mathbb{N}_0$ ) - el Proceso será denominado *Cadena de Markov*; sin hacer más aclaraciones cuando el espacio de estados  $(S, \mathcal{S})$  se conoce. Dicho espacio  $S$  también puede ser, o no, contable.

En parte de la literatura se distingue la Cadena de Markov de un Proceso de Markov según la cardinalidad de  $S$ , más que de  $T$ , y en ese caso se suele aclarar que el proceso (o cadena) es a tiempo discreto o continuo. Nosotros sin embargo no seguiremos esta última convención.

A lo largo de nuestro trabajo nos enfrentaremos con un Proceso de Markov donde el espacio de estados es  $\mathbb{R}_+$ , es decir, no numerable. Como es de esperar, los Procesos de Markov con un espacio de estados que no es discreto presentan una dificultad técnica mucho mayor que cuando éste sí lo es. Sin embargo, comenzaremos por desarrollar los Procesos de Markov

de este último género, ya que las nociones y resultados tienen usualmente un análogo natural (en mayor o menor medida) en el caso más general.

Antes de presentar la definición de un Proceso de Markov en  $S$ , con  $S$  contable, vamos a precisar algunas aclaraciones y definiciones básicas.

En primer lugar, la continuidad del tiempo genera ciertas dificultades técnicas a la hora de intentar calcular probabilidades sobre el comportamiento del proceso. En concreto, la definición de medida (de probabilidad) cuenta con la propiedad de la  $\sigma$ -aditividad sobre uniones disjuntas, es decir  $\mathbb{P}(\bigsqcup A_n) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$ . Pero esta propiedad no se lleva bien con uniones no numerables de conjuntos, que sería el caso si quisieramos calcular la probabilidad de un conjunto tan inocente como  $\{\exists t \geq 0 : X_t = s\}$  para algún  $s \in S$ . Por lo tanto se hace necesario agregar alguna estructura que permita calcular probabilidades como ésta. Por eso consideraremos que el proceso es *càdlàg* (del francés, continuo a derecha y con límite a izquierda).

De esta manera se puede sortear la dificultad antes mencionada porque se pueden reescribir los “conjuntos problemáticos” con uniones, intersecciones, etc. **numerables** de conjuntos. Vale la pena mencionar esto, porque encontrarse con este tipo de obstáculos es usual en Matemática, a medida que uno va complejizando la teoría para hacerla más general; y es interesante -en mi opinión- el desafío de hacer suposiciones lo suficientemente *fuertes* como para sortear los obstáculos, y a la vez lo suficientemente *débiles* como para no perder demasiada generalidad.

El *inicio del Proceso*,  $X_0$  en general se coloca de manera determinística en algún estado  $x \in S$  cualquiera, o bien de manera aleatoria bajo una distribución  $\mu$ . En el primer caso, se suele notar  $\mathbb{P}_x$  y  $\mathbb{E}_x$  a la probabilidad y la esperanza inducidas por condicionar por el conjunto  $\{X_0 = x\}$ .

Para poder tener una comprensión profunda de la evolución de un proceso, se hace necesario introducir la noción de *filtración* de un espacio medible.

**Definición 2.1.** Sea  $(\Omega, \mathcal{F})$  un espacio medible. Una colección de  $\sigma$ -álgebras  $(\mathcal{F}_t)_{t \geq 0}$  contenidas en  $\mathcal{F}$  se dice *filtración* si  $\mathcal{F}_s \subseteq \mathcal{F}_t \forall 0 \leq s \leq t$ .

Dada una filtración, podemos interpretar a  $\mathcal{F}_t$  como la acumulación de toda la información disponible hasta el tiempo  $t$ . En ese sentido resulta natural la siguiente definición

**Definición 2.2.** Dado un proceso estocástico  $X$  definido en  $(\Omega, \mathcal{F})$ , decimos que  $X$  es *adaptado a la filtración*  $(\mathcal{F}_t)_{t \geq 0}$  si para cada  $t \geq 0$ ,  $X_t$  es  $\mathcal{F}_t$ -medible.

Un proceso estocástico  $X$  induce naturalmente una filtración en  $(\Omega, \mathcal{F})$ , dada por  $\mathcal{F}_t^X := \sigma(X_s, s \leq t)$  la menor  $\sigma$ -álgebra tal que  $(X_s)_{s \leq t}$  resultan medibles. Es de inmediata observación que  $X$  es adaptado a esta filtración, que llamamos la *filtración natural*.

## 2.1. Espacio de Estados contable

A lo largo de esta sección,  $S$  denotará un espacio de estados contable.



Observemos ante todo que la “distretitud” de  $S$  tiene como consecuencia que los cambios de estado del Proceso se dan en forma de *saltos*. Más aún, un proceso  $X$  càdlàg en este contexto se traduce en que  $X$  permanece por algún tiempo en cada estado que visita, antes de saltar al siguiente. Es decir, para todo  $\omega \in \Omega$  y  $t \geq 0$  existe algún  $\varepsilon > 0$  tal que  $X_s(\omega) = X_t(\omega) \forall s \in [t, t + \varepsilon)$ .

A continuación, se representa una trayectoria típica de un Proceso càdlàg a un espacio de estados numerable, tomada del Norris:

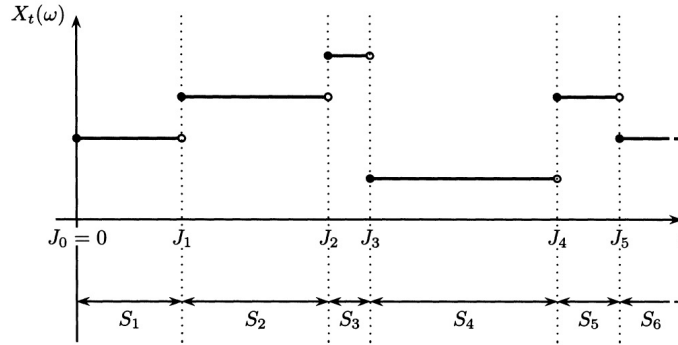


Figura 2.1: Trayectoria de un proceso càdlàg a  $S$  numerable. Los  $(S_n)$  son los *tiempos de espera*, y los  $(J_n)$ , los *tiempos de salto*

**Definición 2.3.** Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  y sea  $X = (X_t)_{t \geq 0}$  un proceso càdlàg a valores en  $S$ . Entonces  $X$  se dice un *Proceso de Markov (homogéneo)* si

$$\begin{aligned} \mathbb{P}(X_{t_{n+1}} = x_{t_{n+1}} | X_{t_0} = x_{t_0}, \dots, X_{t_n} = x_{t_n}) &= \mathbb{P}(X_{t_{n+1}} = x_{t_{n+1}} | X_{t_n} = x_{t_n}) \\ &= \mathbb{P}(X_{t_{n+1}-t_n} = x_{t_{n+1}} | X_0 = x_{t_n}) \end{aligned} \quad (2.1)$$

Donde la homogeneidad se refiere a la última igualdad.

Esta propiedad es la **Propiedad de Markov (Mkv)**. Llamemos  $P_t(x, A) := \mathbb{P}_x(X_t \in A)$  y sean, para  $x, y \in S$ , las *funciones de transición*  $p_t(x, y) := P_t(x, \{y\})$ .

Observar que se puede ver a  $P_t := (p_t(x, y))_{x, y \in S}$  para cada  $t \geq 0$ , como una matriz de  $S \times S$ . Fácilmente se obtiene:

**Propiedad 2.4** (Propiedades de  $P_t$ ). Con la notación anterior,

1.  $P_0 = I$
2.  $p_t(x, y) \geq 0 \quad \forall t, x, y$
3. Fijado  $t$ , para cada  $x \in S$  se tiene que  $(p_t(x, y))_y$  es una distribución en  $S$ . En otras palabras,  $p_t$  es **estocástica** para todo  $t \geq 0$ .
4.  $(p_t)_{t \geq 0}$  verifica la *Propiedad de Semigrupo*, también conocida como las **Ecuaciones de Chapman-Kolmogorov**. Esto es

$$p_{s+t}(x, y) = \sum_{z \in S} p_s(x, z) p_t(z, y) \quad (2.2)$$

**Demostración:** Las primeras 3 afirmaciones son inmediatas. La demostración de la última propiedad consiste en una sencilla aplicación de la Propiedad de Markov *para el caso homogéneo*. En efecto:

$$\begin{aligned} p_{s+t}(x, y) &= \mathbb{P}_x(X_{s+t} = y) = \sum_{z \in S} \mathbb{P}_x(X_{s+t} = y | X_s = z) \mathbb{P}_x(X_s = z) \\ &= \sum_{z \in S} \mathbb{P}_z(X_t = y) \mathbb{P}_x(X_s = z) \end{aligned}$$

donde la última igualdad vale por **(Mkv)**.

Es posible verificar que  $(p_t)_t$  cumple la condición:

$$\lim_{t \rightarrow 0} p_t(x, x) = p_0(x, x)$$

En efecto, llamemos  $J_1 := \inf\{t > 0 : X_t \neq X_0\}$ , entonces, por la continuidad a derecha del Proceso, se tiene que  $J_1 > 0$   $\mathbb{P}_x - a.e.$  para todo  $x \in S$ . Entonces

$$p_t(x, x) \geq \mathbb{P}_x(J_1 > t) \rightarrow 1 \quad (t \rightarrow 0)$$

Cualquier colección de funciones  $(\hat{p}_t(x, y))$  que cumplan con estas propiedades se denomina *semigrupo de probabilidad*, y se puede demostrar que a partir de éstas se puede construir un Proceso de Markov  $(X_t)$  cuyas funciones de transición coinciden con el semigrupo [2][Sección 1.1].

**Observación 2.5.** A partir de la propiedad de Markov, y por un argumento inductivo, es inmediato ver que el proceso  $X$  verifica para todo  $x \in S$  y  $0 = t_0 < t_1 < \dots < t_n$ :

$$\mathbb{P}_x(X_{t_1} = x_{t_1}, \dots, X_{t_n} = x_{t_n}) = p_{t_1 - t_0}(x_{t_0}, x_{t_1}) \dots p_{t_n - t_{n-1}}(x_{t_{n-1}}, x_{t_n})$$

### 2.1.1. El Generador:

Dado un espacio contable  $S$ , un *Generador (de probabilidades)* o *Q-matriz* es una matriz  $Q = (q_{xy})_{x, y \in S}$  tal que

- a)  $q_{xy} \geq 0$  si  $x \neq y$
- b)  $q_x := \sum_{y \neq x} q_{xy} = -q_{xx} < +\infty$

La importancia del generador de probabilidades asociada a un proceso de Markov reside en que da cuenta de la *transición infinitesimal* del Proceso. Es el análogo natural de la matriz  $P$  de transición de 1-paso en el caso de una cadena de Markov (tiempo discreto). De esta manera,  $q_{xy}$  quiere representar *la tasa de salto de  $x$  a  $y$* ; y  $q_x$ , *la tasa de salto desde  $x$* ; que matemáticamente queda expresado por la condición:

$$\left. \frac{dP_t}{dt} \right|_{t=0} = Q$$

o más generalmente,

$$P'_t = QP_t \quad P_0 = I \tag{2.3}$$

Esta ecuación se conoce como la *Ecuación Reversa*. El teorema 2.8.3 del libro de Norris [32] asegura la existencia de una única matriz estocástica  $P_t$  minimal que cumple (2.3) y que cumple la propiedad de semigrupo. En los Procesos más generales, suele presentarse un proceso de Markov por su Generador de Probabilidades (en su versión generalizada), más que por las probabilidades de transición (que también tiene su generalización, los *kernels*, ver Kallenberg [24]).

### 2.1.2. Construcción de Procesos de Markov ( $S$ numerable)

Existen varias formas de construir un Proceso de markov a un espacio de estados numerable. Nosotros elegimos la construcción *por cadena de saltos y tiempos de espera*, de una manera similar a la realizada en Norris [32], con adaptaciones, de manera que se ajuste más claramente como un caso particular en la teoría de *Piecewise Deterministic Markov Processes*, que desarrollamos más abajo.

Esta forma de construir un proceso de Markov para una cantidad numerable de estados se destaca por su sencillez y la claridad para visualizar trayectorias<sup>1</sup>.

Sea  $Q$  un generador de probabilidades, y  $P_t$  la solución de (2.3). Vamos a construir un proceso de Markov  $(X_t)$  con condición inicial  $X_0 = x$  de la siguiente manera: Consideramos  $(\hat{S}_n)_\mathbb{N}$  variables i.i.d. exponenciales de tasa 1. Recordemos que si  $Z \sim \varepsilon(1)$  y  $c > 0$  entonces  $Z/c \sim \varepsilon(c)$ . La estrategia es ir construyendo tramo a tramo la trayectoria de  $(X_t)$ . Tomamos entonces  $J_0 = 0$ ;  $S_1 := \hat{S}_1/q_x$  el tiempo de espera para el 1er salto y  $J_1 := S_1$  el tiempo del 1er salto. Definimos entonces  $X_t = x$  para  $0 \leq t < J_1$  y  $X_{J_1^-} := X_{J_0} = x$ . Luego,

$$\begin{cases} X_{J_1} = y & \text{con probabilidad } \pi_{xy} := q_{xy}/q_x & \text{si } q_x \neq 0, (y \neq x) \\ X_{J_1} = x & \text{en otro caso} \end{cases} \quad (2.4)$$

De esta manera se repite el proceso: Notemos  $Y_i := X_{J_i}$ . Tras  $n$  saltos, condicional a  $Y_n = x$ , definimos  $S_{n+1} := \hat{S}_{n+1}/q_x$  el tiempo de espera para el  $n+1$ -ésimo salto, y  $J_{n+1} = \sum_{i=1}^n S_i$  el tiempo del  $n+1$ -ésimo salto. Definimos  $X_t := x$  para  $J_n \leq t < J_{n+1}$  y  $X_{J_n^-} = x$ ; y finalmente,

$$\begin{cases} X_{J_{n+1}} = y & \text{con probabilidad } \pi_{X_{J_n^-}y} := q_{xy}/q_x & \text{si } q_x \neq 0, y \neq x \\ X_{J_{n+1}} = x & \text{en otro caso} \end{cases} \quad (2.5)$$

La matriz  $(\pi_{xy})$  se dice la *matriz de salto*  $(Y_n)$  es la llamada *cadena de saltos*, y es una cadena de Markov (ver Norris [32]).

### 2.1.3. Medidas invariantes

Cuando se trabaja con procesos que evolucionan en el tiempo, muchas veces es de gran interés entender el comportamiento de ese proceso para tiempos “grandes” para encontrar alguna suerte de equilibrio/estabilidad en el proceso. En ese sentido cobra importancia la búsqueda de *medidas invariantes*.

---

<sup>1</sup>Para ilustrar esto, utilizaremos la misma notación que en la figura 2.1 de esta sección

**Definición 2.6.** Decimos que un proceso de Markov tiene una *medida invariante*  $\pi$  si para todo  $t$  se verifica

$$\mathbb{P}_\pi(X_t = x) = \pi(x) \quad \forall x \in S$$

donde  $\mathbb{P}_\pi(\cdot) := \sum_y \pi(y) \mathbb{P}_y(\cdot)$ . En notación matricial

$$\pi P_t = \pi$$

Si además  $\pi(S) = 1$ ,  $\pi$  se dice *distribución invariante*

Para poder relacionar esta definición con la idea de estabilidad que buscamos, y también para poder hablar asegurar existencia de una distribución así, es preciso definir algunos conceptos.

**Definición 2.7** (Irreducibilidad). Dado un proceso de Markov  $(X_t)$  y  $(P_t)_{t \geq 0}$  sus funciones de transición,

- Dados  $x, y \in S$ , decimos que  $x \rightarrow y$  si

$$p_t(x, y) > 0 \quad \text{para algún } t > 0$$

y decimos que  $x$  se comunica con  $y$  ( lo notamos  $x \leftrightarrow y$ ) si  $x \rightarrow y$  y  $y \rightarrow x$ .

Observar que de esta manera se tiene una relación de equivalencia.

- Un proceso de Markov se dice *irreducible* si y sólo si tiene sólo 1 *clase comunicante* (1 clase de equivalencia bajo  $\leftrightarrow$ ).

**Definición 2.8** (Recurrencia). Con las mismas hipótesis de antes, si  $T_x^+ := \inf\{t \geq J_1(\omega) : X_t = x\}$  es el tiempo de retorno al estado  $x \in S$ , entonces decimos que  $x$  es *recurrente* si  $\mathbb{P}(T_x^+ < +\infty) = 1$ ; y es *recurrente positivo* si además  $m_x := \mathbb{E}_x(T_x^+) < +\infty$ .

En Norris [32, Capítulo 3] se demuestra el siguiente teorema de existencia y unicidad de una distribución invariante.

**Teorema 2.9** (Existencia y Unicidad de Distribución Invariante). *Dado un proceso de Markov  $X$  homogéneo e irreducible, un estado es recurrente positivo si y sólo si todos lo son; y esto es condición necesaria y suficiente para la existencia y unicidad de una distribución invariante.*

**Observación 2.10.** La hipótesis de homogeneidad del proceso  $X$  es una restricción fuerte que resulta necesaria para obtener el resultado de unicidad de la distribución invariante. En casos más generales, como en modelos de sistemas de partículas o en el mismo modelo que nos ocupará en este trabajo donde no hay homogeneidad, no hay resultados generales de unicidad de la distribución invariante.

## 2.2. El espacio de estados Continuo: PDMP

### 2.2.1. Conceptos generales

Queremos ahora tener un espacio de estados más general:  $S = \mathbb{R}_+$ , con su  $\sigma$ -álgebra de Borel  $\mathcal{B}(\mathbb{R}_+)$ , que notaremos simplemente  $\mathcal{B}$ . El espacio de estados puede ser mucho más

general todavía (un *Espacio Polaco*<sup>2</sup>). Sin embargo para esta tesis, no nos interesan espacios tan generales.

Llamemos  $D([0, +\infty), S) := D$  al espacio de funciones càdlàg  $f : [0, +\infty) \rightarrow S$

**Definición 2.11.** Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de Probabilidad y  $X$  un proceso estocástico càdlàg a valores en  $\mathbb{R}_+$ .  $X$  se dice un **Proceso de Markov** si tiene la propiedad de Markov respecto a su filtración natural  $(\mathcal{F}_t)_{t \geq 0}$ . Esto es, si para todo  $A \in \mathcal{B}$  se tiene

$$\mathbb{P}(X_{t+s} \in A | \mathcal{F}_s) = \mathbb{P}(X_{t+s} \in A | X_s) \quad (2.6)$$

o equivalentemente, para toda  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  acotada y medible

$$\mathbb{E}(f(X_{t+s}) | \mathcal{F}_s) = \mathbb{E}(f(X_{t+s}) | X_s) \quad (2.7)$$

Esta formulación, tiene una versión aún más general, por tener trayectorias en  $D$ .

