



UNIVERSIDAD DE BUENOS AIRES  
Facultad de Ciencias Exactas y Naturales  
Departamento de Matemática

## **Análisis estocástico en modelos de redes neuronales biológicas**

Tesis presentada para optar al título de Doctor de la  
Universidad de Buenos Aires en el área Ciencias Matemáticas

**Lic. Pedro Maximiliano Altamirano Carlavan**

Director de tesis: Dr. Matthieu T. Jonckheere  
Consejera de estudios: Dra. Inés Armendariz

Fecha de defensa: 4 de Julio de 2023  
Buenos Aires, 2023

# Análisis estocástico en modelos de redes neuronales biológicas

## Abstract

En esta tesis estudiamos modelos estocásticos que representan el comportamiento del potencial de acción de una red neuronal biológica excitatoria. Estos modelos, que son variaciones del modelo descrito en [RT16], evolucionan mediante procesos de Poisson inhomogéneos que dan cuenta de la aleatoriedad intrínseca de las neuronas y tienen interacciones de carácter *sparse*: cuando una neurona dada efectúa un *spike*, ésta incrementa el potencial de un número fijo de neuronas aleatoriamente elegidas.

En esta tesis mostraremos que esta clase estudiada de modelos para una red neuronal finita siempre termina en la extinción de la actividad neuronal; mientras que, en el límite, la red infinita de neuronas puede sostener la actividad indefinidamente, dependiendo de parámetros relacionados con la intensidad y la cantidad de interacciones en cada *spike*. Mostraremos entonces una transición de fase en términos de la distribución estacionaria de la red infinita, que puede ser no-trivial. De esta manera, modelamos el fenómeno biológico de *persistencia*: La red neuronal, si bien eventualmente muere, puede mostrar actividad para tiempos grandes dependiendo de la población de la red.

El desarrollo de esta tesis resultó en el siguiente trabajo [ACJL23], realizado junto a Matthieu Jonckheere, Roberto Cortez y Lasse Lëskela, y existen varias líneas de trabajo posibles que derivan de lo desarrollado aquí.

**Palabras clave:** *Redes Neuronales, Procesos de McKean-Vlasov, Procesos estocásticos, Límites hidrodinámicos, Propagación del caos*

---

## Stochastic Analysis in biological neural network models

### Abstract

In this thesis we study stochastic models representing the behaviour of the action potential of an excitatory biological neural network. These models, inspired by the one defined in [RT16], evolve through inhomogeneous Poisson processes accounting for the intrinsic randomness of neurons and have sparse-tipe interactions: when a neuron fires, it gives an impulse to the potential of a fixed number of randomly chosen neurons within the network. In this work, we will show that these models always come to extinction when the neural network is finite, whereas the infinite network can sustain activity forever, depending on parameters related to the intensity and the amount of interactions in each spike. We will show a phase transition in terms of the stationary distribution of the infinite network, which could be non-trivial. Consequently, we model the biological phenomenon of *persistence*: Even though the activity of the neural network eventually dies, it can remain active for a long time, depending on the network population.

The development of this thesis resulted in the following work [ACJL23], in collaboration with Matthieu Jonckheere, Roberto Cortez and Lasse L eskela, and several lines of work can be derived from the results presented here.

**Keywords:** *Neural Networks, McKean-Vlasov processes, Stochastic processes, Hydrodynamic limits, Propagation of chaos*

*A mis abuelos; Ana, Blanca, Pedro y Juan.*

# Agradecimientos:

Con esta tesis culmina una larga travesía. Una que podría calificarse de verdadera montaña rusa. Muchos son los pensamientos y emociones que genera el final de esta etapa pero la que domina es un profundo agradecimiento.

Gracias a Dios, por sus constantes regalos y gracias, por esta vida maravillosa que me dio, con todas sus idas y vueltas.

Gracias infinitas a Matt, que me viene acompañando desde la tesis de Licenciatura y luego en esta tesis doctoral. Realmente agradezco su interminable paciencia para conmigo, por saber animarme y sostenerme en los momentos más duros de este camino. Gracias por saber mostrarme y ayudarme a desarrollar mis habilidades como matemático. Gracias por confiar en mí aún cuando yo mismo tenía dificultades para hacerlo: todavía recuerdo una charla que tuvimos en un bar de París en la que con mucha firmeza pero con mucha honestidad y sincera preocupación por mí logró que reaccionara y pusiera todo de mi parte para llegar hasta donde hoy estoy. Matt, simplemente gracias.

Gracias a Paloma, mi novia y futura esposa, que aunque llegó a mi vida en la etapa final de mi doctorado, desde el primer momento fue un sostén indispensable para mí, bancándome, escuchándome, animándome, aguantando con buen humor y mucho cariño mi cansancio y/o mal humor después de horas de trabajo de tesis, y hasta haciendo de oyente para practicar mi defensa aunque no entendiera nada de lo que le decía. Eso es amor del bueno. Llega al final de mi doctorado, al principio de mi camino como investigador y al principio del resto de nuestras vidas. ¡Te amo, Palo!

Gracias a mis padres, Pedro y Ana, y a mis hermanos, Guada, Agus, y Seba, que no sólo me bancaron en estos últimos años sino desde hace 32, siempre ayudándome a ser la mejor versión de mí mismo, sabiendo ser pacientes conmigo y sacándome una buena sonrisa cuando lo necesité. Nunca dejaré de agradecer a Dios por la familia que tengo, y hoy no es la excepción.

Gracias a todos los amigos y conocidos matemáticos que fui encontrando y que me ofrecieron su ayuda desinteresada, enriqueciéndome con su experiencia, sus capacidades, su saber, su ingenio y su tiempo, especialmente a Manu Saenz, Dani Cuesta, Nahuel Soprano, Roberto Cortez, Lasse Lëskela, e Inés Armendáriz, entre tantos otros.



# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Contexto biológico . . . . .	1
1.2. Modelos de redes neuronales . . . . .	3
1.3. Comportamiento límite y teoría de campo-medio . . . . .	4
1.4. Contribuciones de la tesis . . . . .	5
1.5. Estructura de la tesis . . . . .	8
<b>2. Preliminares</b>	<b>9</b>
2.1. Procesos de Poisson . . . . .	9
2.2. Couplings y distancia Wasserstein . . . . .	14
2.3. Cálculo estocástico . . . . .	17
<b>3. Campo medio y propagación del caos</b>	<b>23</b>
3.1. Campo medio y distribuciones empíricas . . . . .	23
3.2. Propagación del Caos y límite hidrodinámico . . . . .	25
3.3. Algunas técnicas para demostrar propagación del caos . . . . .	27
3.4. Propagación del Caos en nuestro modelo . . . . .	31
<b>4. Un modelo de redes neuronales con interacciones sparse</b>	<b>33</b>
4.1. Descripción del modelo . . . . .	33
4.2. Proceso límite de campo medio . . . . .	38
<b>5. Comportamiento límite y persistencia</b>	<b>39</b>
5.1. Comportamiento a tiempos grandes de la red finita de neuronas . . . . .	39
5.2. Transición de fase en el límite de campo medio . . . . .	40
5.3. Propagación del caos . . . . .	42
5.4. Persistencia . . . . .	43
5.5. Pruebas de los resultados . . . . .	44
5.6. Prueba de propagación del caos y persistencia (Teorema 5.3.1) . . . . .	50
<b>6. Simulaciones</b>	<b>63</b>
6.1. El tiempo de extinción . . . . .	63
6.2. Comportamiento de la media empírica: . . . . .	68
6.3. Actividad de una partícula marcada: . . . . .	68
<b>7. Conclusiones</b>	<b>73</b>





# Capítulo 1

## Introducción

El cerebro humano es un órgano de muy alta complejidad, y es el objeto central de estudio multidisciplinar en la Neurociencia. A lo largo del siglo XX y hasta la actualidad, se han realizado numerosos y prolíficos avances que han permitido comenzar a comprender su funcionamiento, su fisiología, sus mecanismos de interacción con otros sistemas del cuerpo humano, entre otros. La importancia de estos avances no sólo radica en una mayor comprensión teórica, sino que también abre las puertas a numerosas aplicaciones en muy variados campos, como la neuroeducación, o el tratamiento de diversas afecciones relacionadas con el sistema nervioso. Por ejemplo, la Organización mundial de la Salud (OMS) ha destinado abundantes recursos para la investigación y manejo de desórdenes mentales.

Es sin embargo incipiente aún el desarrollo en esta área del conocimiento, y ello hace todavía más relevante el trabajo de las distintas disciplinas, las Matemáticas incluídas, para ir dando respuestas a los numerosos interrogantes en este campo.

La unidad funcional básica del cerebro son las *neuronas*, que se interconectan de a millones formando grandes *redes neuronales*. Se les suele añadir el término “*biológicas*” para distinguirlas de las *redes neuronales artificiales*, que son modelos computacionales que se organizan emulando el comportamiento básico de las primeras y que han sido piedra angular del avance en inteligencia artificial.

Esta tesis se centrará en el estudio de un modelo estocástico de redes neuronales biológicas, es decir, un modelo que representa la alta complejidad de las interacciones entre neuronas incorporando dinámicas aleatorias.

### 1.1. Contexto biológico

La corteza cerebral -primordial en funciones como la memoria, la percepción o la consciencia, entre otras- posee entre quince y treinta y tres millones de neuronas [PPSP08] distribuidas en grandes redes de neuronas en un espesor de dos a cuatro milímetros. Estas redes, a su vez, se organizan en una sucesión de “láminas” o capas, llamadas *columnas corticales*, cada una de las cuales está compuesta por poblaciones de neuronas estadísticamente idénticas y fuertemente interconectadas.

Las neuronas (del griego *nêuron* “nervio”, “cuerda”) son células eléctricamente excitables que reciben y transmiten información a través de señales eléctricas y químicas, comunicándose así entre ellas y con otras células, como las células gliales.

Existen diferentes clases de neuronas con fisiologías particulares, pero la estructura celular básica de la mayor parte se compone de *dentritas*, que reciben los impulsos eléctricos de otras neuronas y los trasladan hacia el cuerpo celular; el axón, mediante el cual la neurona transmite sus impulsos (llamados “*potencial de acción*” o “*spike*”<sup>1</sup>) hacia otras neuronas; y finalmente de un *cuerpo celular (soma)*, donde reside el núcleo celular, el ADN de la célula y donde se producen las proteínas que se transportan por las dentritas y el axón. La figura 1.1 representa el esquema típico de una neurona.

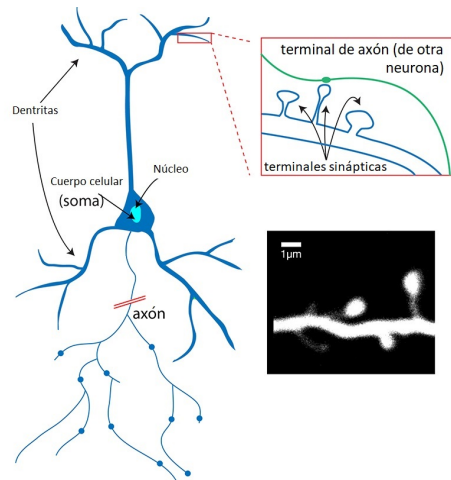


Figura 1.1: Representación esquemática de una neurona (Imagen: Alan Woodruff; De Roo et al / CC BY-SA 3.0 via Commons)

La transmisión de señales electroquímicas entre las neuronas se produce a través de la *sinapsis*: Las neuronas poseen típicamente una diferencia de potencial entre el interior y el exterior de su membrana celular, que en estado de reposo se encuentra alrededor de los  $-70mV$ . Los impulsos que recibe la neurona a través de las dentritas llegan a la membrana a través de neurotransmisores que abren canales de iones (fundamentalmente de Potasio y de Sodio) en ella, generando un crecimiento del potencial (impulso *exitatorio*) o un decrecimiento (impulso *inhibitorio*). Si el impulso excitatorio es lo suficientemente grande, se genera una descarga eléctrica - el potencial de acción - que viaja a través del axón hasta sus terminales; conectadas en general a las dentritas de otras neuronas. Inmediatamente luego de producirse el potencial de acción, la neurona que lo transmitió - neurona *presináptica* - decrece su potencial hasta valores entre  $-80mV$  y  $-70mV$  (levemente por debajo de su potencial de reposo), pasa por un breve período refractario, para luego volver al potencial de reposo.

A partir de las investigaciones de Henry Pickering Bowditch e investigaciones posteriores, se sabe que estas dinámicas eléctricas actúan bajo el *principio de todo-o-nada* [Adr14]. Esto significa que no existen los que podrían llamarse “impulsos fuertes” o “impulsos débiles”; en otras palabras, esto significa que la intensidad de la descarga de una neurona no depende de la intensidad del estímulo que recibe: O bien se produce un potencial de acción, o directamente no emite ningún impulso. Es por este fenómeno que en muchos modelos aparece el concepto de *umbral*, valor para el potencial de membrana a partir del cuál se dispara el impulso sináptico.

<sup>1</sup>también decimos que la neurona se “*dispara*” o produce un “*firing*”

La intensidad del impulso recibido por una neurona post-sináptica, en cambio, sí tiene efecto sobre su frecuencia de “disparo”.

## 1.2. Modelos de redes neuronales

Desde el trabajo pionero de Lapicque [Lap07], la literatura matemática en redes neuronales biológicas ha ido extendiéndose prolíficamente, buscando, entre otras cosas, describir el complicado entramado de interacciones entre las neuronas de una red neuronal, y las interacciones entre éstas últimas. Sin embargo, existen todavía muchas preguntas fundamentales que siguen abiertas, dada la extrema complejidad inherente a este sistema biológico.

Entre estas preguntas se encuentra aquella que busca comprender el fenómeno de la *persistencia*<sup>2</sup>: dado un impulso moderado, la actividad de la red neuronal se sostiene por tiempos sorprendentemente largos. Ha sido observado en diversas regiones del cerebro y en distintos organismos, y se ha postulado como un mecanismo de almacenamiento de memoria de corto plazo [GWL08, MT04, ZS17]. Este fenómeno es uno de nuestros principales intereses desde el punto de vista biológico en esta tesis, y es interesante destacar que al día de hoy no abundan estudios cuantitativos de este fenómeno que ayuden a profundizar en las causas y mecanismos inherentes al mismo.

Recorremos ahora brevemente el amplio espectro de los principales modelos relativos a redes neuronales biológicas: Por un lado, los *modelos basados en conductancia* trabajan en detalle los intercambios químicos entre las neuronas pre-sináptica y post-sináptica cuando se produce la sinápsis. El modelo de Hodgkin-Huxley es paradigmático entre éstos [HH52], y se han realizado varios trabajos bajo este enfoque [BFFT12, Tou14, T<sup>+</sup>14]. En general, estos modelos implican ecuaciones complejas, que en ocasiones específicas pueden ser simplificadas parcialmente [Tou12].

En el otro extremo del espectro están los modelos de tipo “*integrate-and-fire*”. El modelo de Lapicque pertenece a este gran grupo. Estos modelos se encuentran extensamente reseñados en [Bur06a, Bur06b]. Teniendo en cuenta la naturaleza inmediata y estereotipada de los potenciales de acción, estos modelos investigan el problema de la distribución temporal de los impulsos sinápticos, que se modelan como eventos instantáneos. En estos modelos, la magnitud de estudio es el potencial de membrana de cada neurona y se asume que la neurona integra los impulsos recibidos de otras neuronas hasta alcanzar cierto umbral (en general, fijo) que provoca a su vez que ésta emita un potencial de acción. Los modelos “*leaky-integrate-and-fire*” (LIF) tienen adicionalmente en cuenta el decaimiento del potencial de membrana hacia su valor de reposo, que se produce como consecuencia de la “filtración” de iones a través de la membrana celular cuando no hay interacción entre neuronas.

Los primeros trabajos sobre esta familia de modelos han sido de carácter determinístico y abocados al estudio de una neurona aislada. Dicho carácter determinístico no ha sido, sin embargo, impedimento para lograr grandes avances para la época en que fueron publicados, aportando entre otras cosas un primer andamiaje sobre el cual se fueron construyendo modelos más complejos. Por otra parte, hay una aleatoriedad en el comportamiento neuronal que requiere de modelos no-determinísticos para representar mejor al sistema biológico, por

<sup>2</sup>El término se distingue de la *persistencia* en el contexto de *deep learning*, que es una medida compleja que caracteriza y da cuenta de propiedades estructurales de una red neuronal artificial, y permite un proceso más eficiente de entrenamiento de dicha red [RTB<sup>+</sup>18]

ello gran parte de estos modelos introducen estocasticidad en su dinámica. Usualmente, dicha estocasticidad se incorpora a través de procesos Gaussianos externos, que representan el efecto que neuronas exteriores a la red tienen sobre la misma. Esta modelización permite el uso del cálculo de Itô para realizar cálculos y obtener resultados varios. Sin embargo, este enfoque tiene algunas importantes limitaciones. Por ejemplo, estos modelos no tienen en cuenta en general la aleatoriedad intrínseca de las neuronas, que es una característica esencial que no puede ser pasada por alto [RD10]. Algunos modelos recientes sí incluyen algún tipo de aleatoriedad intrínseca introduciendo ruido blanco asociado a cada neurona de la red [DIR<sup>+</sup>15], lo que representa una solución parcial a este problema. Otro aspecto problemático importante de estos modelos tiene que ver con el llamado *efecto avalanche*, provocado por la introducción de un umbral fijo para el potencial de membrana de cada neurona, a partir del cual éstas -de manera determinística- emiten un impulso. Más precisamente, el potencial de membrana *explota* (va a infinito) en el límite de campo-medio (ver por ejemplo [CCP11]), lo cual no representa apropiadamente el comportamiento de grandes poblaciones neuronales.

En este sentido, los *modelos de Poisson Lineales-no lineales* (LNP por sus siglas en inglés) abordan esos problemas de manera más efectiva. El primer modelo LNP fue presentado por Chichilnisky [Chi01]. Estos modelos representan el comportamiento del potencial de membrana utilizando procesos de Poisson no homogéneos de intensidad dependiente del mismo potencial de membrana. Los modelos LNP parecen representar más eficazmente las dinámicas sinápticas de las redes neuronales, muestran un mejor ajuste a los datos experimentales [PPU<sup>+</sup>05, PSP<sup>+</sup>08], y evitan el problema de la explosión del potencial de acción. El modelo de Robert y Touboul [RT16] pertenece a esta familia de modelos, y es la principal inspiración de nuestro marco de trabajo. El modelo presentado en esta tesis es, en efecto, una modificación de [RT16] y pertenece también a esta familia; aunque posee características distintivas que lo separan del modelo de Robert y Touboul, ver la sección 1.4.

### 1.3. Comportamiento límite y teoría de campo-medio

Cuando se modela un sistema de partículas estadísticamente indistinguibles de grandes dimensiones, un abordaje típico es la aproximación de dicho sistema por un modelo “límite” que representa el comportamiento de una partícula genérica del mismo. Este modelo se obtiene mediante la suposición de que el efecto de la interacción de todas las partículas sobre la partícula de referencia puede ser reemplazado por un sólo efecto ponderado, que se aproxima arbitrariamente al comportamiento del sistema original a medida que la dimensión de éste último se acerca a infinito.

El concepto tiene su origen en la termodinámica, a partir de las investigaciones de Weiss y Pierre Curie sobre transición de fase [Wei07] [KPW13] para un sistema ferromagnético de partículas. Esta estrategia probó su efectividad en este sistema y otros sistemas análogos [Kad09], y la idea fundamental luego se extrapoló a otros sistemas no necesariamente termodinámicos, como por ejemplo modelos epidémicos [LBMM07], teoría de colas [BKK<sup>+</sup>92] o modelos de redes estocásticas en general [BRV01], entre las que se encuentran varios modelos de redes neuronales biológicas, incluyendo el modelo presentado en esta tesis.

De esta manera, haciendo uso de este enfoque presentamos un modelo límite que toma la forma de una ecuación diferencial estocástica que podríamos llamar de tipo *Mc-Kean Vlasov*,

en cuanto a que la dinámica del proceso depende en algún sentido de cantidades promedio del proceso mismo.

Para formalizar la idea del modelo de campo-medio como aproximación al modelo estocástico de grandes cantidades de partículas, es necesario demostrar algún tipo de convergencia de éste último al primero, cuando la población tiende a infinito, y típicamente para lograr esto es necesario mostrar que ocurre lo que se conoce como *propagación del caos*. Este concepto esencialmente significa que, si bien hay interacción entre las distintas partículas, y por lo tanto no existe independencia entre ellas, asintóticamente sí la hay (desde un punto de vista probabilístico). Existen distintas estrategias para abordar este problema, ver Capítulo 3. La eficacia de muchas de éstas se apoya en que las interacciones entre las partículas son binarias (es decir, en cada interacción se ven implicadas no más de dos partículas). Esta característica excluye a varios modelos con interacciones no binarias, como es el caso de nuestro modelo.

## 1.4. Contribuciones de la tesis

### 1.4.1. Aspectos originales del modelo de redes neuronales

Inspirado en el modelo propuesto por Robert y Touboul [RT16], proponemos en esta tesis un sistema de partículas que describe el potencial de membrana de una red excitatoria de neuronas y, adicionalmente, un proceso estocástico no lineal de tipo campo-medio modelando el comportamiento de una neurona genérica en una población grande de neuronas. En este sentido el modelo presentado aquí describe, para una población de  $N$  neuronas, el potencial de acción de las mismas a lo largo del tiempo, que notaremos  $\mathbf{X}_t = (X_t^1, \dots, X_t^N)$ . La dinámica es la siguiente: a tiempo  $t$ , cada neurona con potencial  $X_t^i$  se ve afectada por un decaimiento de tipo exponencial hacia su potencial de reposo, que establecemos en cero. Además, ocurren interacciones aleatorias entre las neuronas. Estas interacciones ocurren cuando una neurona emite un potencial de acción o *spike*, momento en el que se resetea de manera inmediata a su potencial de reposo y, simultáneamente, elige de manera aleatoria un número fijo  $\kappa$  de otras neuronas, a las cuales da un impulso excitatorio. Estos eventos se producen con una tasa dependiente del estado de la neurona presináptica.

Puede observarse que, en contraste con [RT16], la dinámica considerada en esta tesis es de carácter *sparse*, en el sentido de que en cada evento de *firing*, la neurona presináptica se comunica sólo con  $k$  otras neuronas y no con todas las de su red. Además, la magnitud del impulso es de orden 1, y no  $1/N$ . Como consecuencia, entre otras, en la ecuación límite de campo-medio el mecanismo de *firing* genera un término no-lineal de salto en lugar de un término de drift, como en [RT16]. Ésta ha sido una elección deliberada para el modelo que aquí estudiamos. Por un lado, permite dinámicas posiblemente más interesantes desde un punto de vista biológico: ver, por ejemplo, [BRC<sup>+</sup>19], donde se plantean esquemas de *sparse-coding* que implican interacciones “esparcidas” a nivel poblacional. En este sentido nuestro modelo es una primera aproximación a una dinámica de interacciones esparcidas (en inglés se las conocen como *sparse interactions*) y una extensión natural para trabajos futuros puede ser el considerar una estructura de grafo subyacente para las interacciones. Por otro lado, este modelo exhibe una transición de fase distinta (y quizás más significativa) relacionada al comportamiento del sistema para tiempos grandes, con una interpretación física natural. Finalmente, difieren también las estrategias matemáticas utilizadas para cuantificar la

decorrelación entre las neuronas (la *propagación del caos*), que en este caso son probablemente generalizables a una clase más grande de modelos similares con interacciones entre partículas no necesariamente binarias.

Para este modelo de sistema de partículas analizamos en primer lugar el comportamiento a tiempos grandes, y mostramos que, para cualquier estado inicial y para cualquier set de parámetros del sistema, el potencial de acción de toda la población tiende a cero. Es decir, la actividad eventualmente se extingue siempre y, ahora sí, dependiendo de ciertos parámetros, la velocidad de extinción puede incluso ser de tipo exponencial.

### 1.4.2. El proceso límite

En gran cantidad de trabajos derivados de perspectivas de física estadística, los modelos de sistemas de partículas suelen tener un proceso de campo-medio límite (cuando la población  $N$  tiende a infinito) asociado, llamado *límite hidrodinámico*. En estos trabajos es usual obtener resultados de convergencia (en algún sentido) del sistema de partículas para tiempos fijos al proceso límite, que incluyen la obtención de la velocidad de convergencia.

Como fue mencionado antes, nuestro sistema de partículas tiene asociado a su vez un proceso límite que sigue una dinámica de tipo McKean–Vlasov, en cuanto a que dichas dinámicas dependen de cantidades promedio del proceso mismo. Este tipo de procesos ha sido extensamente estudiado en las últimas décadas cuando la dinámica de la que provienen es de tipo difusiva, ver por ejemplo el paper seminal [BRTV98] y otros desarrollos posteriores ([Mal01, Tug14] y las referencias allí citadas).

Sin embargo, es necesario notar que en estas referencias los resultados más importantes de velocidad de convergencia a los límites hidrodinámicos respectivos son usualmente obtenidos en el caso de existencia y unicidad de distribución estacionaria y/o se apoyan fuertemente en propiedades de regularidad de la ecuación diferencial límite. En este trabajo no contamos con estas suposiciones y otras típicamente presentes en esos otros mencionados en el párrafo anterior. Por lo tanto, hay toda una familia de técnicas utilizadas allí que no podemos aplicar a nuestro caso.