**Propiedad 2.12.** Si  $X$  es un proceso de Markov con las hipótesis anteriores, entonces para toda  $h : D \rightarrow \mathbb{R}_+$

$$\mathbb{E}(h(X_{t+s}; t \geq 0) | \mathcal{F}_s) = \mathbb{E}(h(X_{t+s}; t \geq 0) | X_s) \quad (2.8)$$

**Observación:** Otra formulación clásica de la propiedad de Markov es aquella que dice que un proceso  $X$  tiene la propiedad de Markov si para todo  $t, s \geq 0$   $X_{t+s}$  es independiente de  $\mathcal{F}_s$  dada  $X_s$ .

Un concepto interesante (y extremadamente útil) es el de **tiempos de parada**. Intuitivamente, se trata de un tiempo aleatorio que no necesita de información futura para detenerse. Por ejemplo, imaginemos un apostador con un capital inicial  $K$  que decide jugar hasta unas 100 veces cierto juego de apuestas. Supongamos que apuesta un monto fijo  $C < K$  cada vez que comienza el juego, y  $X_n$  representa su capital tras  $n$  juegos. Dicho apostador podría decidir retirarse en el tiempo  $T$  en que su capital  $X_T$  sea menor que  $C$  (un apostador bastante moderado, otro quizás apostaría menos para seguir jugando...). Entonces  $T$ , que matemáticamente podemos expresar como  $T := \inf\{n \leq 100 : X_n < C\}$  resulta ser un tiempo de parada, pues siempre, con la información disponible en cada instante, el apostador puede decidir cuando debe frenar. En cambio supongamos que el apostador, no tan moderado, se prometiera a sí mismo abandonar el juego en cuanto llegue a su máxima ganancia de la noche. En ese caso  $T_2$ , que podríamos escribir como  $T_2 := \sup\{n \leq 100 : X_n > X_k \forall k \leq n\}$ , claramente no es un tiempo de parada, porque tras  $n$  juegos, el apostador nunca puede saber si mejorará su capital en caso de seguir jugando (¡raíz del gran vicio del apostador!).

**Definición 2.13** (Tiempos de Parada). Dado un espacio filtrado  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , y  $X$  un proceso càdlàg adaptado,

a)  $T : \Omega \rightarrow \mathbb{R}_+$  se dice *Tiempo de parada* si

$$\{T \leq t\} \in \mathcal{F}_t \quad \forall t \geq 0$$

---

<sup>2</sup>Espacio topológico 2do contable completamente metrizable.

b)  $\mathcal{F}_T := \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t \ \forall t \geq 0\}$

c) La variable aleatoria  $\mathcal{F}_T$ -medible  $X_T$  se define como  $X_T(\omega) := X_{T(\omega)}(\omega)$

Con este concepto podemos formular una propiedad interesante y de inmensa utilidad teórica y práctica es la *Propiedad Fuerte de Markov*.

**Definición 2.14.** Un Proceso  $X$  como en la definición anterior tiene la *Propiedad Fuerte de Markov* si para todo  $T$  tiempo de parada y para todo  $f : S \rightarrow \mathbb{R}_+$  acotada y medible se tiene

$$\mathbb{E}(f(X_{T+s})|\mathcal{F}_s) = \mathbb{E}(f(X_{T+s})|X_s) \quad (2.9)$$

**Observación 2.15.** Todo Proceso de Markov  $X$  cumple la igualdad anterior para todo  $T$  que asuma a lo sumo numerables valores. En particular las cadenas de Markov tienen la Propiedad Fuerte de Markov.

**Observación 2.16.** Notar que la Propiedad Fuerte de Markov resulta ser una generalización de la Propiedad de Markov estándar, pues si  $X$  verifica (2.9) entonces puedo tomar  $T \equiv t$  y se obtiene automáticamente (2.6)

## 2.2.2. Procesos de Markov Determinísticos a Trozos

Esta sección se apoya en la teoría desarrollada en *Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models*, de M. Davis [14], y en *Piecewise deterministic Markov process — recent results* [4], de Azais et al.; entre otras referencias.

Vamos a presentar una familia de Procesos, llamados **Procesos de Markov Determinísticos a Trozos**. Estos procesos, si bien no agotan los Procesos de Markov en toda su generalidad, son de mucha riqueza y variedad tanto teórica como en sus aplicaciones, que se extienden a Finanzas, biología [45], física, etc.(Ver también [31]).

Los **PDMP** se caracterizan por presentar una evolución determinística en el tiempo, salvo en una secuencia creciente de tiempos aleatorios  $(J_n)$ , llamados *tiempos de salto* en donde el proceso cambia abruptamente de estado mediante un mecanismo aleatorio, para continuar con la dinámica determinística.

Observemos que dentro de esta clase de Procesos de Markov quedan cubiertos aquellos con espacio  $S$  de estados discreto. En ese caso - y utilizando la notación de la construcción en 2.1.2 - la dinámica inter-salto es simplemente mantenerse en ese mismo estado, los tiempos de salto son los  $(J_n)$ , y el mecanismo aleatorio está dado por la matriz de salto  $(\pi_{xy})$ .

Los **PDMP**'s, tal como los describimos, quedan determinados por 3 características locales:

1. *El flujo que determina la dinámica inter-salto:* Éste se expresa mediante

$$\phi : \mathbb{R}_+ \times S \rightarrow S \text{ con } \phi(0, x) = x$$

$\phi(t, x)$  representa el estado del proceso a tiempo  $t$ , con estado inicial  $x$ ;

2. La **tasa de salto** que determina la distribución de los tiempos de salto: Representada por  $\lambda : S \rightarrow \mathbb{R}_+$ , y con la condición de que exista un  $\varepsilon > 0$  tal que

$$\int_0^\varepsilon \lambda(\phi(s, x)) ds < +\infty$$

Esta condición asegura que  $\mathbb{P}(J_k < J_{k+1}) = 1$ , teniendo en cuenta que la distribución del primer tiempo de salto  $J_1$  suele ser el de una *exponencial generalizada*,

$$\mathbb{P}_x(J_1 > t) = \exp\left(-\int_0^t \lambda(\phi(s, x)) ds\right) \quad (2.10)$$

3. La **medida de transición**  $Q : S \times \mathcal{B}(S) \rightarrow [0, 1]$  en función de la cual se elige el nuevo estado en los tiempos de salto.

De esta manera se puede construir la trayectoria de un **PDMP**  $(X_t)_{t \geq 0}$  con estado inicial  $X_0 = x$  del siguiente modo:

- $X_t := \phi(t, x)$  para  $0 \leq t < J_1$
- Definimos  $X_{J_1^-} := \phi(J_1, x)$ . Habiendo llegado a  $J_1$ , el *estado post-salto*  $Y_1 := X_{J_1}$  se elige como

$$\mathbb{P}(Y_1 \in A | J_1) := Q(X_{J_1^-}, A)$$

- Recomenzamos el Proceso desde  $Y_1$ .

**Propiedad 2.17** (Propiedad (Fuerte) de Markov). Un Proceso así construido cumple con la Propiedad de Markov. Más aún, cumple con la Propiedad Fuerte de Markov.

En Davis [14, Sec 3] se encuentra la demostración, para caso general donde el espacio de estados es una unión disjunta  $\bigcup_{v \in C} M_v$  con  $C$  un conjunto contable de índices y cada  $M_v$  es un abierto de  $\mathbb{R}^N$ .

## Capítulo 3

# Modelos de Redes Neuronales

### 3.1. Contexto biológico

El cerebro es el principal órgano del cuerpo humano, y el centro del sistema nervioso. Su morfología particular - *arrugada* - es tal que maximiza la superficie en el volumen de que dispone, dentro del cráneo. De esta manera, la *corteza cerebral* - que juega un rol fundamental en la memoria, la percepción, el razonamiento y la conciencia, entre otras funciones - posee alrededor de entre 15 y 33 mil millones de neuronas [40] en un espesor de 2 a 4 milímetros; distribuidas en grandes poblaciones de redes de neuronas. Estas poblaciones se organizan en una serie de capas o “láminas”, llamadas *columnas corticales*, cada una de las cuales se compone de poblaciones de neuronas estadísticamente idénticas y fuertemente interconectadas.

Las neuronas son células eléctricamente excitables que reciben y transmiten información a través de señales eléctricas y químicas. Existen diferentes tipos de neuronas, pero la estructura celular básica de la mayor parte se compone de un *cuerpo celular (soma)*; las *dendritas*, que reciben los impulsos eléctricos y los trasladan hacia el cuerpo celular; y un *axon*, mediante el cual la neurona transmite sus impulsos hacia otras neuronas, u otras células. La figura 3.1 representa el esquema típico de una neurona.

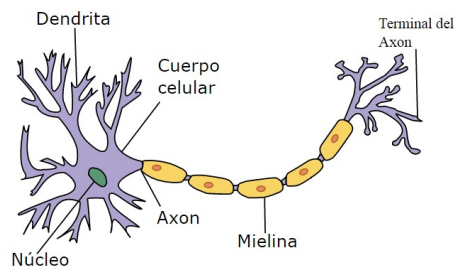


Figura 3.1: Representación esquemática de una Neurona

La transmisión de las señales electroquímicas entre las neuronas se da a través de las *sinapsis*: Las neuronas poseen típicamente una diferencia de potencial entre el interior y el exterior de su membrana celular, que en su estado de reposo se encuentra alrededor de los  $-70mV$ . Los impulsos que recibe la neurona a través de las dendritas llegan a la membrana a través de



neurotransmisores que abren canales de iones (de Potasio y de Sodio, fundamentalmente) en dicha membrana, generando un crecimiento del potencial (impulso *exitatorio*) o un decrecimiento (impulso *inhibitorio*). Si el impulso exitatorio es lo suficientemente grande, se genera una descarga eléctrica - llamada potencial de acción o *spike*, en inglés - que viaja a través del axón hasta sus terminales; en general las dentritas de otras neuronas. Inmediatamente tras producirse el potencial de acción, la neurona que lo transmitió - neurona *presináptica* - decrece su potencial hasta valores entre  $-80$  y  $-70mV$ , un poco menores a su potencial de reposo, pasando por un breve período refractario, para reestablecerse en su potencial de reposo.

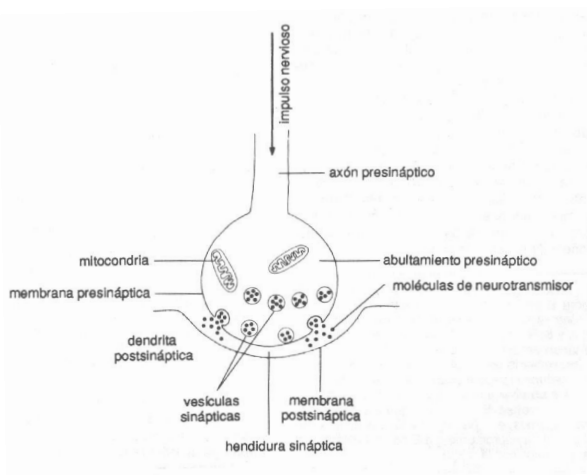


Figura 3.2: Sinapsis

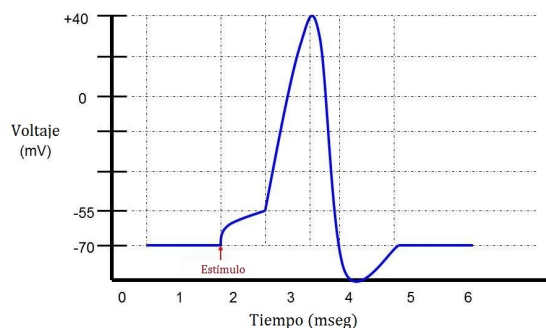


Figura 3.3: Esquema de la evolución de un Potencial de acción en el tiempo

Es importante mencionar que, basados en las investigaciones de Henry Pickering Bowditch y posteriores, se sabe que en estas dinámicas eléctricas que se producen en las redes neuronales, las neuronas actúan bajo el *principio de todo-o-nada* [1]. Esto significa que no existen lo que podrían llamarse “impulsos fuertes” o “impulsos débiles”; en otras palabras, la intensidad de la descarga de una neurona no depende de la intensidad del estímulo que recibe: O bien efectúa un potencial de acción, o directamente no emite ningún impulso.

La intensidad del impulso recibido por una neurona post-sináptica, en cambio, sí tiene efecto sobre la *tasa de disparo* de ésta, es decir, sobre su frecuencia de “disparo”. Este hecho

empírico se verá también reflejado en el modelo.

### 3.2. Los modelos *integrate-and-fire*

Los modelos del tipo *integrate-and-fire* han sido, y siguen siendo, de inmensa importancia para la profundización en el entendimiento del comportamiento de la actividad neuronal de una neurona aislada, aunque los progresos experimentales (empíricos y computacionales) trasladaron el interés a modelos con poblaciones (pequeñas o grandes) de neuronas; con la consiguiente aparición de otros modelos, y también de la adaptación de los existentes. En el caso redes de neuronas piramidales, interneuronas o neuronas de tipo Purkinje, dada su enorme cantidad de sinapsis, se han utilizado teoremas límites del estilo campo-medio para concretar esas adaptaciones. (cf. [46] [22]).

Existe una vastísima literatura en relación a modelos (**LIF**), determinísticos y también incorporando un factor estocástico. Para una reseña muy completa y metódica, ver el trabajo de Sacerdote y Giraudo [46]; de allí extrajimos mucha información en estos párrafos. También aportan en este sentido los trabajos de Burkitt [9] y [10].

Tanto los modelos *integrate-and-fire* (**IF**) como los *leaky-integrate-and-fire* (**LIF**) poseen varios supuestos y características en común, a saber:

- Los modelos describen matemáticamente la evolución del potencial de una neurona en función del tiempo. Su foco se centra en el estudio de intervalos inter-descarga (**IID**), o lo que es lo mismo, en los tiempos de descarga; buscando poder dar una caracterización de los mismos.
- Para ello, dichos modelos explotan la relativa “lentitud” de los tiempos entre cada descarga, en relación al tiempo que tardan en efectuarse los cambios de potencial debido a la emisión de un spike. Por eso, dicha emisión es reducida a un evento instantáneo en el que la neurona se “resetea” a su estado de reposo. Esta simplificación también implica pasar por alto los mecanismos biológico-químicos que intervienen en la sinapsis; esto es, la apertura de los canales iónicos con el consiguiente intercambio de iones, origen de los cambios de potencial bruscos en la neurona. Sin embargo, estos cambios en el voltaje durante un potencial de acción son un fenómeno altamente estereotipado, lo que sumado a la mencionada rapidez relativa a los (**IID**), hace esta suposición razonable.
- La neurona suma linealmente los impulsos sinápticos que recibe del exterior o de otras neuronas (algunos modelos más complejos incorporan *sinapsis de conductancia*, en que la integración no es lineal); que pueden ser inhibitorios o excitatorios, y que suelen ir acompañados de un peso correspondiente a su *eficacia sináptica*.
- La descarga se produce cuando el potencial alcanza cierto umbral **fijo** (o a lo sumo representado por una función determinística del tiempo), que suele situarse en  $V_{th} := -55$  mV (según datos experimentales).
- En el caso de los modelos (**LIF**), se incorpora el supuesto de que, sin recibir impulsos externos, la neurona tiene un efecto de decaimiento o pérdida (“leakage”) exponencial de voltaje hacia el valor de reposo.

El primer modelo *integrate-and-fire* - embrionario para prácticamente todo el estudio posterior en la materia - fue el introducido por Louis Lapicque en 1907 [26]. Desde el punto de vista heurístico, su modelo se basa en la comparación del funcionamiento de una neurona con el de un circuito con un resistor y un capacitor en paralelo, sujeto a una corriente. Por ello matemáticamente, Lapicque arma su modelo derivando en el tiempo la Ley de Capacitancia, obteniendo esencialmente una ecuación de la forma

$$\begin{cases} \frac{dv}{dt} = R I(t) & \text{si } v(t) < v_{th} \\ v(t) = v_r & \text{si } v(t^-) = v_{th} \end{cases} \quad (3.1)$$

Donde  $v(t)$  representa el voltaje de la neurona a tiempo  $t$ ,  $R$  es la resistencia de la membrana celular,  $v_{th}$  es el potencial umbral, y  $v_r$  es el potencial de reposo.

A pesar de ser un modelo muy limitado - aún en su forma “leaky”  $\frac{dV}{dt} = RI(t) - V(t)$  - este modelo fue seminal en el desarrollo de otros modelos en la misma línea y en plena vigencia, entre los que se encuentra el modelo presentado por Robert y Touboul que utilizaremos aquí.

El modelo de Lapicque y sus primeras modificaciones son determinísticas; sin embargo biológicamente se ha visto que las neuronas desarrollan una actividad irregular, con “ruido”, aún aisladas de una red; y por eso la necesidad de incorporar el elemento estocástico si se pretende captar la dinámica neuronal con mayor realismo. El primer modelo en introducir aleatoriedad se debe a Gerstein y Mandelbrot [20], en la forma de un paseo aleatorio representando descargas provocadas por neuronas exteriores a la red considerada; y por lo tanto, no teniendo en cuenta aún la aleatoriedad intrínseca de los tiempos de descarga de las propias neuronas de la red.

Sin embargo, este modelo contribuyó también a consolidar el estudio de modelos de tipo (**LIF**) estocásticos, en especial a través del uso de ecuaciones diferenciales estocásticas. Un modelo clásico de tipo (**LIF**) se puede escribir por lo tanto de la siguiente manera<sup>1</sup>:

$$\begin{cases} dX_i(t) = -(X_i(t) - v_r)dt + \sum_{j \neq i} w_{ji}(t)dD_j(t) + dI_i(t) & \text{si } X_i(t) < v_{th} \\ dD_i(t) = 1, X_i(t) = v_r & \text{si } X_i(t^-) = v_{th} \end{cases} \quad (3.2)$$

Donde

- $(w_{ji}(\cdot))$  representa la fuerza o eficacia sináptica de la neurona  $j$  a la neurona  $i$ .
- $D_j(t)$  es la cantidad de spikes de la neurona  $j$  hasta el tiempo  $t$ .
- $I_i(t)$  es el input proveniente del exterior del sistema, hacia la neurona  $i$ .

¿Donde suele colocar la aleatoriedad la literatura? Desde el mencionado trabajo de Gerstein y Mandelbrot en adelante, la literatura propone colocar la aleatoriedad en sus modelos (**LIF**) en el input externo. Lo común es proponer

$$dI_i(t) = dB_i(t)$$

Con  $B_i$  un proceso Gaussiano, [41], [29], [35]. Se trata de un proceso de *Ornstein–Uhlenbeck*, que resulta estacionario, Gaussiano y Markoviano (el único no trivial dentro de estas categorías de manera simultánea).

---

<sup>1</sup>Cambiamos la notación, como puede apreciarse, para una comparación más agradable a la vista con el modelo que proponemos, en la ecuación (3.3)

Esta descripción tiene sus ventajas. Podemos mencionar su relativa sencillez desde el punto de vista de la complejidad matemática, que permite resoluciones analíticas parciales o incluso completas en algunos casos sencillos, y en consecuencia la obtención de resultados explícitos y en gran medida relevantes, pues estos modelos mantienen suficiente complejidad como para retener muchas características esenciales del procesamiento neuronal. Más aún, con el avance computacional existen métodos de simulación y ajuste de parámetros para muchos de estos modelos. En consecuencia se ha podido profundizar en el conocimiento de la dinámica eléctrica de una neurona. Ver por ejemplo Brunel [8] o Pillow et al. [41].

Sin embargo, existen también limitaciones importantes por parte de estos modelos. En primer lugar, distintas investigaciones muestran que la aleatoriedad inherente a cada neurona no sólo *no* es una característica despreciable en la dinámica de descargas de las redes neuronales, sino que cumplirían un rol fundamental en los procesos de toma de decisiones, la memoria de corto plazo, la percepción, y la atención (el libro de Rolls y Deco [15], es una fuente bastante completa al respecto).

Otra desventaja que presentan estos modelos es la condición sobre el valor umbral  $v_{th}$  fijo. En general, esto dirige el estudio de los tiempos de spike al estudio de la distribución de

$$\tau_{th} := \inf\{t > 0 : X_t > v_{th}\}$$

Es decir, el *primer tiempo de paso* (*first passage time*, en inglés) del potencial ( $X_t$ ) a través de  $v_{th}$  (también podría leerse como el *hitting time* sobre el conjunto  $(v_{th}, +\infty)$  en un contexto Markoviano).

El problema con esta condición es, entre otros, el efecto *blow-up* del sistema: Está probado en más de un trabajo que, para redes neuronales de impulsos exitatorios, la solución del sistema *explota* en tiempo finito en el límite de *campo-medio* (o sea, en una red neuronal infinita, ver abajo); es decir la tasa promedio de disparo de la red se va a infinito (intuitivamente, esto se interpreta como que una gran proporción de las neuronas en la red se disparan al mismo tiempo, provocando una “avalancha”). Entre otros trabajos, se pueden consultar los de Cáceres, Carrillo y Perthame [12]; Cáceres, Schneider y Granados [21]; Delarué [18], en donde se muestra que, bajo una buena elección de parámetros (relacionados con la eficacia sináptica proporcional) sí es posible alguna solución para todo tiempo, sin explosión.

Lo expuesto apunta a la conclusión de que un modelo clásico de tipo **(LIF)** quizás no sea el modelo más adecuado en el contexto de una gran población de neuronas en donde se quiera tener en cuenta su *ruido* intrínseco. Para hacer frente a estas limitaciones por lo tanto, otro modelo estocástico que resulta natural para representar la aleatoriedad en este contexto es el que resulta de considerar, en lugar de un proceso Gaussiano externo al sistema, una serie de Procesos de Poisson no-homogéneos cuyas tasas son funciones no necesariamente lineales del input recibido. El primero en proponer un modelo en esta forma - el *Linear non-linear Poisson Model* (**LNPM**) - fue Chihilnisky [13], y ha mostrado éxito en especial desde el punto de vista computacional, y por su relativa simplicidad para obtener estimadores de sus parámetros. Ver Pillow et al [41], Paninski [39].

Lo dicho hasta ahora motiva un modelo como el propuesto por Robert y Touboul - que exploramos en esta tesis - que se apoya tanto en este modelo como en los modelos **(LIF)** para sus puntos de partida, buscando ser superador. Representa el estudio de una red neuronal *exitatoria*, y la dinámica es la descrita en la introducción (1.1). Adaptando la notación para

compararla con (3.2), tendríamos

$$dX_i(t) = -(X_i(t) - v_r)dt + \sum_{j \neq i} W_{ji} d\mathcal{N}_{b(X_j(t^-) - v_r)}^j(t) - X_i(t^-) d\mathcal{N}_{b(X_i(t^-) - v_r)}^i(t) \quad (3.3)$$

Con  $v_r = 0$  y  $d\mathcal{N}_z^i(t)$  un proceso de Poisson de tasa  $z$ , y  $(W_{ij}(t))$  variables aleatorias independientes e idénticamente distribuidas (i.i.d.).

Observar que si quisiéramos utilizar la suposición del valor umbral, eso se enmarcaría en este modelo mediante  $b(x) := \mathbf{1}_{[v_{th}, +\infty)}(x) \cdot (+\infty)$ , donde  $\mathbf{1}_A$  es la función indicadora del conjunto  $A$ , siguiendo la convención “ $0 \cdot +\infty = 0$ ”. La aleatoriedad intrínseca de las neuronas queda capturada por un lado a través de los procesos de Poisson y también mediante las variaciones aleatorias de las eficacias sinápticas  $W$ . Notar además que el valor  $b(0)$  representa la tasa de descarga de una neurona en su potencial de reposo, y puede ser interpretado entonces como el ruido externo, que de esta manera queda modelado como un Proceso de Poisson homogéneo de tasa  $b(0)$ .

Queda expuesta la motivación para el modelo propuesto en [43], que pasamos a describir con más detalle.

### 3.3. El modelo estocástico

El modelo consiste en un proceso  $(X_i(t))$  en el que el estado  $X_i(t)$  representa la diferencia de potencial eléctrico entre las paredes externa e interna de la membrana celular de la neurona  $i$  a tiempo  $t$ . En ausencia de descargas, esta diferencia de potencial evoluciona en el tiempo decayendo exponencialmente - bajo un régimen determinístico -, según la ecuación diferencial:

$$dX_i(t) = -X_i(t) dt$$

Este comportamiento representa el decaimiento - a falta de estímulo - hacia un Potencial de Reposo  $V_R$ , que declaramos igual a 0.

Cada neurona produce descargas de acuerdo a un proceso de Poisson no homogéneo cuya tasa es una función no-negativa, no-decreciente del voltaje  $X_i(t)$ . Cuando la neurona  $j$  descarga, ésta resetea su potencial al valor de equilibrio  $V_R$ ; y transmite a cada una de las neuronas de la red una carga excitatoria  $W_{ji}$ , modelada como una variable aleatoria no negativa con ley  $W$ . De esta manera, tras una descarga de la Neurona  $i$ , se tiene

$$\begin{cases} X_i(t) = 0 \\ X_j(t) = X_j(t^-) + W_{ij} \quad j \neq i \end{cases}$$

Donde  $W_{ij} \sim W$  i.i.d.

Equivalentemente, puede describirse este proceso en términos de la ecuación diferencial

estocástica (3.3), que puede escribirse como

$$\begin{aligned}
dX_i(t) = & -X_i(t)dt + \sum_{j \neq i} \int_{\mathbb{R}_+^2} z_j \mathbb{1}_{\{0 \leq u \leq b(X_j(t-))\}} \mathcal{N}_j(du, dz, dt) \\
& - X_i(t^-) \int_{\mathbb{R}_+^2} \mathbb{1}_{\{0 \leq u \leq b(X_i(t-))\}} \mathcal{N}_i(du, dz, dt)
\end{aligned} \tag{3.4}$$

Donde los  $(\mathcal{N}_j)$  son procesos de Poisson independientes con medida de intensidad dada por  $du \otimes \mathbb{W}^j(dz) \otimes dt$ , con

$$\mathbb{W}^j(dz) = \bigotimes_{k=1}^{j-1} W(dz_k) \otimes \delta_0(dz_j) \otimes \bigotimes_{k=j+1}^N W(dz_k)$$

(La medida correspondiente a la descarga de la Neurona  $j$  sobre las demás Neuronas de la Red)

En su forma integral, la ecuación anterior toma la forma

$$\begin{aligned}
X_i(t) = & X_i(0) - \int_0^t X_i(s) ds + \mathbb{E}(W) \sum_{j \neq i} \int_0^t b(X_j(s)) ds \\
& - \int_0^t X_i(s) b(X_i(s)) ds + M_i(t)
\end{aligned} \tag{3.5}$$

Donde

$$\begin{aligned}
M_i(t) = & \sum_{j \neq i} \int_{s=0}^t \left[ \int_{u=0}^{b(X_j(s-))} \int_{z=0}^{+\infty} z_i \mathcal{N}_j(du, dz, ds) - \mathbb{E}(W_1) du \right] ds \\
& - \int_{s=0}^t \left[ X_i(s^-) \int_{u=0}^{b(X_i(s-))} \int_{z=0}^{+\infty} \mathcal{N}_i(du, dz, ds) - 1 du \right] ds
\end{aligned} \tag{3.6}$$

es la Martingala local asociada. Para más detalle, [43][Proposición 13] y [44]

Nos interesa dar otro enfoque más a esta dinámica: el de un **PDMP**. Como dijimos en el capítulo anterior, este enfoque enriquece el conocimiento del Proceso, entre otras cosas porque da una idea muy clara del comportamiento de sus trayectorias. Tomemos  $S := \mathbb{R}_+^N$ , y supongamos que a tiempo  $t = 0$  tenemos  $X_0 = x \in \mathbb{R}_+^N - \{0\}$ . Entonces hasta la primer descarga del sistema, su estado es determinístico y vale

$$X_t = \phi(x, t) := e^{-t}x = (x_1 e^{-t}, \dots, x_N e^{-t})$$

La tasa de salto del sistema  $\lambda(t)$  de la ecuación(2.10) queda determinada por la igualdad

$$J_1 := \inf\{\tau_i : i \in \{1, \dots, N\}\}$$

donde  $\tau_i$  es el tiempo de espera de la neurona  $i$ , definida - recordemos - a partir de un proceso de Poisson no-homogéneo de tasa  $\lambda_i(t) = b(X_i(t))$  mientras no hay spikes de otras neuronas;

y por lo tanto  $\tau_1, \dots, \tau_N$  son variables aleatorias independientes, exponenciales generalizadas de tasa  $\lambda_i$  respectivamente; y se concluye que  $\lambda(t) := \sum_i \lambda_i(t)$ . Finalmente

$$\mathbb{P}(Y_1 \in A | J_1) = \mathbb{P}(\xi_i(X_{J_1^-} + (W_{i1}, \dots, W_{iN})) \in A) \quad \text{si } \{J_1 = \tau_i\}$$

Con  $\xi_i(x_1, \dots, x_i, \dots, x_N) := (x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_N)$

Observemos que, dada la condición inicial  $X(0) = (x_1, \dots, x_N)$ , las variables  $\tau_i$  tienen la siguiente distribución:

$$\mathbb{P}(\tau_i > t) = \exp\left(-\int_0^t b(x_i e^{-s}) ds\right) = \exp\left(-\int_{x_i e^{-t}}^{x_i} \frac{b(z)}{z} dz\right) \quad (3.7)$$

Por lo tanto, haciendo uso de la independencia de estas variables, el *primer instante de descarga del sistema*  $\tau := \min_i \tau_i$  tiene distribución

$$\mathbb{P}(\tau > t) = \exp\left(-\int_0^t \sum_{i=1}^N b(x_i e^{-s}) ds\right) \quad (3.8)$$

Finalmente, damos el **generador** del proceso (cf. [30] [19]). Para una función  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  suave, se tiene

$$\mathcal{L}f = \sum_{i=1}^N \left[ b(x_i) \int_{\mathbb{R}_+^{N-1}} [f(x + \alpha_i(x, w_1, \dots, w_{N-1})) - f(x)] \bigotimes_{i=1}^{N-1} dW(w_i) \right] - \langle \nabla f(x), x \rangle$$

Con

$$\alpha_i(x, w_1, \dots, w_{N-1}) = \begin{cases} x_j + w_j & j < i \\ -x_i & j = i \\ x_j + w_{j-1} & j > i \end{cases}$$

## Capítulo 4

# Redes Neuronales Finitas

Comenzamos con la siguiente observación:

### Observación 4.1. Propiedades de extinción de la red

Teniendo en cuenta el marco biológico del modelo, es natural asumir que  $b(x)$  -la tasa de descarga de una neurona - es positiva no-decreciente; ya que a mayor voltaje son mayores las chances de ocurra un spike. Si el estado inicial de una neurona es  $x$  y durante el intervalo de tiempo  $[0, t]$  no ocurren descargas, entonces  $X_i(t) = xe^{-t}$  y en particular, su tasa de descarga será  $b(xe^{-t})$ . Esta tasa, a falta de estímulo, decrece a  $b(0)$ , que corresponde a la tasa de descarga con potencial de reposo, y que como observamos con anterioridad, puede considerarse como la representación del ruido externo. Si  $b(0) = 0$ , entonces habrá probabilidad positiva de que no ocurran más spikes. Más aún, si  $X(0) = (x_i)_{i=1\dots N}$  es lo suficientemente chico, es de esperar que no ocurra ningún spike. Los siguientes resultados mostrarán que el comportamiento de  $b(x)$  cerca del 0 condicionan la extinción -o no- de la red neuronal.

### 4.1. Comportamiento asintótico en el tiempo.

El siguiente lema se debe a Robert y Touboul, en [43]. Cuando sea éste el caso, notaremos (R-T).

**Lema 4.2** (R-T). *Si se tiene la siguiente condición:*

$$\int_0^1 \frac{b(s)}{s} ds = +\infty$$

*Entonces todo nodo con condición inicial  $x > 0$  descarga con probabilidad 1.*

**Observación:** Siendo  $b(x)$  positiva no-decreciente,  $b(0) > 0$  implica la condición del Lema; y bajo condiciones de regularidad mínimas para  $b$ , vale la recíproca.

#### **Demostración:**

Observemos que la condición dada por la hipótesis, también vale en el intervalo  $[0, a]$  con



$a > 0$  arbitrario, por ser  $b$  no-decreciente.

Supongamos que para algún  $i \in \mathbb{I}_N$ ,  $X_i(0) = x > 0$ . Sea  $\tau_i := \inf\{t > 0 : X_i(t) = 0\}$  el momento de la 1er descarga de la neurona  $i$ .

Entonces se tiene

$$\begin{aligned}\mathbb{P}(\tau_i > t) &= \mathbb{E}\left(e^{-\int_0^t b(X_i(s))ds}\right) \leq \exp\left(-\int_0^t b(xe^{-s})ds\right) \\ &= \exp\left(-\int_{xe^{-t}}^x \frac{b(z)}{z} dz\right)\end{aligned}$$

La desigualdad vale porque  $X_i(s) \geq xe^{-s}$  (pues mientras  $s < \tau_i$ ,  $X_i$  sólo puede incrementar su potencial por las descargas de otras neuronas) y porque  $b$  es no-decreciente.

Considerando los 2 extremos de la desigualdad, y haciendo tender  $t \rightarrow +\infty$  se obtiene

$$\mathbb{P}(\tau_i = +\infty) = 0$$

□

Con este panorama, la propiedad de ergodicidad en este caso parece posible mientras las neuronas no descarguen demasiado rápido, como en el modelo **(LIF)** analizado en Cáceres et al. [12].

Ahora demostraremos un resultado técnico que estima el tiempo de retorno a un compacto específico, dependiente de la dimensión de la red y de la eficacia sináptica.

**Lema 4.3** (Estimación de Tiempo de Retorno (R-T)). *Existe una constante  $C_0 > 0$ , dependiente sólo de la dimensión  $N$  y de  $\mathbb{E}(W)$ , tal que si*

$X(0) = x = (x_i) \notin [0, C_0]^N$  y

$$T_0 := \inf\left\{t > 0 : X(t) \in [0, C_0]^N\right\}$$

entonces

$$\mathbb{E}_x(T_0) \leq \|x\| := \sum_{i=1}^N x_i \tag{4.1}$$

### Demostración:

Sea  $\|\cdot\|_1$  la norma 1 en  $\mathbb{R}^N$ . Si  $S(t) := \|X(t)\|_1 = \sum_{i=1}^N X_i(t)$ , entonces a partir de (3.5) y (3.6),

$$\begin{aligned}S(t) &= S(0) - \int_0^t S(s)ds + \mathbb{E}(W)(N-1) \int_0^t \sum_{i=1}^N b(X_i(s))ds \\ &\quad - \int_0^t \sum_{i=1}^N X_i(s)b(X_i(s)) ds + \sum_{i=1}^N M_i(t) \\ &= S(0) + \int_0^t \sum_{i=1}^N \left[ (N-1)\mathbb{E}(W)b(X_i(s)) - X_i(s)[1 + b(X_i(s))] \right] ds + \sum_{i=1}^N M_i(t)\end{aligned}$$

Sea

$$F := \left\{ x \in \mathbb{R}_+^N : \sum_{i=1}^N \left( (N-1)\mathbb{E}(W)b(x_i) - x_i[1+b(x_i)] \right) \geq -1 \right\}$$

Luego  $F \subseteq [0, C_0]^N$  y es un compacto, con  $C_0 = N(N-1)\mathbb{E}(W) + 1$  (Ver lema(7.1) en el apéndice).

Consideremos  $X(0) = x \notin F$  y sea

$$T_F := \inf \{ t > 0 : X(t) \in F \}$$

Entonces se tiene que  $T_0 \leq T_F$  y  $T_F$  es un tiempo de parada. Luego

$$\begin{aligned} \mathbb{E}(S(T_F \wedge t)) - S(0) &= \mathbb{E} \left( \int_0^{T_F \wedge t} \sum_{i=1}^N \{ (N-1)\mathbb{E}(W)b(X_i(s)) - X_i(s)[1+b(X_i(s))] \} ds \right) \\ &\leq -\mathbb{E}(T_F \wedge t) \end{aligned}$$

Y finalmente  $\mathbb{E}(T_F) \leq S(0)$  se obtiene tomando límite y acotando.

□

Ahora demostraremos Propiedades de estabilidad del Proceso de Markov ( $X(t)$ ), bajo el contexto de Procesos de Markov Harris Ergódicos. Una introducción a este tema se puede ver en Nummelin [33] y Asmussen [3]. El enunciado es de (R-T); la demostración del primer punto fue completada por nosotros, mientras que fue corregida demostración de la última parte del segundo punto.

**Teorema 4.4** (Comportamiento asintótico).

1. Si se cumple la condición:

$$\int_0^1 \frac{b(s)}{s} ds < +\infty \quad (4.2)$$

Entonces  $\mathbb{P}(\tau_{ext} < +\infty) = 1$ ; con  $\tau_{ext}$  el tiempo de extinción del sistema (el tiempo aleatorio en que no ocurren más spikes). En particular, se tiene

$$\lim_{t \rightarrow +\infty} (X_i(t), i = 1 \dots N) = \mathbf{0}$$

y  $\delta_0$  - la masa de Dirac en 0 - es la única distribución invariante.