De la misma manera, es extensa la literatura abocada a sistemas de partículas con interacciones binarias (de salto) con conservación de la energía, junto a sus ecuaciones límite respectivas. Ver por ejemplo el trabajo pionero [Kac56], o bien [MM13] y sus referencias para trabajos posteriores. En el contexto descrito, muchos son los resultados obtenidos en cuanto a convergencia al equilibrio y propagación del caos. En contraste, la dinámica de nuestro modelo involucra en cada interacción el salto simultáneo de tres o más partículas (neuronas), sin conservación de la energía (nisiquiera en promedio). Es por eso que la ecuación límite obtenida es de naturaleza cualitativamente distinta. Esto tiene fuertes consecuencias para los resultados de propagación del caos o velocidad de convergencia, que se vuelven más difíciles de obtener y fuertemente dependientes de los parámetros del modelo. En efecto, muchas de las técnicas clásicas utilizadas en modelos de sistemas de partículas que provienen del ámbito físico no pueden ser aplicadas en este caso, al menos de manera directa. Nuestro estudio de propagación del caos, por otro lado, se apoya en técnicas recientes de acoplamiento para interacciones binarias desarrolladas en [CF16], que son adaptadas para este contexto y extendidas para su aplicación al modelo aquí presentado.

En este contexto entonces, analizamos el comportamiento de una neurona típica de la red

neuronal cuando la población tiende a infinito ( $N \rightarrow \infty$ ) a tiempo fijo, y mostramos que éste converge al proceso límite, que de igual manera llamaremos *proceso no lineal*. Para probar la propagación del caos utilizamos, como dijimos arriba, técnicas modernas de acoplamiento inspiradas en [CF16] y adaptadas correspondientemente.

También analizamos en detalle al proceso límite, mostrando que, además de estar bien definido, experimenta una *transición de fase* dependiente de los parámetros del modelo: si la transmisión de potencial de acción es lo suficientemente fuerte (en términos de intensidad y frecuencia), el proceso no lineal se mantiene activo por siempre; mientras que de otra manera, decae al estado de reposo.

### 1.4.3. Fenómeno de persistencia

Finalmente, mostramos que en nuestro modelo se da un fenómeno de *persistencia* dependiente de los parámetros del modelo: si la transmisión de potencial es lo suficientemente fuerte, entonces aunque la red finita eventualmente se extingue para cualquier tamaño  $N$  de población, se mantiene activa por un tiempo de orden al menos  $\log(N)$ . En otro caso, la actividad se extingue a una velocidad independiente de  $N$ .

Es interesante observar la criticalidad del fenómeno de persistencia en este modelo: en el capítulo 5 probaremos que en este fenómeno la dependencia de los parámetros es indirecta y a través de un único parámetro  $\theta$ , que determina si la actividad de la población de neuronas sobrevive (al menos por un tiempo logarítmico en la población) o bien se extingue exponencialmente. Vamos a explicar este fenómeno de criticalidad con un poco más de detalle: Existen cuatro parámetros para nuestro modelo  $N$ -poblacional y su límite (ver 4.1): la *tasa de decaimiento*  $\mu > 0$ , que determina la velocidad de decaimiento hacia el potencial de reposo; la *constante de proporción de la tasa de firing*  $\gamma > 0$ , que regula la intensidad con la que se produce un firing, en relación al potencial de membrana de las neuronas; el *rango de alcance*  $\kappa \in \mathbb{N}$ , que representa la cantidad de neuronas excitadas al producirse un spike; y la *magnitud de impulso*  $\rho > 0$  que recibe cada neurona seleccionada en un evento de firing. Heurísticamente, es razonable la expectativa de que todos y cada uno de estos parámetros debe jugar un rol condicionante en la fenomenología general del sistema, al menos cuantitativamente si no de manera cualitativa. Por ejemplo, fijados  $\rho$ ,  $\gamma$ ,  $\mu$ , es razonable esperar que a mayor rango de alcance  $\kappa$ , encontremos mayor actividad a lo largo del tiempo en el modelo finito, incluso quizás evitando la extinción. Los teoremas 5.1.1 y 5.2.2 nos confirman, por un lado, que efectivamente  $\kappa$  tiene (desde el punto de vista heurístico) el efecto esperado sobre el sistema; pero este efecto sólo se hace efectivo a través de *una función* de este parámetro y del resto  $\theta = \theta(\kappa, \gamma, \rho, \mu)$  (ver (5.1)), y más aún, el comportamiento asintótico en el tiempo sólo experimenta un cambio cualitativo en función de  $\kappa$  en el caso infinito, mientras que en el caso finito el efecto es sólo cuantitativo, más precisamente, determina la cota superior para la velocidad de extinción. Un análisis análogo podemos hacer para el resto de los parámetros con similares conclusiones.

Es interesante también señalar el paralelismo existente con los procesos de contacto. Éstos son procesos que generalmente modelan la propagación de un virus dentro de una población<sup>3</sup>, donde los individuos eventualmente pueden recuperarse y reinfectarse, y donde la dinámica

<sup>3</sup>otras aplicaciones incluyen por ejemplo, propagación de información y en general, diversos modelos de crecimiento de poblaciones.

básica consiste en que un individuo infectado puede infectar a sus “vecinos” y/o recuperarse con sendas tasas. En estos procesos el objetivo básico es estudiar si la “actividad” del virus eventualmente se extingue o no, y a qué tasa. De manera análoga a nuestro caso, en los procesos de contacto con población infinita, la pregunta crucial es si existe o no probabilidad positiva de *supervivencia* para la actividad del virus; mientras que en el caso finito generalmente se observa extinción casi segura del virus, siendo la velocidad de extinción la pregunta principal, cuya respuesta suele traducirse en una transición de fase dependiente fundamentalmente de la tasa de infección. Históricamente, los primeros estudios trabajaron con poblaciones infinitas (por ejemplo, [HL78] donde se prueba existencia de un parámetro  $\tau$  crítico en  $\mathbb{Z}^d$ , o [BG90], donde se obtiene extinción segura en el caso crítico) para luego abordar modelos de contacto con población finita, como en [CD09], donde se prueba que, en un modelo con nodos con distribución de grado polinomial, para toda tasa de infección positiva existe metaestabilidad.

Consecuentemente, en ambos procesos (el que trabajamos aquí y el de contacto) la respuesta a la pregunta por la *persistencia* se infiere a partir de resultados de tipo *transición de fase*, que en nuestro caso se explicita en la dicotomía ya mencionada: o tarda un tiempo de orden  $\log(N)$  para extinguirse, o es independiente de  $N$ .

Estos elementos por tanto representan un aporte interesante y original al estudio del fenómeno biológico de *persistencia* que presentamos en la sección anterior 1.2. Como mencionamos allí, los estudios cuantitativos sobre este fenómeno no abundan y creemos que lo aquí presentado puede tener impacto positivo en esta dirección de investigación. Adicionalmente, el paralelismo con los procesos de contacto genera como consecuencia vínculos que pueden ir más allá de la heurística, sirviendo no sólo como fuente mutua de inspiración sino también estableciendo potenciales vínculos matemáticos que den lugar a resultados en una familia de modelos que sean provechosos en la otra.

## 1.5. Estructura de la tesis

La estructura de la tesis es la siguiente. En el Capítulo 2, proporcionamos una breve introducción a los procesos de Poisson y su cálculo estocástico asociado. En el Capítulo 3 presentamos el concepto central de propagación del caos y límite hidrodinámico. Luego, en el Capítulo 4 presentamos el modelo estocástico, aportando tres formas equivalentes. En el Capítulo 5 analizamos el comportamiento asintótico en el tiempo del sistema finito de partículas y de su proceso no lineal asociado, el cual, mostramos, exhibe una transición de fase. Asimismo, mostramos el fenómeno de propagación del caos, que nos indica que una neurona típica del sistema finito puede ser bien aproximada, en una población grande, por el comportamiento del proceso no lineal. Las pruebas a cada resultado se encuentran al final de este capítulo. En el capítulo 6 realizamos varias simulaciones que ilustran y refuerzan los resultados obtenidos, así como también permiten realizar algunas conjeturas. Finalmente, en el capítulo 7, compartimos algunas conclusiones y trazamos posibles líneas de trabajo futuro.



## Capítulo 2

# Preliminares

En los siguientes dos capítulos vamos a definir los principales objetos que vamos a utilizar durante la tesis para trabajar con nuestro modelo de redes neuronales biológicas. Consecuentemente, comenzaremos dando en este capítulo una introducción breve de los procesos de Poisson y el cálculo estocástico asociado, y también estableceremos la notación que utilizaremos a lo largo de la tesis. El concepto de *propagación del caos* merece un capítulo aparte y consecuentemente será tratado en el siguiente apartado.

Cuando queramos hacer explícita la distinción entre elementos vectoriales y escalares, notaremos por ejemplo  $\mathbf{X}$  en lugar de  $X$ . Notamos  $I_N$  al conjunto  $\{1, \dots, N\}$ . También, como es usual, dado  $x \in E$  notaremos por  $\mathbb{P}_x$  a la medida de probabilidad dado que  $X_0 = x$  c.s. (*casi seguramente*, es decir, con probabilidad 1) y  $\mathbb{E}_x$  a la esperanza respecto de  $\mathbb{P}_x$ . Si  $X$  es una variable aleatoria, notamos por  $\mathcal{L}(X)$  a la ley o distribución de dicha variable. Dada una medida  $\mu$ , notaremos en la forma  $d\mu$  o bien  $\mu(dx)$  para integrar respecto de dicha medida. Asimismo, notaremos  $\mu(f) = \int f d\mu$  a la integral de  $f$  respecto a  $\mu$ .  $\mathbb{P}_\mu$  nota a  $\int \mathbb{P}_x d\mu$  cuando esto sea posible. La convergencia en distribución la notamos  $\mu_N \rightarrow \mu$  y significa que  $\int \phi(x) \mu_N(dx) \rightarrow \int \phi(x) \mu(dx)$  ( $N \rightarrow +\infty$ ) para toda  $\phi \in C_b(E)$ , es decir,  $\phi$  continua y acotada sobre  $E$ .

En la sección 2.1, realizaremos un breve repaso de los conceptos básicos de los procesos de Poisson y algunas propiedades útiles, así como también definiremos los procesos de Markov determinísticos a trozos. Luego en la sección 2.2, definiremos la distancia Wasserstein, y mostraremos su relación con los acoplamientos. Finalmente en la sección 2.3 repasaremos algunas ideas fundamentales del cálculo estocástico asociado a procesos puntuales de Poisson.

### 2.1. Procesos de Poisson

Los procesos de Poisson son una herramienta muy útil para modelar fenómenos con una fuerte componente aleatoria, sin un patrón claro, que a la vez poseen (entre otras cosas) una propiedad de independencia estadística como característica distintiva que resulta altamente provechosa desde el punto de vista matemático.

Recordemos que una variable aleatoria  $Z$  a valores en  $\mathbb{N}_0$  tiene distribución  $\mathcal{P}(\lambda)$  (es decir,

Poisson de parámetro  $\lambda > 0$ ) si  $\mathbb{P}[Z = k] = \frac{\lambda^k}{k!} e^{-\lambda}$ , lo cual implica que  $\mathbb{E}[Z] = \lambda$ . Una de las propiedades más distintivas de esta distribución es su  $\sigma$ -aditividad: dada una colección (finita o contable) de variables aleatorias independientes  $(X_i)_{i \in I}$  cada una con distribución  $\mathcal{P}(\lambda_i)$ , entonces  $S := \sum_I X_i$  tiene distribución  $\mathcal{P}(\sum_I \lambda_i)$  (con la convención de que  $Z \sim \mathcal{P}(+\infty)$  implica  $\mathbb{P}[Z = +\infty] = 1$ ). Esta propiedad se trasladará luego a los procesos puntuales de Poisson.

### 2.1.1. Procesos puntuales de Poisson

En esta sección, nos basaremos principalmente en [Kin92] y en [KS12].

#### Procesos puntuales de Poisson

Informalmente, un *proceso puntual de Poisson* (PPP) puede pensarse como una colección aleatoria de puntos en cierto espacio  $E$ . De esta manera, el proceso de Poisson homogéneo sería sencillamente un PPP en el espacio  $S = \mathbb{R}_+$ . Alternativamente, podemos pensar en estos procesos como en medidas de contar aleatorias.

Nos limitaremos a definir estos procesos en subconjuntos de  $\mathbb{R}^d$ , que son los de interés en el contexto de esta tesis, aunque esencialmente todos los resultados aquí mencionados son generalizables a espacios medibles más generales.

**Definición 2.1.1.** Sea  $S \subseteq \mathbb{R}^d$ . Un proceso puntual de Poisson en  $S$  es una colección contable aleatoria  $\mathcal{N}$  de puntos en  $S$  que cumple:

- i. Para toda colección  $A_1, \dots, A_n \in \mathcal{B}(S)$  disjuntos, las aplicaciones  $\mathcal{N}(A_i) := \#(\mathcal{N} \cap A_i)$  son variables aleatorias independientes y
- ii.  $\mathcal{N}(A) \sim \mathcal{P}(\mu(A))$  donde  $\mu(A) := \mathbb{E}[\mathcal{N}(A)]$

Por cómo fue definida,  $\mu$  resulta una medida sobre  $S$ , que llamamos la *intensidad* del proceso, que muchas veces está dada por una densidad  $\lambda(\mathbf{x})$  (que también suele denominarse intensidad), de manera tal que

$$\mu(A) = \int_A \lambda(\mathbf{x}) \, d\mathbf{x} \quad (2.1)$$

Cuando  $d = 1$  y  $S$  representa el tiempo, también suele denominarse la *tasa* del proceso. Otro caso particular destacable es cuando  $\lambda(\mathbf{x}) \equiv \lambda$ , lo cual implica  $\mu(A) = \lambda|A|$ , y decimos que el proceso es *homogéneo* (observar que, definido de esta manera, incluye a nuestro primer ejemplo, el proceso de Poisson homogéneo en  $\mathbb{R}_+$ ).

Observar también que si  $\mathcal{N}$  es un PPP en  $S$ , entonces la aplicación  $\omega \rightarrow \mathcal{N}_\omega(\cdot)$  es una medida de contar aleatoria. En la definición, hemos adoptado la convención adicional de que  $Z \sim \mathcal{P}(0) \implies \mathbb{P}[Z = 0] = 1$ .

El PPP tiene muchas propiedades interesantes que hacen de varios de sus cálculos un trabajo relativamente simple. El *teorema de superposición* es una de las más importantes, y es una consecuencia casi inmediata de la  $\sigma$ -aditividad de la distribución de Poisson. La prueba de este teorema puede encontrarse en [Kin92, Sección 2.2].

**Teorema 2.1.1.** (Teorema de superposición, Sección 2.2 en [Kin92])

Sean  $\mathcal{N}_1, \mathcal{N}_2, \dots$  una colección contable de PPP independientes en  $S$  y para cada  $k \in \mathbb{N}$  sea  $\mu_k$  la intensidad del proceso  $\mathcal{N}_k$ .

Entonces la superposición  $\mathcal{N} = \sum_k \mathcal{N}_k := \bigcup_k \mathcal{N}_k$  resulta un PPP en  $S$  con intensidad  $\mu = \sum_k \mu_k$

Un resultado relacionado, de prueba inmediata por definición, es el siguiente:

**Teorema 2.1.2.** (Teorema de restricción, Sección 2.2 en [Kin92])

Sea  $\mathcal{N}$  un PPP en  $S$  con intensidad  $\mu$ , y sea  $A$  un subconjunto medible de  $S$ . Entonces la restricción  $\mathcal{N}^A := \mathcal{N} \cap A$  resulta un PPP en  $A$  con intensidad  $\mu^A$  la restricción sobre  $A$ .

La segunda gran propiedad de estos procesos es el Teorema de Mapeo, que permite generar procesos de Poisson en espacios  $T \subseteq \mathbb{R}^l$  a partir de un PPP en  $S$  y una función medible  $f : S \rightarrow T$  con escasas condiciones adicionales.

**Teorema 2.1.3.** (Teorema de Mapeo, Sección 2.3 en [Kin92])

Sea  $\mathcal{N}$  un PPP en  $S$  con intensidad  $\sigma$ -finita  $\mu$ , y sea  $f : S \rightarrow T$  una función medible tal que la medida inducida  $\mu^*(B) := \mu(f^{-1}(B))$  no posee átomos. Entonces  $f(\mathcal{N})$  resulta un PPP en  $T$  con intensidad  $\mu^*$ .

Para ilustrar la utilidad de este teorema, extraemos el ejemplo dado en [Kin92, Sección 2.3]: Sea  $\mathcal{N}$  un PPP en  $\mathbb{R}^d$  con tasa  $\lambda(\mathbf{x})$ . Dado  $k < d$ , sea  $f = \pi_k$  la proyección sobre las primeras  $k$  coordenadas. Entonces, definiendo

$$\lambda^*(x_1, \dots, x_k) := \int_{\mathbb{R}^{d-k}} \lambda(\mathbf{x}) dx_{k+1} \dots dx_d \quad (2.2)$$

se tiene

$$\mu^*(B) = \int_{B \times \mathbb{R}^{d-k}} \lambda(\mathbf{x}) d\mathbf{x} = \int_B \lambda^*(x_1, \dots, x_k) dx_1 \dots dx_k \quad (2.3)$$

En este caso, si la integral (2.2) converge, estamos bajo las condiciones del teorema anterior y  $\pi_k(\mathcal{N})$  es un PPP con tasa  $\lambda^*$  en  $\mathbb{R}^k$ . Es decir, la proyección de un proceso de Poisson es un proceso de Poisson.

Es natural preguntarse si efectivamente existen procesos de Poisson con medidas de intensidad dadas. La respuesta a esa pregunta es afirmativa, con pocas restricciones para la medida elegida.

**Teorema 2.1.4.** (Teorema de existencia de PPP, Sección 2.5 en [Kin92] )

Sea  $\mu$  una medida  $\sigma$ -finita, no atómica, en  $S$ . Entonces existe un Proceso Puntual de Poisson con intensidad  $\mu$ .

### Teorema de Coloreo y Procesos de Poisson Marcados

Si una variable aleatoria  $Z$  tiene distribución  $\mathcal{P}(\mu)$ ,  $\mu > 0$  y se tiene una segunda variable aleatoria  $Y$  con distribución condicional a  $Z$  de tipo binomial  $B(Z, p)$ , entonces se puede mostrar que  $Y$  y  $Z - Y$  son independientes, Poisson, de parámetros  $p\mu$  y  $(1-p)\mu$  respectivamente. Este resultado tiene una importante consecuencia sobre los procesos de Poisson.

**Teorema 2.1.5.** (Teorema de Coloreo, Sección 5.1 en [Kin92])

Sea  $\mathcal{N}$  un PPP en  $S$  con medida de intensidad  $\mu$ . Supongamos que los puntos de  $\mathcal{N}$  se colorean con un color  $i \in \{1, \dots, k\}$  con probabilidad  $p_i$  de manera independiente entre sí. Si  $\mathcal{N}_i$  es el conjunto de puntos coloreados con  $i$ , entonces  $(\mathcal{N}_i)_{i=1 \dots k}$  forma una colección de PPP independientes con intensidad  $\mu_i := p_i \mu$ .

Observar que este resultado es consistente con el teorema de superposición, teniendo en cuenta que  $\mathcal{N}$  es la superposición de los  $\mathcal{N}_i$ .

Este teorema, si bien es interesante, puede resultar muy acotado en su poder de aplicación. La teoría de *procesos de Poisson marcados* generaliza este resultado:

Supongamos que tenemos un proceso de Poisson  $\mathcal{N}$  en  $S$ . La idea es que a cada punto  $X$  del proceso queremos asociarle una *marca*: una variable aleatoria  $m_X \in M$ , donde  $M \subseteq \mathbb{R}^l$  (más en general, puede ser un espacio medible), y donde la distribución sólo dependa -eventualmente- de  $X$ , de tal manera que  $m_X$  y  $m_{X'}$  sean independientes. Definiendo el conjunto aleatorio  $\mathcal{M} := \{(X, m_X); X \in \mathcal{N}\} \subseteq S \times M$ . El teorema de Marcado dice que este conjunto aleatorio es efectivamente un PPP en  $S \times M$ .

**Teorema 2.1.6.** (Teorema de Marcado, Sección 5.1 en [Kin92])

Sea  $\mathcal{N}$  PPP sobre  $S$  con intensidad  $\mu$  y  $\{p(x, \cdot)\}_{x \in S}$  una colección de probabilidades sobre  $M$  tal que  $p(\cdot, A)$  resulta medible para todo  $A \in \mathcal{B}(M)$ . Entonces el conjunto aleatorio  $\mathcal{M}$  definido arriba resulta un PPP sobre  $S \times M$  con medida de intensidad

$$\nu(C) := \iint_{(x,m) \in C} p(x, dm) \mu(dx) \quad (2.4)$$

Observar que, por el teorema de Mapeo (2.1.3), si tomamos  $f = \pi_S$  recuperamos exactamente el proceso  $\mathcal{N}$  original; mientras que si tomamos  $f = \pi_M$  obtenemos un PPP sobre  $M$  con medida de intensidad  $\hat{\nu} = \nu(S \times \cdot)$ . Observar también que si  $M = \{1, \dots, k\}$  y  $p_i(x) \equiv p_i$ , es decir, la colección de probabilidades en  $M$  es independiente de la posición  $x \in S$ , recuperamos el teorema de Coloreo. Otro caso particular y frecuente es cuando tomamos  $S = \mathbb{R}_+$ , donde usualmente (de nuevo) interpretamos  $S$  como el tiempo, y donde a tiempos poissonianos de intensidad  $\mu$  ocurren eventos en  $M$  guiados por la colección  $\{p(t, \cdot)\}_{t \in S}$ . En este caso, el proceso  $\mathcal{N}$  puede ser identificado con la sucesión  $(\tau_n, m_n)$  que representa los tiempos del proceso y sus marcas asociadas.

### 2.1.2. Procesos de Markov determinísticos a trozos

Esta sección se apoya principalmente en [ABG<sup>+</sup>14] y en [Dav18].

Los *Procesos de Markov determinísticos a trozos* o *PDMP* (*Piecewise deterministic Markov Processes*), originalmente presentados por [Dav84], forman una familia de procesos estocásticos càdlàg que cubre una amplísima variedad de procesos de Markov no difusivos, con aplicación a una igualmente amplia cantidad de modelos: biológicos [RTK15], sistemas dinámicos con control aleatoriamente perturbados [dVCD13], o teoría de colas [Dav84] entre otros.

Como su nombre lo indica, un PDMP es un proceso que mantiene una dinámica determinística dada por un *flujo*  $\phi$  hasta que se produce un salto con una *tasa*  $\lambda$  (generalmente dependiente del estado del proceso) de tipo Poisson, y pasa instantáneamente a un nuevo estado elegido por un *núcleo de transición*  $Q$ , para luego continuar evolucionando según la dinámica dada por el flujo hasta el siguiente salto. De esta manera, todo PDMP queda determinado por estas tres *características locales*:  $\phi, \lambda$  y  $Q$ .

Más formalmente, sea  $E$  un subconjunto de  $\mathbb{R}^N$  (que es suficiente en el contexto de esta tesis). Definimos ahora:

- El *flujo* es una familia uniparamétrica de funciones continuas  $\phi : E \times \mathbb{R}_+ \rightarrow E$  que cumplen la propiedad de semigrupo, es decir:

- i)  $\phi(\cdot, 0) \equiv \text{id}_E$
- ii)  $\phi(x, t + s) = \phi(\phi(x, s), t)$

La interpretación es que  $\phi(x, t)$  representa la posición del proceso a tiempo  $t$ , habiendo partido desde  $x \in E$  a tiempo cero.

- La *tasa de salto* es una función medible  $\lambda : E \rightarrow \mathbb{R}_+$  tal que

$$\forall x \in E, \exists \varepsilon = \varepsilon(x) > 0 / \quad \Lambda(x, \varepsilon) := \int_0^\varepsilon \lambda(\phi(s, x)) \, ds < +\infty \quad (2.5)$$

Esta condición busca asegurar la no multiplicidad de saltos. Es decir, si  $\{T_k\}_k$  es la sucesión de tiempos de salto del proceso,  $\mathbb{P}_x[T_k < T_{k+1}] = 1$ , sabiendo que (como se verá en la construcción) la distribución de  $T_1$  viene dada por

$$\mathbb{P}_x[T_1 > t] = e^{-\Lambda(x, t)} \quad (2.6)$$

Observar que esta condición no asegura la no acumulación de los tiempos de salto, i.e.  $T_k \rightarrow +\infty$  *c.s.*

- El *núcleo de transición*  $Q : E \times \mathcal{B}(E) \rightarrow [0, 1]$  es un *núcleo de Markov* en el sentido de que  $Q(\cdot, A)$  es  $\mathcal{B}(E)$ -medible para todo  $A \in \mathcal{B}(E)$  y  $Q(x, \cdot)$  es una medida de probabilidad sobre  $\mathcal{B}(E)$ .

$Q$  debe interpretarse como el mecanismo que le indica al proceso a dónde saltar desde un punto  $x$  dado, cuando la tasa de salto lo requiere. Pedimos además que  $Q(x, E \setminus \{x\}) = 1$  para evitar saltos “triviales”.