2. Si  $b(0) > 0$ , y  $\exists K : W \leq Ka.e.$ , entonces  $(X(t))$  es Harris Ergódico.

**Demostración:**

1. Supongamos que la ecuación (4.2) se cumple. Sin pérdida de generalidad,  $X(0) = x \neq 0$ . Si  $x \in \mathcal{C} := [0, C_0]^N$ , sea  $\tau := \inf \tau_i$  el momento de la 1er descarga del sistema; y sean  $E_{x_i}^i$  variables aleatorias independientes “pseudo-exponenciales” es decir,

$$\mathbb{P}(E_{x_i}^i > t) = \exp \left( - \int_0^t b(x_i e^{-u}) du \right)$$

Entonces la distribución de  $\tau$  puede pensarse como el resultado de colocar  $N$  relojes independientes  $E_{x_i}^i$  y tomar el instante en que suena el primero. Matemáticamente:

$$\begin{aligned}\mathbb{P}(\tau > t) &= \mathbb{P}(\text{mín } E_{x_i}^i > t) = \prod_{i=1}^N \mathbb{P}(E_{x_i}^i > t) \\ &= \exp\left(-\int_0^t \sum_{i=1}^N b(x_i e^{-u}) du\right) \geq \exp\left(-\int_0^t N b(C_0 e^{-u}) du\right) \\ &= \exp\left(-N \int_{C_0 e^{-t}}^{C_0} \frac{b(s)}{s} ds\right)\end{aligned}$$

Entonces se verifica

$$\mathbb{P}(\tau = +\infty) \geq \exp\left(-N \int_0^{C_0} \frac{b(s)}{s} ds\right) := \eta^N$$

Ahora, si  $x \notin \mathcal{C}$  por el lema anterior 4.3,  $(X(t))$  vuelve a  $\mathcal{C}$  con probabilidad 1. Más aún, **en tiempo integrable**. Luego, si  $\tau_{ext} \in [0, +\infty]$  es el *tiempo de extinción del sistema* y  $\mathcal{Z} :=$  Cantidad de spikes del sistema dentro de  $\mathcal{C}$ ; entonces se tiene que  $\{\tau_{ext} < \infty\} = \{\mathcal{Z} < \infty\}$ , y por lo tanto  $\mathbb{P}(\tau_{ext} < \infty) = \mathbb{P}(\mathcal{Z} < \infty)$ . Pero  $\mathcal{Z} \leq \mathcal{Z}_\eta$  a.e. con  $\mathcal{Z}_\eta \sim Ge(\eta^N)$ . Entonces  $\mathbb{P}(\mathcal{Z} < \infty) \geq \mathbb{P}(\mathcal{Z}_\eta < \infty) = 1$

2. Sabemos que  $b(0) > 0$ . La idea es mostrar que  $\mathcal{C}$  es un “*Regeneration Set*” para la cadena de saltos  $Y_n = (Y_{n,i})_{i \in \mathbb{I}_N}$ , (es decir,  $Y_n := X_{s_n}$  con  $s_n$  los tiempos de salto de  $X(t)$ ). Un *regeneration set* (R.S.) es un conjunto al que un proceso con “renovación” vuelve recurrentemente. En algún sentido es la generalización de los estados recurrentes en el caso de espacio de estados numerable. De hecho la construcción de una medida invariante a partir de los ciclos entre visitas a un (R.S.) es análoga a lo que ocurre con la visita a un estado recurrente. Matemáticamente,  $R \in \mathcal{B}$  se dice un *Regeneration Set* si es *recurrente* (i.e. para todo  $x \in \mathbb{R}_+$ ,  $\mathbb{P}_x(\tau(R) < +\infty) = 1$ ; con  $\tau(R) := \inf\{t > \varepsilon_0 > 0 : X(t) \in R\}$ ) y para algún  $r > 0$ , la probabilidad  $\mathbb{P}_x(X(r) \in \cdot)$  posee un *componente común*: existe algún  $\delta \in (0, 1)$  y una medida de probabilidad  $\mu$  en  $\mathcal{B}(\mathbb{R}_+)$  tal que

$$\mathbb{P}_x(X(r) \in F) \geq \delta \mu(F) \quad \forall x \in R, F \in \mathcal{B}(\mathbb{R}_+)$$

Ver Asmussen para más detalles [3]

Tomemos para cada  $k \in \mathbb{N}$ ,  $E_k^i \sim \mathcal{E}(b(0))$  y  $(E_k^{i,x}, x > 0)$  i.i.d. -independientes de  $E_k^i$ -con distribución

$$\mathbb{P}(E_k^{i,x} > t) = \exp\left(-\int_0^t [b(xe^{-u}) - b(0)] du\right)$$

y elegidas de tal modo que  $E^{i,x} \leq E^{i,y}$  a.e. si  $x \geq y$ . Entonces se tiene que si  $X(0) = (x_i) = 0$ , para cada  $i$  la distribución de la variable aleatoria  $E_k^{i,x_i} \wedge E_k^i$  se corresponde con el tiempo de espera de un proceso de Poisson no-homogéneo de tasa  $\lambda_i(t) = b(x_i e^{-t})$ , es decir

$$\mathbb{P}(E_k^{i,x_i} \wedge E_k^i > t) = \exp\left(-\int_0^t b(x_i e^{-u}) du\right)$$

que es la distribución del primer spike de la neurona  $i$ , con condición inicial  $x_i$ , sin spikes de otras neuronas hasta el tiempo  $t$ . De esta manera, la afirmación de  $\mathcal{C}$  como *R.S.* nacerá de observar que cuando la neurona  $i$  emite un spike por su componente  $b(0)$  (no dependiente de la posición), la cadena de saltos “olvida” la posición  $x_i$ , y deja de depender de ella. Volveremos a utilizar una idea análoga en la Proposición 4.14.

Sea  $f \geq 0$  boreliana en  $\mathbb{R}_+^N$ . Si  $x \in \mathcal{C}$ . Entonces  $x_i \leq C_0$  y vale

$$\mathbb{E}_x(f(Y_1)) \geq \mathbb{E} \left( f(Y_1) \mathbf{1}_{\{E_1^i \leq \min_i E_1^{i, x_i} \wedge E_1^i\}} \right) \geq \mathbb{E}(f(Y_1) \mathbf{1}_{A_1})$$

con  $A_1 := \{E_1^i \leq \min_i E_1^{i, C_0} \wedge E_1^i\}$ . Observar que  $\mathbb{P}(A_1) > 0$ ; y que en  $A_1$  tenemos

$$Y_1 = (Y_{1,1}, Y_{1,2}, \dots, Y_{1,N}) = (0, x_2 e^{-E_1^1} + W_{12}, \dots) := Z_1$$

que no depende de  $x_1$ . De la misma manera, observando que  $Y_{1,i} \leq C_0 + K$  ( $\forall i$ )  $\Rightarrow$

$$\mathbb{E}_x(f(Y_2)) \geq \mathbb{E}(f(Z_2) \mathbf{1}_{A_1 \cap A_2})$$

donde  $A_2 := \{E_2^i \leq \min_i E_2^{i, C_0} \wedge E_2^i\}$ , evento con probabilidad positiva.

Siguiendo este razonamiento, obtendremos finalmente que

$$\mathbb{E}_x(f(Y_N)) \geq \mathbb{E}(f(Z_N) \mathbf{1}_{\mathcal{A}}) \quad \forall x \in \mathcal{C} \quad (4.3)$$

donde  $Z_N$  y  $\mathcal{A} := A_1 \cap A_2 \dots \cap A_N$  **NO** dependen de  $x \in \mathcal{C}$  y además  $\mathbb{P}(\mathcal{A}) > 0$ . Por lo tanto,  $(Y_n)$  cumple la condición característica de una *Harris-Cadena* ( ver Asmussen [3], página 198) con el conjunto  $\mathcal{C}$ ,  $r = N$ , y  $\lambda = \mathbb{P}(Z_N \in \cdot \cap \mathcal{A})$ .

Falta ver que  $\mathcal{C}$  es recurrente, y que existe la medida ergódica. Para ello, basta ver que si

$$T_0^+ := \inf \{t > 0 : X(t) \in \mathcal{C} \wedge (\exists u \leq t : X(u) \notin \mathcal{C})\}$$

es el momento de la primera vuelta a  $\mathcal{C}$ , entonces

$$\sup_{x \in \mathcal{C}} \mathbb{E}_x(T_0^+) < +\infty \quad (4.4)$$

(ver Asmussen [3], Teorema 3.2, página 200). Para  $x \in \mathcal{C}$ , el primer instante en que  $X(t)$  está en  $y \notin \mathcal{C}$ , se tiene que  $y_i \leq C_0 + K$ . Entonces, por la propiedad Fuerte de Markov

$$\mathbb{E}_x(T_0^+) \leq \mathbb{E}_x(\Gamma_{\mathcal{C}}) + \sup_{\substack{y \notin \mathcal{C} \\ \|y\|_{\infty} \leq C_0 + K}} \mathbb{E}_y(T_0) \stackrel{(*)}{\leq} R_0 + N(C_0 + K)$$

Donde

a)  $\Gamma_{\mathcal{C}} := \inf \{t > 0 : X(t) \notin \mathcal{C}\}$  el primer instante en que  $X(t)$  sale de  $\mathcal{C}$ .

b)  $(*)$  vale pues

- (4.3) implica que  $\mathbb{P}(Y_N \notin \mathcal{C}) \geq p_0 := \mathbb{P}(\{Z_N \in \mathcal{C}^c\} \cap \mathcal{A})$  entonces puedo mayorar  $\mathbb{E}(\Gamma_{\mathcal{C}})$  por

$$\mathbb{E} \left( \sum_{j=1}^{N \cdot \mathcal{G}} \mathcal{E}_j \right), \text{ con } \mathcal{G} \sim Ge(p_0) \text{ y } (\mathcal{E}_j) \sim \mathcal{E}(b(0)) \text{ i.i.d.}$$

- El segundo término de la suma se puede acotar por  $N(C_0 + K)$  haciendo uso directo del Lema 4.3.

□

## 4.2. ¿Cuándo muere la Red?

En la sección anterior vimos que, en caso de tener ausencia de ruido externo (matemáticamente, cuando  $b(0) = 0$ ), entonces *con probabilidad 1 la actividad neuronal muere* 4.4. Este resultado nos sugiere preguntas interesantes y además muy naturales que no están planteadas, al menos de modo explícito, en el trabajo de Robert y Touboul. Esencialmente la pregunta es: ¿Cuánto tarda en morir la Red Neuronal?

O de una forma levemente más formal ¿Cuál es la esperanza del tiempo de extinción de la red?.

Más aún, esto trae consigo otras preguntas: ¿Es un tiempo integrable? Y en caso afirmativo ¿Se puede calcular explícitamente, o al menos dar una cota? Además, ¿cómo depende de la dimensión?.

En esta sección, de producción original, queremos esbozar algunos resultados que apuntan a responder estas preguntas. Son casos sencillos, quizás no demasiado relevantes de manera directa desde el punto de vista biológico (donde las redes son de inmenso tamaño), pero sí representan pasos en esa dirección, con la idea de poder hacer más adelante estimaciones para casos más generales.

El abordaje de estas cuestiones es tanto teórico, como computacional. En este sentido, hicimos uso del programa Octave para realizar algunos algoritmos que permiten verificar las estimaciones, y también aportar intuición sobre qué resultados deberíamos esperar. Esto resulta especialmente útil cuando se maneja una gran cantidad de variables con una correlación que dificulta mucho (o vuelve virtualmente imposible) el cálculo explícito, o aún una estimación no trivial.

Enlistemos entonces una serie de suposiciones a priori:

- a) *Condición inicial simétrica:*  $X(0) = (x, \dots, x) \in \mathbb{R}_+^N$
- b) *Eficacias sinápticas constantes:*  $W \equiv c > 0$  a.e.
- c) *La tasa de descarga es la función identidad:*  $b(\cdot) = Id$
- d)  $N = 2$

Como regla general para los cálculos, no se hace uso de las hipótesis, al menos hasta que sea estrictamente necesario. Para facilitar la comprensión, será útil un poco de notación:

- $\tau_{ext}$  denotará el *instante del último spike del sistema*. Consideraremos esta variable como el *tiempo de extinción del sistema*. Nuestro objetivo es poder calcular explícitamente su esperanza, o al menos poder dar una cota aceptable.

- El instante del primer spike, para  $X_i(0) = x$ , se notará  $\tau_i^x$  y se define

$$\begin{cases} \tau_i^x := \inf\{t > 0 : X_i(t) = 0\} & (\text{si } x > 0) \\ \tau_i^0 := +\infty \end{cases}$$

- Dado  $X \in \mathbb{R}_+^N$ ,  $\tau^X := \min_i \tau_i^{x_i}$  es el instante del primer spike del sistema.

- Si  $Z$  es una v.a. , entonces notamos  $Z' := Z \cdot \mathbf{1}_{\{Z < +\infty\}}$

Hagamos algunas observaciones. Resultará muy útil contrastarlas con la figura 4.2, para entenderlas mejor. En ella se representa una trayectoria típica dadas las hipótesis que hemos asumido, realizada a partir de un algoritmo en Octave que simula la dinámica para  $N$  neuronas y eficacias sinápticas con distribución  $W \sim U[c, c + l]$ , en el caso particular en que  $l = 0, c = x = 1$  (en principio,  $x$  no tiene por qué valer igual a  $c$ ).

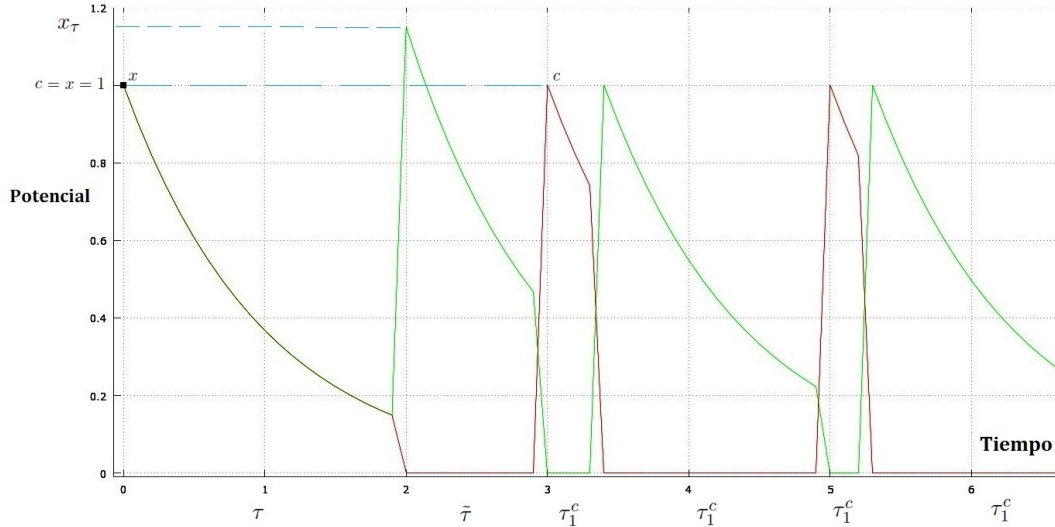


Figura 4.1: Una trayectoria típica de  $(X_1(t), X_2(t))$

En primer lugar, la distribución de  $(\tau_i^x)$  NO depende de  $i$ , sino sólo de  $x$ , y en consecuencia fijado  $x$  se tiene  $\tau_i^x \sim \tau_1^x$ . En este caso tan particular, tenemos 3 “tipos” de tiempos de espera para los spikes del sistema. El primer tiempo de espera tiene punto inicial simétrico, y es muy sencillo de calcular. Al primer instante de descarga del sistema  $\tau^{X(0)}$  lo denotaremos simplemente  $\tau$ .

A partir del segundo tiempo de espera, sólo tengo una neurona “viva”, y por las definiciones de  $\tau_i^x$  y  $\tau^X$ , los tiempos de espera del sistema serán de la forma  $\tau_1^{(\cdot)}$  (aquí también estamos usando implícitamente la Propiedad de Markov). El segundo tiempo de espera, depende entonces del valor de  $\tau$  y tiene la forma

$$\tilde{\tau} := \tau_1^{x_\tau}, \text{ donde } x_\tau := (xe^{-\tau} + c)\mathbf{1}_{\{\tau < +\infty\}} = xe^{-\tau} + c\mathbf{1}_{\{\tau < +\infty\}} \quad (4.5)$$

La idea es que el punto inicial para la segunda descarga  $x_\tau$  es una variable aleatoria que representa el estado post-salto de la neurona que “sobrevive” en la primer descarga del sistema.

Finalmente observemos que, tras el segundo salto, el proceso entra en una situación “estacionaria” en la que los tiempos de espera sólo dependen de los impulsos recibidos por las eficacias sinápticas (similar a lo que ocurre con la variable aleatoria  $Z_N$  en la demostración de 4.4[Parte b]<sup>1</sup>), que en este caso son constantes, y por lo tanto, todos son de la forma  $\tau_1^c$  i.i.d’s. Es pertinente aclarar que para llegar a esta última conclusión estamos usando fuertemente el hecho de tener sólo 2 neuronas, pues esto determina que, al emitir un spike alguna de ellas, sólo existe 1 posibilidad a lo sumo y es que eventualmente descargue la otra, y por lo tanto el tiempo de espera para el siguiente spike del sistema se identifica con el tiempo para la neurona “viva”. En un caso más general, desaparece esta restricción y tendremos que utilizar otra estrategia, probablemente una cota mediante un acoplamiento.

La idea es simplemente sumar estos tiempos de espera (quizás con alguna modificación técnica), de manera de obtener

$$\tau_{ext} = \tau + \tilde{\tau} + \sum \tau_1^x$$

Donde en el último término quiero sumar los tiempos de espera que me den un valor finito. Para poder realizar algo así necesitaremos algunos cuidados técnicos, dado que estas variables aleatorias tienen probabilidad positiva de tomar el valor  $+\infty$ . Esencialmente, si alguna de estos tiempos da infinito, queremos “retocarla” para que ella y todas las variables posteriores valgan cero.