Con estas características locales, se puede asegurar la existencia (ver [Dav18]) de un espacio filtrado  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \{\mathbb{P}_x\})$  donde puede construirse trayectorialmente el proceso  $(X_t)$  desde  $x \in E$  de manera inductiva así: Sea  $T_0 \equiv 0$ ,  $X_0 = x$ , luego

1. Definimos la variable aleatoria  $J_1$  con distribución dada por (2.6) y  $T_1 := T_0 + J_1$
2. Si  $T_1 = \infty$ , entonces el proceso sigue el flujo para  $t > T_0$ , i.e.  $X_t = \phi(x, t)$ ;  $t > T_0$ . Caso contrario, definimos  $Z_1$  con distribución dada por  $Q(\phi(x, T_1), \cdot)$ .

3. Definimos la trayectoria para  $t \in (0, T_1]$  como

$$X_t = \begin{cases} \phi(x, t) & t < T_1 \\ Z_1 & t = T_1 \end{cases} \quad (2.7)$$

4. Comenzando desde  $X_{T_1} = Z_1$ , repetimos el procedimiento

Observar que naturalmente queda definida una *cadena de Markov*  $(\Phi_n)_n$  dada por  $\Phi_n := (Z_n, J_n)$  con  $\Phi_0 := (x, 0)$

Una suposición usual (que nuestro modelo cumple) será que para todo punto inicial  $x \in E$ , y para todo  $t \in \mathbb{R}_+$  vale  $\mathbb{E}_x[N_t] < +\infty$ , donde  $N_t = \sum_k \mathbf{1}(T_k \leq t)$  lo cual implica la no acumulación de saltos.

## 2.2. Couplings y distancia Wasserstein

En esta sección nos basaremos principalmente en [Lin92] y [FG00]. Para la distancia Wasserstein también usamos [Dud18] entre otros.

En muchos contextos es necesario con frecuencia comparar medidas de probabilidad, y el *acoplamiento* resulta una herramienta sumamente valiosa y eficiente para perseguir este objetivo. Construir un *acoplamiento* o *coupling* entre dos medidas de probabilidad significa esencialmente construir variables aleatorias, siguiendo sendas distribuciones, ***mediante el uso del mismo mecanismo aleatorio***.

Este concepto fue introducido en 1938 por Doeblin con el objetivo de probar la convergencia al equilibrio de una cadena de Markov. En los años subsiguientes, este método fue retomado por varios autores, y ha probado ser muy prolífico en la búsqueda de tiempos de convergencia (o *mixing times*), así como también en la comparación entre procesos donde hay cierta relación de monotonía, entre otras aplicaciones.

A modo de ejemplo, sea  $\mathcal{M}$  un PPP en  $\mathbb{R}^2$  homogéneo de tasa  $\mu = 1$  y sean  $\lambda_1 \leq \lambda_2$ . Entonces si definimos  $\mathcal{N}_i(I) := \mathcal{M}(I \times [0, \lambda_i])$ ;  $i = 1, 2$ , es inmediato (por lo expuesto en la sección 2.1.1) ver que son procesos de Poisson en  $\mathbb{R}$  homogéneos de tasa  $\lambda_i$  y que  $\mathcal{N}_1(I) \leq \mathcal{N}_2(I)$  *c.s.* para todo intervalo  $I \in \mathbb{R}$ . Análogamente, definiendo  $\hat{\mathcal{N}}_1(I) := \mathcal{M}(I \times [0, \lambda_1])$  y  $\hat{\mathcal{N}}_2(I) := \mathcal{M}(I \times [\lambda_1, \lambda_1 + \lambda_2])$ , obtenemos nuevamente 2 procesos de Poisson en  $\mathbb{R}$  homogéneos de tasas  $\lambda_1$  y  $\lambda_2$ , pero esta vez independientes. Observemos que usando esta misma idea, se puede probar (ver [FG00]) que dada una función continua  $\lambda : \mathbb{R} \rightarrow \mathbb{R}_+$  se puede construir un PPP  $\mathcal{N}$  inhomogéneo de tasa  $\lambda(t)$  en  $\mathbb{R}$  a partir de un proceso  $\mathcal{M}$  en  $\mathbb{R}^2$  como el anterior, definiendo  $\mathcal{N}(I) := \mathcal{M}(\text{Graph}_\lambda(I))$  donde  $\text{Graph}_\lambda(I) := \{(t, y) \in \mathbb{R}^2 / t \in I, y \in [0, \lambda(t)]\}$ . Estas construcciones son interesantes porque ilustran lo “artesanales” que pueden ser las construcciones de acoplamientos, lejos de una receta general, si bien algunas ideas se repiten en ciertos contextos.

Existen diversas versiones de definición formal de acoplamiento o *coupling*. Nosotros damos la siguiente (ver [Lin92]):

**Definición 2.2.1.** Sea  $(\Omega, \mathcal{F})$  un espacio medible y  $\mu, \nu$  dos medidas de probabilidad en dicho espacio. Un *acoplamiento* de  $\mu$  y  $\nu$  es un par de variables aleatorias  $(X, Y)$  en  $(\Omega \times \Omega, \mathcal{F} \otimes \mathcal{F}, \Gamma)$  donde

$$\Gamma(X \in A) = \mu(A) \quad \text{y} \quad \Gamma(Y \in A) = \nu(A) \quad (2.8)$$

es decir,  $(X, Y)$  tiene distribuciones marginales  $\mu$  y  $\nu$  respectivamente.

Una definición alternativa, esencialmente igual que la anterior, es definir el coupling directamente como la medida de probabilidad  $\Gamma$  de la definición anterior. Asimismo, se define el coupling entre dos variables aleatorias  $X_1$  y  $X_2$  simplemente como el coupling entre sus distribuciones  $\mu_1$  y  $\mu_2$  respectivamente. Observar que la definición no implica que distribución conjunta de  $(X, Y)$  sea  $\mu \otimes \nu$ . Es decir, no tienen por qué ser independientes (y en general no se busca que lo sean).

Como dijimos anteriormente, los acoplamientos buscan comparar el comportamiento de dos -o más- variables aleatorias (o más generalmente, elementos aleatorios). Es por esto que resulta natural la necesidad de medir de alguna manera la diferencia entre medidas de probabilidad/variables aleatorias. De esta manera, hay más de una definición de distancia que se utiliza con frecuencia en la literatura.

La *distancia en variación total* entre dos distribuciones  $\mu$  y  $\nu$  se define como

$$\|\mu - \nu\|_{\text{tv}} := \sup_{A \in \mathcal{F}} |\mu(A) - \nu(A)|. \quad (2.9)$$

Existe una íntima relación entre la variación total y el acoplamiento entre dos probabilidades, como ilustra la siguiente desigualdad. Sea  $(Z_\mu, Z_\nu)$  un acoplamiento arbitrario de  $\mu$  y  $\nu$ . Entonces

$$\|\mu - \nu\|_{\text{tv}} = \sup_{A \in \mathcal{F}} |\Gamma(Z_\mu \in A) - \Gamma(Z_\nu \in A)|, \quad (2.10)$$

pero dado  $A \in \mathcal{F}$ ,

$$\begin{aligned} \Gamma(Z_\mu \in A) - \Gamma(Z_\nu \in A) &\leq \Gamma(Z_\mu \in A, Z_\nu \notin A) \\ &\leq \Gamma(Z_\mu \neq Z_\nu), \end{aligned}$$

y análogamente  $\Gamma(Z_\nu \in A) - \Gamma(Z_\mu \in A) \leq \Gamma(Z_\mu \neq Z_\nu)$ , por lo tanto

$$\|\mu - \nu\|_{\text{tv}} \leq \inf_{\mathcal{C}(\mu, \nu)} \{\Gamma(Z_\mu \neq Z_\nu)\} \quad (2.11)$$

Donde  $\mathcal{C}(\mu, \nu)$  denota el conjunto de todos los couplings entre  $\mu$  y  $\nu$ . Notar que para que esto tenga sentido, necesitamos que  $\{Z_\mu = Z_\nu\}$  sea medible. Basta con que el espacio de estados de las variables sea un *espacio polaco*  $E$ , de manera tal que la diagonal  $\Delta = \{(z, z') \in E \times E / z = z'\} \in \mathcal{B}(E \times E)$ , entonces la condición se cumple.

La desigualdad (2.11) resulta de hecho una igualdad (ver [Lin92, Teorema 5.2]), y todo acoplamiento que realiza la igualdad se llama *acoplamiento maximal*. Notemos que un acoplamiento maximal maximiza la probabilidad de que  $Z_\mu$  y  $Z_\nu$  coincidan.

## La distancia Wasserstein

Existe otra métrica para las probabilidades estrechamente relacionada con los acoplamientos: La *distancia Wasserstein* ha sido denominada así por primera vez por Dobrushin en 1970 a partir del trabajo de Wasserstein [Vas69]. Esta métrica sin embargo fue definida antes por Kantorovich [Kan39], motivo por la que a veces es denominada también “distancia Kantorovich” o “métrica Kantorovich”.

Sea  $(E, d)$  un espacio Polaco con su  $\sigma$ -álgebra  $\mathcal{B}(E)$  y sea  $\mathcal{P}(E)$  el conjunto de las probabilidades en  $E$ . Dado  $1 \leq p < \infty$ , sea  $\mathcal{P}_p(E)$  el conjunto de las probabilidades en  $E$  tal que  $g_y := d(\cdot, y)$  tiene  $p$ -momento finito para algún  $y \in E$  (equivalentemente, para todo  $y$ ).

**Definición 2.2.2.** Dadas  $\mu, \nu \in \mathcal{P}_p(E)$  definimos la distancia  $p$ -Wasserstein entre  $\mu$  y  $\nu$  como

$$W_p(\mu, \nu) := \inf_{\mathcal{C}(\mu, \nu)} \left[ \int d(x, y)^p d\Gamma(x, y) \right]^{1/p}. \quad (2.12)$$

Observar que podemos escribirla de manera equivalente como

$$W_p(\mu, \nu) := \inf_{\substack{X \sim \mu \\ Y \sim \nu}} [ \mathbb{E} [d(X, Y)^p] ]^{1/p}. \quad (2.13)$$

Para  $p = 1$ , la distancia 1-Wasserstein (o directamente distancia Wasserstein) queda:

$$W_1(\mu, \nu) := \inf_{\mathcal{C}(\mu, \nu)} \int d(x, y) d\Gamma(x, y) = \inf_{\substack{X \sim \mu \\ Y \sim \nu}} \mathbb{E} [d(X, Y)]. \quad (2.14)$$

Nosotros nos enfocaremos en esta última distancia, que es la que utilizaremos en este trabajo. Notar que para nosotros  $E$  será en general un subconjunto de  $\mathbb{R}^n$  con la distancia usual, o bien con la *distancia 1 normalizada*, es decir,  $d(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{i=1}^n |x_i - y_i|$ , que es una distancia usual en el contexto de propagación del caos.

Que la ecuación (2.14) define una métrica no es trivial, ver [Dud18]. Si el espacio  $(E, d)$  es Polaco (como en nuestro caso), se puede probar que  $(\mathcal{P}_1(E), W_1)$  es también Polaco (ver por ejemplo [Bog18, Teorema 3.2.2]). Además, existe un *plan óptimo*  $\Gamma_0$  que realiza el ínfimo, ver [Bog18, Teorema 3.2.7].

Una forma de entender esta métrica es viéndolo como un problema de transporte óptimo. Informalmente, podemos explicarlo así: supongamos que tenemos una “pila de arena” distribuida en el espacio  $E$  como  $\mu$ , es decir que podemos pensar que en torno al sitio  $x$  hay masa “ $\mu(x) dx$ ”; y queremos desplazarla para formar una pila distribuida como  $\nu$ . Observar que al tener ambas la misma masa, no hay creación ni destrucción de masa. Un plan de transporte podría ser descrito mediante una función  $\Gamma(x, y)$  que indica cuánta masa transportar de  $x$  a  $y$ . Tiene sentido entonces pedir que  $\int \Gamma(x, y) dx = \nu(y)$  y  $\int \Gamma(x, y) dy = \mu(x)$ , es decir, que la cantidad de arena desplazada *hacia* el sitio  $y$  debe ser  $\nu(y)$  y la cantidad desplazada *desde*  $x$  debe ser  $\mu(x)$  respectivamente. Esto se corresponde con la idea intuitiva de que  $\Gamma$  representa la distribución conjunta con marginales  $\mu$  y  $\nu$ . Finalmente, supongamos que el transporte de masa de  $x$  a  $y$  tiene un costo  $c(x, y)$  que elegimos equivalente a la distancia  $d(x, y)$  entre los puntos. Por lo tanto el costo total de transportar la pila  $\mu$  en la pila  $\nu$  vía el plan  $\Gamma$  equivale a



$\iint d(x, y) \Gamma(x, y) dx dy = \iint d(x, y) d\Gamma(x, y)$ . Si nos interesa el plan óptimo, tomamos entonces el ínfimo sobre los planes  $\hat{\Gamma}$  que cumplen las condiciones que cumplía  $\Gamma$ , y esto coincide con  $W_1(\mu, \nu)$ .

Un ejemplo trivial que reafirma esta idea intuitiva es el caso de dos distribuciones degeneradas  $\mu = \delta_{x_0}$  y  $\nu = \delta_{y_0}$ . En ese caso, el único acoplamiento posible es  $\delta_{(x_0, y_0)}$  y por lo tanto  $W_1(\mu, \nu) = \|x_0 - y_0\|_2$ , es decir, el costo óptimo de transportar toda la masa de  $x_0$  en  $y_0$  es exactamente la distancia entre ambos puntos. El valor para  $p > 1$  es el mismo en este caso.

Esta métrica está íntimamente relacionada con la convergencia en distribución.

**Teorema 2.2.1.** (Metrización de  $\mathcal{P}_1(E)$ , Teorema 5.9 en [Vil09])

*Dada una secuencia de distribuciones  $(\mu_N)$  en  $\mathcal{P}_1(E)$  y otra distribución  $\mu \in \mathcal{P}_1(E)$  entonces son equivalentes*

I.  $W_1(\mu_N, \mu) \rightarrow 0$

II.  $\mu_N \rightarrow \mu \quad y \quad \mu_N(d(x_0, x)) \rightarrow \mu(d(x_0, x)) \quad \forall x_0 \in E$

### 2.3. Cálculo estocástico

Esta sección se basa principalmente en [Mie06], [Bic02] y en [Oks13].

Las ecuaciones diferenciales clásicas son ecuaciones donde la solución es una función (o familia de funciones), y donde ésta se encuentra relacionada con derivadas de ella misma. Es innecesario remarcar lo prolífica que ha sido esta herramienta dentro y fuera de las matemáticas, entre otras cosas, para modelar fenómenos de la realidad a partir de conocimientos sobre su dinámica.

Una ecuación diferencial ordinaria de primer orden suele tener la forma general

$$\frac{dx}{dt} = f(x(t), t)$$

para alguna función  $f$  adecuada. A veces una ecuación diferencial de este tipo se expresa mediante su versión integral, esto es

$$x(t) = x_0 + \int_0^t f(x(t), t) dt \tag{2.15}$$

que se suele notar de forma compacta como

$$dx = f dt . \tag{2.16}$$

Las ecuaciones diferenciales estocásticas (SDE por sus siglas en inglés) agregan a la dinámica un componente aleatorio, que permite agregar comportamientos complejos o “caóticos” que agregan realismo a muchos modelos matemáticos con ecuaciones diferenciales. En este caso, las soluciones a SDE’s resultan procesos estocásticos. Tomamos el siguiente ejemplo de [Oks13] para ilustrar esto: Consideremos el modelo clásico simple de crecimiento poblacional

$$\frac{dx}{dt} = a(t)x(t); \quad x(0) = x_0 \quad (2.17)$$

donde  $x(t)$  es el tamaño de la población a tiempo  $t$ , y  $a(t)$  es la tasa relativa de crecimiento a tiempo  $t$ . En este caso la solución, que tiene la forma  $x(t) = x_0 \cdot \exp(\int_0^t a(s)ds)$ , representa una trayectoria determinística. Sin embargo, no tenemos por qué tener información exacta sobre  $a$ , que podría estar afectada por efectos ambientales complejos, de manera que en lugar de tener en la ecuación anterior una función  $a$  determinística, tenemos  $A(t) = \alpha(t) + \text{“ruido”}$ . De esta manera, la solución a la nueva ecuación diferencial ya no será una trayectoria determinística  $x(t)$  sino más bien un proceso estocástico  $X_t$ , que puede interpretarse como una probabilidad sobre el espacio de trayectorias posibles.

De esta manera, vamos a dar un significado a ecuaciones diferenciales de la forma

$$dX = f dt + g dM \quad \equiv \quad X_t = X_0 + \int_0^t f dt + \int_0^t g dM, \quad (2.18)$$

donde la variación estocástica está condensada en el último término de cada expresión, siendo  $M$  un proceso estocástico y consecuentemente la segunda integral resulta una *integral estocástica*. Daremos un sentido más preciso a estos conceptos en breve.

La primer clase de ecuaciones diferenciales estocásticas, que es también la más estudiada y desarrollada, es aquella donde la componente aleatoria está dada por *Movimientos Brownianos*. La integral estocástica asociada  $\int_0^t g dB$  se denomina *integral de Itô* y el estudio de esta clase de SDE's se denomina *cálculo de Itô*. Kiyoshi Itô, apoyándose en el trabajo de Wiener, fue el primero en aportar una teoría sólida para estas integrales, y para el cálculo integro-diferencial asociado. Heurísticamente, el ruido  $dB$  busca modelar un gran número de perturbaciones infinitesimales, aleatorias, e independientes, y el teorema central del límite nos indica que la distribución debería ser la de una normal; busca también reflejar que la aleatoriedad sea estacionaria en el tiempo, y mantenga independencia en intervalos de tiempo disjuntos. Estas imposiciones sobre el componente aleatorio es consistente con la elección de *Movimientos Brownianos*.

Procesos dados por este tipo de ruido son por tanto de tipo esencialmente difusivo y tienen numerosas aplicaciones en física, ingeniería, finanzas, y también dentro de las matemáticas, como por ejemplo en teoría de potencial, entre tantos otros. El cálculo de Itô asociado es robusto y está ampliamente estudiado, y esto contribuye a su uso frecuente al momento de modelar fenómenos.

Sin embargo, esto no agota todo el ecosistema de ecuaciones diferenciales con aleatoriedad. Muchos fenómenos presentan un ruido que no se comporta como el dado por el *Movimiento Browniano*. En particular, las soluciones de SDE's dadas por el cálculo de Itô pueden ser en general utilizadas ellas mismas como el componente aleatorio que modifica a otro sistema [Bic02].

### 2.3.1. Integrales estocásticas respecto a medidas de Poisson

Otra fuente de aleatoriedad para integrales estocásticas está dada por las medidas aleatorias de Poisson. Éstas permiten modelar sistemas donde la componente aleatoria provoca cambios abruptos en el estado del sistema, “saltos”, como por ejemplo la señal de un contador Geiger, o un modelo de redes neuronales describiendo la evolución temporal del potencial intermembrana, como es el caso de este trabajo.

Para una teoría de SDE’s con medidas de salto generales, ver [Bic02]. Nosotros nos enfocaremos en el caso particular de los procesos puntuales de Poisson vistos anteriormente en 2.1.1, basándonos sobre todo en [Mie06, Capítulo 7]. Definimos entonces la integral sobre la medida aleatoria de Poisson trayectorialmente:

**Definición 2.3.1.** Sea  $\mathcal{N}$  un PPP sobre  $S$  con intensidad  $\mu$ . Sea  $f : S \rightarrow \mathbb{R}_+$  medible. Entonces definimos la integral de  $f$  respecto a  $\mathcal{N}$  como

$$\mathcal{N}(f)(\omega) := \left( \int f(x) \mathcal{N}(dx) \right) (\omega) . \quad (2.19)$$

La proposición 7.2.1 de [Mie06] nos asegura que efectivamente  $\mathcal{N}(f)$  resulta una variable aleatoria y relaciona su valor esperado con la intensidad del proceso:

**Proposición 2.3.1.** (Proposición 7.2.1 en [Mie06])

Dados  $\mathcal{N}$  y  $f$  como en la definición anterior,  $\mathcal{N}(f)$  resulta una variable aleatoria y

$$\mathbb{E}[\exp(-\mathcal{N}(f))] = \exp \left( - \int_S (1 - e^{-f(x)}) \mu(dx) \right) . \quad (2.20)$$

Más aún, si  $f : S \rightarrow \mathbb{R}$  es medible y  $\mu$ -integrable entonces  $\mathcal{N}(f)$  es una variable aleatoria,  $f$  es  $\mathcal{N}$ -integrable c.s. y vale

$$\mathbb{E}[\exp(i\mathcal{N}(f))] = \exp \left( \int_S (e^{if(x)} - 1) \mu(dx) \right) \quad (2.21)$$

*Observación 2.3.1.* Observar que a partir del resultado anterior, podemos obtener el momento de  $\mathcal{N}(f)$  para cualquier  $f \geq 0$  o  $\mu$ -integrable: reemplazando  $f$  por  $af$  en la fórmula (2.20), diferenciando, haciendo tender  $a \rightarrow 0$ , y haciendo uso de convergencia monótona (si  $f \geq 0$ ) o mayorada (si  $f$  es  $\mu$ -integrable); obteniendo de esta manera

$$\mathbb{E}[\mathcal{N}(f)] = \int_S f(x) \mu(dx) . \quad (2.22)$$

Un *proceso de Lévy*  $X$  es un proceso estocástico a valores en  $\mathbb{R}$  con incrementos independientes y estacionarios. Por ejemplo, el Proceso de Poisson homogéneo es un proceso de Lévy.

Nos interesa aplicar este concepto de integral estocástica vía procesos de Poisson en el contexto de generar procesos estocásticos a partir de ecuaciones integro-diferenciales. Es por eso pertinente el siguiente resultado previo: Dado  $(E, \mathcal{B}(E))$ , sea  $G$  una medida  $\sigma$ -finita sobre  $E$ . Consideremos la medida producto  $\nu = dt \otimes G(dx)$  en  $\mathbb{R}_+ \times E$ , con  $dt$  la medida usual de Lebesgue en  $\mathbb{R}_+$ .

**Proposición 2.3.2.** (Proposición 7.3.1 en [Mie06])

Dado un proceso de Poisson  $\mathcal{N}$  con medida de intensidad  $\mu$  como la definida arriba, y dada  $f : E \rightarrow \mathbb{R}$   $G$ -integrable, entonces el proceso

$$\mathcal{N}_t(f) := \int_0^t \int_E f(x) \mathcal{N}(dx, ds), \quad t \geq 0 \quad (2.23)$$

resulta un proceso de Lévy. Más aún, el proceso

$$\mathcal{M}_t(f) := \int_0^t \int_E f(x) \mathcal{N}(dx, ds) - t \int_E f(x) G(dx), \quad t \geq 0 \quad (2.24)$$

resulta una martingala respecto a la filtración  $\mathcal{F}_t := \sigma(\mathcal{N}([0, s] \times A); A \in \mathcal{B}(E))$ ,  $t \geq 0$ .

Podemos generalizar un poco más este resultado. En primer lugar, consideremos un proceso marcado en  $\mathbb{R}_+ \times E$  con medida  $\nu = \mu(dt) \otimes G(dx)$ . Tomando la filtración anterior, definimos  $\hat{\mathcal{P}}(\mathcal{F}_t)$  como la  $\sigma$ -álgebra  $\mathcal{P}(\mathcal{F}_t) \otimes \mathcal{B}(E)$ . Dada una aplicación  $H : \mathbb{R}_+ \times \Omega \times E \rightarrow \mathbb{R}$ , la llamamos proceso  $\mathcal{F}_t$ -predecible indexado por  $E$  si es  $\hat{\mathcal{P}}(\mathcal{F}_t)$ -medible. Claramente, esta  $\sigma$ -álgebra está generada por aplicaciones de la forma:

$$C_t(\omega) \mathbf{1}(x \in A)$$

donde  $A \in \mathcal{B}(E)$  y  $C$  es un proceso  $\mathcal{F}_t$ -predecible. Finalmente, definamos  $q(dt, dx) := \mathcal{N}(dt, dx) - \nu(dt, dx)$ .

**Teorema 2.3.1.** (Teorema T3 y Corolario C4 en [Bré81])

Sean  $\mathcal{N}$  y  $H$  definidas arriba,  $H$  no-negativa, entonces

$$\mathbb{E} \left[ \int_0^{+\infty} \int_E H \mathcal{N}(ds, dx) \right] = \mathbb{E} \left[ \int_0^{+\infty} \int_E H \nu(ds, dx) \right]. \quad (2.25)$$

Más aún, para toda  $H$  tal que para todo  $t \geq 0$  vale  $\mathbb{E} \left[ \int_0^{+\infty} \int_E H \nu(ds, dx) \right] < +\infty$  entonces

$$\int_0^t \int_E H q(ds, dx) := \int_0^t \int_E H \mathcal{N}(ds, dx) - \int_0^t \int_E H \nu(ds, dx) \quad (2.26)$$

resulta una  $(\mathcal{F}_t)$ -martingala.

### 2.3.2. Ecuaciones diferenciales estocásticas

Con estos elementos, nos interesa definir procesos estocásticos que sean solución de ecuaciones de la forma

$$X_t = X_0 + \int_0^t a(s, X_s) ds + \int_0^t \int_E b(s, X_s, z) \mathcal{N}(ds, dz). \quad (2.27)$$

Es común en la literatura encontrar la notación diferencial

$$dX_t = a(t, X_t) dt + \int_E b(t, X_t, z) \mathcal{N}(dt, dz); \quad X(0) = X_0. \quad (2.28)$$

Analicemos heurísticamente la ecuación (2.28): el diferencial  $dX_t = X_{t+dt} - X_t$  denota el crecimiento infinitesimal del proceso  $X$  en el intervalo  $[t, t + dt)$ . Si durante este intervalo no

ocurren eventos del proceso  $\mathcal{N}$  entonces  $\mathcal{N}(dt, dz) = 0$  y el incremento para  $X$  es  $a(t, X_t) dt$ . Si ocurre un evento con marca  $Z$ , entonces el incremento es  $a(t, X_t) dt + b(t, X_t, z)$ , y observar que el primer término es despreciable respecto al segundo.

El estudio de estos objetos es vasto y abarca una gran riqueza de modelos. Algunas referencias para una revisión completa son [Bré81], [SM12] o en un contexto más general [Bic02]. En esta tesis abordaremos la existencia y unicidad de una de estas ecuaciones.



## Capítulo 3

# Campo medio y propagación del caos

Como explicamos en el capítulo introductorio, la teoría de campo medio surge a partir de estudiar grandes sistemas de partículas estadísticamente indistinguibles que interactúan entre sí mediante mecanismos microscópicos que se asumen conocidos. Los modelos de grandes sistemas de partículas son utilizados tanto para describir modelos físicos como para aproximar un modelo continuo.

Intrínsecamente asociados a este contexto están aquellos conceptos referidos al *límite de campo medio* (a veces denominado *límite hidrodinámico*) y la *propagación del caos*, que guardan entre sí una íntima relación. El propósito de este capítulo es en primer lugar introducir dichos conceptos, y luego realizar una breve revisión de diversas técnicas utilizadas para abordar la propagación del caos en modelos concretos, en particular introduciendo la técnica que abordamos en este trabajo, que se apoya en las ideas de [CF16], inspirados a su vez en [FM16] y [FGM08].

Para el desarrollo de este capítulo, nos apoyamos principalmente en la reseña [CD21].

En la sección 3.1 introducimos los conceptos de *límite hidrodinámico* y *medida empírica* y su rol en la teoría de campo medio. Luego en la sección 3.2 definimos distintas nociones de *propagación del caos* y su relación con el límite hidrodinámico. En la sección 3.3, realizamos una breve reseña de las técnicas utilizadas en la literatura para probar propagación del caos; y finalmente en la sección 3.4 describimos brevemente la estrategia utilizada en este trabajo.

### 3.1. Campo medio y distribuciones empíricas

Si un sistema de partículas de tamaño  $N$  está representado por un proceso  $X^N = (X_t^N)_{t \geq 0}$ , entonces la aproximación de campo medio consiste “tomar límite” (en algún sentido) en  $N$  de la secuencia de procesos  $(X^N)_N$  para obtener información del comportamiento típico de las partículas, así como también descubrir lo que ampliamente se denomina *transición de fase*,

es decir, grandes cambios cualitativos en el comportamiento del sistema.

La literatura a veces distingue entre *límite hidrodinámico* y *límite de campo medio* [Spo12]. Esencialmente, en el segundo caso no existe localidad de interacciones: potencialmente cualquier partícula puede interactuar con otra, y los parámetros de interacción se escalan a medida que la población crece, funcionando de esta manera como escalamiento espacial. Por otro lado, es usual en todos los casos realizar eventualmente un escalamiento temporal del proceso: lo fundamental es que el escalamiento espacio-temporal elegido resulte en un desplazamiento promedio de orden 1 por unidad de tiempo de la escala temporal elegida.

Para ejemplificar, consideremos un sistema de  $N$  partículas en  $\Gamma \subseteq \mathbb{R}^d$  con interacciones binarias representadas por una aplicación  $H$  y un componente estocástico representando otras fuerzas azarosas actuando sobre cada partícula, como por ejemplo, interacciones de cada partícula con el medio en que se encuentra. Una forma general para esta dinámica es

$$dY_t^i = \frac{1}{N} \sum_{j \neq i} H(Y_t^i, Y_t^j) dt + \sqrt{2\sigma} dW_t^i \quad i \in \{1, \dots, N\} \quad (3.1)$$

donde  $W^i$  son movimientos Brownianos independientes. El factor  $1/N$  es el *escalamiento de campo medio* y es el que formalmente mantiene de orden 1 la dinámica total en el sistema. La ecuación de arriba abarca un amplio espectro de fenómenos, por ejemplo, tomemos  $Y_t^i = (X_t^i, V_t^i) \in E := \Gamma \times \mathbb{R}^d$  representando la posición y velocidad de una partícula a tiempo  $t$ , consideremos  $W^i = (0, B^i)$  donde  $B^i$  son movimientos Brownianos en  $\mathbb{R}^d$  y  $H(Y^i, Y^j) = (V^i, K(X^i - X^j))$  donde el *núcleo*  $K$  representa las fuerzas de interacción newtonianas entre 2 cuerpos (un núcleo usual es el *núcleo de Poisson*  $K(x) = C_d x \|x\|^{-d}$ , donde  $C_d$  es una constante dependiente del fenómeno físico y la dimensión). En este caso, se trata del clásico modelo newtoniano para  $N$  partículas indistinguibles y la ecuación (3.1) queda de la forma

$$dX_t^i = V_t^i dt ; \quad dV_t^i = \frac{1}{N} \sum_{j \neq i} K(X_t^i - X_t^j) dt + \sqrt{2\sigma} dB_t^i \quad i \in \{1, \dots, N\} \quad (3.2)$$

y lo que se espera, al menos formalmente, es que cuando la cantidad de partículas  $N$  va a infinito, las ecuaciones que gobiernan la dinámica para el sistema de partículas “tienden” a una ecuación diferencial continua. En este caso, la candidata a *límite de campo medio* es la ecuación de McKean-Vlasov

$$\partial_t f + v \cdot \nabla_x f + (K * \rho) \cdot \nabla_v f = \sigma \Delta_v f \quad (3.3)$$

donde  $f = f(t, x, v)$  es la distribución de partícula y  $\rho = \rho(t, x)$  es la densidad espacial de partículas. Ver [JW17].

### 3.1.1. La medida empírica

Dijimos que esperamos cierta “convergencia” cuando la cantidad de partículas  $N$  va a infinito, por lo que en particular, el espacio de estados del sistema crece con  $N$  ¿En qué sentido es esta convergencia entonces? Un objeto de interés que cuantifica esta convergencia es la llamada *medida empírica*:



**Definición 3.1.1.** Sea  $\mathbf{x} = (x_1, x_2, \dots, x_N) \in E^N$ . Definimos la medida empírica asociada a  $\mathbf{x}$  como

$$\bar{\mathbf{x}} = \mu_{\mathbf{x}}^N := \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \quad (3.4)$$

La medida empírica entonces define, para cada  $\mathbf{x}$  fijo, una distribución sobre  $E$  tal que para cada  $A \in \mathcal{B}(E)$ ,  $\mu_{\mathbf{x}}^N(A)$  representa la proporción de puntos  $x_i$  que se encuentran en  $E$ . De esta manera, asociada a la medida empírica tenemos una aplicación llamada de igual manera  $p_N := \mu_{(\cdot)}^N : E^N \rightarrow \mathcal{P}(E)$ . Notamos al conjunto de estas medidas (que coinciden con la imagen de  $p_N$ ) como  $\widehat{\mathcal{P}}_N(E) \subseteq \mathcal{P}(E)$ . Observar también que existe una identificación entre  $\mu_{\mathbf{x}}^N$  y la clase  $\tilde{\mathbf{x}}$  de todos los puntos en  $E^N$  bajo el grupo de permutaciones sobre  $I_N$ .

Si  $\mathbf{X}^N$  es una variable aleatoria en  $E^N$  (que simplemente notaremos  $\mathbf{X}$ ), entonces podemos considerar la medida empírica  $\bar{\mathbf{X}} = \mu_{\mathbf{X}}^N$  asociada a  $\mathbf{X}$ , que resulta de componer  $p_N \circ \mathbf{X}$ . Suponemos aquí que  $\mathbf{X}$  es *intercambiable*, en el sentido de que la distribución conjunta  $\mathbf{X}$  es invariante bajo las permutaciones sobre  $I_N$ , lo cual va en línea con el estudio de un sistema de partículas indistinguibles. Observar que de esta manera obtenemos una *medida aleatoria de probabilidad*  $\bar{\mathbf{X}} \in \mathcal{P}(\mathcal{P}(E))$ , cuya ley es el *push-forward* de la distribución  $f^N$  de  $\mathbf{X}$  vía la aplicación  $p_N$ . Observar también que fijado  $A \in \mathcal{B}(E)$ ,  $\mathbb{E} [\bar{\mathbf{X}}(A)(\omega)] = \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N \delta_{X_i}(A) \right] = \mathbb{P} [X_1 \in A]$ , donde la última igualdad vale por intercambiabilidad. Más aún, para toda  $\phi \in C_b(E)$ , vale

$$\mathbb{E} [\bar{\mathbf{X}}(\phi)] = f^{1,N}(\phi) \quad (3.5)$$

donde  $f^{k,N} \in \mathcal{P}(E^k)$  nota la  $k$ -ésima marginal de  $f^N$ . De esto se ve que a partir de la medida empírica  $\bar{\mathbf{X}}_t$  se puede reconstruir la ley de una partícula individual. Más aún, se puede ver que caracteriza la ley conjunta de  $\mathbf{X}_t^N$ .

## 3.2. Propagación del Caos y límite hidrodinámico

En la búsqueda de describir el límite de la dinámica de  $\mathbf{X}_t^N$  cuando  $N \rightarrow +\infty$ , varias nociones pueden surgir, pero generalmente todas convergen en torno a un objeto *no lineal* que describe el comportamiento *promedio* del sistema.

Una idea natural es estudiar el límite en  $N$  de las marginales  $f_t^{k,N}$  de  $\mathbf{X}_t^N$  para todo  $k \in \mathbb{N}$ . En este sentido, el concepto central de *propagación del caos* consiste en que para todo  $t \in \mathbb{R}_+$ , existe  $f_t \in \mathcal{P}(E)$  tal que

$$\forall k \in \mathbb{N}; \quad f_t^{k,N} \longrightarrow f_t^{\otimes k} \quad (N \rightarrow +\infty) \quad (3.6)$$

para la convergencia débil de medidas de probabilidad, siempre que (3.6) valga para  $t = 0$ ; y donde  $\mu^{\otimes k}$  denota una medida en  $\mathcal{P}(E^k)$  con marginales i.i.d. de distribución  $\mu$ . A  $f_t^N$  (y por extensión, a  $\mathbf{X}_t^N$ ) con esta propiedad a veces se la llama  $f_t$ -caótica. Por lo tanto, esta definición de propagación del caos nos dice que si  $f_0^N$  es  $f_0$ -caótica, entonces para todo tiempo  $t$  posterior  $f_t^N$  es  $f_t$ -caótica.

Puede probarse que (3.6) es equivalente a la convergencia en distribución

$$\bar{\mathbf{X}}_t \longrightarrow f_t \quad (N \rightarrow +\infty) \quad (3.7)$$

Ver [Szn91, Proposición 2.2].

La primera definición nos indica que bajo propagación del caos, tomando cualquier cantidad  $k$  de partículas de entre las  $N$ , su comportamiento se vuelve asintóticamente independiente con  $N$  estadísticamente hablando. Esto condensa la idea proveniente de la física de que para sistemas grandes de partículas, las correlaciones entre dos (o más) partículas dadas se vuelven despreciables. Si miramos entonces al sistema como un todo, lo que se observa es un comportamiento ponderado, en lugar de las trayectorias correlacionadas específicas de cada partícula. Este nivel de descripción es lo que en la física estadística se denomina *escala mesoscópica*.

Por lo expuesto cobra suma relevancia la descripción del límite hidrodinámico, esto es, el objeto límite  $f_t$ , que será típicamente la solución de una SDE no lineal. En muchos casos, el programa consiste en

- Probar la propagación del caos.
- Identificar el límite  $f_t$  como la solución de una ecuación no lineal específica.
- Probar la buena definición y unicidad de solución para la ecuación no lineal.

Cabe mencionar que en cuanto a propagación del caos se refiere, muchas veces probar la mera convergencia es informativo aunque no lo suficiente, y por lo tanto se recurre a la búsqueda de estimaciones cuantitativas de la *velocidad de convergencia*, es decir, estimaciones específicas que aporten cotas de convergencia dependiendo explícitamente de la cantidad  $N$  de partículas y del tiempo  $t$ . En este sentido es útil estudiar la propiedad de ser  $f$ -caótica de una secuencia  $(f^N)$  en términos de la distancia Wasserstein, perspectiva introducida en [HM14]:

- *Caos  $p$ -Wasserstein marginal:*  $\forall k \in \mathbb{N}$

$$W_p(f^{k,N}, f^{\otimes k}) \longrightarrow 0 \quad (N \rightarrow +\infty)$$

- *Caos  $p$ -Wasserstein empírico:* Si  $X^N \sim f^N$  y  $\bar{\mathbf{X}} = \mu_{X^N}^N$

$$\mathbb{E} \left[ W_p(\bar{\mathbf{X}}, f) \right] \longrightarrow 0 \quad (N \rightarrow +\infty)$$

Cuando  $p = 1$  y se tienen momentos acotados, estas nociones son equivalentes [CD21].

Ver también [MM13] para algunas definiciones de propagación del caos cuantitativa.

Existe también una versión *trayectorial* de la definición de propagación del caos, y bajo esta perspectiva, el objeto límite  $f_I; (I = [0, T]; T \leq +\infty)$  es típicamente la solución de un problema de martingala no-lineal. Es, como claramente se intuye, una noción más fuerte de propagación del caos.

### 3.3. Algunas técnicas para demostrar propagación del caos

Las primeras intuiciones en torno a las profundas interrelaciones entre la dinámica micro y macroscópica en un gran sistema de partículas ya estaban presentes en tiempos de Maxwell, fuertemente ligadas al contexto de la termodinámica. Boltzmann postuló, no sin cierta resistencia de la comunidad científica de su época, que dada la miríada de partículas presentes en un gas, sería ineficiente e ineficaz pretender dar cuenta de la dinámica de cada partícula individual, y que los esfuerzos deberían enfocarse en una descripción estadística del sistema. La intuición (el *Stosszahlansatz*, como es llamado con frecuencia) en la que Boltzmann se apoyaba consistía esencialmente en la suposición de que dos partículas del sistema tomadas al azar deberían ser estadísticamente independientes, a pesar de multitud de interacciones. Derivó la ahora llamada *ecuación de Boltzmann*, que describe la evolución temporal de la distribución de una partícula típica. El primer cimiento estaba colocado.

El desarrollo formal de una teoría de propagación del caos se remonta a los trabajos seminales de Kac con la ecuación de Boltzmann: en su esfuerzo por responder el llamado sexto problema de Hilbert (donde el trabajo de Boltzmann y el *Stosszahlansatz* toman rigor matemático) propuso un modelo más sencillo de sistema de partículas bajo la forma de un proceso de Markov, del que pudo probar que se comportaba en el límite como la ecuación de Boltzmann espacialmente homogénea, siempre que la “propiedad de Boltzmann” (así es como él llamó a la convergencia en (3.6)) se “propague en el tiempo” [Kac56]. Mc Kean luego extendió este resultado a un modelo de colisión de partículas más realista, el llamado modelo de Maxwell con *cutoff* [MJ67], y más tarde diversos resultados en la misma dirección se replicaron en varios modelos incluso con tasas de colisión no acotadas, como los modelos de Maxwell sin *cutoff* o los modelos de *esferas rígidas* (ver por ejemplo, [MM13, MMW15, BGS17, CD21]).

Siguiendo los pasos de Kac, en todos estos trabajos han habido numerosos esfuerzos para abordar el estudio de la propagación del caos para diversas clases de modelos de partículas, que exceden el ámbito de la teoría cinética, buscando presentar estrategias que alberguen cierto grado de generalidad. En esta sección nos proponemos hacer un breve racconto de algunas de técnicas “clásicas”. Es interesante mencionar que la elección de estrategia en cada caso está condicionada por varios aspectos del modelo estudiado.

En primer lugar, el conocimiento que se tiene del proceso no lineal límite puede ser determinante. Muchas estrategias “históricas” se apoyan en características y propiedades del proceso límite para llevar a cabo la propagación del caos. Ejemplos de esto son [Kac56], donde se explota una expansión en serie explícita de la solución de la ecuación de Liouville para el problema de Boltzmann; [McK69] donde se utiliza una propiedad de “completitud”; o bien [Oel84], donde se utiliza una formulación vía un problema de martingala y se aplican criterios de compacidad. En todos estos casos, la clave para el *trade-off* entre el sistema de partículas y el proceso no lineal es representarlos dentro de un marco común que permita comparar ambos procesos.

Otro aspecto a tener en cuenta es la clase de noción de propagación de caos. De entre ellas, la elección entre caos “puntual” y aquél “trayectorial” suele ser condicionante. La segunda es

una noción más fuerte, mientras que es cualitativamente más accesible obtener estimaciones cuantitativas para la primer clase.

## Métodos de acoplamiento

En un primer grupo se encuentran los *métodos de acoplamiento*. Normalmente estos métodos se apoyan en una representación del sistema de partículas como una SDE. Por este motivo, suelen requerir por un lado cierta regularidad a nivel microscópico, y por otro lado un resultado de buena definición para el sistema límite presentado también como SDE no lineal. Conducen normalmente a resultados cuantitativos de propagación del caos. Un ejemplo de aplicación de una estrategia de este tipo la encontramos en [DMGLP15].

En general, la estrategia consiste en conseguir una secuencia de procesos  $(\mathbf{Y}_t^N)_t$  con ley  $f_t^{\otimes N} \in \mathcal{P}(E^N)$  que cumpla una desigualdad del estilo

$$\frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \sup_{t \leq T} d_E(X_t^i, Y_t^i)^p \right] \leq \varepsilon(N, T) \quad (3.8)$$

donde  $\varepsilon(N, T)$  tiende a cero con  $N$  y  $f_t$  es la ley del proceso no lineal límite. En ese caso se puede conseguir un resultado de propagación del caos cuantitativo en distancia  $W_p$  [CD21, Lema 4.3].

En muchos contextos que lo permiten (procesos difusivos de McKean-Vlasov, procesos de salto *mean-field*, ver [CD21, Sección 2.2]), la forma más frecuente y simple de acoplamiento es la que se conoce como *acoplamiento sincrónico*, que consiste sencillamente en construir  $N$  procesos no lineales utilizando, según sea el caso, los mismos movimientos brownianos o PPP's que el  $N$ -sistema de partículas utiliza; y donde el proceso no lineal resulta de tomar límite formal en la ecuación para  $X_t^i$ , reemplazando la dependencia en  $\mu_{\mathbf{X}_t}^N$  por  $f_t$ .

En algunos modelos de salto, o en los modelos de Boltzmann, para los cuales los saltos se producen en tiempos discretos, las ideas de acoplamiento sincrónico son complementadas con el concepto de *saltos óptimos*, donde los saltos del proceso original de  $N$  partículas y el acoplado están vinculados por un *plan de transferencia* óptimo en el sentido de Wasserstein, minimizando el error en los tiempos de salto, y utilizando ideas de sincronicidad (eventualmente) en los intervalos inter-salto. La existencia de tal plan de transferencia no siempre es trivial ni garantizado, y requiere de cierto trabajo.

El método que adoptamos en este trabajo para probar propagación del caos se encuentra dentro de esta clase de metodologías de acoplamiento; y en particular, la estrategia desarrollada comparte ideas con el último tipo mencionado. En [CF16], de donde tomamos inspiración, se desarrolla una estrategia que hace un uso provechoso de acoplamiento(s) mediante saltos óptimos para obtener resultados cuantitativos. Tanto en [CF16] como en nuestro trabajo se destaca la novedad de estos acoplamientos (que implican el paso por un acoplamiento de partículas no independientes) y el uso de recientes estimaciones agudas para medidas empíricas. Más aún, en nuestro trabajo se agrega la novedad de lidiar con interacciones

no-binarias, lo que agrega un grado de complejidad a lo realizado en [CF16]. En la siguiente sección 3.4 ampliamos sobre esta estrategia.

### Métodos de compacidad

Una segunda clase de estrategias consiste en *métodos de compacidad*: a grandes rasgos, estos métodos consisten en probar *tightness* para la medida empírica, que implica relativa compacidad gracias al teorema de Prokhorov; y algún resultado de unicidad de límite. Tienen la ventaja de tener un amplio rango de aplicabilidad (modelos difusivos, de saltos, de Boltzmann, e incluso mixtos); como contracara, la naturaleza misma de éstos lleva a resultados no cuantitativos de propagación del caos. Históricamente asociado al estudio de la ecuación de Boltzmann espacialmente homogénea [Tan83],[Szn84], [Wag96] o más recientemente [FM01]; aplicaciones de esta estrategia en otros modelos incluyen dinámicas difusivas de McKean-Vlasov [Szn86], [Gär88] y hasta modelos mixtos de salto/difusivos, como [GM97] o [Mél96]. Generalmente, se apoyan en propiedades del generador del proceso, aunque no siempre es el caso, como por ejemplo en [RT16], donde representan el sistema de partículas y el proceso límite como SDE's.

Estos métodos reducen el problema a uno de compacidad para una secuencia de probabilidades sobre un espacio que ya no depende de  $N$ , pero que por otro lado es en sí mismo un espacio de probabilidad, lo cual implica procedimientos más intrincados para manejar este espacio. El programa aquí consiste esencialmente en **a)** probar *tightness* para la medida empírica, que implica relativa compacidad vía el teorema de Prokhorov, es decir, podemos extraer una subsucesión convergente; **b)** mostrar que todo “punto” límite sigue la dinámica correspondiente al proceso no-lineal esperado y **c)** unicidad de solución para dicha dinámica.

Existe un abordaje, relativamente reciente y algo distinto a lo descripto, que sin embargo cae dentro de este paraguas para el caso de modelos de McKean-Vlasov llamados *sistemas de gradientes*, en los que la ecuación límite cumple una dinámica llamada *ecuación de medio granular*: Explota la teoría de *flujos de gradiente*, caracterizando como tales a la ley de  $\mu_{X_t^N}^N$  y a su límite, vistos como curvas continuas a valores en  $\mathcal{P}(\mathcal{P}(E))$ , espacio donde se utiliza un argumento de compacidad. Puede encontrarse en [CDP20].

La carencia de resultados cuantitativos por parte de esta clase de métodos es el principal motivo por el cual elegimos abordar la propagación del caos para este modelo con otros métodos.

### Método abstracto de cálculo diferencial para probabilidades

Un tercer abordaje que podemos mencionar consiste en un enfoque más bien abstracto, que debemos a [MM13], [MMW15], inspirados en Grünbaum [Grü71]: El objetivo aquí es obtener estimaciones cuantitativas de propagación del caos para la *ley* de la medida empírica del sistema de  $N$  partículas, vista como un proceso a valores en  $\mathcal{P}(E)$ . Requiere de definir nociones de semigrupo y de generador del proceso empírico del sistema de partículas y del proceso límite, así como un andamiaje de cálculo diferencial para funciones actuando sobre

$\mathcal{P}(E)$ , tarea que abunda en dificultades técnicas ya que el espacio  $\mathcal{P}(E)$  no resulta un espacio vectorial sino tan sólo un espacio métrico. Estas nociones y estructuras proporcionan el marco adecuado para probar propagación del caos vía *estimaciones de consistencia*, es decir, estimaciones funcionales entre el generador  $G^N$  asociado al sistema de partículas y el asociado al proceso límite,  $Q$ ; y por otro lado, a través de *estimaciones de estabilidad* del flujo del proceso límite en función de la data inicial. El resultado es un método con gran nivel de abstracción, aplicable a varios modelos luego de encuadrarlos dentro de este marco y probar que se cumplen las estimaciones mencionadas. Un ejemplo de juguete donde se aplica este método puede encontrarse en [CD21, Sección 4.3.2].

### Clan de ancestros

Queremos finalizar esta sección haciendo breve mención a una estrategia conocida como *clan de ancestros*, que intentamos aplicar a una versión discretizada (en el espacio de estados) de nuestro modelo, inspirándonos en [AFG11] y [FM07].

La técnica se apoya en primer lugar en una construcción explícita, para cada  $i \in \{1, \dots, N\}$ , de dos procesos  $(\omega_i^D, \omega_i^S)$ , que permiten construir de manera inductiva la dinámica de no-interacción (el *Decaimiento* en nuestro caso, que en la versión discretizada estaba asociada a un *proceso de nacimiento y muerte*) y la dinámica de interacción (los eventos de *Spike* en nuestro modelo) respectivamente. Los procesos de interacción  $(\omega_i^S)_i$  son procesos de Poisson Marcados, donde las marcas indican las posiciones “post-salto”.

El segundo elemento clave es la construcción de *etiquetas*  $(\psi^i(T))_{i=1\dots N}$  que registran, para cada neurona  $i$ , todas las posibles influencias de otras neuronas debido a un spike en el intervalo  $[0, T]$ . La idea es mostrar que la probabilidad de que haya habido interacción entre 2 partículas dadas sea de orden  $1/N$ , más precisamente, se busca una cota de la forma

$$\mathbb{P} \left[ \psi^i(T) \cap \psi^j(T) \neq \emptyset \right] \leq \frac{K}{N} e^{Ct} \quad \text{para algún } K, C > 0 \quad (3.9)$$

Ver por ejemplo [AFG11, Lema 2.1]. Este resultado permite luego probar la decorrelación asintótica de las partículas.