Tenemos, al igual que en (3.8), y utilizando las hipótesis a y c

$$\mathbb{P}(\tau > t) = \exp\left(-Nx \int_0^t e^{-s} ds\right) = e^{Nx(e^{-t}-1)} \quad (4.6)$$

Y en particular

$$\mathbb{P}(\tau_1^x > t) = e^{x(e^{-t}-1)} \quad \text{y} \quad \mathbb{P}(\tau = +\infty) = e^{-Nx} = \mathbb{P}(\tau' = 0)$$

De lo que se desprende también que  $\mathbb{P}(\tau_1^x > 0) = 1$  y  $\mathbb{P}(\tau_1^x = +\infty) = e^{-x}$

Procedamos entonces a calcular la esperanza de  $\tau_{ext}$ . Para ello nos interesará la distribución de  $\tau'$ , pues en caso de tener  $\tau = +\infty$ , la situación es que ninguna neurona disparó, y en ese caso quisiéramos decir que el tiempo de extinción es 0. Dicha distribución vale:

$$\begin{aligned} \mathbb{P}(\tau' > t) &= \mathbb{P}(t < \tau < +\infty) = \mathbb{P}(\tau > t) - \mathbb{P}(\tau = +\infty) = e^{Nx(e^{-t}-1)} - e^{-Nx} \\ \mathbb{P}(\tau' > t) &= e^{-Nx}(e^{Nx e^{-t}} - 1) \end{aligned} \quad (4.7)$$

En relación al segundo tiempo de espera,  $\tilde{\tau}$ , recordemos que queremos que valga cero si ella (o  $\tau$ ) vale infinito. Por lo tanto el tiempo que necesitamos es

$$\tilde{\tau}' \mathbf{1}_{\{\tau < +\infty\}} = \tilde{\tau}'$$

Donde la igualdad vale porque, por la definición de  $\tilde{\tau}$ ,  $\{\tilde{\tau} < +\infty\} \subseteq \{\tau < +\infty\}$

En cuanto a los tiempos de espera tras el segundo spike, la idea es esperar en cada “ciclo” un tiempo  $\tau_1^c$ , siempre y cuando haya dado un valor finito; y dejar de sumar en cuanto obtenga  $+\infty$ . Este mecanismo puede pensarse así:

---

<sup>1</sup>Más aún, una idea similar aparece en la demostración de la primer proposición de la siguiente sección

- Tomo una v.a. geométrica  $G$  que me diga el instante del primer “éxito”:  $\tau_1^c = +\infty$ . Esto significa que  $G \sim Ge(p)$ ; con  $p = \mathbb{P}(\tau_1^c = +\infty) = e^{-c}$ .
- De esta manera,  $G(\omega) - 1$  me dice cuantos tiempos de espera *finitos a.e.* debo sumar. Para conseguir esto último vamos a considerar las variables aleatorias i.i.d.  $(\hat{\tau}_1^c(k))_{k \in \mathbb{N}}$  con ley  $\hat{\tau}_1^c$  independientes de  $G$ , donde

$$\begin{aligned} \mathbb{P}(\hat{\tau}_1^c > t) &:= \mathbb{P}(\tau_1^c > t | \tau_1^c < +\infty) = \frac{\mathbb{P}(\tau_1^c > t) - \mathbb{P}(\tau_1^c = +\infty)}{1 - \mathbb{P}(\tau_1^c = +\infty)} \\ &= \frac{e^{c(e^{-t}-1)} - e^{-c}}{1 - e^{-c}} \\ \mathbb{P}(\hat{\tau}_1^c > t) &= \frac{e^{ce^{-t}} - 1}{e^c - 1} \end{aligned} \quad (4.8)$$

Ahora sí, podemos afirmar que

$$\begin{aligned} \mathbb{E}(\tau_{ext}) &= \mathbb{E}\left(\tau' + \hat{\tau}' + \left[\sum_{k=1}^{G-1} \hat{\tau}_1^c(k)\right] \mathbf{1}_{\{\tau < +\infty\}}\right) \\ &= \mathbb{E}(\tau') + \mathbb{E}(\hat{\tau}') + \mathbb{E}\left(\left[\sum_{k=1}^{G-1} \hat{\tau}_1^c(k)\right]; \tau < +\infty\right) \end{aligned} \quad (4.9)$$

Observemos que el último termino se puede reescribir así

$$\begin{aligned} \mathbb{E}\left(\left[\sum_{k=1}^{G-1} \hat{\tau}_1^c(k)\right]; \tau < +\infty\right) &= \mathbb{E}\left(\mathbb{E}\left(\left[\sum_{k=1}^{G-1} \hat{\tau}_1^c(k)\right] \mathbf{1}_{\{\tau < +\infty\}} \middle| \tau\right)\right) \\ &= \mathbb{E}\left(\sum_{k=1}^{G-1} \hat{\tau}_1^c(k)\right) \mathbb{P}(\tau < +\infty) \end{aligned} \quad (4.10)$$

La última igualdad vale por propiedades básicas de la esperanza condicional, utilizando que  $(\hat{\tau}_1^c(k))$ ,  $G$  son independientes de  $\tau$ , y que ésta última es - trivialmente-  $\sigma(\tau)$ -medible.

Por ahora nos vamos a concentrar en calcular explícitamente la cantidad  $\mathbb{E}\left(\sum_{k=1}^{G-1} \hat{\tau}_1^c(k)\right)$ , que tiene interés en sí mismo, puesto que si consideramos la suposición

$$a') X(0) = (c, 0)$$

entonces  $\mathbb{E}(\tau_{ext}) = \mathbb{E}\left(\sum_{k=1}^{G-1} \hat{\tau}_1^c(k)\right)$ . Calculemos entonces,

$$\mathbb{E}\left(\sum_{k=1}^{G-1} \hat{\tau}_1^c(k)\right) \stackrel{*}{=} \mathbb{E}(G-1)\mathbb{E}(\hat{\tau}_1^c) = (e^c - 1)\mathbb{E}(\hat{\tau}_1^c)$$

Donde \* vale porque  $G$  es independiente de  $(\hat{\tau}_1^c(k))$ , condicionando sobre  $G$ . Finalmente,

$$\mathbb{E}(\hat{\tau}_1^c) = \int_0^{+\infty} \mathbb{P}(\hat{\tau}_1^c > t) dt = \frac{1}{e^c - 1} \int_0^{+\infty} e^{ce^{-t}} dt = \frac{1}{e^c - 1} \int_0^{+\infty} \frac{e^a - 1}{a} da$$



Por lo tanto tenemos que

$$\mathbb{E} \left( \sum_{k=1}^{G-1} \hat{\tau}_1^c(k) \right) = \int_0^{+\infty} \frac{e^a - 1}{a} da = \sum_{k=1}^{+\infty} \frac{c^k}{k!k} \leq e^c - 1 \quad (4.11)$$

La integral no tiene un resultado cerrado propiamente dicho. Se puede presentar como  $\lim_{\varepsilon \rightarrow 0^+} (Ei(a) - \log(a))|_{\varepsilon}^c$ , donde  $Ei(x) := \int_{-\infty}^x \frac{e^t}{t} dt$  es la *integral exponencial*. Por ejemplo, para  $c = 1$  se tiene, numéricamente  $\mathbb{E}(\tau_{ext}) \approx 1,3179$ . En el apéndice, las figuras 7.1 y 7.2 muestran un histograma de tiempos de extinción, y un histograma de la media muestral del tiempo de extinción (el promedio de  $n$  variables aleatorias i.i.d., en este caso  $\tau_{ext}$ ) respectivamente; para  $c = 1$ , realizados en Octave.

Volvamos a la hipótesis a. Nos falta calcular entonces  $\mathbb{E}(\tau')$  y  $\mathbb{E}(\tilde{\tau}')$ .

$$\begin{aligned} \mathbb{E}(\tau') &= \int_0^{+\infty} \mathbb{P}(\tau' > t) dt = \frac{1}{e^{Nx}} \int_0^{+\infty} e^{Nxe^{-t}} - 1 dt \\ &= \frac{1}{e^{Nx}} \int_0^{Nx} \frac{e^a - 1}{a} da \stackrel{*}{=} \frac{1}{e^2} \int_0^2 \frac{e^a - 1}{a} da \approx 0,4986 \end{aligned} \quad (4.12)$$

Donde en \* usamos  $N = 2$  y elegimos adicionalmente  $x = 1$ . Con estas mismas suposiciones, y llamando  $\tau_1^{x+c} \mathbf{1}_{\{\tau_1^{x+c} < +\infty\}} := \tilde{\tau}$ :

$$\begin{aligned} \mathbb{E}(\tilde{\tau}') &\leq \mathbb{E}(\tilde{\tau}) = \int_0^{+\infty} (1 - F_{\tilde{\tau}}) dt = \int_0^{+\infty} e^{-(x+c)} (e^{(x+c)e^{-t}} - 1) dt \\ &\stackrel{x+c=2}{=} \frac{1}{e^2} \int_0^2 \frac{e^a - 1}{a} da \approx 0,4986 \end{aligned} \quad (4.13)$$

De esta manera,

$$\mathbb{E}(\tau_{ext}) \leq \mathbb{E}(\tau') + \mathbb{E}(\tilde{\tau}) + \mathbb{E} \left( \sum_{k=1}^{G-1} \hat{\tau}_1^c(k) \right) \approx 2(0,4986) + 1,3179 = 2,3151$$

Al igual que hicimos en el caso más simple, en las figuras 7.3 y 7.4 se muestran histogramas análogos realizados en Octave.

Estos son algunos cálculos básicos que emprendimos; quedan preguntas pendientes, en función de las cuales algunos algoritmos y cálculos se han realizado, y no se incluyen en esta tesis, a la espera de poder corroborar su validez con más tiempo y cuidado, antes de presentarlos. Entre ellos, por ejemplo, se tiene un algoritmo para calcular el estado de la red a tiempos grandes. A partir de este cómputo, se podría obtener información del estado de la red, cuando ésta todavía no murió; que puede aportar intuición para el estudio de las posibles distribuciones cuasi-estacionarias.

### 4.3. El Proceso independiente del Estado

El caso de la tasa de descarga  $b(x)$  constante cae en el segundo caso del teorema 4.4 que demostramos más arriba. Resulta ser muy ilustrativo, pues podemos dar explícitamente la ley

de su única medida invariante; y, como comentan Robert y Touboul en [43] es muy informativo pues en este caso particular no ocurre explosión, es decir, no se acumulan infinitos spikes en finito tiempo. Desde el punto de vista biológico es muy importante que esto no ocurra en el modelo, pues no es lo que sucede en el fenómeno real, y los modelos **(LIF)** en general no logran resolver (recordar [12] por ejemplo). Más aún, por comparación con este caso, se pueden probar resultados similares para casos más generales [43].

**Proposición 4.5** (Distribución invariante (Caso  $b(x) = b > 0$ )-(R-T)). *Si la tasa de disparo es constantemente  $b > 0$ , entonces la distribución invariante de  $(X(t))$  viene dada por la ley del vector  $X = (X_1, X_2, \dots, X_N)$  con:*

$$X_i = \sum_{j \neq i} \sum_{k \geq 1} W_{ji}^k e^{-T_j^k} \mathbf{1}_{\{T_j^k \leq T_i^1\}} \quad i \in \mathbb{I}_N \quad (4.14)$$

Donde  $W_{ji}^k \sim W$  i.i.d.  $\forall i, j, k$  y  $(T_j^k)_{k \in \mathbb{N}} \sim \mathcal{PP}_{\mathbb{R}_{\geq 0}}(b)$  independientes ( $j \in \mathbb{I}_N$ )

**Demostración:** La siguiente prueba, basada en la presentada por Robert y Touboul, se apoya en el argumento de *Acoplamiento desde el Pasado*. Una buena referencia para este tipo de argumento se encuentra en “Markov Chains and Mixing Times” de Levin et al. [28]. En esencia, este tipo de argumento supone que el proceso se viene corriendo desde el infinito, o desde hace mucho tiempo, lo suficiente como para asegurar que éste ya alcanzó en el presente la medida estacionaria. Luego busca analizar la dinámica del proceso particular se tiene entre manos, para determinar cuál es el estado presente que ya está distribuido como la estacionaria.

Sean, para cada  $j \in \mathbb{I}_N$ ,  $\mathcal{N}_j = (\tau_j^k)_{k \in \mathbb{Z}}$  un proceso de Poisson en  $\mathbb{R}$  con tasa  $b$  (independientes) representando los tiempos en que se dispara la neurona  $j$ ; y tal que  $\tau_j^{-1}$  es el último instante antes de  $t = 0$ .

Sea  $T := \inf_j \{\tau_j^{-1}\}$ . Observar que, por la homogeneidad del Proceso de Poisson, un proceso  $\hat{X}$  con la misma dinámica y distribución inicial  $\hat{X}(0) \sim X(T)$  verifica que  $\hat{X}(-T) \sim X(0)$ . Observe también que para  $t \in (\tau_i^{-1}, 0)$ ,  $X(t)$  sólo depende de los spikes de las demás neuronas. En el instante  $\tau_i^{-1}$  podemos decir que el proceso *colapsa* todas las trayectorias en 0 para la coordenada  $i$ . Esto implica que  $X(0)$  no depende de  $T$ , y por lo tanto  $X(0) \sim X$  la distribución invariante.

Ahora, si la neurona  $j$  se dispara en el instante  $s \in (\tau_i^{-1}, 0)$  entonces su contribución a la neurona  $i$  es  $W_{ji}e^s$ . Llamando  $T_j^k := -\tau_j^{-k}$ ,  $k \in \mathbb{N}$  y entonces  $(T_j^k)_k \sim \mathcal{PP}_{\mathbb{R}_{\geq 0}}(b)$  y  $X$  tiene la ley 4.14.  $\square$

**Proposición 4.6** (Transformada de Laplace de la distribución invariante (R-T)). *Si  $b(x) \equiv b > 0$ , la transformada de Laplace de una neurona en el equilibrio,  $\tilde{X}_1(\xi) := \mathbb{E}(e^{-\xi X_1})$  viene dada por la ecuación*

$$\tilde{X}_1(\xi) = \int_0^{+\infty} \exp\left(-b(N-1) \int_0^x (1 - \tilde{W}(\xi e^{-u})) du\right) b e^{-bx} dx \quad (4.15)$$

**Demostración:** Para simplificar notación, llamemos  $T := T_1^1$ . Siguiendo la notación de la proposición anterior,  $\mathcal{M} := \mathcal{N}_2 + \dots + \mathcal{N}_N = (s_n)$  es un proceso de Poisson homogéneo de

tasa  $b(N - 1)$

$$\tilde{X}_1(\xi) = \mathbb{E} \left( \mathbb{E} \left( \prod_{n \geq 1} e^{-\xi W_n e^{-s_n} \mathbf{1}_{\{s_n \leq T\}}} \middle| (s_n), T \right) \right)$$

Por independencia condicional, esto es igual a

$$\begin{aligned} &= \mathbb{E} \left( \prod_{n \geq 1} \mathbb{E} \left( e^{-\xi W_n e^{-s_n} \mathbf{1}_{\{s_n \leq T\}}} \middle| (s_n), T \right) \right) = \mathbb{E} \left( \prod_{n \geq 1} \left[ \mathbb{E} \left( e^{-\xi W_n e^{-s_n} \mathbf{1}_{\{s_n \leq T\}}} \right) \right] \right) \\ &= \mathbb{E} \left( \prod_{s_n \leq T} \left[ \mathbb{E} \left( e^{-\xi W_n e^{-s_n}} \right) \right] \right) \end{aligned}$$

Ahora, si definimos  $g(u) := -\log(\tilde{W}(\xi e^{-u})) \equiv e^{-g(u)} = \tilde{W}(\xi e^{-u})$ , entonces se tiene

$$\begin{aligned} \tilde{X}_1(\xi) &= \mathbb{E} \left( \left[ e^{-\sum_{s_n \leq T} g(s_n)} \right] \right) \\ &= \mathbb{E} \left( e^{-\int_0^T g(u) \mathcal{M}(du)} \right) = \int_0^{+\infty} \mathbb{E} \left( e^{-\int_0^t g(u) \mathcal{M}(du)} \right) b e^{-bt} dt \\ &= \int_0^{+\infty} \exp \left( -b(N-1) \int_0^{+\infty} (1 - e^{-g(u)}) du \right) b e^{-bt} dt \end{aligned}$$

Donde la tercera igualdad es por ser  $T \sim \mathcal{E}(b)$ ; y la cuarta igualdad es un resultado conocido, cuya demostración puede verse en [42]. Queda demostrada la proposición.  $\square$

Como consecuencias de este resultado, obtuvimos

1.  $\mathbb{P}(X_1 = 0) = \frac{1}{N}$  (La probabilidad que la Neurona 1 sea la última en dispararse)
2.  $\mathbb{E}(X_1) = \frac{b}{b+1}(N-1)\mathbb{E}(W)$

**Demostración:** En efecto, el primer resultado se obtiene a partir del siguiente cálculo general para cualquier Transformada de Laplace

$$\begin{aligned} \widehat{X}_1(\xi) &= \mathbb{E} \left( e^{-\xi X_1} \right) = \mathbb{E} \left( e^{-\xi X_1}; X_1 = 0 \right) + \mathbb{E} \left( e^{-\xi X_1}; X_1 > 0 \right) \\ &= \mathbb{P}(X_1 = 0) + \mathbb{E} \left( e^{-\xi X_1}; X_1 > 0 \right) \rightarrow \mathbb{P}(X_1 = 0) \quad (\xi \rightarrow +\infty) \end{aligned} \quad (4.16)$$

Entonces

$$\begin{aligned}\mathbb{P}(X_1 = 0) &= \int_0^{+\infty} \exp\left(-b(N-1) \int_0^x (\mathbb{P}(W > 0)) du\right) b e^{-bx} dx \\ &= b \int_0^{+\infty} \exp(-bNx) dx = \frac{1}{N}\end{aligned}\quad (4.17)$$

El segundo resultado proviene de considerar que  $\left.\frac{\partial \widehat{X}}{\partial \xi}(\xi)\right|_0 = -\mathbb{E}(X)$  y manipulando entonces (4.15) de manera análoga al caso anterior.  $\square$

La siguiente proposición, enunciada en [43], tiene esbozada la idea de la demostración, que nosotros demostramos de manera completa.

**Proposición 4.7** (Comportamiento asintótico de una Red Neuronal Estacionaria (R-T)). *Si para cada  $N \in \mathbb{N}$  el valor de cada spike tiene distribución  $W^N := V/N$  para alguna v.a.  $V \in L^1$  y  $(X_i^N)_{i \in \mathbb{I}_N}$  es el vector con la distribución estacionaria de la Red de  $N$  neuronas; entonces:*

1.  $X_1^N \xrightarrow{\mathcal{D}} X^\infty$  ( $N \rightarrow +\infty$ ) donde  $X^\infty$  tiene la siguiente distribución:

$$f_\infty(u) = \frac{1}{\mathbb{E}(V)} \left(1 - \frac{u}{b \mathbb{E}(V)}\right)^{b-1} \quad \text{para } u \in [0, b\mathbb{E}(V)]$$

2. Para cada par  $i \neq j \in \mathbb{I}_N$ , las variables aleatorias  $X_i^N$  y  $X_j^N$  son asintóticamente independientes cuando  $N$  tiende a infinito.