Resulta interesante discutir qué características inherentes a nuestro modelo hicieron intrínsecamente inaplicable esta estrategia, al menos si se quiere mantener la esencia de lo aplicado en [AFG11], o [FM07]: Ambos trabajos estudian en primer lugar un modelo de Fleming-Viot con tasa de absorción acotada, algo que no sucede en nuestro modelo. Sin embargo esta última característica es redimible mediante un proceso de truncamiento. La característica fundamental de estos modelos, que no poseía el nuestro y que resulta insalvable, es que son modelos de tipo *Nanbu*, es decir, en cada interacción sólo una partícula se ve afectada. Esto resulta ser fundamental para poder afirmar que dado un subconjunto  $L$  de  $\{1, \dots, N\}$  puedo construir las etiquetas  $(\psi^i(T))_{i \in L}$  tan sólo con los procesos  $(\omega_i^S)_{i \in L}$  y que para subconjuntos disjuntos  $L$  y  $\hat{L}$ , hay independencia para  $\{\psi^i(T) = L, X_T^i = x\}$  y  $\{\psi^i(T) = \hat{L}, X_T^i = y\}$ . Sin embargo, cuando las interacciones involucran más de una neurona ya no es posible demostrar (3.9) de esta manera ya que los procesos  $\psi^i$  crecen cualitativamente más que en los otros modelos. Una variante de esta estrategia fue construir  $\psi^i(T)$  como el conjunto de etiquetas de las neuronas

que la neurona  $i$  eligió en cada spike de  $i$  (es decir, si la neurona  $i$  fue elegida por la neurona  $j$  que emitió spike, no agrego  $j$  a  $\psi^i$  sino  $i$  a  $\psi^j$ ), pero esta elección impide independencia de los conjuntos  $\{\psi^i(T) = L, X_T^i = x\}$ , ya que no se puede seguir afirmando que estos dependen tan sólo de  $(\omega_i^D, \omega_i^S)_{i \in L}$ .

### 3.4. Propagación del Caos en nuestro modelo

Para demostrar la propagación del caos en el caso del modelo presentado en este trabajo, llevamos a cabo una estrategia de acoplamiento óptimo, con características análogas a las expuestas en la sección anterior. Más precisamente, nuestra técnica está inspirada en el programa propuesto en [CF16], aunque algunas características de nuestro modelo no hacen posible una aplicación directa de dicho programa y por lo tanto nuestra estrategia tiene componentes originales que hacen posible aplicar este camino a nuestro caso.

Procedemos a explicar el esquema de esta estrategia. Como dijimos antes, la misma se basa en realizar una serie de acoplamientos con optimalidad en los mecanismos de interacción: Comenzamos realizando un acoplamiento óptimo  $\mathbf{Z}_t = (Z_t^1, \dots, Z_t^N) \in \mathbb{R}_+^N$  con marginales  $f_t$ ; la optimalidad proviene de resultados de transporte óptimo de medidas entre la empírica de  $\mathbf{Z}_t$  y  $f_t$  (ver [CF16, Lema 3]) y de una escritura inteligente del proceso como SDE, eligiendo los saltos de  $Z_t^i$  para que queden cerca de  $X_t^i$ . Sin embargo, lograr eso en procesos donde las interacciones afectan a más de una partícula implican que  $\mathbf{Z}_t$  no tendrá coordenadas independientes. En consecuencia construimos un segundo acoplamiento auxiliar  $\mathbf{V}_t$  con ley  $f_t^{\otimes N}$ , desacoplando los saltos en cada coordenada, mediante una (re)construcción de los PPP involucrados. Dicha construcción nos permite demostrar un desacoplamiento uniforme en el tiempo para las  $k$ -marginales de  $\mathbf{Z}_t$ .

En la literatura, se suele hacer distinción entre modelos de tipo *Nanbu*, donde en cada interacción *una* sola partícula cambia su estado; y modelos de tipo *Bird*, en donde cada interacción cambia el estado de *dos* partículas del sistema, siendo el primer caso más simple que el segundo. La estrategia de [CF16] fue pensada para un modelo de partículas de Kac, con interacciones de este segundo tipo. Nuestro modelo, por otra parte, sostiene interacciones en las que intervienen  $\kappa + 1$  partículas, donde  $\kappa$  es un parámetro del modelo con valor mayor (o igual) a uno, por lo que las interacciones no son (salvo cuando  $\kappa = 1$ ) binarias. Más aún, en cada interacción, digamos entre las neuronas  $i$  y  $j_1, \dots, j_\kappa$ , una de ellas (la neurona  $i$  que emite el spike) sostiene un comportamiento distinguido al resto. Esto resulta en que no es posible aplicar de manera directa la estrategia de [CF16], principalmente porque la construcción del acoplamiento auxiliar está apoyado fuertemente en la “binariedad” de las interacciones. Por lo tanto, la construcción de  $\mathbf{V}_t$  para nuestro modelo es readaptada, tomando los aspectos esenciales de la construcción original y aplicándolos a nuestro caso.

La construcción explícita de los acoplamientos (en especial  $\mathbf{V}_t$ ) en el cuerpo principal de este trabajo fue realizada para  $\kappa = 2$ , mientras que la construcción de los acoplamientos para  $\kappa$  general fue relegada al apéndice A. El motivo es que en este caso, cada interacción involucra a *tres* neuronas y en consecuencia los PPP asociados se pueden representar dentro de  $\mathbb{R}^3$  (ver 5.1), lo cual permite una visualización espacial que hace más sencillo comprender

la elección de los PPP de manera de hacer posible el desacople de las coordenadas. Una vez comprendido el caso  $\kappa = 2$ , es relativamente sencillo realizar la generalización a cualquier valor del parámetro  $\kappa$  realizada en el apéndice. El cálculo posterior que se realiza en el Lema 5.6.3 tampoco reviste dificultad mayor que el realizado para  $\kappa = 2$ , y éste último resulta menos técnico y cargado de notación.

Cabe destacar que los resultados de propagación del caos que obtenemos siguiendo este camino son cuantitativos (en  $N$  y en  $t$ ) con una dependencia razonable en  $t$ , llegando a obtenerse cotas uniformes en el tiempo para cierto subconjunto de los parámetros del modelo.



## Capítulo 4

# Un modelo de redes neuronales con interacciones sparse

Como explicamos en el capítulo introductorio 1.4.3, es destacable en nuestro modelo (tanto finito como infinito) la forma en que se hace efectiva la criticalidad del sistema de manera indirecta a través de una función de los parámetros. A lo largo de las distintas secciones de este capítulo se definen dichos parámetros y se explicita su interpretación y algunos cálculos que permiten profundizar heurísticamente en el rol que juega cada uno de ellos en el comportamiento del sistema.

En la sección 4.1 introducimos el modelo matemático que describe el sistema con población finita  $N$ . Presentamos tres descripciones equivalentes, cada una de las cuales posee características y ventajas propias. Luego, en la sección 4.2 presentamos el modelo de aproximación de campo medio para grandes poblaciones de neuronas, a través de una ecuación diferencial de tipo McKean-Vlasov.

### 4.1. Descripción del modelo

Describimos un modelo de red neuronal biológica como un sistema de partículas donde el potencial de cada neurona es representado como un proceso estocástico a valores en  $\mathbb{R}_+ = [0, \infty)$ . Presentamos tres descripciones de este sistema de partículas, cada uno de los cuales aporta una perspectiva útil y distintiva: como un PDMP (Sección 4.1.1), como la solución a una ecuación diferencial estocástica dada por procesos de Poisson (Sección 4.1.2), y como un proceso construido a partir de una cadena de Markov a tiempo discreto, obtenida al instanciar el sistema en los tiempos de spikes (Sección 4.1.3).

Cuatro parámetros caracterizan el modelo: la *tasa de decaimiento*  $\mu > 0$ , que determina la velocidad de decaimiento hacia el potencial de reposo, que tomamos como 0; la *constante de proporción de la tasa de firing*  $\gamma > 0$ , que regula la intensidad con la que se produce un firing, en relación al potencial de membrana de las neuronas; el *rango de alcance*  $\kappa \in \mathbb{N}$ , que representa la cantidad de neuronas excitadas al producirse un spike; y la *magnitud de impulso*  $\rho > 0$  que recibe cada neurona seleccionada en un evento de firing.

Asumiremos siempre que la condición inicial  $\mathbf{X}_0 = (X_0^1, \dots, X_0^N)$  es una colección de

copias i.i.d. de una variable aleatoria  $Z_0 \in \mathbb{R}_+$  dada. Como consecuencia de esto y de la dinámica del sistema,  $\mathbf{X}_t = (X_t^1, \dots, X_t^N)$  resultará intercambiable para todo  $t \geq 0$ .

Describimos entonces un sistema de  $N$  neuronas con potencial de membrana  $X_t^1, \dots, X_t^N \geq 0$  a tiempo  $t$ . Cada neurona tiene tasa de decaimiento  $\mu$  y tasa de disparo  $\gamma$ . En ausencia de spikes, el potencial de cada neurona decae al potencial de reposo de acuerdo con la ecuación  $\frac{d}{dt}X_t^i = -\mu X_t^i$ . Los spikes de una neurona  $i$  se producen de acuerdo a eventos de un proceso de Poisson no homogéneo con intensidad  $\gamma X_t^i$ , es decir, dependiente del estado de dicha neurona, independientemente de las demás neuronas. Cuando una neurona emite un spike, salta instantáneamente a cero y elige  $\kappa \geq 1$  otras neuronas de manera aleatoria uniforme, y éstas elevan su potencial en  $\rho > 0$  unidades.

Observar que, a diferencia de [RT16], no hay escalamiento (en  $N$ ) en ninguno de los parámetros del modelo  $\kappa, \rho, \mu, \gamma$ .

#### 4.1.1. Definición como un PDMP

Tal como describimos el comportamiento de  $\mathbf{X}_t = (X_t^1, \dots, X_t^N)$  en los párrafos anteriores, las trayectorias de  $\mathbf{X}_t$  resultan un PDMP con espacio de estados  $\mathbb{R}_+^N$ , un flujo dado por  $\phi : (\mathbf{x}, t) \mapsto \mathbf{x}e^{-\mu t}$ , tasa de salto  $\lambda : \mathbf{x} \mapsto \gamma \|\mathbf{x}\|$  donde  $\|\mathbf{x}\| = \sum_i |x_i|$ , y núcleo de transición

$$Q : (\mathbf{x}, A) \mapsto \sum_i \sum_{K:K \not\ni i} \frac{x^i}{\|\mathbf{x}\|} \binom{N-1}{\kappa}^{-1} \delta_{\mathbf{x}-x^i e_i + \rho e_K}(A)$$

donde la segunda suma se realiza sobre subconjuntos  $K \subset \{1, \dots, N\}$  con  $\kappa$  elementos,  $e_i$  denota el  $i$ -ésimo vector unitario y  $e_K = \sum_{j \in K} e_j$ .

#### 4.1.2. Descripción como solución de una SDE

Alternativamente, la trayectoria del sistema dada una condición inicial puede representarse como la única solución casi segura de un sistema de ecuaciones diferenciales estocásticas dirigidas por procesos de Poisson: Para cada  $i = 1, \dots, N$ , el comportamiento de  $X_t^i$  puede ser descrito por

$$\begin{aligned} dX_t^i &= -\mu X_t^i dt - X_t^i \sum_{K:K \not\ni i} \int_0^\infty \mathbf{1}(u \leq \gamma X_t^i) \mathcal{N}_K(dt, du, (i-1, i)) \\ &\quad + \rho \sum_{K:K \ni i} \int_0^\infty \int_0^N \mathbf{1}(u \leq \gamma X_t^{\lceil \xi \rceil}) \mathcal{N}_K(dt, du, d\xi), \end{aligned} \tag{4.1}$$

donde todas las sumas que involucran  $K$  son consideradas respecto a subconjuntos de tamaño  $\kappa$ , y  $\mathcal{N}_K(dt, du, d\xi)$ ,  $K \subset \{1, \dots, N\}$ , son medidas aleatorias de Poisson independientes entre sí sobre el espacio  $\mathbb{R}_+ \times \mathbb{R}_+ \times (0, N]$  con medida de intensidad común

$$\binom{N-1}{\kappa}^{-1} dt du \mathbf{1}([\xi] \notin K) d\xi.$$

Cada átomo  $(T, U, \Xi)$  de la medida aleatoria de Poisson  $\mathcal{N}_K$  corresponde a un posible firing donde las neuronas con etiqueta en  $K$  son excitadas:  $T \geq 0$  es el tiempo del spike, la marca

$U \geq 0$  es el medio para modelar la no-homogeneidad correspondiente a la tasa de firing, y  $\Xi \in (0, N]$  se utiliza para elegir uniformemente a la neurona  $i = \lceil \Xi \rceil \notin K$  que emite el spike. Observar que las medidas  $\mathcal{N}_K(dt, du, (i-1, i])$  del segundo término son una proyección sobre  $\mathbb{R}_+ \times \mathbb{R}_+$  de la restricción de los átomos de  $\mathcal{N}_K$  que caen en  $\mathbb{R}_+ \times \mathbb{R}_+ \times (i-1, i]$ , con intensidad  $\binom{N-1}{\kappa}^{-1} dt du$ .

Notar que, en principio, se podría objetar que es posible una definición de este sistema más simple, utilizando una colección de procesos de Poisson i.i.d  $\mathcal{N}_K^i(dt, du)$  correspondientes al evento de firing de la neurona  $i$  donde las neuronas del conjunto  $K \not\ni i$  son excitadas por  $i$ . El motivo para obtener  $i$  mediante una variable aleatoria uniforme  $\xi$  es que ésta será utilizada en la construcción mediante couplings durante la demostración de propagación del caos. Más precisamente, si  $f_t$  es la ley del proceso no lineal a tiempo  $t$ , un punto clave de la construcción consiste en samplear  $f_t$  de manera tal de quedar “cerca” de  $X_t^{\lceil \xi \rceil}$  de (4.1), y la manera de lograrlo es utilizando  $\xi$  como el input de una función cuidadosamente elegida, que requiere que  $\xi$  sea continua y no discreta. Ver sección 5.6, en particular ver (5.20).

### 4.1.3. Construcción a partir de una cadena de Markov a tiempo discreto

Damos ahora una tercera construcción del modelo usando una cadena de Markov para los tiempos de spike, que resulta especialmente oportuna para simular el sistema.

Un sistema comenzado en el estado  $\mathbf{x} = (x^1, \dots, x^N)$  evoluciona de la siguiente manera: Observemos que por la independencia de las medidas de Poisson, la tasa total de firing en el estado  $\mathbf{X}_t$  equivale a  $\gamma \|\mathbf{X}_t\|$ , y mientras no ocurran spikes, la suma de los potenciales de membrana decaen de acuerdo a  $\|\mathbf{X}_t\| = \|\mathbf{x}\| e^{-\mu t}$ . Por lo tanto, la probabilidad de que no ocurran spikes en el intervalo de tiempo  $[0, t]$  es la misma que la probabilidad de que un proceso de Poisson no homogéneo de intensidad  $\lambda_t = \gamma \|\mathbf{x}\| e^{-\mu t}$  no tenga saltos durante  $[0, t]$ . Como ese número de saltos en el intervalo  $[0, t]$  está distribuido de acuerdo a una distribución de Poisson con media  $\int_0^t \lambda_s ds = (\gamma/\mu) \|\mathbf{x}\| (1 - e^{-\mu t})$ , deducimos que la distribución del primer tiempo de firing  $\tau$  está caracterizado por

$$\mathbb{P}(\tau > t) = e^{-(\gamma/\mu)\|\mathbf{x}\|(1-e^{-\mu t})}, \quad 0 \leq t < \infty. \quad (4.2)$$

en particular, el sistema no emite spikes con probabilidad  $\mathbb{P}(\tau = \infty) = e^{-(\gamma/\mu)\|\mathbf{x}\|}$ . Otra observación importante es que la proporción de los potenciales de las neuronas se mantiene constante en ausencia de spikes, por lo tanto, dado  $\tau < \infty$ , la etiqueta  $I$  de la primer neurona en emitir un spike se distribuye de acuerdo a

$$\mathbb{P}(I = i) = \frac{x^i}{\|\mathbf{x}\|}, \quad i = 1, \dots, N. \quad (4.3)$$

Usando (4.2)–(4.3), podemos definir formalmente la trayectoria del sistema de  $N$  neuronas comenzando en  $\mathbf{X}_0 \in \mathbb{R}_+^N$ .

Notamos por  $\hat{\mathbf{X}}_k$  al estado del sistema inmediatamente después del  $k$ -ésimo spike, y por  $\hat{T}_k \in [0, \infty]$  al tiempo de dicho spike, con  $\hat{T}_k = \infty$  indicando que han habido menos de  $k$  spikes en total. Las variables aleatorias  $(\hat{\mathbf{X}}_k, \hat{T}_k)$ ,  $k = 0, 1, \dots$ , pueden ser entonces recursivamente

definidas de la siguiente manera: Definimos  $(\hat{\mathbf{X}}_0, \hat{T}_0) = (\mathbf{X}_0, 0)$ . Para cada  $k \geq 1$ , tomamos  $\mathbf{x} = \hat{\mathbf{X}}_{k-1}$  e instanciamos el  $k$ -ésimo tiempo de espera  $\tau_k$  hasta el  $k$ -ésimo spike según la distribución (4.2). Si  $\tau_k = \infty$ , definimos  $(\hat{\mathbf{X}}_k, \hat{T}_k) = (0, \infty)$ . De lo contrario, instanciamos la etiqueta  $I_k$  de la neurona que emite el  $k$ -ésimo spike según la distribución (4.3), y luego sorteamos un conjunto de neuronas excitadas  $K_k$  uniformemente de entre los subconjuntos de  $\{1, \dots, N\} \setminus \{I_k\}$  de tamaño  $\kappa$ . Luego establecemos

$$(\hat{\mathbf{X}}_k, \hat{T}_k) = ((\mathbf{x} - x^{I_k} e_{I_k})e^{-\mu\tau_k} + \rho e_{K_k}, \hat{T}_{k-1} + \tau_k) \quad (4.4)$$

por lo tanto la trayectoria continua a trozos del sistema se define por

$$\mathbf{X}_t = \hat{\mathbf{X}}_k e^{-\mu(t-\hat{T}_k)} \quad \text{para} \quad \hat{T}_k \leq t < \hat{T}_{k+1}. \quad (4.5)$$

Las ecuaciones (4.4)–(4.5) nos proporcionan una forma sencilla de simular las trayectorias del sistema. Sortear según la distribución (4.2) puede realizarse mediante la elección  $\tau = -\frac{1}{\mu} \log(1 - \frac{\mu}{\gamma\|\mathbf{x}\|} \xi)_+$  donde  $\xi$  posee distribución exponencial de media 1 y con la convención  $\log 0 := -\infty$ , de manera que  $\tau = \infty$  cuando  $\xi \geq \frac{\gamma\|\mathbf{x}\|}{\mu}$ . También podemos deducir que  $\lim_{t \rightarrow \infty} \mathbf{X}_t = 0$  si y sólo si la cantidad de spikes del sistema es finito, lo cual es lo que se espera de este sistema.

En las figuras 4.1 y 4.2 se representan a modo ilustrativo las trayectorias del potencial de membrana de 3 neuronas en una población de 200 neuronas con  $\gamma = \rho = \mu = 1$ ,  $\kappa = 2$  y con  $\kappa = 3$  respectivamente.

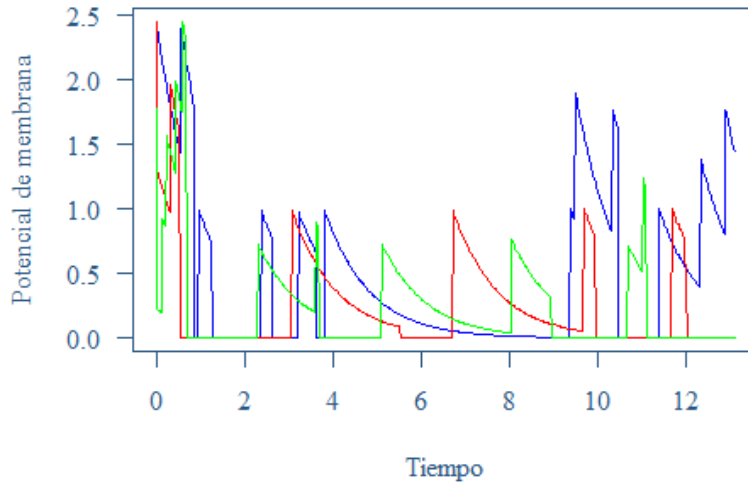


Figura 4.1: Trayectoria de 3 neuronas en una población de  $N = 200$  neuronas, con  $\gamma = \rho = \mu = 1$ ,  $\kappa = 2$

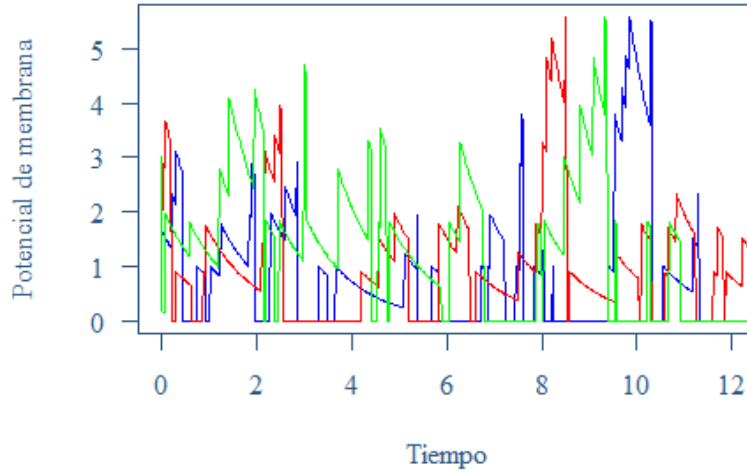


Figura 4.2: Trayectoria de 3 neuronas en una población de  $N = 200$  neuronas, con  $\gamma = \rho = \mu = 1$ ,  $\kappa = 3$

En las figuras 4.3 y 4.4 también se ilustran las trayectorias de 3 neuronas en una población de 200 neuronas. La distinta elección de parámetros muestra un comportamiento cualitativamente distinto de la actividad neuronal: en el primer caso ( $\gamma = 0,5$ ,  $\rho = \mu = 1$ ,  $\kappa = 2$ , figura 4.3) se advierte una extinción de la actividad, mientras que en el segundo caso ( $\gamma = 1,5$ ,  $\rho = \mu = 1$ ,  $\kappa = 2$ , figura 4.4) la actividad parece sostenerse en el tiempo.

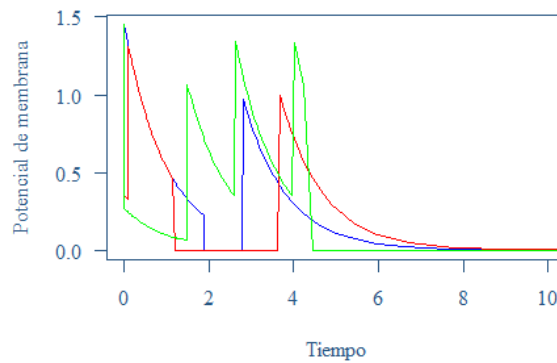


Figura 4.3: Trayectoria de 3 neuronas en una población de  $N = 200$  neuronas, con  $\gamma = 0,5$ ,  $\rho = \mu = 1$ ,  $\kappa = 2$

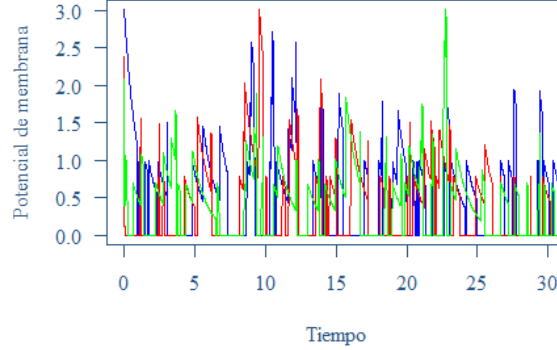


Figura 4.4: Trayectoria de 3 neuronas en una población de  $N = 200$  neuronas, con  $\gamma = 1,5$ ,  $\rho = \mu = 1$ ,  $\kappa = 2$

## 4.2. Proceso límite de campo medio

Cuando la cantidad de neuronas  $N$  es muy grande, podemos derivar heurísticamente una aproximación de campo medio para el sistema de la siguiente manera: fijemos una neurona de referencia  $i$ . Como dijimos, en ausencia de spikes, el potencial de membrana decae exponencialmente según  $\frac{d}{dt}X_t^i = -\mu X_t^i$ , y la neurona emite un spike con tasa  $\gamma X_t^i$ . Cuando cualquier otra neurona emite un spike, la neurona  $i$  es excitada con probabilidad  $\binom{1}{\kappa-1} \binom{N-2}{\kappa}^{-1} = \frac{\kappa}{N-1}$ . Ahora, asumiendo que el promedio empírico  $\frac{1}{N} \sum_j X_t^j$  de los potenciales está “cerca” de la esperanza  $\mathbb{E}[X_t^i]$ , la tasa promedio bajo la cual la neurona  $i$  es excitada es aproximadamente

$$\gamma \frac{\kappa}{N-1} \sum_{j \neq i} X_t^j \approx \gamma \kappa \mathbb{E}[X_t^i].$$

Esto nos sugiere que la evolución de una neurona referencia en un sistema grande de neuronas puede ser aproximado por la solución de la siguiente ecuación diferencial estocástica de tipo McKean-Vlasov

$$\begin{aligned} dZ_t = & -\mu Z_t dt - Z_t \int_0^\infty \mathbf{1}(u \leq \gamma Z_t) \mathcal{N}(dt, du) \\ & + \rho \int_0^\infty \mathbf{1}(u \leq \gamma \mathbb{E}[Z_t]) \mathcal{M}(dt, du), \end{aligned} \quad (4.6)$$

donde  $\mathcal{N}(dt, du)$  y  $\mathcal{M}(dt, du)$  son medidas aleatorias de Poisson independientes sobre  $\mathbb{R}_+^2$  con medidas de intensidad  $dt du$  y  $\kappa dt du$ , respectivamente.

Notar que el tercer término de (4.6), resultante del mecanismo de firing “local” de este modelo, es de salto. Como mencionamos anteriormente, ésta es una diferencia clave respecto al modelo en [RT16] y a otros similares, donde el mecanismo de firing es global y producen en contraste un término continuo de deriva. Notar también que la tasa a la cual  $Z_t$  salta positivamente depende de su propia esperanza; y es en este sentido que el proceso es no lineal de tipo McKean Vlasov.

## Capítulo 5

# Comportamiento límite y persistencia

En este capítulo desarrollamos los principales resultados y sus demostraciones respectivas. Como introdujimos en el capítulo 1, en estos resultados observaremos un conjunto de fenómenos análogos a los del proceso de contacto: en el caso finito existe una extinción segura, cuyo tiempo de extinción depende de un parámetro  $\theta$ , mientras que en el caso infinito se observa una transición de fase, dada por el mismo parámetro, que determina la supervivencia o no del proceso. Estos dos fenómenos quedan ligados a través de los resultados de propagación del caos y persistencia.

En la sección 5.1 desarrollamos los resultados relacionados a la red finita de neuronas. En la siguiente sección 5.2, demostramos la transición de fase de la población infinita. En la sección 5.3 probamos la propagación del caos. Finalmente, en la sección 5.4 demostramos el fenómeno de persistencia para este modelo.

### 5.1. Comportamiento a tiempos grandes de la red finita de neuronas

En primer lugar analizamos el comportamiento a largo plazo del sistema finito. El teorema 5.1.1 más abajo nos dice que en un sistema finito de neuronas, el potencial de cada una de ellas converge a cero, independientemente del estado inicial y de los parámetros del modelo. Este resultado implica que la única distribución invariante del proceso de Markov  $\mathbb{R}_+^N$ -valuado es la delta de Dirac en  $\mathbf{0} \in \mathbb{R}_+^N$ . Sin embargo, la naturaleza de esta convergencia sí varía drásticamente dependiendo de los parámetros del sistema.

Vamos a definir un *parámetro de reproducción*, que utilizaremos repetidamente en futuras referencias:

$$\theta = \kappa(1 - e^{-\rho\gamma/\mu}) \tag{5.1}$$

Esta constante dependiente de los parámetros del sistema juega un papel fundamental en nuestro estudio. En la sección 5.2 damos una interpretación de esta magnitud.

**Teorema 5.1.1** (Comportamiento a largo plazo de la red finita de neuronas). *Para todo estado inicial, el proceso  $\mathbf{X}_t$  converge casi seguramente a  $\mathbf{0}$ . Más aún, cuando  $\theta < 1$ , la convergencia de cualquier neurona  $i$  ocurre exponencialmente rápido (e independiente de  $N$ ), de acuerdo a*

$$\mathbb{E}[\omega(X_t^i, 0)] \leq \mathbb{E}[\omega(X_0^i, 0)] e^{-(1-\theta)\mu t} \quad (5.2)$$

para la métrica

$$\omega(x, y) = 1 - e^{-(\gamma/\mu)|x-y|}. \quad (5.3)$$

Es notoria esta propiedad de decaimiento global descrita en el teorema 5.1.1, ya que vale para valores arbitrariamente grandes de la tasa de disparo  $\gamma$ . Para apreciar mejor esta característica, notemos que en un sistema ligeramente modificado donde el potencial de la neurona que emite el spike permanece constante en lugar de ir a cero, el proceso de Markov asociado  $(\mathbf{Y}_t)_{t \geq 0}$  satisface  $\frac{d}{dt} \mathbb{E}[\|\mathbf{Y}_t\|] = (\rho\kappa\gamma - \mu) \mathbb{E}[\|\mathbf{Y}_t\|]$ , por lo tanto el sistema modificado permanece acotado en media sólo si  $\gamma < \frac{\mu}{\rho\kappa}$ . Observar también que si la neurona que emite el spike sólo transmite impulso a una neurona (es decir, si  $\kappa = 1$ ) entonces, independientemente de los demás parámetros del sistema, se tiene  $\theta < 1$ ; y en particular, según este teorema, se encuentra en las condiciones donde ocurre (5.2).

## 5.2. Transición de fase en el límite de campo medio

Cuando la cantidad  $N$  de neuronas tiende a infinito, se tiene la expectativa de que cada neurona de la red converja un proceso límite no trivial, dado por (4.6). El siguiente resultado confirma la buena definición de la SDE (4.6), que representa la aproximación de campo medio del sistema.

**Teorema 5.2.1** (Buena definición del límite de campo medio). *Para todo estado inicial  $Z_0 \in \mathbb{R}_+$ , existe única solución fuerte  $(Z_t)_{t \geq 0}$  para (4.6).*

Uno de los principales resultados en este trabajo es la descripción de una transición de fase para el proceso no-lineal límite, caracterizado por el parámetro de reproducción  $\theta$  definido por (5.1).

**Teorema 5.2.2** (Transición de fase en el límite de campo medio). *Supongamos  $\mathbb{E}[Z_0] > 0$  y  $\mathbb{E}[Z_0^2] < \infty$ . Entonces  $(Z_t)_{t \geq 0}$  es uniformemente integrable, y se da la siguiente transición de fase:*

- Si  $\theta < 1$ , entonces  $\mathbb{E}[Z_t] \rightarrow 0$  y  $\int_0^\infty \mathbb{E}[Z_t] dt < \infty$ .
- Si  $\theta = 1$ , entonces  $\mathbb{E}[Z_t] \rightarrow 0$  y  $\int_0^\infty \mathbb{E}[Z_t] dt = \infty$ .
- Si  $\theta > 1$ , entonces  $\inf_{t \geq 0} \mathbb{E}[Z_t] > 0$ .

Más aún, para  $\theta < 1$ , se tiene  $Z_t \rightarrow 0$  casi seguramente, y la convergencia en esperanza es exponencial para la métrica  $\omega$  definida en (5.3), según la desigualdad

$$\mathbb{E}[\omega(Z_t, 0)] \leq \mathbb{E}[\omega(Z_0, 0)] e^{-(1-\theta)\mu t}. \quad (5.4)$$



Observar que, como consecuencia de este teorema, **si  $\theta > 1$ , el proceso límite exhibe comportamiento no trivial a tiempos grandes** (más precisamente, ocurre si y solo si  $\theta > 1$ ), lo cual es destacable, ya que, también bajo  $\theta > 1$ , el sistema finito decae a cero indefectiblemente con el tiempo.

Una explicación heurística para este fenómeno es la siguiente: con razonamiento análogo a (4.2), observamos que  $p = 1 - e^{-\rho\gamma/\mu}$  equivale a la probabilidad de que una neurona que fue excitada desde su estado de reposo va a emitir un spike eventualmente. En un sistema grande donde una gran proporción de neuronas está en reposo, el número de neuronas excitadas por un spike que a su vez eventualmente emitirá un spike se puede aproximar por una variable aleatoria binomialmente distribuida con probabilidad de éxito  $p$  y media  $\theta = \kappa p$ . Un proceso de Galton–Watson asociado sobrevivirá entonces si y sólo si  $\theta > 1$ . Por lo tanto:

- Cuando  $\theta > 1$ , los spikes tienen suficiente frecuencia e intensidad para autosostenerse perpetuamente.
- Si  $\theta < 1$ , entonces la frecuencia de los spikes es cada vez menor a medida que avanza el tiempo, y eventualmente el proceso se extingue, por el decaimiento exponencial o por un reseteo final.
- En el caso crítico  $\theta = 1$ , es sumamente interesante que, si bien  $\mathbb{E}[Z_t] \rightarrow 0$ , el proceso  $Z_t$  *nunca muere realmente*: es cierto que los spikes sí se vuelven cada vez más infrecuentes, pero dado que  $\int_0^\infty \mathbb{E}[Z_t] dt = \infty$ , éstos nunca realmente acaban de sucederse.

*Observación 5.2.1.* El modelo de Robert y Touboul exhibe también una transición de fase similar, pero la condición donde esto ocurre es distinta a la nuestra. En efecto, en su modelo para el caso de una función lineal  $b(x)$  para la tasa de disparo, la transición de fase ocurre cuando la cantidad  $\theta_c = \kappa\rho\gamma/\mu$  es mayor a 1, ver [RT16, Sección 7.2.1] (hemos usado aquí la notación de este trabajo; notar que, dado que el parámetro  $\kappa$  no tiene equivalente en su modelo, hemos interpretado  $\kappa\rho$  como la magnitud del impulso exitatorio en su modelo, antes de reescalar por  $1/N$ ). Como  $\theta \leq \theta_c$ , vemos que el fenómeno de persistencia en nuestro modelo requiere un mayor valor para  $\rho\gamma/\mu$ , es decir, la magnitud del impulso exitatorio por la razón entre la tasa de disparo y la de decaimiento. Esta diferencia es una consecuencia de la naturaleza *sparse* del mecanismo de interacción considerado en este trabajo, en contraste con [RT16]; y es de las características más distintivas e interesantes de nuestro modelo. Por otro lado, nuestras cotas superiores para la red finita de neuronas (Teoremas 5.3.1 y 5.4.1) indican que  $\theta_c$  tiene relevancia también para nuestro caso.

*Observación 5.2.2.* El Teorema 5.2.2 implica que si  $\mathbb{E}[Z_0] > 0$  y  $\theta > 1$ , entonces todo límite en distribución de  $Z_t$  satisface  $\mathbb{E}[Z_\infty] > 0$ . Nuestra conjetura es que para cada  $\theta > 1$ , el proceso  $Z_t$  admite *exactamente dos distribuciones estacionarias*: la trivial (delta de Dirac en 0), que sólo se alcanza cuando la condición inicial es cero, o sea  $Z_0 = 0$  casi seguramente; y una segunda distribución no trivial, a la que  $Z_t$  se acerca siempre que  $\mathbb{E}[Z_0] > 0$ . La construcción en la sección 4.1.3 da una forma simple de simular aproximadamente esta distribución.

Como consecuencia del Teorema 5.2.2, obtenemos el siguiente resultado, que da cuenta de la proporción límite de neuronas en el potencial de reposo para la red infinita.

**Corolario 5.2.1.** *El proceso  $Z_t$  satisface para todo  $t \geq 0$*

$$\mathbb{P}(Z_t = 0) = \frac{1}{\kappa} + \left( \mathbb{P}(Z_0 = 0) - \frac{1}{\kappa} \right) e^{-\gamma\kappa \int_0^t \mathbb{E}[Z_s] ds}. \quad (5.5)$$

*Como consecuencia,  $\lim_{t \rightarrow \infty} \mathbb{P}(Z_t = 0) = \frac{1}{\kappa}$  cuando  $\mathbb{E}[Z_0] > 0$  y  $\theta \geq 1$ .*

### 5.3. Propagación del caos

El siguiente resultado confirma que el comportamiento de una neurona típica del sistema resulta bien aproximada por la SDE no lineal de campo medio (4.6), cuando el número de neuronas  $N \rightarrow \infty$ . Las cotas superiores sobre la velocidad de convergencia están expresadas en función de otra constante fundamental,

$$\theta_c = \kappa\rho\gamma/\mu, \quad (5.6)$$

que domina al parámetro de reproducción definido en (5.1) bajo la desigualdad  $\theta \leq \theta_c$ . Ver también la observación 5.2.1. El subíndice  $c$  en  $\theta_c$  es por “caos”.

Recordemos que, dadas las medidas de probabilidad  $\nu, \tilde{\nu}$  en  $\mathbb{R}$ , su *distancia 1-Wasserstein* puede ser definida en este contexto como

$$W_1(\nu, \tilde{\nu}) = \inf \mathbb{E}[|Y - \tilde{Y}|],$$

donde el ínfimo es tomado sobre todos los *acoplamientos* de  $\nu$  y  $\tilde{\nu}$ , siendo esta métrica completa en el espacio de probabilidades con primer momento finito, ver la sección 2.2. Notar que *cualquier* acoplamiento induce una cota superior inmediata para  $W_1$ ; éste es un hecho que utilizaremos seguido.

**Teorema 5.3.1** (Propagación del caos). *Sea  $\kappa = 2$  y  $\mathbb{E}[(X_0^1)^2] < \infty$ . Entonces existe una constante  $C > 0$  tal que para todo  $N$  y para todo  $t \in \mathbb{R}_+$ , la distancia 1-Wasserstein esperada entre  $\bar{\mathbf{X}}_t = \frac{1}{N} \sum_{i=1}^N \delta_{X_i}$  y  $f_t = \mathcal{L}(Z_t)$  satisface*

$$\mathbb{E} \left[ W_1(f_t, \bar{\mathbf{X}}_t) \right] \leq \frac{C}{N^{1/3}} \times \begin{cases} 1, & \theta_c < 1, \\ 1 + t, & \theta_c = 1, \\ e^{(\theta_c - 1)\mu t}, & \theta_c > 1. \end{cases}$$

Observar que el Teorema 5.3.1 prueba la propagación del caos en términos de la distancia  $W_1$  *esperada*, ya que la medida empírica es una medida en el espacio  $\mathcal{P}(\mathbb{R})$ , es decir una medida aleatoria, y por lo tanto una forma equivalente de enunciar este resultado sería en términos de la medida 1-Wasserstein  $W_1$  sobre el espacio de probabilidades en  $\mathbb{R}$  con primer momento finito tomando a  $W_1$  como la distancia  $d$  para ese espacio. Nuevamente, nos remitimos a 2.2.

El Teorema 5.3.1, por lo tanto, muestra la propagación del caos vía la convergencia de la medida empírica al proceso no lineal límite. La tasa de convergencia  $N^{-1/3}$  en la cota superior parece razonable, teniendo en cuenta que la convergencia de la medida empírica de una sucesión de variables i.i.d. (es decir el caso “ideal”) hacia su ley común ocurre a tasa

$N^{-1/2}$  en términos de la distancia  $W_1$  esperada, ver [FG15] y (5.25). Observar que en el caso  $\theta_c < 1$  la cota es además uniforme en el tiempo.

Probaremos el Teorema 5.3.1 en la Sección 5.6 usando un argumento de acoplamiento óptimo inspirado en [CF16], siguiendo las ideas desarrolladas en la Sección 3.4: Vamos a reescribir la SDE (4.1) de una neurona en la red finita, y la SDE de campo medio (4.6), de una manera ligeramente distinta que nos permita acoplar estos procesos de manera natural y que éstos resulten más accesibles para realizar cálculos.

*Observación 5.3.1.* Por simplicidad el resultado está enunciado para  $\kappa = 2$ , pero es extensible a  $\kappa$  general. Para un  $\kappa \in \mathbb{N}$  fijo arbitrario, o incluso un  $\kappa \in \{1, \dots, N-1\}$  aleatorio con esperanza fija y varianza acotada, la idea principal de la prueba sigue funcionando. Sin embargo, la prueba general requeriría de un aumento significativo de notación y tratamiento técnico cauteloso que sólo hace al desarrollo de la demostración extremadamente técnico y difícil de seguir. Más aún, dicha incorporación no agrega prácticamente nada relevante al resultado del teorema. Por eso trataremos explícitamente sólo el caso  $\kappa = 2$ .

## 5.4. Persistencia

Combinando los resultados previos, podemos ahora enunciar el siguiente teorema, que arroja luz sobre el fenómeno de *persistencia*, en términos del parámetro de reproducción  $\theta$  definido en (5.1) y la constante  $\theta_c$  definida por (5.6).

**Teorema 5.4.1** (Persistencia). *Sea  $\kappa = 2$  y  $\mathbb{E}[(X_0^1)^2] < \infty$ . Si  $\mathbb{E}[X_0^1] > 0$ , entonces la red finita satistace la siguiente dicotomía:*

- *Si  $\theta > 1$ , entonces existen constantes  $c > 0$ ,  $\tilde{c} \in \mathbb{R}$ , y  $\epsilon > 0$  (independientes de  $N$ ) tal que para todo  $N \geq 1$ ,*

$$\mathbb{E}[X_t^1] \geq \epsilon \quad \text{para todo} \quad t \leq \tilde{c} + c \log N.$$

- *Si  $\theta < 1$ , entonces para la métrica  $\omega$  definida por (5.3), tenemos*

$$\mathbb{E}[\omega(X_t^1, 0)] \leq \mathbb{E}[\omega(X_0^1, 0)] e^{-(1-\theta)\mu t}.$$

*Si, adicionalmente, asumimos que  $\theta_c < 1$ , entonces*

$$\mathbb{E}[X_t^1] \leq \mathbb{E}[X_0^1] e^{-(1-\theta_c)\mu t}.$$

El Teorema 5.4.1 afirma que en el caso  $\theta > 1$ , la actividad de la red finita persiste por un tiempo de orden (al menos)  $\log N$ ; pero si  $\theta < 1$ , entonces la actividad de la red finita se extingue a una tasa independiente de  $N$ .

*Observación 5.4.1.* Creemos que  $\log N$  es una cruda estimación por debajo del tiempo que le lleva a la red para decaer. Algunas simulaciones sugieren que cuando  $\theta > 1$ , la actividad de la red persiste bastante más tiempo. Más aún, para  $N$  razonablemente grande y  $\theta > 1$  suficientemente lejos de 1, la medida empírica de la red parece estabilizarse en torno a una distribución estacionaria no trivial; lo cual sugiere que el tiempo de decaimiento crece a gran velocidad con  $N$ , quizás exponencialmente. Este comportamiento también se observa en [RT16, Section 7.2.1], aunque la condición en la que la transición de fase ocurre es diferente, ver la Observación 5.2.1.

## 5.5. Pruebas de los resultados

En esta sección proporcionamos pruebas detalladas de nuestros resultados.

Observar que, midiendo el potencial de membrana en unidades de la magnitud del impulso, podemos establecer  $\rho = 1$ , y midiendo el tiempo en unidades de la vida media de la neurona dividido  $\log(2)$ , podemos tomar  $\mu = 1$ . En consecuencia, para simplificar notación frecuentemente asumiremos  $\rho = \mu = 1$  sin pérdida de generalidad.

En primer lugar, deducimos una fórmula simple para la esperanza de una función del proceso  $(\mathbf{X}_t)_{t \geq 0}$ , que será utilizado en varias ocasiones. Fijemos una  $\phi : \mathbb{R}^N \rightarrow \mathbb{R}$ , y sea  $F_t = \phi(\mathbf{X}_t)$ . En ausencia de spikes,  $F_t$  evoluciona de acuerdo a

$$\frac{d}{dt}F_t = \sum_i \partial_i \phi(\mathbf{X}_t) \frac{d}{dt}X_t^i = -\mu \sum_i X_t^i \partial_i \phi(\mathbf{X}_t).$$

Si, en cambio, la neurona  $i$  se dispara y exita un conjunto  $K$  de neuronas en el instante  $t$ , entonces  $F_t$  cambia de acuerdo a

$$F_t - F_{t-} = \phi(\mathbf{X}_{t-} + \rho e_K - X_{t-}^i e_i) - \phi(\mathbf{X}_{t-})$$

donde recordamos que  $e_K := \sum_{j \in K} e_j$ . De estas afirmaciones se sigue que

$$\begin{aligned} dF_t &= -\mu \sum_i X_t^i \partial_i \phi(\mathbf{X}_t) dt \\ &+ \sum_i \sum_{K \not\ni i} \int_{\mathbb{R}_+} \mathbf{1}(u \leq \gamma X_t^i) \left( \phi(\mathbf{X}_t + \rho e_K - X_t^i e_i) - \phi(\mathbf{X}_t) \right) \mathcal{N}_{i,K}(dt, du) \end{aligned}$$

donde  $\mathcal{N}_{i,K}(dt, du)$  es notación para  $\mathcal{N}_K(dt, du, (i-1, i])$ . Tomando esperanza, tenemos que

$$\begin{aligned} \frac{d}{dt} \mathbb{E}[F_t] &= -\mu \sum_i \mathbb{E} \left[ X_t^i \partial_i \phi(\mathbf{X}_t) \right] \\ &+ \binom{N-1}{\kappa}^{-1} \sum_i \sum_{K \not\ni i} \gamma \mathbb{E} \left[ X_t^i \left( \phi(\mathbf{X}_t + \rho e_K - X_t^i e_i) - \phi(\mathbf{X}_t) \right) \right] \end{aligned} \quad (5.7)$$

### 5.5.1. Comportamiento a tiempos grandes para la red finita de neuronas (Teorema 5.1.1)

*Prueba del teorema 5.1.1.* La existencia de solución fuerte a la ecuación (4.1) se obtiene truncando la tasa de disparo y haciendo tender el parámetro de truncamiento a infinito, como en la prueba del Lema 5.5.1. La unicidad es estándar.

Ahora procedemos a mostrar que  $\|\mathbf{X}_t\| = \sum_{i=1}^N X_t^i \rightarrow 0$  casi seguramente. Veamos en primer lugar que  $H_t := e^{-(\gamma/\mu)\|\mathbf{X}_t\|}$  es una submartingala con respecto a la filtración  $(\mathcal{F}_t)_{t \geq 0}$  generada por el proceso de Markov marcado  $(\mathcal{N}^i)_{i=1}^N$ , donde

$$\mathcal{N}^i := \sum_{K: K \not\ni i} \mathcal{N}_K(dt, du, (i-1, i]).$$

En ausencia de spikes, cada neurona evoluciona de acuerdo a  $\frac{d}{dt}X_t^i = -\mu X_t^i$ , por lo tanto  $\frac{d}{dt}\|\mathbf{X}_t\| = -\mu\|\mathbf{X}_t\|$  y  $\frac{d}{dt}H_t = -(\gamma/\mu)H_t \frac{d}{dt}\|\mathbf{X}_t\| = \gamma\|\mathbf{X}_t\|H_t$ . Si, en cambio, la neurona  $i$  se dispara en el instante  $t$ , vemos que  $\|\mathbf{X}_t\| - \|\mathbf{X}_{t^-}\| = \rho\kappa - X_t^i$ , y como consecuencia

$$H_t - H_{t^-} = \left( e^{-(\gamma/\mu)(\rho\kappa - X_t^i)} - 1 \right) H_{t^-}.$$

Se sigue que para todo  $0 \leq s \leq t$ ,

$$\begin{aligned} H_t - H_s &= \int_s^t \gamma \|\mathbf{X}_r\| H_r dr \\ &+ \sum_i \int_s^t \int_{\mathbb{R}_+} \mathbf{1}(u \leq \gamma X_{r^-}^i) \left( e^{-(\gamma/\mu)(\rho\kappa - X_{r^-}^i)} - 1 \right) H_{r^-} \mathcal{N}^i(dr, du), \end{aligned}$$

de donde concluimos que

$$H_t - H_s \geq \int_s^t \gamma \|\mathbf{X}_r\| H_r dr - \sum_i \int_s^t \int_{\mathbb{R}_+} \mathbf{1}(u \leq \gamma X_{r^-}^i) H_{r^-} \mathcal{N}^i(dr, du) \quad (5.8)$$

casi seguramente. Notando que  $\sum_i \int_{\mathbb{R}_+} \mathbf{1}(u \leq \gamma X_r^i) du = \gamma \|\mathbf{X}_r\|$ , y recordando que la medida de intensidad de  $\mathcal{N}^i(dr, du)$  equivale a  $dr du$ , vemos que

$$\mathbb{E} \left[ \sum_i \int_s^t \int_{\mathbb{R}_+} \mathbf{1}(u \leq \gamma X_{r^-}^i) H_{r^-} \mathcal{N}^i(dr, du) \middle| \mathcal{F}_s \right] = \mathbb{E} \left[ \int_s^t \gamma \|\mathbf{X}_r\| H_r dr \middle| \mathcal{F}_s \right].$$

La esperanza condicional del lado derecho de (5.8) entonces vale cero, y como consecuencia  $\mathbb{E}[H_t - H_s | \mathcal{F}_s] \geq 0$  casi seguramente. Por lo tanto  $H_t$  es una submartingala acotada, y por el teorema de convergencia de Doob,  $H_t \rightarrow H_\infty$  casi seguramente, donde  $H_\infty \in [0, 1]$  es una variable aleatoria. Por lo tanto  $\|\mathbf{X}_t\| \rightarrow -\frac{\mu}{\gamma} \log H_\infty \in [0, \infty]$  casi seguramente.