**Demostración:** Demostraremos ambas afirmaciones de manera simultánea.

Sean  $(V_{rm}^k) \sim V$  i.i.d. y  $\mathcal{N}_m = (T_m^k)_{k \in \mathbb{N}} \sim \mathcal{PP}_{\mathbb{R}_+}(b)$  como en la proposición anterior. Para simplificar notación, llamemos  $T_m := T_m^1$ . Por simetría, podemos considerar  $(X_1^N)$  y  $(X_2^N)$ . Para  $m = 1, 2$  y  $\tilde{m} := 3 - m$ , sea

$$Y_m^N := \sum_{r \geq 3} \sum_{k \geq 1} \frac{1}{N} V_{rm}^k e^{-T_r^k} \mathbf{1}_{\{T_r^k \leq T_m\}} = X_m^N - \sum_{k \geq 1} \frac{1}{N} V_{\tilde{m}m}^k e^{-T_m^k} \mathbf{1}_{\{T_m^k \leq T_m\}} \quad (4.18)$$

Observemos que así definidas, la distribución de  $(Y_1^N, Y_2^N)$  es arbitrariamente cercana a la de  $(X_1^N, X_2^N)$ ; más aún  $(X_1^N, X_2^N) - (Y_1^N, Y_2^N) \xrightarrow{c.s.} \mathbf{0}$

Procediendo de manera análoga a la proposición anterior, sea  $\nu := \mathcal{N}_3 + \mathcal{N}_4 + \dots + \mathcal{N}_N := (s_n) \sim \mathcal{PP}_{\mathbb{R}_+}(b(N-2))$ . Entonces podemos escribir:

$$e^{-\xi_1 Y_1^N} = \exp\left(-\frac{\xi_1}{N} \sum_{n \geq 1} V_n e^{-s_n} \mathbf{1}_{\{s_n \leq T_1\}}\right) \quad \wedge \quad e^{-\xi_2 Y_2^N} = \exp\left(-\frac{\xi_2}{N} \sum_{n \geq 1} \hat{V}_n e^{-s_n} \mathbf{1}_{\{s_n \leq T_1\}}\right)$$

De esta manera

$$\begin{aligned}
\mathbb{E} \left( e^{-\xi_1 Y_1^N - \xi_2 Y_2^N} \right) &= \mathbb{E} \left( \mathbb{E} \left( \prod_{n \geq 1} e^{-\frac{\xi_1}{N} V_n e^{-s_n} \mathbf{1}_{\{s_n \leq T_1\}}} e^{-\frac{\xi_2}{N} \hat{V}_n e^{-s_n} \mathbf{1}_{\{s_n \leq T_2\}}} \middle| (s_n), T_1, T_2 \right) \right) \\
&= \mathbb{E} \left( \prod_{n \geq 1} \mathbb{E} \left( e^{-\frac{\xi_1}{N} V_n e^{-s_n} \mathbf{1}_{\{s_n \leq T_1\}}} \middle| (s_n), T_1 \right) \prod_{n \geq 1} \mathbb{E} \left( e^{-\frac{\xi_2}{N} \hat{V}_n e^{-s_n} \mathbf{1}_{\{s_n \leq T_2\}}} \middle| (s_n), T_2 \right) \right) \\
&= \mathbb{E} \left( \prod_{n \geq 1} \tilde{V} \left( \frac{\xi_1}{N} e^{-s_n} \mathbf{1}_{\{s_n \leq T_1\}} \right) \prod_{n \geq 1} \tilde{V} \left( \frac{\xi_2}{N} e^{-s_n} \mathbf{1}_{\{s_n \leq T_2\}} \right) \right)
\end{aligned}$$

Al igual que en 4.6 definimos  $e^{-g_m(u)} := \tilde{V} \left( \frac{\xi_m}{N} e^{-u} \mathbf{1}_{\{u \leq T_m\}} \right)$   $m = 1, 2$ . Entonces se tiene

$$\begin{aligned}
\mathbb{E} \left( e^{-\xi_1 Y_1^N - \xi_2 Y_2^N} \right) &= \mathbb{E} \left( e^{-\sum_{n \geq 1} g_1(s_n) + g_2(s_n)} \right) = \mathbb{E} \left( e^{-\int_0^{+\infty} g_1(u) + g_2(u) \nu(du)} \right) \\
&= \mathbb{E} \left( \exp \left( -b(N-2) \int_0^{+\infty} \left[ 1 - \tilde{V} \left( \frac{\xi_1}{N} e^{-u} \mathbf{1}_{\{u \leq T_1\}} \right) \tilde{V} \left( \frac{\xi_2}{N} e^{-u} \mathbf{1}_{\{u \leq T_2\}} \right) \right] du \right) \right) \quad (4.19)
\end{aligned}$$

Donde en la última igualdad utilizamos nuevamente la proposición 1.5 de Robert [42] con  $f = g_1 + g_2$  y medida de intensidad  $\mu = b(N-2) | \cdot |$ , siendo  $| \cdot |$  la medida de Lebesgue; y  $\tilde{V}(s) = \mathbb{E}(e^{-sV})$

Ahora, como  $\tilde{V}(s) = 1 - s\mathbb{E}(V) + o(s)$  ( $s \rightarrow 0$ ); entonces

$$\left( 1 - \tilde{V} \left( \frac{\xi_1}{N} e^{-u} \mathbf{1}_{\{u \leq T_1\}} \right) \tilde{V} \left( \frac{\xi_2}{N} e^{-u} \mathbf{1}_{\{u \leq T_2\}} \right) \right) \simeq \frac{\mathbb{E}(V)}{N} e^{-u} (\xi_1 \mathbf{1}_{\{u \leq T_1\}} + \xi_2 \mathbf{1}_{\{u \leq T_2\}})$$

A medida que  $N$  tiende a  $+\infty$ .

Por lo tanto,

$$\begin{aligned}
&\lim_{N \rightarrow +\infty} \mathbb{E} \left( e^{-\xi_1 Y_1^N - \xi_2 Y_2^N} \right) = \\
&\lim_{N \rightarrow +\infty} \mathbb{E} \left( \exp \left( -b(N-2) \int_0^{+\infty} \left[ \frac{\mathbb{E}(V)}{N} e^{-u} (\xi_1 \mathbf{1}_{\{u \leq T_1\}} + \xi_2 \mathbf{1}_{\{u \leq T_2\}}) \right] du \right) \right) \\
&\lim_{N \rightarrow +\infty} \mathbb{E} \left( \exp \left[ -b\mathbb{E}(V) \frac{(N-2)}{N} \xi_1 (1 - e^{-T_1}) \right] \right) \mathbb{E} \left( \exp \left[ -b\mathbb{E}(V) \frac{(N-2)}{N} \xi_2 (1 - e^{-T_2}) \right] \right) \\
&= H(\xi_1) H(\xi_2) \quad (4.20)
\end{aligned}$$

Con

$$H(\xi) = \mathbb{E} \left( \exp \left[ -b\mathbb{E}(V) \xi (1 - e^{-T_1}) \right] \right)$$

Observar que en la última igualdad usamos que  $T_1 = T_1^1$  y  $T_2 = T_2^1$  son i.i.d. (con

distribución exponencial de tasa  $b$ ). Por último,

$$\begin{aligned} H(\xi) &= \int_0^{+\infty} \exp[-b\mathbb{E}(V)\xi(1 - e^{-t})] b e^{-bt} dt \\ &= \int_0^{b\mathbb{E}(V)} e^{-\xi x} \frac{1}{\mathbb{E}(V)} \left(1 - \frac{x}{b\mathbb{E}(V)}\right)^{b-1} dx \end{aligned} \quad (4.21)$$

haciendo el cambio de variable  $x = b\mathbb{E}(V)(1 - e^{-t})$ .

Con ello queda demostrada la independencia asintótica, con la densidad dada en la proposición.  $\square$

Como hemos anticipado antes, este es un primer resultado en la dirección de campo-medio; y que da intuición sobre los resultados que pueden surgir al considerar redes neuronales con dimensión acercándose a infinito. Observar que el reescalamiento que nos proporcionó un límite no trivial fue el de orden  $1/N$ . Esto permite sospechar, como efectivamente ocurrirá, que ése será el reescalamiento que puede aportar resultados de campo-medio en casos más generales.



## Capítulo 5

# Redes Neuronales Infinitas: Resultados Asintóticos

Como se observó en la introducción, la población de neuronas se cuenta de a cientos de miles en pocos milímetros en la corteza cerebral. Esto hace relevante el estudio del régimen asintótico del proceso regido por el modelo que presentamos.

Al respecto, lo que ocurre con este modelo es que bajo ciertas condiciones sobre el crecimiento de la función  $b(x)$  - que determina la tasa de disparo- y escalando correctamente los spikes ( $W_{ij}$ ) que generan las neuronas; se obtiene un resultado de tipo **campo-medio**.

Los resultados de campo-medio surgen en el estudio de sistemas dinámicos con un gran número de *componentes o nodos* (neuronas, partículas, etc.) interactuando entre sí, pudiendo haber eventualmente alguna interacción con el exterior de ese sistema y cuya dinámica quiere conocerse. A medida que el número de componentes crece, resolver ese sistema teniendo en cuenta cada interacción individual puede volverse extremadamente complejo. Es por esto que la teoría de campos medios transforma el problema N-dimensional a uno de dimensión 1, considerando que la interacción que recibe el nodo es una suerte de *campo promedio*, más que el conjunto de todas las interacciones con el resto de los nodos del sistema.

### 5.1. Existencia y unicidad de Proceso límite:

De esta manera, consideramos para  $N$  grande que las variables aleatorias ( $W_{ij}^N$ ) tienen la distribución de  $(V_{ij}/N)$  con los  $V_{ij} \sim V$  i.i.d. (al igual que en la Proposición 4.7 del Capítulo anterior). Así, tenemos la ecuación (1.2). Esta ecuación diferencial estocástica tiende en distribución, bajo ciertas condiciones, al siguiente Proceso de McKean-Vlasov:

$$dY(t) = -Y(t)dt + \mathbb{E}(V)\mathbb{E}(b(Y(t)))dt - Y(t^-)\mathcal{N}_{b(Y(t^-))}(dt) \quad (5.1)$$

No es trivial que esta ecuación define efectivamente un único proceso. Este es un resultado probado en Robert y Touboul [43] formulado de la siguiente manera:

**Teorema 5.1.** *Si  $b : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  es una función no-decreciente, no acotada,  $C^1$  y tal que*



$\forall x \geq 0$

$$b'(x) \leq \alpha b(x) + c \quad \text{con } 0 \leq \alpha \leq \frac{1}{3\mathbb{E}(V)} \quad \text{y } c > 0 \quad (5.2)$$

Entonces  $\forall T, x_0 > 0$  existe un único Proceso estocástico càdlàg  $(Z(t))_{t \leq T}$  satisfaciendo la ecuación (5.1), con condición inicial  $Z(0) = x_0$

## 5.2. Convergencia al Proceso de McKean-Vlasov:

Recordemos que, para  $N$  neuronas, la dinámica está dada por:

$$\begin{aligned} dX_i^N(t) = & -X_i^N(t)dt + \frac{1}{N} \sum_{j \neq i} \int_{\mathbb{R}_+^2} z_i \mathbb{1}_{\{0 \leq u \leq b(X_j^N(t-))\}} \mathcal{N}_j(du, dz, dt) \\ & - X_i^N(t^-) \int_{\mathbb{R}_+^2} \mathbb{1}_{\{0 \leq u \leq b(X_i^N(t-))\}} \mathcal{N}_i(du, dz, dt) \end{aligned} \quad (5.3)$$

Donde los  $(\mathcal{N}_j)$  son procesos de Poisson independientes con medida de intensidad dada por  $du \otimes V(dz) \otimes dt$ , con  $V(dz)$  la distribución de  $V$ .

Con este panorama, para probar la Convergencia al Proceso de McKean-Vlasov Robert y Touboul demuestran que bajo condiciones apropiadas - a desarrollar más abajo - la *distribución empírica*  $(\Lambda_N(t))$  converge en distribución a la distribución de dicho proceso (5.1), donde

$$\langle \Lambda_N(t), \phi \rangle := \frac{1}{N} \sum_{i=1}^N \phi(X_i^N(t)) \quad \forall \phi \text{ continua en } \mathbb{R}_+ \quad (5.4)$$

Para ello se asume lo siguiente:

### Suposiciones MF:

(a) *Condición de Crecimiento de b:*

$b: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  es una función no-decreciente, no acotada,  $C^1$  y tal que  $\forall x \geq 0$  se cumple 5.2

(b) *Soporte Acotado de  $V(dx)$ :*

$\exists K > 0: V([0, K_V]) = 1 \equiv V \leq K_V$  a.e.

(c) *Sobre la Condición inicial:*

Las variables aleatorias  $(X_i^N(0))_{i \in \mathbb{N}}$  son i.i.d. con ley  $L_0$  de soporte acotado.

### Observación 5.2.

1. La condición (a) implica un lento crecimiento en el infinito, i.e.  $\forall a > 0, \exists x_a, R_a: \frac{b(x+a)}{b(x)} \leq R_a$  si  $x \geq x_a$ .
2. Se asume  $b(0) = 0$  para probar convergencia, puesto que el caso  $b(0) > 0$  resulta más sencillo para analizar desde el punto de vista de Campo-Medio, pues las neuronas se “actualizan” a una tasa mínima  $b(0)$  independiente de los estados; y por lo tanto hay un tiempo maximal entre 2 spikes de un nodo, también independiente del estado.

La estrategia para demostrar esto es la siguiente:

En primer lugar, a partir de la representación del proceso como Ecuación Diferencial estocástica - y haciendo uso de un resultado sobre Martingalas locales asociadas a Procesos de Poisson Marcados, cf. [43, Proposición 13] - se obtiene una escritura de la medida empírica de la forma

$$\begin{aligned}
\langle \Lambda_N(t), f \rangle &= \langle \Lambda_N(0), f \rangle - \int_0^t \langle \Lambda_N(u), (\cdot) f'(\cdot) \rangle du \\
&+ N \int_0^t \int_{\mathbb{R}_+} \left\langle \Lambda_N(u), \left( f \left( \cdot + \frac{v}{N} \right) - f(\cdot) \right) \right\rangle \langle \Lambda_N(u), b \rangle V(dv) du \\
&- \int_0^t \int_{\mathbb{R}_+} \left\langle \Lambda_N(u), \left( f \left( \cdot + \frac{v}{N} \right) - f(\cdot) \right) b(\cdot) \right\rangle V(dv) du \\
&- \int_{\mathbb{R}_+} \langle \Lambda_N(u), (f(0) - f(\cdot)) b(\cdot) \rangle du + M_f^N(t)
\end{aligned} \tag{5.5}$$

con  $M_f^N(t)$  con su proceso previsible asociado conocido.

Luego se prueba una estimación para el primer momento escalado de la tasa de disparo  $b$ : se muestra que con alta probabilidad, la tasa total de disparo del sistema (reescalada) permanece, para cualquier intervalo finito, acotada en norma infinito. Más precisamente, para  $k = 1$  y para algún  $k > 1$  se verifica

$$\lim_{N \rightarrow +\infty} \mathbb{P}(\|\langle \Lambda_N(t), b^k \rangle\|_\infty^T \geq C_0) = 0 \quad \text{con} \|F\|_\infty^T := \sup_{t \in [0, T]} |F(t)| \tag{5.6}$$

Los métodos empleados en la demostración de este resultado (ver [43, Proposición 5]), se pueden adaptar para obtener que para toda función  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ ,  $C^1$  con soporte compacto, se verifica que  $\mathbb{E}((M_f^N(t))^2) \sim O(1/N)$ , y en particular se tiene que  $M_f^N(t)$  converge a 0 en distribución para la norma infinito en intervalos finitos de tiempo. Esto, junto a (5.6) permitirán demostrar que la secuencia  $(\Lambda_N(t))$  es tight.

La demostración de *tightness* de  $(\Lambda_N(t))$  hace uso del Teorema 3.7.1 de Dawson [16], que dice que basta con probar, para cualquier  $f$ ,  $C^1$  con soporte compacto la secuencia  $(\langle \Lambda_N(t), f \rangle)$  es tight con la topología de la norma infinito sobre conjuntos compactos. Esto último se lleva a cabo vía los módulos de continuidad de  $(\langle \Lambda_N, f \rangle)$ :

$$w_{\Lambda_N, f}(\delta) := \sup\{|\langle \Lambda_N(s), f \rangle - \langle \Lambda_N(s'), f \rangle| : 0 \leq s \leq s' \leq t, |s - s'| < \delta\}$$

Finalmente, haciendo uso de la ecuación (5.5) y de un lema -que establece que para cualquier límite  $(\Lambda(t))$  de una subsucesión  $(\Lambda_{N_k}(t))$  se verifica  $\lim_{k \rightarrow +\infty} (\langle \Lambda_{N_k}(t), b \rangle) = (\langle \Lambda(t), b \rangle)$  - se prueba un resultado de unicidad del proceso límite  $(\Lambda(t))$  que da como consecuencia la convergencia al Proceso de McKean - Vlasov (5.1)

### 5.3. Análisis de las distribuciones invariantes

**Notación:** En lo siguiente  $\pi$  denotará una distribución en  $\mathbb{R}_+$  invariante para el proceso de McKean-Vlasov.

Observemos que si  $X(0) \sim \pi$ , entonces  $X_t \sim \pi$ , y por lo tanto  $t \rightarrow \mathbb{E}(b(X_t))$  es una aplicación constante. Luego, si  $\alpha := \mathbb{E}(V)\mathbb{E}(b(X(0)))$  y  $(X(t))$  es el proceso estacionario, entonces es solución de

$$dY(t) = (\alpha - Y(t))dt - Y(t^-)\mathcal{N}_{b(Y(t^-))}(dt) \quad (5.7)$$

Llamemos  $\tau := \inf\{t > 0 : Y(t) = 0 \wedge (\exists u \leq t : Y(u) \neq 0)\}$  *el primer instante de spike de la neurona*. Si  $Y(0) = 0$ , entonces para  $t < \tau$ ,  $(Y(t))$  resulta determinístico y se comporta como la solución de la EDO

$$\begin{cases} x'(t) = \alpha - x(t) \\ x(0) = 0 \end{cases}$$

Es decir, para  $t < \tau$ ,  $Y(t) = \alpha(1 - e^{-t})$ . Más aún,  $(Y(t))$  cae en el marco de los *Procesos de Renovación*, y por lo tanto podemos hacer uso de sus resultados. Observemos también que  $\tau$  es el tiempo de espera de un proceso de Poisson no homogéneo de tasa  $b(Y(t))$ , i.e.  $\mathbb{P}(\tau > t) = \exp(-\int_0^t b(x(u)) du)$ .

Entonces, la distribución invariante (de existir) puede expresarse así

$$\pi(f) = \frac{1}{\mathbb{E}(\tau)} \mathbb{E} \left( \int_0^\tau f(x(u)) du \right) \quad (5.8)$$

para toda función  $f$  continua con soporte compacto en  $[0, \alpha)$ . Por el teorema de Fubini

$$\begin{aligned} \mathbb{E} \left( \int_0^\tau f(x(u)) du \right) &= \int_0^\infty f(x(u)) \mathbb{P}(\tau > u) du \\ &= \int_0^\infty f(x(u)) \exp \left( - \int_0^u b(x(a)) da \right) du \\ &\stackrel{(*)}{=} \int_0^\alpha \frac{f(v)}{\alpha - v} \exp \left( - \int_0^v \frac{b(s)}{\alpha - s} ds \right) dv \end{aligned} \quad (5.9)$$

Donde en (\*) usamos el cambio de variable  $v = x(u)$  y  $s = x(a)$ .