Análogamente, vemos que

$$\begin{aligned} \|\mathbf{X}_t\| - \|\mathbf{X}_s\| &= - \int_s^t \mu \|\mathbf{X}_r\| dr \\ &+ \sum_i \int_s^t \int_{\mathbb{R}_+} \mathbf{1}(u \leq \gamma X_{r^-}^i) (\rho\kappa - X_{r^-}^i) \mathcal{N}^i(dr, du). \end{aligned}$$

y tomando esperanza, queda

$$\mathbb{E}[\|\mathbf{X}_t\|] - \mathbb{E}[\|\mathbf{X}_s\|] = - \int_s^t \mu \mathbb{E}[\|\mathbf{X}_r\|] dr + \mathbb{E} \left[ \sum_i \int_s^t \gamma X_r^i (\rho\kappa - X_r^i) dr \right],$$

por lo tanto

$$\frac{d}{dt} \mathbb{E}[\|\mathbf{X}_t\|] = (\gamma\rho\kappa - \mu) \mathbb{E}[\|\mathbf{X}_t\|] - \gamma \mathbb{E} \left[ \sum_i (X_t^i)^2 \right].$$

Por la desigualdad de Jensen,  $\mathbb{E}[\sum_i (X_t^i)^2] \geq N^{-1} \mathbb{E}[\|\mathbf{X}_t\|^2] \geq N^{-1} (\mathbb{E}[\|\mathbf{X}_t\|])^2$ , y se sigue que  $y_t = \mathbb{E}[\|\mathbf{X}_t\|]$  satisface

$$\frac{d}{dt} y_t \leq (\gamma\rho\kappa - \mu) y_t - \gamma N^{-1} y_t^2.$$

La desigualdad de arriba muestra que  $\frac{d}{dt} y_t \leq 0$  siempre que  $y_t \geq N(\rho\kappa - \mu/\gamma)$ . Consecuentemente,  $\sup_{t \geq 0} y_t < \infty$ . El lema de Fatou implica que  $\mathbb{E}[\lim_{t \rightarrow \infty} \|\mathbf{X}_t\|] \leq \liminf_{t \rightarrow \infty} \mathbb{E}[\|\mathbf{X}_t\|] <$

$\infty$ . En particular,  $\lim_{t \rightarrow \infty} \|\mathbf{X}_t\| < \infty$  casi seguramente.

Ahora afirmamos que existen casi seguramente finitos saltos del sistema. En efecto, recordemos que  $\|\mathbf{X}_t\| \rightarrow -\frac{\mu}{\gamma} \log H_\infty \in [0, \infty)$  casi seguramente y supongamos por absurdo que existe un evento  $S$  de probabilidad positiva en el que hay infinitos saltos. Como el número de saltos en todo intervalo acotado es finito casi seguramente, entonces  $\|\mathbf{X}_t\|$  no puede converger en  $S$ , por lo tanto no podría converger casi seguramente.

Ahora, como los saltos eventualmente se terminan, el decaimiento exponencial implica que  $\lim_{t \rightarrow \infty} \|\mathbf{X}_t\| = 0$  casi seguramente.

Asumiendo  $\theta < 1$ , probaremos ahora (5.2). Aplicando (5.7) a  $\phi(\mathbf{x}) = e^{-(\gamma/\mu)x_i}$ , y observando que  $\binom{N-1}{\kappa}^{-1} \binom{N-2}{\kappa-1} = \frac{\kappa}{N-1}$ , se ve que

$$\begin{aligned} \frac{d}{dt} \mathbb{E} \left[ e^{-(\gamma/\mu)X_t^i} \right] &= \gamma \mathbb{E} \left[ X_t^i e^{-(\gamma/\mu)X_t^i} \right] \\ &\quad + \gamma \mathbb{E} \left[ X_t^i \left( 1 - e^{-(\gamma/\mu)X_t^i} \right) \right] \\ &\quad + \binom{N-1}{\kappa}^{-1} \sum_{j \neq i} \sum_{K \ni i, K \not\ni j} \gamma \mathbb{E} \left[ X_t^j \left( e^{-(\gamma/\mu)(X_t^i + \rho)} - e^{-(\gamma/\mu)X_t^i} \right) \right] \\ &= \gamma \mathbb{E} \left[ X_t^i \right] - \gamma \left( 1 - e^{-\rho\gamma/\mu} \right) \frac{\kappa}{N-1} \sum_{j \neq i} \mathbb{E} \left[ X_t^j e^{-(\gamma/\mu)X_t^i} \right] \\ &\geq \gamma \mathbb{E} \left[ X_t^i \right] - \gamma \left( 1 - e^{-\rho\gamma/\mu} \right) \frac{\kappa}{N-1} \sum_{j \neq i} \mathbb{E} \left[ X_t^j \right] \\ &= \gamma(1 - \theta) \mathbb{E} \left[ X_t^i \right], \end{aligned}$$

gracias a la intercambiabilidad. Dado que  $(\gamma/\mu)X_t^i \geq 1 - e^{-(\gamma/\mu)X_t^i}$ , se sigue que

$$\frac{d}{dt} \mathbb{E} \left[ e^{-(\gamma/\mu)X_t^i} \right] \geq \mu(1 - \theta) \left( 1 - \mathbb{E} \left[ e^{-(\gamma/\mu)X_t^i} \right] \right).$$

Por lo tanto,  $g_t = 1 - \mathbb{E}[e^{-(\gamma/\mu)X_t^i}]$  satisface  $\frac{d}{dt} g_t \leq -\mu(1 - \theta)g_t$ . La desigualdad buscada (5.2) se sigue entonces del lema de Grönwall.  $\square$

### 5.5.2. Buena definición del proceso límite (Teorema 5.2.1)

El teorema 5.2.1 se prueba siguiendo una estrategia análoga a [RT16, Theorem 2]. Por simplicidad, en esta sección asumimos  $\rho = \mu = 1$ . Comenzamos por *linearizar* la SDE fijando una función para la tasa de disparos: dada una función continua  $r : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , consideremos la siguiente SDE, con condición inicial  $Y_0^r = Z_0$ :

$$dY_t^r = -Y_t^r dt - Y_t^r \int_0^\infty \mathbf{1}(u \leq \gamma Y_t^r) \mathcal{N}(dt, du) + \int_0^\infty \mathbf{1}(u \leq \gamma r_t) \mathcal{M}(dt, du), \quad (5.9)$$

donde  $\mathcal{N}$  y  $\mathcal{M}$  son como en (4.6). Notamos  $\|r\|_T := \sup_{t \in [0, T]} |r_t|$ .

**Lema 5.5.1.** *Existe una única solución fuerte para (5.9).*

*Prueba del Lema 5.5.1.* La prueba de unicidad es clásica. Para probar existencia, fijamos un horizonte temporal  $T > 0$ , una función continua  $r : [0, T] \rightarrow \mathbb{R}_+$ , un umbral  $M > 0$ , y

consideramos la SDE truncada

$$\begin{aligned} dY_t^{r,M} &= -Y_t^{r,M} dt - Y_t^{r,M} \int_0^\infty \mathbf{1}(u \leq \min(\gamma Y_t^{r,M}, M)) \mathcal{N}(dt, du) \\ &\quad + \int_0^\infty \mathbf{1}(u \leq \gamma r_t) \mathcal{M}(dt, du), \end{aligned} \quad (5.10)$$

con  $Y_0^{r,M} = Z_0$ . Como las tasas son ahora acotadas, hay existencia y unicidad para (5.10). Consideremos ahora el evento  $B_M = \{\gamma Y_t^{r,M} \leq M : \forall t \in [0, T]\}$ , que crece con  $M$ , y observemos que para  $\tilde{M} \geq M$  tenemos  $Y^{r,M} \equiv Y^{r,\tilde{M}}$  sobre  $B_M$ . Por lo tanto, podemos definir  $Y_t^r = \lim_{M \rightarrow \infty} Y_t^{r,M}$  en el evento  $B_\infty = \bigcup_{M>0} B_M$ . Es fácil ver que  $(Y_t^r)$  es solución de (5.9) en  $B_\infty$ . Falta ver que  $\mathbb{P}(B_\infty) = 1$ , para lo cual es suficiente con probar que  $\lim_{M \rightarrow \infty} \mathbb{P}(B_M^c) = 0$ . En efecto, como  $\|r\|_T < \infty$ , no es difícil obtener una cota como  $\mathbb{P}(B_M^c) \leq C_T/M$  para alguna constante  $C_T > 0$ ; por ejemplo, estudiar la solución de (5.9) sin “reseteos” ni decaimiento exponencial y comparar con  $Y_t^{r,M}$ . Esto prueba existencia y unicidad para (5.9) on  $[0, T]$  para todo  $T > 0$ , que puede ser fácilmente extendido a  $[0, \infty)$ .  $\square$

Ahora consideremos el funcional  $\mathcal{A} : C(\mathbb{R}_+, \mathbb{R}_+) \rightarrow C(\mathbb{R}_+, \mathbb{R}_+)$  definido por

$$(\mathcal{A}r)_t = \mathbb{E}[Y_t^r] \quad \text{donde } (Y_t^r) \text{ resuelve (5.9)}. \quad (5.11)$$

**Lema 5.5.2.** *El funcional  $\mathcal{A}$  tiene un único punto fijo.*

*Prueba del lema 5.5.2.* Fijemos un horizonte temporal  $T > 0$ , y sean  $r, q \in C([0, T], \mathbb{R}_+)$ . Sean  $(Y_t^r), (Y_t^q)$  soluciones fuertes para (5.9) con funciones de tasa de disparo  $r$  y  $q$ , respectivamente, y con  $Y_0^r = Y_0^q = Z_0$ . Sea  $h_t = \mathbb{E}[|Y_t^r - Y_t^q|]$ , entonces para cada  $0 \leq s < t$ , se tiene:

$$h_t - h_s = - \int_s^t h_v dv + \mathbb{E} \int_s^t \int_0^\infty \Delta_v \mathcal{N}(dv, du) + \mathbb{E} \int_s^t \int_0^\infty \Delta_v \mathcal{M}(dv, du), \quad (5.12)$$

donde  $\Delta_v = |Y_v^r - Y_v^q| - |Y_v^r - Y_v^q|$ . Definamos  $m_v := \min(Y_v^r, Y_v^q)$  y  $M_v := \max(Y_v^r, Y_v^q)$ . Tras un evento de firing de  $Y^r$  ó  $Y^q$  (o de ambos), el proceso respectivo vuelve a 0. Por lo tanto, para todo átomo  $(v, u)$  de  $\mathcal{N}$ , tenemos

$$\Delta_v = \begin{cases} -(M_v - m_v) & \text{si } u \leq \gamma m_v, \\ m_v - (M_v - m_v) & \text{si } \gamma m_v < u \leq \gamma M_v, \\ 0 & \text{si } u > \gamma M_v. \end{cases}$$

Consecuentemente, el segundo término de (5.12) vale:

$$\begin{aligned} &\mathbb{E} \int_s^t \int_0^\infty \Delta_v \mathcal{N}(dv, du) \\ &= \mathbb{E} \int_s^t [-\gamma m_v (M_v - m_v) + \gamma (M_v - m_v) (2m_v - M_v)] dv \\ &= -\gamma \mathbb{E} \int_s^t (Y_v^r - Y_v^q)^2 dv \leq 0. \end{aligned} \quad (5.13)$$

Por otro lado, tras un evento excitatorio de  $Y^r$  ó  $Y^q$  (o de ambos), el proceso respectivo incrementa su valor en  $\rho = 1$  unidad. Por lo tanto para todo átomo  $(v, u)$  de  $\mathcal{M}$ , tenemos:

$$\Delta_v \begin{cases} = 0 & \text{si } u \leq \gamma \min(r_v, q_v), \\ \leq 1 & \text{si } \gamma \min(r_v, q_v) < u \leq \gamma \max(r_v, q_v), \\ = 0 & \text{si } u > \gamma \max(r_v, q_v). \end{cases}$$

Por lo tanto, el tercer término de (5.12) vale:

$$\mathbb{E} \int_s^t \int_0^\infty \Delta_v \mathcal{M}(dv, du) \leq \kappa \int_s^t \gamma |r_v - q_v| dv.$$

De (5.12) y estas desigualdades, obtenemos:

$$\partial_t h_t \leq -h_t + \gamma \kappa |r_t - q_t|.$$

Por lo tanto,

$$|(\mathcal{A}r)_t - (\mathcal{A}q)_t| \leq h_t \leq \gamma \kappa \int_0^t |r_s - q_s| e^{-(t-s)} ds \leq \gamma \kappa \int_0^t |r_s - q_s| ds.$$

Iterando esta desigualdad, obtenemos

$$\|\mathcal{A}^n r - \mathcal{A}^n q\|_T \leq \frac{(\gamma \kappa T)^n}{n!} \|r - q\|_T$$

para todo  $n \in \mathbb{N}$ . Por lo tanto, para  $n$  suficientemente grande,  $\mathcal{A}^n$  es una contracción en  $(C([0, T], \mathbb{R}_+), \|\cdot\|_T)$ , entonces  $\mathcal{A}$  tiene un único punto fijo en  $C([0, T], \mathbb{R}_+)$ . La extensión a  $C(\mathbb{R}_+, \mathbb{R}_+)$  es directa.  $\square$

*Prueba del teorema 5.2.1.* Definimos  $Z_t = Y_t^r$ , donde  $(Y_t^r)$  es la única solución fuerte para (5.9) con función de tasa de disparo  $r \in C(\mathbb{R}_+, \mathbb{R}_+)$  elegida como el único punto fijo del funcional  $\mathcal{A}$  dado por (5.11). Por la definición de  $\mathcal{A}$ , tenemos  $\mathbb{E}[Z_t] = r_t$ , por lo tanto  $(Z_t)$  resuelve (4.6).

La unicidad es directa: sea  $Z$  y  $\tilde{Z}$  soluciones fuertes para (4.6) con  $Z_0 = \tilde{Z}_0$ . Si definimos  $r_t := \mathbb{E}[Z_t]$  y  $\tilde{r}_t := \mathbb{E}[\tilde{Z}_t]$ , entonces  $Z = Y^r$  y  $\tilde{Z} = Y^{\tilde{r}}$  son soluciones fuertes para (5.9) y es inmediato probar que tanto  $r$  como  $\tilde{r}$  son puntos fijos para  $\mathcal{A}$ , por lo tanto  $r = \tilde{r}$  por el Lema 5.5.2, y consecuentemente  $Z = Y^r = Y^{\tilde{r}} = \tilde{Z}$  por el Lema 5.5.1.  $\square$

### 5.5.3. Transición de fase para el límite de campo medio (Teorema 5.2.2)

*Prueba del Teorema 5.2.2.* El Lema 5.6.4 abajo implica que  $\sup_t \mathbb{E}[Z_t^2] < \infty$  cuando asumimos  $\mathbb{E}[Z_0^2] < \infty$ , por lo tanto  $(Z_t)_{t \geq 0}$  es uniformemente integrable.

Estudiamos ahora la transición de fase. Analicemos la evolución de  $h_t = \mathbb{E} \left[ e^{-(\gamma/\mu)Z_t} \right]$ . Observemos que  $h_t \in [0, 1]$ . En ausencia de spikes,  $e^{-(\gamma/\mu)Z_t}$  evoluciona según  $\frac{d}{dt} e^{-(\gamma/\mu)Z_t} = e^{-(\gamma/\mu)Z_t} (-\gamma/\mu) \frac{d}{dt} Z_t = -\gamma Z_t e^{-(\gamma/\mu)Z_t}$ . En un evento de reseteo,  $e^{-(\gamma/\mu)Z_t}$  aumenta su valor en  $1 - e^{-(\gamma/\mu)Z_t^-}$ , y en el evento de excitación de la neurona, crece una magnitud de  $(e^{-\rho\gamma/\mu} - 1)e^{-(\gamma/\mu)Z_t^-}$ . Consecuentemente,

$$\begin{aligned} d(e^{-(\gamma/\mu)Z_t}) &= \gamma Z_t e^{-(\gamma/\mu)Z_t} dt \\ &+ (1 - e^{-(\gamma/\mu)Z_t^-}) \int_{\mathbb{R}_+} \mathbf{1}(u \leq \gamma Z_t^-) \mathcal{N}(dt, du) \\ &+ (e^{-\rho\gamma/\mu} - 1) e^{-(\gamma/\mu)Z_t^-} \int_{\mathbb{R}_+} \mathbf{1}(u \leq \gamma \mathbb{E}[Z_t^-]) \mathcal{M}(dt, du), \end{aligned}$$

donde  $\mathcal{N}(dt, du)$  y  $\mathcal{M}(dt, du)$  son como en (4.6). Tomando esperanza, vemos que los términos  $\mathbb{E}[\gamma Z_t e^{-(\gamma/\mu)Z_t}]$  se cancelan, obteniendo entonces

$$\frac{d}{dt} h_t = \gamma r_t - \gamma \theta r_t h_t, \tag{5.14}$$



donde  $r_t = \mathbb{E}[Z_t]$ , y recordamos que  $\theta = \kappa(1 - e^{\rho\gamma/\mu})$ . Esta ecuación diferencial es resuelta por

$$h_t = e^{-\gamma\theta R_t} h_0 + \left(1 - e^{-\gamma\theta R_t}\right) \frac{1}{\theta},$$

donde  $R_t = \int_0^t r_s ds$ . Como  $R_t$  es no decreciente, se sigue que  $h_t$  converge a

$$h_\infty = e^{-\gamma\theta R_\infty} h_0 + \left(1 - e^{-\gamma\theta R_\infty}\right) \frac{1}{\theta}, \quad (5.15)$$

donde  $R_\infty = \int_0^\infty r_s ds \in [0, \infty]$ . Observemos que  $h_t \in [0, 1]$  para todo  $t \geq 0$ , por lo que también  $h_\infty \in [0, 1]$ . Probemos ahora que  $h_\infty = \min(1, \frac{1}{\theta})$ . Ante todo, realizamos algunas observaciones:

1.  $h_t$  y  $h_\infty$  son combinaciones convexas de  $h_0 < 1$  (ya que  $\mathbb{E}[Z_0] > 0$ ) y  $\frac{1}{\theta}$ .
2. Si  $R_\infty = \infty$ , entonces  $h_\infty = \frac{1}{\theta}$ . La afirmación recíproca es también cierta cuando  $h_0 \neq \frac{1}{\theta}$  (que es el caso cuando  $\theta \leq 1$ , porque  $h_0 < 1 \leq \frac{1}{\theta}$ ).
3. Si  $R_\infty < \infty$ , entonces  $h_\infty = 1$ . Eso sucede porque  $R_\infty < \infty$  implica  $\lim_{n \rightarrow \infty} \mathbb{E}[Z_{t_n}] = 0$  para alguna secuencia  $t_n \rightarrow \infty$ . Por lo tanto, la desigualdad  $e^{-x} \geq 1 - x$  implica  $h_\infty = \lim_{n \rightarrow \infty} \mathbb{E}[e^{-(\gamma/\mu)Z_{t_n}}] \geq \lim_{n \rightarrow \infty} \mathbb{E}[1 - (\gamma/\mu)Z_{t_n}] = 1$ , por lo tanto  $h_\infty = 1$ .

Consecuentemente,

- Cuando  $\theta < 1$ , se tiene necesariamente  $R_\infty < \infty$ , porque de otra manera la observación 2 implicaría  $h_\infty = \frac{1}{\theta} > 1$ . Luego, la observación 3 implica  $h_\infty = 1$ .
- Cuando  $\theta = 1$ , entonces  $h_\infty = 1$  independientemente del valor de  $R_\infty$ , gracias a las observaciones 2 y 3. Pero usando la recíproca en la observación 2, podemos deducir que  $R_\infty = \infty$ .
- Cuando  $\theta > 1$ , tenemos  $\frac{1}{\theta} < 1$  y la observación 1 implica que  $h_\infty < 1$ . En consecuencia, por la observación 3,  $R_\infty = \infty$ , que a su vez implica  $h_\infty = \frac{1}{\theta}$  por la observación 2.

Esto prueba que  $h_\infty = \min(1, \frac{1}{\theta})$ ; y simultáneamente, que  $\int_0^\infty \mathbb{E}[Z_t] dt < \infty$  cuando  $\theta < 1$  y  $\int_0^\infty \mathbb{E}[Z_t] dt = \infty$  cuando  $\theta = 1$ . Por lo tanto,  $\lim_{t \rightarrow \infty} \mathbb{E}[e^{-(\gamma/\mu)Z_t}] = 1$  para  $\theta \leq 1$ , lo que implica que  $\mathbb{E}[\omega(Z_t, 0)] \rightarrow 0$  para la métrica  $\omega(x, y) = 1 - e^{-(\gamma/\mu)|x-y|}$ . De esto se deduce que  $1 - e^{-(\gamma/\mu)Z_t} \rightarrow 0$  en probabilidad, y por lo tanto  $Z_t \rightarrow 0$  en probabilidad. Como  $(Z_t)_{t \geq 0}$  es uniformemente integrable, se sigue que  $\mathbb{E}[Z_t] \rightarrow 0$ .

Para el caso  $\theta < 1$ , probemos ahora que  $\lim_t Z_t = 0$  casi seguramente: como  $\int_0^\infty r_t dt < \infty$ , podemos deducir que sólo finitos átomos  $(t, u)$  de  $\mathcal{M}$  satisfacen  $u \leq \gamma r_t$ , lo cual implica que el instante del último spike de  $Z_t$  es finito casi seguramente; tras ese evento,  $Z_t$  convergerá a cero debido al decaimiento exponencial o a un reseteo final.

Todavía en el caso  $\theta < 1$ , probemos ahora (5.4): observando que  $h_t \leq 1$  y  $(\gamma/\mu)Z_t \geq 1 - e^{-(\gamma/\mu)Z_t}$ , de (5.14) obtenemos

$$\frac{d}{dt} h_t \geq (1 - \theta)\gamma r_t \geq (1 - \theta)\mu(1 - h_t).$$

Por lo tanto,  $g_t = 1 - h_t$  satisface  $\frac{d}{dt} g_t \leq -(1 - \theta)\mu g_t$ , de lo que se deduce que  $g_t \leq g_0 e^{-(1-\theta)\mu t}$ . Esto confirma (5.4) y concluye la prueba de las afirmaciones para los casos  $\theta < 1$  y  $\theta = 1$ .

Finalmente, tratamos el caso  $\theta > 1$ : usando la observación 1, vemos que

$$\sup_{t \geq 0} h_t \leq \max \left( h_0, \frac{1}{\theta} \right) < 1.$$

Como  $(\gamma/\mu)\mathbb{E}[Z_t] \geq 1 - h_t$ , se sigue que  $\inf_{t \geq 0} \mathbb{E}[Z_t] > 0$ . Esto completa la prueba.  $\square$

*Prueba del Corolario 5.2.1.* Sea  $p_t = \mathbb{P}(Z_t = 0) = \mathbb{E}[\mathbf{1}(Z_t = 0)]$ . Claramente, en un reseteo la indicadora  $\mathbf{1}(Z_t = 0)$  se incrementa en  $1 - \mathbf{1}(Z_{t^-} = 0)$ ; en un evento de spike, se incrementa en  $-\mathbf{1}(Z_{t^-} = 0)$ , mientras que el drift no afecta a  $p_t$ . Por lo tanto,

$$\frac{d}{dt} p_t = \mathbb{E}[(1 - \mathbf{1}(Z_t = 0))\gamma Z_t] - \mathbb{E}[\mathbf{1}(Z_t = 0)\gamma \kappa r_t] = \gamma r_t - \gamma \kappa r_t p_t.$$

donde  $r_t = \mathbb{E}[Z_t]$ . Esta ecuación diferencial la resuelve

$$p_t = \frac{1}{\kappa} + \left( p_0 - \frac{1}{\kappa} \right) e^{-\gamma \kappa \int_0^t r_s ds},$$

que es exactamente (5.5). Más aún, gracias al Teorema 5.2.2, sabemos que  $\int_0^\infty r_s ds = \infty$  cuando  $\mathbb{E}[Z_0] > 0$  y  $\theta \geq 1$ . Tomando límite, obtenemos  $\lim_{t \rightarrow \infty} p_t = \frac{1}{\kappa}$  en este caso.  $\square$

## 5.6. Prueba de propagación del caos y persistencia (Teorema 5.3.1)

Asumimos  $\kappa = 2$ , y por simplicidad, en esta sección asumimos  $\mu = 1$  y  $\rho = 1$ . Notamos  $f_t = \mathcal{L}(Z_t)$ , donde  $Z_t$  es la única solución para (4.6). Sean  $\mathcal{N}(dt, du)$  y  $\hat{\mathcal{M}}(dt, du, dz)$  medidas aleatorias de Poisson independientes sobre  $\mathbb{R}_+^2$  y  $\mathbb{R}_+^3$  con respectivas medidas de intensidad  $dt du$  y  $2dt du f_t(dz)$ , ambos independientes de  $Z_0$ . Consideremos la SDE

$$dZ_t = -Z_t dt - Z_t \int_0^\infty \mathbf{1}(u \leq \gamma Z_t) \mathcal{N}(dt, du) + \int_0^\infty \int_0^\infty \mathbf{1}(u \leq \gamma z) \hat{\mathcal{M}}(dt, du, dz). \quad (5.16)$$

**Lema 5.6.1.** *La Ecuación (5.16) admite una única solución fuerte que tiene la misma ley que la solución a (4.6).*

*Demostración.* (5.16) puede reducirse a (4.6) en un sentido fuerte: definimos  $\mathcal{M}$  como la medida puntual sobre  $\mathbb{R}_+^2$  con átomos  $(t, ur_t/z)$  por cada átomo  $(t, u, z)$  de  $\hat{\mathcal{M}}$  con  $z > 0$ , donde  $r_t = \int_0^\infty z f_t(dz)$ . Es directo verificar que  $\mathcal{M}(dt, du)$  es una medida de Poisson aleatoria con intensidad  $2dt du$  y que (5.16) es exactamente (4.6). La afirmación entonces se deduce de esta reducción y del Teorema 5.2.1.  $\square$

Reescribamos ahora la dinámica (4.1) de la siguiente manera: Como el tamaño de  $K$  es  $\kappa = 2$ , notamos a los procesos de Poisson ( $\mathcal{N}_K$ ) con una notación más cómoda  $(\mathcal{N}_{ij})_{1 \leq i < j \leq N}$ , donde cada  $\mathcal{N}_{ij}$  tiene medida de intensidad:

$$\frac{2}{(N-1)(N-2)} dt du \mathbf{1}([\xi] \notin \{i, j\}) d\xi \quad (5.17)$$

y notemos  $\mathcal{N}_{ij} = \mathcal{N}_{ji}$  siempre que  $i > j$ . Como dijimos antes, observamos que  $\mathcal{N}_{ij}$  provee la aleatoriedad para todos los saltos posibles donde son las neuronas  $i$  y  $j$  las que son excitadas

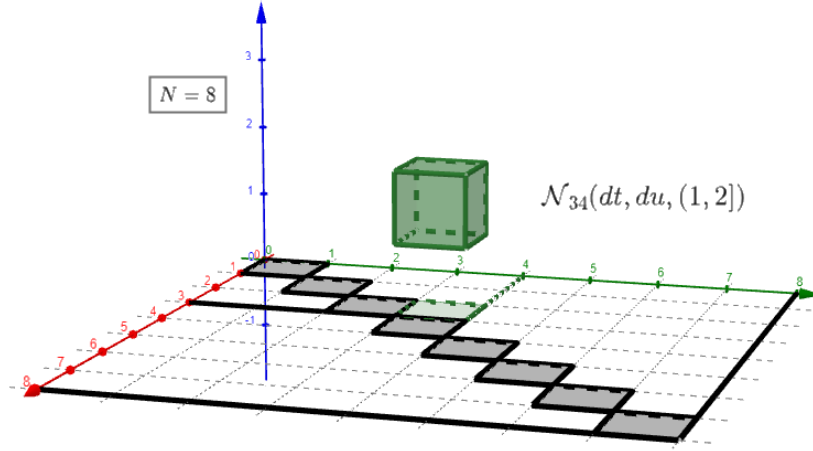


Figura 5.1: Representación del Proceso de Poisson  $\mathcal{N}_{34}(dt, du, (1, 2])$  con  $N = 8$ .

simultáneamente en un evento de spike, y la partícula responsable de este evento, que es la reseteada, corresponde a  $[\xi] \in \{1, \dots, N\} \setminus \{i, j\}$ . Como fue adelantado en el capítulo 3.4,  $\mathcal{N}_{ij}$  admite una representación geométrica en  $\mathbb{R}^3$  que resulta útil para entender las ideas detrás de los argumentos que siguen. En la figura 5.1, realizamos una representación heurística en  $\mathbb{R}^3$  del proceso de Poisson  $\mathcal{N}_{34}(dt, du, (1, 2])$ , que codifica las interacciones donde la neurona  $k = 2$  se disparó, excitando a las neuronas  $i = 3$  y  $j = 4$ . La heurística consiste en identificar al proceso de Poisson  $\mathcal{N}_{ij}(dt, du, (k-1, k])$  con un cubo  $\mathcal{C}_{ij}^k := [i-1, i] \times [j-1, j] \times [k-1, k]$ .