$\pi$  resulta una distribución si y sólo si

$$\mathbb{E}(\tau) = \int_0^\alpha \frac{1}{\alpha - v} \exp \left( - \int_0^v \frac{b(s)}{\alpha - s} ds \right) dv < +\infty \quad (5.10)$$

Además, tiene soporte compacto en  $[0, \alpha)$ . Por lo tanto estamos en condiciones de enunciar el siguiente teorema, enunciado por (R-T) y demostrado por nosotros:

**Teorema 5.3.** *[Distribución invariante del Proceso de McKean-Vlasov] Si  $C(\beta) < +\infty$ , la distribución invariante del Proceso de McKean-Vlasov dado por (5.1) tiene densidad*

$$g_\beta(u) = \frac{1}{C(\beta)(\beta\mathbb{E}(V) - u)} \exp \left( - \int_0^u \frac{b(s)}{\beta\mathbb{E}(V) - s} ds \right) \quad (5.11)$$

con  $u \in [0, \beta\mathbb{E}(V))$ , y donde

$$C(\beta) = \int_0^{\beta\mathbb{E}(V)} \frac{1}{\beta\mathbb{E}(V) - u} \exp\left(-\int_0^u \frac{b(s)}{\beta\mathbb{E}(V) - s} ds\right) du \quad (5.12)$$

y  $\beta$  es la solución de la ecuación

$$\beta C(\beta) = 1 - \exp\left(-\int_0^{\beta\mathbb{E}(V)} \frac{b(s)}{\beta\mathbb{E}(V) - s} ds\right). \quad (5.13)$$

**Demostración:** Del desarrollo anterior (5.9) se puede deducir fácilmente (5.11) y (5.12). Para obtener (5.13) tomaremos  $\beta := \mathbb{E}(b(X(0)))$  y de la misma manera llamemos  $C(\beta)$  a  $\mathbb{E}(\tau)$  en función de  $\beta$ . Se tiene entonces que, de existir  $\pi$ , se verifica:

$$\begin{aligned} \beta &= \pi(b) \\ \beta C(\beta) &= \pi(b)C(\beta) \\ &\stackrel{(5.9)}{=} \mathbb{E}\left(\int_0^\tau b(x(u)) du\right) \\ &= \int_0^{\beta\mathbb{E}(V)} \frac{b(v)}{\beta\mathbb{E}(V) - v} \exp\left(-\int_0^v \frac{b(s)}{\beta\mathbb{E}(V) - s} ds\right) dv \\ \beta C(\beta) &= \left[-\exp\left(-\int_0^u \frac{b(s)}{\beta\mathbb{E}(V) - s} ds\right)\right] \Big|_0^{\beta\mathbb{E}(V)} \end{aligned}$$

□

**Observación 5.4.** Si  $b(\beta\mathbb{E}(V)) > 0$  (equivalentemente, si  $\beta > 0$  o  $b(0) > 0$  - suponiendo que  $b(x) > 0$  para  $x > 0$ ), el segundo término de la expresión a la derecha de la igualdad en (5.13) desaparece, y la ecuación se reduce a  $\beta C(\beta) = 1$ . Además, podemos simplificar la expresión de  $C(\beta)$  mediante los cambios de variable  $x = u/\beta\mathbb{E}(V)$  y  $y = s/\beta\mathbb{E}(V)$  :

$$C(\beta) = \int_0^1 \frac{1}{1-x} \exp\left(-\int_0^x \frac{b(\beta\mathbb{E}(V)y)}{1-y} dy\right) dx$$

Esta expresión simplificada hará más sencillo el análisis de del comportamiento de la aplicación  $\varphi(\beta) := \beta C(\beta)$ , con el fin de caracterizar el número de posibles distribuciones estacionarias.

### 5.3.1. Cantidad de Distribuciones Estacionarias

El siguiente Lema mostrará que la cantidad de posibles distribuciones estacionarias para el Proceso de McKean-Vlasov 5.1 depende de manera crítica del comportamiento de  $b(x)$  cerca de 0.

**Lema 5.5 (R-T).** *Para toda función  $b$  satisfaciendo las suposiciones (MF) se tiene:*

1.

$$\lim_{\beta \rightarrow +\infty} \varphi(\beta) = +\infty$$

2. Si

$$\lim_{x \rightarrow 0} \frac{b(x)}{x} := \lambda \in [0, +\infty] \text{ entonces} \quad (5.14)$$

$$\lim_{\beta \rightarrow 0} \varphi(\beta) = \frac{1}{\lambda \mathbb{E}(V)}$$

**Demostración:**

1. Llamemos  $\alpha := \beta \mathbb{E}(V)$ , y observemos que, si definimos  $B(t) := \int_0^t b(y) dy$ , entonces  $B(t) \leq tb(t)$  (pues  $b$  es ↗).

La estrategia será ver que  $\forall C > 0$ ,

$$\lim_{\beta \rightarrow +\infty} \varphi(\beta) \geq \lim_{\beta \rightarrow +\infty} \psi(\delta(\beta)) = C$$

Con  $\psi$  y  $\delta$  a determinar. Sea  $\delta \in (0, 1)$  arbitrario. Entonces

$$\varphi(\beta) = \beta \int_0^1 \frac{1}{1-x} \exp\left(-\int_0^x \frac{b(\alpha y)}{1-y} dy\right) dx \geq \beta \int_0^\delta \frac{1}{1-x} \exp\left(-\int_0^x \frac{b(\alpha y)}{1-y} dy\right) dx$$

Para seguir acotando, buscamos cota superior para  $\int_0^x \frac{b(\alpha y)}{1-y} dy$ :

$$\int_0^x \frac{b(\alpha y)}{1-y} dy \leq \int_0^\delta \frac{b(\alpha y)}{1-y} dy \leq \frac{1}{1-\delta} \frac{1}{\alpha} B(\alpha\delta) \leq \frac{\delta}{1-\delta} b(\alpha\delta)$$

De esta manera,

$$\varphi(\beta) \geq \beta e^{-\frac{\delta}{1-\delta} b(\alpha\delta)} \int_0^\delta \frac{1}{1-x} dx = -\beta \log(1-\delta) e^{-\frac{\delta}{1-\delta} b(\alpha\delta)} := \psi(\delta)$$

Como  $\delta$  es arbitrario en  $(0, 1)$ , dado  $C > 0$ , sea  $\delta(\beta) := C/\beta$  para  $\beta > C$ . Entonces

$$\varphi(\beta) \geq \left[-\beta \log\left(1 - \frac{C}{\beta}\right)\right] \left[\exp\left(-\frac{C}{\beta - C} b(C\mathbb{E}(V))\right)\right] \xrightarrow{\beta \rightarrow +\infty} C$$

2. Veamos ahora cómo se comporta  $\varphi(\beta)$  en cero, en función de  $\lambda$ :

a)  $\lambda = 0$  : Sea  $\delta > 0$ . Entonces existe un  $\beta(\delta)$  tal que si  $\beta < \beta(\delta)$ , entonces  $b(\beta \mathbb{E}(V)y) < \delta \beta \mathbb{E}(V)y$  ( con  $y \leq x \leq 1$  ).

Entonces si  $\beta < \beta(\delta)$  :

$$\begin{aligned}
\varphi(\beta) &\geq \beta \int_0^1 \frac{1}{1-x} \exp\left(-\beta \mathbb{E}(V) \delta \int_0^x \frac{y}{1-y} dy\right) dx \\
&= \beta \int_0^1 \frac{1}{1-x} \exp(\beta \mathbb{E}(V) \delta [x + \log(1-x)]) dx \\
&= \beta \int_0^1 e^{(\beta \mathbb{E}(V) \delta x)} (1-x)^{(\delta \beta \mathbb{E}(V) - 1)} dx \\
&= \beta \left[ \frac{1}{\delta \beta \mathbb{E}(V)} + \int_0^1 e^{(\beta \mathbb{E}(V) \delta x)} (1-x)^{(\delta \beta \mathbb{E}(V))} dx \right] \\
&= \frac{1}{\delta \mathbb{E}(V)} + \beta \int_0^1 e^{(\beta \mathbb{E}(V) \delta x)} (1-x)^{(\delta \beta \mathbb{E}(V))} dx
\end{aligned}$$

Por lo tanto, tomando límite en  $\beta$  y por convergencia mayorada,

$$\lim_{\beta \rightarrow 0} \varphi(\beta) \geq \frac{1}{\delta \mathbb{E}(V)} \quad (5.15)$$

Y como  $\delta$  es arbitrario, se tiene  $\lim_{\beta \rightarrow 0} \varphi(\beta) = +\infty$

b)  $\lambda = +\infty$  : Sea  $K > 0$ . Entonces existe un  $\beta(K)$  tal que  $\beta < \beta(K)$  implica  $b(\beta \mathbb{E}(V)y) > K \beta \mathbb{E}(V)y$ .

Luego, de manera totalmente análoga al procedimiento anterior (5.15), se tiene:

$$\varphi(\beta) \leq \frac{1}{K \mathbb{E}(V)} + \beta \int_0^1 e^{(\beta \mathbb{E}(V) K x)} (1-x)^{(K \beta \mathbb{E}(V))} dx$$

Por lo tanto,

$$\lim_{\beta \rightarrow 0} \varphi(\beta) \leq \frac{1}{K \mathbb{E}(V)} \quad (5.16)$$

Donde  $K$  es arbitrario, lo que da como resultado  $\lim_{\beta \rightarrow 0} \varphi(\beta) = 0$

c)  $\lambda \in (0, +\infty)$ : Mediante idéntico procedimiento a los 2 casos anteriores, es fácil ver que, para todo  $\delta > 0$  suficientemente pequeño se tiene

$$\frac{1}{(\lambda + \delta) \mathbb{E}(V)} \leq \lim_{\beta \rightarrow 0} \varphi(\beta) \leq \frac{1}{(\lambda - \delta) \mathbb{E}(V)}$$

□

Como consecuencia inmediata podemos enunciar la siguiente proposición de [43].

**Proposición 5.6** (Cantidad de Soluciones Estacionarias de la Ecuación de Campo-Medio).  
Sea  $\lambda$  como antes (5.14). Entonces se tiene, para el Proceso de McKean-Vlasov (5.1):

- Si  $\lambda \in (1/\mathbb{E}(V), +\infty]$ , entonces existe al menos una solución estacionaria no trivial.

- Si  $\lambda \in [0, 1/\mathbb{E}(V))$  entonces  $\delta_0$  es una solución estacionaria de (5.1), y soluciones estacionarias no triviales existen si y sólo si el mínimo absoluto de  $\varphi(\beta)$  es menor a 1.

La proposición, que caracteriza la cantidad de Distribuciones Estacionarias posibles para el Proceso de McKean-Vlasov, proviene simplemente de conocer cuántas soluciones tiene la ecuación  $\beta C(\beta) = 1$ .

**Conclusión:** De estos resultados podemos concluir que el Proceso de McKean-Vlasov puede tener eventualmente varias soluciones estacionarias no-triviales. Es natural por tanto preguntarse por su *estabilidad* (en el sentido de cuán “atractivas” son las soluciones para el proceso; se precisará más abajo). Recordemos que, en el caso en que  $b(0) = 0$ , el Teorema 4.4 nos dice que la solución trivial  $\delta_0$  es la única solución estacionaria para el proceso de redes finitas, y éstas convergen en el tiempo hacia esa solución. Estos hechos sugieren un fenómeno de cuasi-estacionariedad antes de la absorción, a medida que la red finita crece, que podría ser muy informativa desde el punto de vista matemático y también biológico.

### 5.3.2. Estabilidad de la Solución Trivial

Sabemos que si  $b(0) = 0$ , entonces  $\delta_0$  es solución de la ecuación de Campo-Medio. Pero dada la existencia de otras posibles soluciones no-triviales, según el comportamiento de  $b(x)$  cerca del cero, surge la pregunta acerca de qué tipo de solución “elige” el sistema. El siguiente resultado da una respuesta al respecto.

**Proposición 5.7** (Estabilidad de la solución Trivial). *Con la notación anterior, y  $b(0) = 0$*

- a. Si  $\lambda \in [0, 1/\mathbb{E}(V))$ , entonces  $\delta_0$  es a.e. exponencialmente estable. Es decir, existe  $\varepsilon > 0$  y  $A_\varepsilon > 0$  suficientemente pequeños tal que para toda condición inicial con soporte en  $[0, A_\varepsilon]$

$$\limsup_{t \rightarrow +\infty} \frac{\log(X_t)}{t} < -\varepsilon \quad (5.17)$$

o equivalentemente

$$\mathbb{P}(\exists s_0 = s_0(\omega) : t \geq s_0 \Rightarrow X_t < e^{-\varepsilon t}) = 1 \quad (5.18)$$

- b. Si  $\lambda \in (1/\mathbb{E}(V), +\infty]$ , la solución trivial es inestable en probabilidad. Específicamente, existe  $A > 0$  tal que para cualquier condición inicial  $X_0$  con  $\mu_0 := \mathbb{E}(X_0) > 0$  existe  $t_0 > 0$  tal que

$$\mathbb{P}\left(\sup_{t \in [0, t_0]} X_t > A\right) > 0 \quad (5.19)$$

**Observación 5.8.** Si definimos la cantidad  $\rho := \lambda \mathbb{E}(V)$ , los casos se traducen en  $\rho \in (0, 1)$  y  $\rho \in (1, +\infty]$  respectivamente. Esta cantidad puede verse como la tasa de excitación de la red. Observemos también que el resultado para  $\rho < 1$  (relacionado con el caso 2 de 5.14) es bastante fuerte, mientras que el resultado para  $\rho > 1$  (caso 1 de 5.14) es considerablemente más débil: En el primer caso, casi cualquier trayectoria converge exponencialmente rápido a 0, siempre que se comience “suficientemente cerca de 0”; mientras que en el segundo caso, contamos con una probabilidad estrictamente positiva de estar, en un tiempo finito, “alejado” de 0.

**Demostración:**

- a. Como los saltos dados por el Proceso son todos negativos, una solución  $(X_t)$  del mismo tendrá la cota superior

$$X_t \leq X_0 + \int_0^t [-X_s + \mathbb{E}(V)\mathbb{E}(b(X_s))] ds \quad (5.20)$$

Definamos  $X_t^* := \text{essSup } X_t = \inf\{u > 0 : X_t \leq u \text{ a.e.}\}$  Entonces por el Lema de Gronwall<sup>1</sup> y la monotonicidad de  $b$

$$X_t \leq X_0 e^{-t} + \int_0^t e^{s-t} \mathbb{E}(V)\mathbb{E}(b(X_s)) ds \leq X_0^* e^{-t} + \int_0^t e^{s-t} \mathbb{E}(V)b(X_s^*) ds$$

Entonces podemos concluir que

$$X_t^* \leq X_0^* e^{-t} + \int_0^t e^{s-t} \mathbb{E}(V)b(X_s^*) ds \quad (5.21)$$

Además, como  $\rho < 1$ , podemos tomar un  $\varepsilon > 0$  tal que

$$\mathbb{E}(V) \frac{b(x)}{x} < 1 - \varepsilon \quad \equiv \quad -x + \mathbb{E}(V)b(x) < -\varepsilon x$$

Ahora, suponiendo que  $X_0^* < \varepsilon$ , mostremos que el supremo esencial de  $X_t$  nunca supera  $\varepsilon$ , más aún, tiende a cero. Definamos el tiempo (determinístico)

$$T_0 := \inf\{t > 0 : X_t^* > \varepsilon\}$$

Entonces en  $[0, T_0)$  tenemos

$$X_t^* \leq X_0^* e^{-t} + (1 - \varepsilon) \int_0^t e^{s-t} X_s^* ds \quad (5.22)$$

Multiplicando todo por  $e^t$  y utilizando nuevamente el Lema de Gronwall

$$X_t^* \leq X_0^* e^{-\varepsilon t} < A_\varepsilon e^{-\varepsilon t} < A_\varepsilon$$

Esto implica

- a)  $T_0 = +\infty$

Pues si fuera finito entonces la última desigualdad indica que el proceso debería dar un salto positivo en  $T_0$ .

- b)  $X_t^*$  converge exponencialmente hacia cero.

- b. Observemos que si  $X_t$  es solución de la ecuación de McKean-Vlasov (5.1), entonces

$$X_t = X_0 + \int_0^t -X_s + \mathbb{E}(V)\mathbb{E}(b(X_s)) - X_s b(X_s) ds + M_t \quad (5.23)$$

---

<sup>1</sup>ver apéndice 7



with  $M_t$  la martingala local asociada. Por tanto, tomando esperanza y definiendo  $\mu_t := \mathbb{E}(X_t)$  se tiene

$$\begin{aligned}\mu_t &= \mu_0 + \int_0^t -\mu_s + \mathbb{E}(V)\mathbb{E}(b(X_s)) - \mathbb{E}[X_s b(X_s)] ds \\ &= \mu_0 + \int_0^t \left( \int_{\mathbb{R}_+} -x + \mathbb{E}(V)b(x) - xb(x) p_s(dx) \right) ds\end{aligned}\quad (5.24)$$

Con  $X_s \sim p_s(dx)$ .

Como  $\rho > 1$ , puedo tomar un  $\varepsilon > 0$  y  $A = A(\varepsilon) > 0$  suficientemente pequeños, tal que si  $x < A$ , entonces

$$\mathbb{E}(V)\frac{b(x)}{x} - b(x) \geq 1 + \varepsilon \equiv -x + \mathbb{E}(V)b(x) - xb(x) \geq \varepsilon x$$

Entonces, tomando  $X_0$  tal que  $X_0^* < A$  y  $T_0$  como antes, se tiene en  $[0, T_0)$

$$\mu_t \geq \mu_0 + \int_0^t \left( \int_{\mathbb{R}_+} \varepsilon x p_s(dx) \right) ds = \mu_0 + \varepsilon \int_0^t \mu_s ds$$

entonces, por Lema de Gronwall para  $f(s) = -\mu_s$

$$\begin{aligned}\mu_0 e^{\varepsilon t} &\leq \mu_t \leq X_t^* \leq A \\ t &\leq \frac{1}{\varepsilon} \log \left( \frac{A}{\mu_0} \right) =: t_0 \text{ entonces } T_0 \leq t_0\end{aligned}$$

Y por lo tanto,

$$\exists r \in [T_0, t_0] : X_r^* > A \text{ (por la definicion de infimo)}$$

luego, por la definicion de  $X_r^*$

$$\mathbb{P}(X_r \leq A) < 1 \equiv \mathbb{P}(\sup_{[0, t_0]} X_s > A) \geq \mathbb{P}(X_r > A) > 0$$

□

### Estabilidad: Casos particulares

A modo de cierre, enunciamos un resultado más específico para la estabilidad de la solución trivial y la cantidad de soluciones no estacionarias, para funciones de tasa de disparo afines, es decir de la forma  $b(x) = mx + d$ . Notar que en este caso

$$\frac{b(x)}{x} = m + \frac{d}{x} = \begin{cases} m & \text{si } d = 0 \\ \xrightarrow{x \rightarrow 0} +\infty & \text{si } d > 0 \end{cases}$$

**Proposición 5.9** ((R-T)). Si  $\rho := \lambda \mathbb{E}(V)$ , y  $b(x) = \lambda x + d$  entonces

1. Para  $d = 0$  :

- a) Si  $\rho < 1$ , entonces  $\delta_0$  es la única solución estacionaria, y es estable exponencialmente casi seguramente.
- b) Si  $\rho > 1$   $\delta_0$  es inestable en probabilidad y existe además una única solución no trivial estacionaria.

2. Para  $d > 0$ : Existe una única solución estacionaria y es no trivial.

La demostración de esta proposición es consecuencia de la aplicación de los resultados (5.6), (5.7) y de mostrar que para una tasa de descarga afín, la función  $\varphi(\beta)$  es creciente en todos los casos. Ver [43]

Asumiendo este resultado, vemos que en el caso en que la tasa de excitación  $\rho$  es lo suficientemente grande, la solución estacionaria trivial deja de ser “atractiva”, y emerge una solución no trivial adicional.



## Capítulo 6

# Conclusiones

En cuanto a lo expuesto en la tesis relativo a las Redes Neuronales, un hecho sobresaliente es la posibilidad de la aparición de medidas estacionarias no triviales en el límite hidrodinámico (el límite de campo-medio) cuando en el caso finito sólo se observaba una única distribución invariante. Como mencionamos en el ejemplo de la tasa afín, y de lo que se desprende de la proposición 5.6, si bien cuando  $b(0) = 0$  sabemos que la Red eventualmente morirá, vemos que bajo ciertas condiciones en el límite esta solución no es atractiva, y más aún, existen otras medidas que podrían atraer a la Red. Como se observó anteriormente, esto podría dar margen a un fenómeno cuasi-estacionario, es decir, a una estabilidad local antes de morir. A la hora de estudiar la distribución condicionada, habrá que definir con cuidado qué entendemos por el evento "ser absorbido" que el estado absorbente en este marco,  $0 \in \mathbb{R}^N$  no es un estado al que se llegue, a menos que la condición inicial sea justamente 0. Probablemente bastará para obtener información el considerar alguna pequeña "caja" en torno al 0, que dependerá de la dimensión.



# Capítulo 7

## Apéndice A

### 7.1. Algunos resultados

Aquí demostraremos algunos lemas y resultados que utilizamos en la tesis.

**Lema 7.1.** *Sea*

$$F := \left\{ x \in \mathbb{R}_+^N : \sum_{i=1}^N \left( (N-1)\mathbb{E}(W)b(x_i) - x_i[1+b(x_i)] \right) \geq -1 \right\}$$

Entonces  $F \subseteq [0, C_0]^N$  es un compacto, con  $C_0 = N(N-1)\mathbb{E}(W) + 1$

**Demostración:** Que es compacto es inmediato. Veamos la inclusión. Supongamos, sin pérdida de generalidad, que  $x_1 > C_0$ ; y notemos  $\alpha_N := (N-1)\mathbb{E}(W)$ . Queremos ver entonces que

$$\sum_{i=1}^N \left( \alpha_N b(x_i) - x_i[1+b(x_i)] \right) < -1 \quad (7.1)$$

que reacomodando términos equivale a

$$1 + \sum_{i=2}^N \phi(x_i) < -\phi(x_1) \quad (7.2)$$

Donde  $\phi(x) := b(x)(\alpha_N - x) - x$ . Es inmediato ver que

$$\begin{cases} \text{Si } x > \alpha_N \Rightarrow \phi \text{ es estrictamente decreciente} \Rightarrow \phi(x) < \phi(\alpha_N) = -\alpha_N \\ \text{Si } x \leq \alpha_N \Rightarrow \phi(x) \leq \alpha_N b(C_0) \end{cases}$$

Como  $x_1 > C_0 > \alpha_N$ , para probar (7.2) basta con ver que

$$1 + \sum_{i=2}^N \phi(x_i) < -\phi(C_0) = b(C_0)(C_0 - \alpha_N) + C_0$$

En efecto, el término de la izquierda puede acotarse por  $1 + (N - 1)\alpha_N b(C_0)$  ; y

$$\begin{aligned} 1 + (N - 1)\alpha_N b(C_0) &< b(C_0)(C_0 - \alpha_N) + C_0 \Leftrightarrow \\ 0 &< b(C_0)(C_0 - N\alpha_N) + (C_0 - 1) \end{aligned}$$

Y la última igualdad vale pues los 2 términos son positivos.

□

**Lema 7.2** (Gronwall). *Supongamos que  $x : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  satisface la siguiente desigualdad diferencial*

$$x'(t) \leq g(t)x(t) + h(t) \quad (7.3)$$

*Con  $g$  continua y  $h$  localmente integrable. Entonces se tiene*

$$x(t) \leq x(0)e^{G(t)} + \int_0^t e^{G(t)-G(s)}h(s) ds \quad (7.4)$$

donde  $G(t) = \int_0^t g(a)da$ .

De esta manera el resultado en sale de considerar  $x(t) = X(t)$ ,  $g(t) = -1$  y  $h(t) = \mathbb{E}(V)\mathbb{E}(b(X(t)))$

**Demostración:** Consideramos el factor integrante  $e^{-G(t)}$ , y calculamos

$$(x(t)e^{-G(t)})' = x'(t)e^{-G(t)} - x(t)g(t)e^{-G(t)} \leq e^{-G(t)} (g(t)x(t) + h(t) - g(t)x(t)) = e^{-G(t)}h(t)$$

Integrando esta desigualdad, se obtiene

$$x(t)e^{-G(t)} - x(0) \leq \int_0^t e^{-G(s)}h(s) ds$$

□

## 7.2. Histogramas de la Sección 4.2

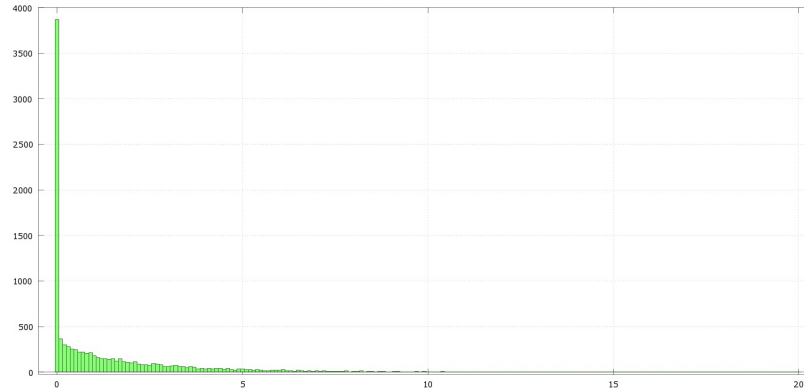


Figura 7.1: Histograma de tiempos de extinción, realizado con 40000 ensayos en Octave, con  $c=1$

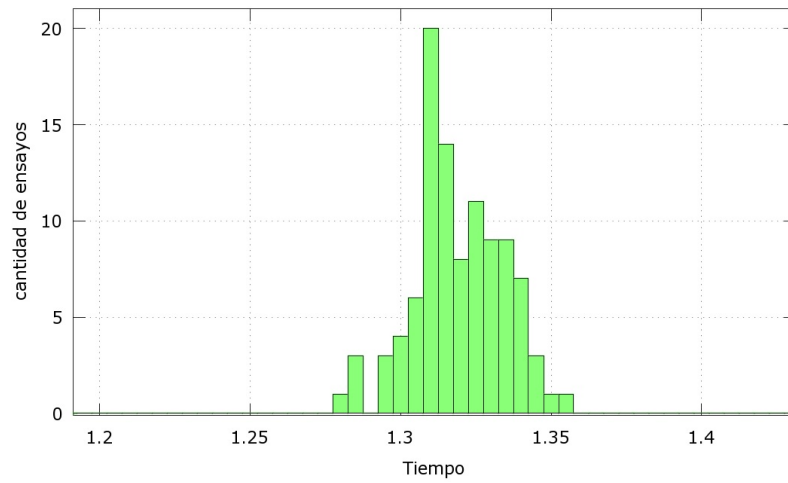


Figura 7.2: Histograma de la media muestral  $\bar{\tau}_{ext}^n$  (el promedio de  $n$  variables i.i.d.) de tiempos de extinción, con  $n=10000$ , realizado con 100 ensayos en Octave, con  $c=1$



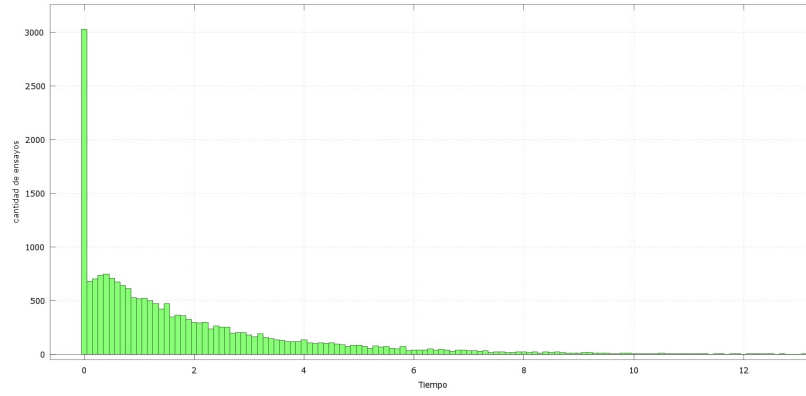


Figura 7.3: Histograma de tiempos de extinción con condición inicial simétrica  $c=1$ , realizado con 20000 ensayos en Octave

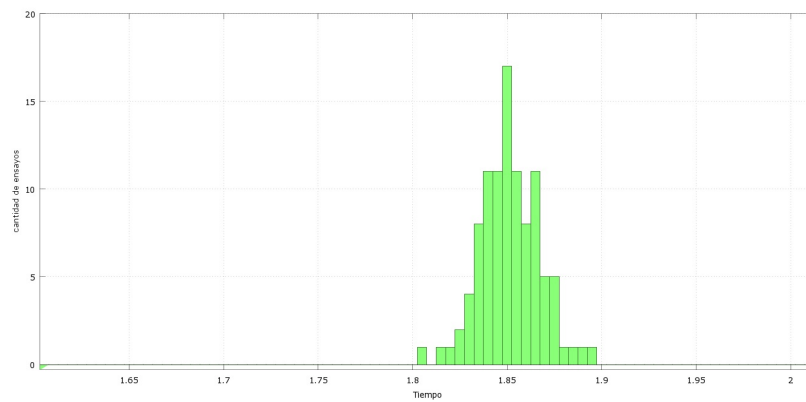


Figura 7.4: Histograma de la media muestral  $\bar{\tau}_{ext}^n$  de tiempos de extinción, con  $n=10000$ , realizado con 100 ensayos en Octave, con  $x=c=1$

# Bibliografía

- [1] E. Adrian. The all-or-none principle in nerve. *The Journal of physiology*, 47(6):460, 1914.
- [2] W. Anderson. *Continuous-time Markov chains: An applications-oriented approach*. Springer Science & Business Media, 2012.
- [3] S. Asmussen. *Applied probability and queues*, volume 51. Springer Science & Business Media, 2008.
- [4] R. Azaïs, J. Bardet, A. Génadot, N. Krell, and P. Zitt. Piecewise deterministic markov process—recent results. In *ESAIM: Proceedings*, volume 44, pages 276–290. EDP Sciences, 2014.
- [5] F. Baccelli, F. Karpelevich, M. Kelbert, A. Puhalskii, A. Rybko, and Yu. Suhov. A mean-field limit for a class of queueing networks. *Journal of statistical physics*, 66(3-4):803–825, 1992.
- [6] J. Baladron, D. Fasoli, O. Faugeras, and J. et al Touboul. Mean-field description and propagation of chaos in networks of hodgkin-huxley and fitzhugh-nagumo neurons. *The Journal of Mathematical Neuroscience*, 2(1):10, 2012.
- [7] A. Barabási, E. Ravasz, and T. Vicsek. Deterministic scale-free networks. *Physica A: Statistical Mechanics and its Applications*, 299(3):559–564, 2001.
- [8] N. Brunel. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of computational neuroscience*, 8(3):183–208, 2000.
- [9] A. Burkitt. A review of the integrate-and-fire neuron model: I. homogeneous synaptic input. *Biological cybernetics*, 95(1):1–19, 2006.
- [10] A. Burkitt. A review of the integrate-and-fire neuron model: II. inhomogeneous synaptic input and network properties. *Biological cybernetics*, 95(2):97–112, 2006.
- [11] M. Caceres and B. Perthame. Beyond blow-up in excitatory integrate and fire neuronal networks: refractory period and spontaneous activity. *Journal of theoretical biology*, 350:81–89, 2014.
- [12] M. J. Cáceres, J. A. Carrillo, and B. Perthame. Analysis of nonlinear noisy integrate & fire neuron models: blow-up and steady states. *J. Math. Neurosci*, 1(7), 2011.
- [13] EJ. Chichilnisky. A simple white noise analysis of neuronal light responses. *Network: Computation in Neural Systems*, 12(2):199–213, 2001.

- [14] M. Davis. Piecewise-deterministic markov processes: A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 353–388, 1984.
- [15] M. Dawkins. The noisy brain: Stochastic dynamics as a principle of brain function, 2010.
- [16] year=1993 publisher=Springer Dawson, D. *Measure-valued Markov processes*.
- [17] A. De Masi, A. Galves, E. Löcherbach, and E. Presutti. Hydrodynamic limit for interacting neurons. *Journal of Statistical Physics*, 158(4):866–902, 2015.
- [18] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré. Global solvability of a networked integrate-and-fire model of mckean–vlasov type. *The Annals of Applied Probability*, 25(4):2096–2133, 2015.
- [19] S. Ethier and T. Kurtz. *Markov processes: characterization and convergence*, volume 282. John Wiley & Sons, 2009.
- [20] G. Gerstein and B. Mandelbrot. Random walk models for the spike activity of a single neuron. *Biophysical journal*, 4(1 Pt 1):41, 1964.
- [21] Granados, M. Cáceres, and R. Schneider. Blow-up, steady states and long time behaviour of excitatory-inhibitory nonlinear neuron models. *arXiv preprint arXiv:1602.05604*, 2016.
- [22] R. Jolivet, A. Roth, F. Schürmann, W. Gerstner, and W. Senn. Special issue on quantitative neuron modeling. *Biological cybernetics*, 99(4):237–239, 2008.
- [23] L. Kadanoff. More is the same; phase transitions and mean field theories. *Journal of Statistical Physics*, 137(5-6):777–797, 2009.
- [24] O. Kallenberg. *Foundations of modern probability*. Springer Science & Business Media, 2006.
- [25] M. Kochmański, T. Paszkiewicz, and S. Wolski. Curie–weiss magnet—a simple model of phase transition. *European Journal of Physics*, 34(6):1555, 2013.
- [26] L. Lapique. Recherches quantitatives sur l’excitation électrique des nerfs traitée comme une polarisation. *J. Physiol. Pathol. Gen*, 9(1):620–635, 1907.
- [27] J-Y Le Boudec, D. McDonald, and J. Munding. A generic mean field convergence result for systems of interacting objects. In *Quantitative Evaluation of Systems, 2007. QEST 2007. Fourth International Conference on the*, pages 3–18. IEEE, 2007.
- [28] D. Levin, Y. Peres, and E. Wilmer. *Markov chains and mixing times*. American Mathematical Soc., 2009.
- [29] T. Lewis and J. Rinzel. Dynamics of spiking neurons connected by both inhibitory and electrical coupling. *Journal of computational neuroscience*, 14(3):283–309, 2003.
- [30] T. Liggett. *Continuous time Markov processes: an introduction*, volume 113. American Mathematical Soc., 2010.

- [31] F. Malrieu. Some simple but challenging markov processes. *arXiv preprint arXiv:1412.7516*, 2014.
- [32] J. Norris. *Markov chains*. Number 2008. Cambridge university press, 1998.
- [33] E. Nummelin. *General irreducible Markov chains and non-negative operators*, volume 83. Cambridge University Press, 2004.
- [34] World Health Organization. *The World Health Report 2001: Mental health: new understanding, new hope*. World Health Organization, 2001.
- [35] S. Ostojic, N. Brunel, and V. Hakim. Synchronization properties of networks of electrically coupled neurons in the presence of noise and heterogeneities. *Journal of computational neuroscience*, 26(3):369–392, 2009.
- [36] K. Pakdaman, B. Perthame, and D. Salort. Dynamics of a structured neuron population. *Nonlinearity*, 23(1):55, 2009.
- [37] K. Pakdaman, B. Perthame, and D. Salort. Relaxation and self-sustained oscillations in the time elapsed neuron network model. *SIAM Journal on Applied Mathematics*, 73(3):1260–1279, 2013.
- [38] K. Pakdaman, B. Perthame, and D. Salort. Adaptation and fatigue model for neuron networks and large time asymptotics in a nonlinear fragmentation equation. *The Journal of Mathematical Neuroscience*, 4(1):1–26, 2014.
- [39] L. Paninski. Maximum likelihood estimation of cascade point-process neural encoding models. *Network: Computation in Neural Systems*, 15(4):243–262, 2004.
- [40] DP. Pelvig, H. Pakkenberg, AK. Stark, and B. Pakkenberg. Neocortical glial cell numbers in human brains. *Neurobiology of aging*, 29(11):1754–1762, 2008.
- [41] J. Pillow, L. Paninski, V. Uzzell, E. Simoncelli, and EJ. Chichilnisky. Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. *The Journal of Neuroscience*, 25(47):11003–11013, 2005.
- [42] P. Robert. *Stochastic networks and queues*, volume 52. Springer Science & Business Media, 2013.
- [43] P. Robert and J. Touboul. On the dynamics of random neuronal networks. *arXiv preprint arXiv:1410.4072*, 2014.
- [44] LCG. Rogers and D. Williams. Diffusions, markov processes, and martingales. vol. 2. itô calculus. reprint of the second (1994) edition. cambridge mathematical library, 2000.
- [45] R. Rudnicki and M. Tyran-Kamińska. Piecewise deterministic markov processes in biological models. In *Semigroups of Operators-Theory and Applications*, pages 235–255. Springer, 2015.
- [46] L. Sacerdote and M. Giraudo. Stochastic integrate and fire models: a review on mathematical methods and their applications. In *Stochastic biomathematical models*, pages 99–148. Springer, 2013.

- [47] J. Touboul. Mean-field equations for stochastic firing-rate neural fields with delays: Derivation and noise-induced transitions. *Physica D: Nonlinear Phenomena*, 241(15):1223–1244, 2012.
- [48] J. Touboul. Spatially extended networks with singular multi-scale connectivity patterns. *Journal of Statistical Physics*, 156(3):546–573, 2014.
- [49] J. et al. Touboul. Propagation of chaos in neural fields. *The Annals of Applied Probability*, 24(3):1298–1328, 2014.
- [50] P. Weiss. L’hypothèse du champ moléculaire et la propriété ferromagnétique. *Journal de Physique*, 6(661):36, 1907.