Con estos elementos básicos, para  $i \in \{1, \dots, N\}$  definimos las medidas aleatorias de Poisson

$$\mathcal{N}^i(dt, du) = \sum_{\substack{k, j \neq i \\ j < k}} \mathcal{N}_{jk}(dt, du, (i-1, i]) \quad (5.18)$$

$$\mathcal{M}^i(dt, du, d\xi) = \sum_{j \neq i} \mathcal{N}_{ij}(dt, du, d\xi). \quad (5.19)$$

Observemos que  $\mathcal{N}^i$  codifica los instantes de spike de la neurona  $i$ , y  $\mathcal{M}^i$  codifica los instantes en que la neurona  $i$  es excitada. En las figuras 5.2 y 5.3 se encuentran a modo ilustrativo las representaciones heurísticas de  $\mathcal{N}^3$  y  $\mathcal{M}^4$  respectivamente, para  $N = 8$ . Observar que éstas no son más que la unión de los cubos  $\mathcal{C}_{ij}^k$  correspondientes.

Claramente, la intensidad de  $\mathcal{N}^i$  es

$$\sum_{\substack{j < k \\ k, j \neq i}} \frac{2}{(N-1)(N-2)} dt du = dt du,$$

mientras que la intensidad de  $\mathcal{M}^i$  es

$$\sum_{j \neq i} \frac{2}{(N-1)(N-2)} dt du d\xi \mathbf{1}([\xi] \notin \{i, j\}) = \frac{2}{N-1} dt du d\xi,$$

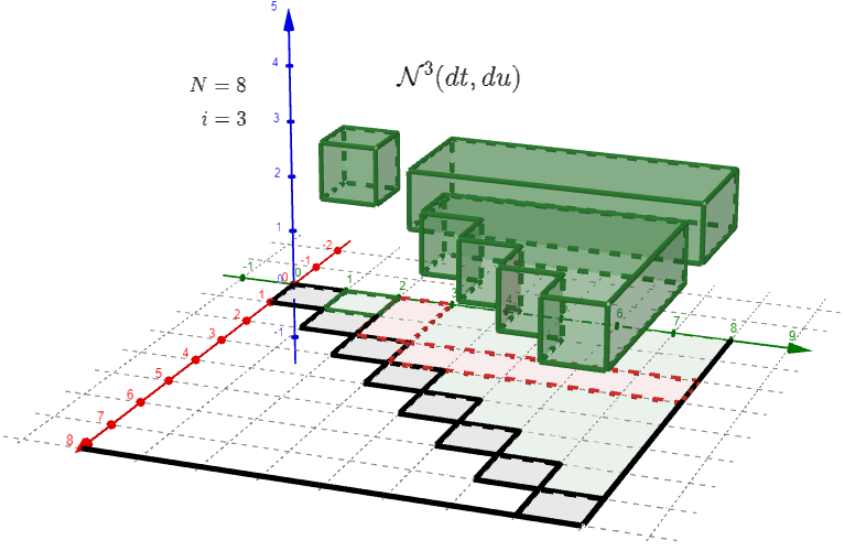


Figura 5.2: Representación del Proceso de Poisson  $\mathcal{N}^3$  con  $N = 8$ .

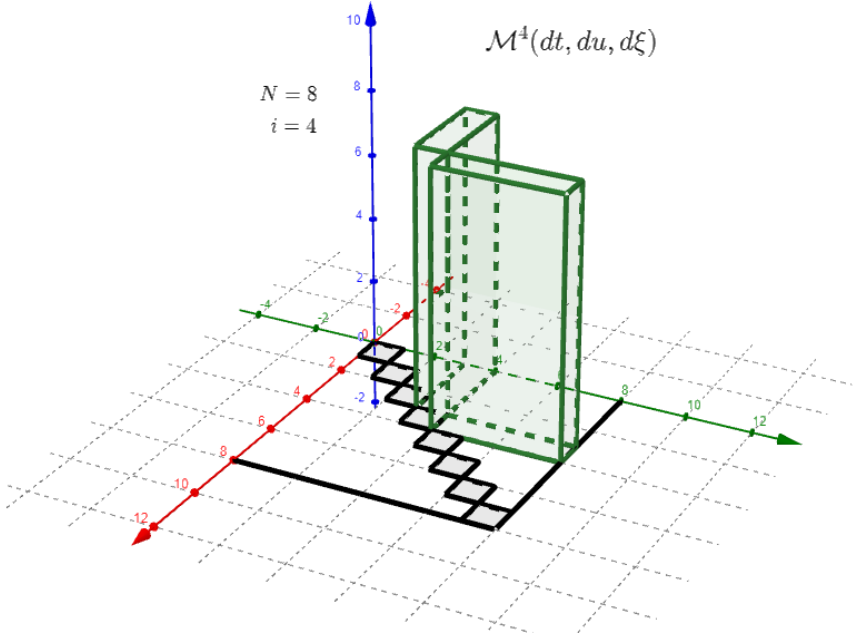


Figura 5.3: Representación del Proceso de Poisson  $\mathcal{M}^4$  con  $N = 8$

pues  $\sum_{j \neq i} \mathbf{1}([\xi] \notin \{i, j\}) = (N - 2)\mathbf{1}([\xi] \neq i)$ .

Por lo tanto, (4.1) puede expresarse como

$$\begin{aligned} dX_t^i &= -X_t^i dt - X_t^i \int_0^\infty \mathbf{1}(u \leq \gamma X_t^i) \mathcal{N}^i(dt, du) \\ &\quad + \int_0^\infty \int_0^N \mathbf{1}(u \leq \gamma X_t^{[\xi]}) \mathcal{M}^i(dt, du, d\xi), \end{aligned} \quad (5.20)$$

que tiene un aspecto muy similar a (5.16). La única diferencia es que en el tercer término de (5.16) la variable  $z$  es una instancia de  $f_t = \mathcal{L}(Z_t)$ , mientras que en (5.20) esto es reemplazado por  $X_t^{[\xi]}$ , que elige aleatoriamente una neurona  $j \neq i$ ; esto es,  $X_t^{[\xi]}$  es una instancia sorteada a partir de la medida empírica (aleatoria)  $\bar{\mathbf{X}}_t^i$ . Acá, utilizamos las notaciones

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \delta_{x^i} \quad \text{y} \quad \bar{\mathbf{x}}^i = \frac{1}{N-1} \sum_{j \neq i} \delta_{x^j}.$$

Motivados por esto, para cada  $i$  introducimos un mapa medible  $(t, \mathbf{z}, u) \mapsto F_t^i(\mathbf{z}, u)$  de  $\mathbb{R}_+ \times \mathbb{R}_+^N \times (0, N]$  en  $\mathbb{R}_+$  tal que

$$\mathcal{L}(F_t^i(\mathbf{z}, U)) = f_t$$

y

$$\mathbb{E} \left[ \left| F_t^i(\mathbf{z}, U) - z^{\lceil U \rceil} \right| \right] = W_1(f_t, \bar{\mathbf{z}}^i) \quad (5.21)$$

para todo  $t \in \mathbb{R}_+$ , para todo  $\mathbf{z} \in \mathbb{R}_+^N$ , y para cada variable aleatoria  $U$  uniformemente distribuida en  $(0, N] \setminus (i-1, i]$ . Observando que  $\mathcal{L}(z^{\lceil U \rceil}) = \bar{\mathbf{z}}^i$ , vemos que el par  $(F_t^i(\mathbf{z}, U), z^{\lceil U \rceil})$  constituye un acoplamiento óptimo entre  $f_t = \mathcal{L}(Z_t)$  y la medida empírica  $\bar{\mathbf{z}}^i = \frac{1}{N-1} \sum_{j \neq i} \delta_{z^j}$ . Ver [CF16, Lema 3] para la prueba de existencia de un mapa así. Más aún, ese mismo resultado asegura que dicho mapa satisface

$$\mathbb{E} \int_{j-1}^j \phi(F_t^i(\mathbf{Y}, \xi)) d\xi = \int \phi(z) f_t(dz) \quad (5.22)$$

para todo  $j \in \{1, \dots, N\}$ ,  $j \neq i$ , para toda función medible  $\phi : \mathbb{R} \rightarrow \mathbb{R}$ , y para todo vector intercambiable  $\mathbf{Y} \in \mathbb{R}^N$ .

Ahora, especificamos nuestro acoplamiento definiendo  $\mathbf{Z}_t = (Z_t^1, \dots, Z_t^N)$  como la solución para

$$\begin{aligned} dZ_t^i &= -Z_t^i dt - Z_t^i \int_0^\infty \mathbf{1}(u \leq \gamma Z_t^i) \mathcal{N}^i(dt, du) \\ &\quad + \int_0^\infty \int_{(0, N]} \mathbf{1}(u \leq \gamma F_t^i(\mathbf{Z}_t, \xi)) \mathcal{M}^i(dt, du, d\xi), \end{aligned} \quad (5.23)$$

comenzando en  $Z_0^i = X_0^i$  (recordar que  $X_0^1, \dots, X_0^N$  son copias i.i.d. de  $Z_0$ ). Acá  $\mathcal{N}^i$  y  $\mathcal{M}^i$  son las mismas medidas de Poisson que aparecen en (5.20). Gracias al Lema 5.6.1, la ecuación (5.23) admite una única solución fuerte, y es un proceso no lineal. Sin embargo, estos procesos **no son independientes** porque tienen saltos simultáneos. No obstante, su dependencia se va disipando a medida que  $N$  crece, de acuerdo al siguiente resultado:

**Lema 5.6.2.** *Existe una constante  $C > 0$  tal que*

$$\mathbb{E} \left[ W_1(f_t, \bar{\mathbf{Z}}_t) \right] \leq \frac{C}{N^{1/3}}.$$

*Lo mismo vale para  $\bar{\mathbf{Z}}_t^i$  en lugar de  $\bar{\mathbf{Z}}_t$ , para todo  $i \in \{1, \dots, N\}$ .*

Para demostrar este lema, precisemos algunas notaciones. En primer lugar, dado  $n \in \mathbb{N}$  y  $\nu \in \mathcal{P}(\mathbb{R})$ , notamos

$$\varepsilon_n(\nu) := \mathbb{E}[W_1(\bar{\mathbf{Y}}, \nu)],$$

donde  $\mathbf{Y} = (Y^1, \dots, Y^n)$  es una colección de variables independientes y  $\nu$ -distribuidas.

En segundo lugar, dadas  $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$ , consideraremos la métrica 1-Wasserstein con la distancia 1 normalizada en  $\mathbb{R}^n$  (ver 2.2), es decir:

$$W_1(\mu, \nu) = \inf_{\substack{\mathbf{X} \sim \mu \\ \mathbf{Y} \sim \nu}} \mathbb{E} \left[ \frac{1}{n} \sum_{i=1}^n |X^i - Y^i| \right]$$

Finalmente,  $\mathcal{L}^n(\mathbf{Y})$  denota la ley de las primeras  $n$  coordenadas de un vector aleatorio intercambiable  $\mathbf{Y}$  de  $\mathbb{R}^N$ .

*Prueba del Lema 5.6.2.* Notemos  $m_r(\nu) := \int_{\mathbb{R}} |z|^r \nu(dz)$  para cualquier  $\nu \in \mathcal{P}(\mathbb{R})$ . Mediante aplicación directa de [CF16, Lemma 7], vale lo siguiente para todo  $n \leq N$  y  $t \geq 0$ :

$$\begin{aligned} \mathbb{E}[W_1(\bar{\mathbf{Z}}_t, f_t)] &\leq \frac{q}{N} \frac{n}{N} \left( W_1(f_t^{\otimes n}, \mathcal{L}^n(\mathbf{Z}_t)) + \varepsilon_n(f_t) \right) \\ &\quad + \frac{l}{N} \left( W_1(f_t^{\otimes l}, \mathcal{L}^l(\mathbf{Z}_t)) + \varepsilon_l(f_t) \right) \\ &\leq W_1(f_t^{\otimes n}, \mathcal{L}^n(\mathbf{Z}_t)) + \varepsilon_n(f_t) + C \frac{n}{N} m_1(f_t) \end{aligned} \tag{5.24}$$

donde  $q \in \mathbb{N}$  y  $l \in \mathbb{N}$  son respectivamente el cociente y el resto de dividir por  $n$ ; y donde hemos acotado  $W_1(f_t^{\otimes l}, \mathcal{L}^l(\mathbf{Z}_t))$  y  $\varepsilon_l(f_t)$  por  $2m_1(f_t)$ ,  $\frac{qn}{N}$  por 1, y  $l$  por  $n$ .

Basta entonces con encontrar cotas adecuadas para los tres términos. Observemos que:

- El Lema 5.6.3 más abajo afirma que existe una constante  $C > 0$  tal que  $W_1(f_t^{\otimes n}, \mathcal{L}^n(\mathbf{Z}_t)) \leq C \frac{n}{N}$  para todo  $n \in \{2, \dots, N\}$
- El Lema 5.6.4 implica que  $m_1(f_t), m_2(f_t), m_3(f_t)$  están uniformemente acotados.
- Haciendo uso de [FG15, Teorema 1] con  $p, d = 1$ , se tiene la siguiente cota para cualquier  $\nu \in \mathcal{P}(\mathbb{R})$ :

$$\varepsilon_n(\nu) = C(r) m_r^{1/r}(\nu) \left( n^{-1/2} + n^{\frac{r-1}{r}} \right) \leq C'(r) m_r^{1/r}(\nu) n^{-1/2} \tag{5.25}$$

para cualquier  $r > 2$ .

- En particular, usando  $r = 3$  se sigue que  $\varepsilon_n(f_t) \leq C n^{-1/2}$

Por lo tanto, tomando  $n = \lceil N^{2/3} \rceil$ . Es inmediato que  $n/N \sim N^{-1/3}$  y de igual manera  $n^{-1/2} \sim N^{-1/3}$ . En consecuencia, de (5.24) se sigue que

$$\mathbb{E}[W_1(\bar{\mathbf{Z}}_t, f_t)] \leq W_1(f_t^{\otimes n}, \mathcal{L}^n(\mathbf{Z}_t)) + \varepsilon_n(f_t) + C \frac{n}{N} m_1(f_t) \leq \frac{\hat{C}}{N^{1/3}}$$

y se tiene el resultado para  $\bar{\mathbf{Z}}_t$ . Para  $\bar{\mathbf{Z}}_t^i$  la demostración es análoga, trabajando con el vector  $\mathbf{Z}_t$  sin la  $i$ -ésima coordenada. □

Procedemos ahora a demostrar los lemas 5.6.3 y 5.6.4, que completan la demostración del lema 5.6.2.

**Lema 5.6.3.** *Existe una constante  $C > 0$  tal que para todo  $n \in \{2, \dots, N\}$ , y  $t \geq 0$ ,*

$$W_1(f_t^{\otimes n}, \mathcal{L}^n(\mathbf{Z}_t)) \leq C \frac{n}{N}.$$

*Prueba del Lema 5.6.3.* Fijamos un entero  $2 \leq n \leq N$ . La idea es definir un segundo acoplamiento  $\mathbf{V}_t = (V_t^1, \dots, V_t^n)$  tal que  $\mathcal{L}(\mathbf{V}_t) = f_t^{\otimes n}$  (es decir,  $\mathbf{V}_t$  es una colección independiente de  $n$  procesos no lineales), y luego probar que la ley de este nuevo acoplamiento está cerca, en distancia Wasserstein, de  $\mathcal{L}^n(\mathbf{Z}_t)$ .

Para lograr esto, la idea en la construcción de  $\mathbf{V}_t$  es que, siempre que más de un proceso de  $(Z_t^1, \dots, Z_t^n)$  experimenta un salto simultáneamente (correspondiente a una interacción entre las primeras  $n$  neuronas), vamos a excluir parte de esa interacción en la SDE que define a  $\mathbf{V}_t$  de manera que sólo uno de los procesos experimenta un salto y de esta forma asegurar la independencia. Para compensar los saltos “borrados”, añadiremos otros saltos de fuentes independientes.

Fijemos una neurona  $i \in \{1, \dots, n\}$ . A partir de la expresión de  $Z_t^i$  en (5.23), observemos que su dinámica está determinada por los procesos  $\mathcal{N}^i$  y  $\mathcal{M}^i$ , que codifican los eventos de firing y de excitación de la neurona  $i$ , respectivamente. Definiremos  $V_t^i$  a través de una SDE análoga a la de  $Z_t^i$ , utilizando el mismo  $\mathcal{N}^i$ , pero construyendo cuidadosamente un proceso  $\mathcal{R}^i$  para los eventos de excitación, de manera de asegurar la independencia entre los procesos de  $\mathbf{V}_t$ .

Recordemos la expresión de  $\mathcal{M}^i$  en (5.19):

$$\begin{aligned} \mathcal{M}^i(dt, du, d\xi) &= \sum_{j \neq i} \mathcal{N}_{ij}(dt, du, d\xi) \\ &= \sum_{j > n} \mathbf{1}(\xi > n) \mathcal{N}_{ij}(dt, du, d\xi) + \sum_{j \neq i} \mathbf{1}((\xi \leq n) \vee (j \leq n)) \mathcal{N}_{ij}(dt, du, d\xi) \end{aligned} \tag{5.26}$$

donde  $(\mathcal{N}_{ij})_{i < j}$  son procesos de Poisson independientes sobre  $\mathbb{R}_+^2 \times (0, N]$  y  $\mathcal{N}_{ij} = \mathcal{N}_{ji}$  si  $j < i$ . Observemos además que cada evento donde la neurona  $i \in \{1, \dots, n\}$  es excitada está asociado a un proceso  $\mathcal{N}_{ij}$  con marca  $\xi \in (k-1, k]$  donde  $k \neq i$ . Por lo tanto, para asegurar la independencia en  $\mathbf{V}_t$ , debemos reemplazar por fuentes independientes siempre que  $j \leq n$  ó  $k \leq n$ , condición correspondiente al segundo término en (5.26).

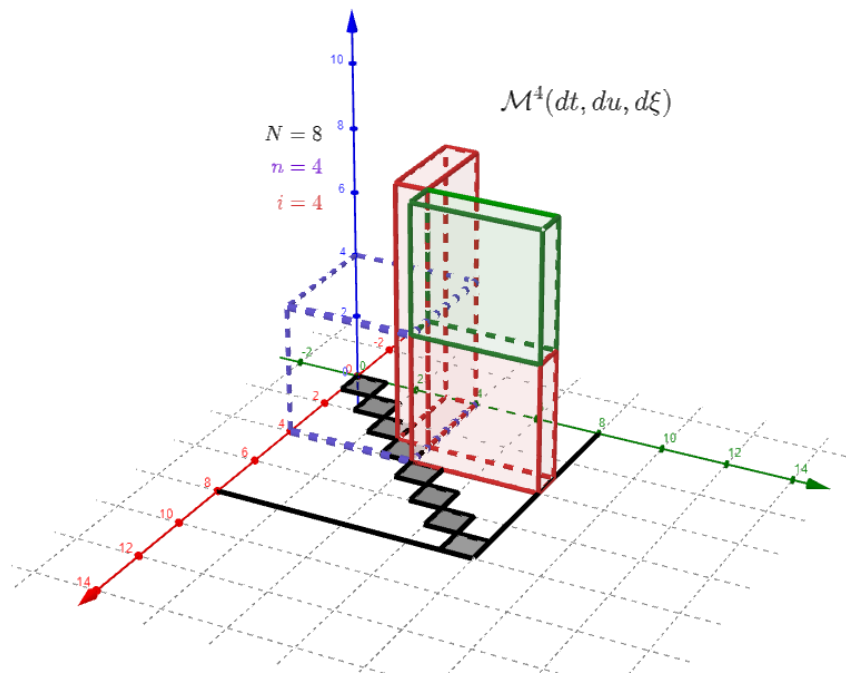


Figura 5.4: Representación del proceso de Poisson  $\mathcal{M}^4$  con  $N = 8$ ,  $n = 4$  y  $i = 4$ . El “bloque” verde representa la restricción de  $\mathcal{M}^4$  que no interactúa con otras neuronas de  $\{1, \dots, n\}$ , mientras que los “bloques” rojos representan las que sí.

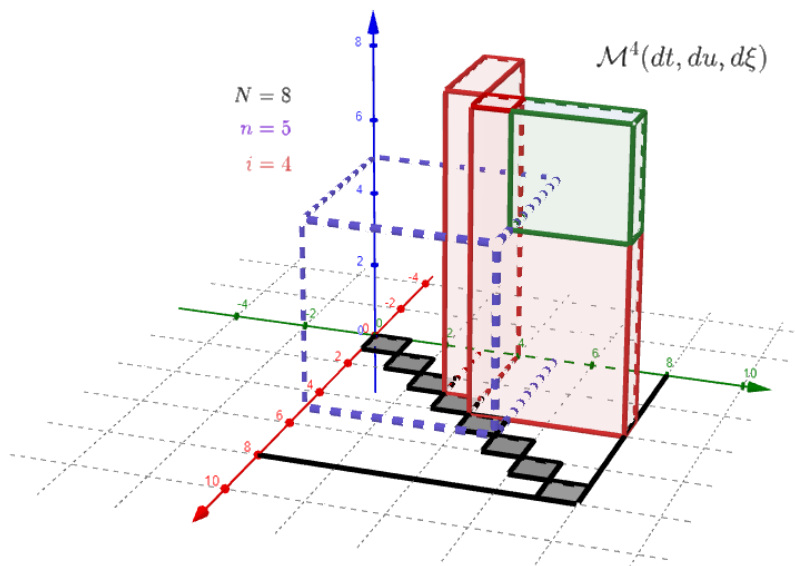


Figura 5.5: Representación del proceso de Poisson  $\mathcal{M}^4$  con  $N = 8$ ,  $n = 5$  y  $i = 4$ . El “bloque” verde representa la restricción de  $\mathcal{M}^4$  que no interactúa con otras neuronas de  $\{1, \dots, n\}$ , mientras que los “bloques” rojos representan las que sí.



Las figuras 5.4 y 5.5 ilustran esto (comparar 5.4 con 5.3).

En consecuencia, definimos:

$$\mathcal{R}^i(dt, du, d\xi) = \sum_{j>n} \mathbf{1}(\xi > n) \mathcal{N}_{ij}(dt, du, d\xi) + \sum_{j \neq i} \mathbf{1}((\xi \leq n) \vee (j \leq n)) \tilde{\mathcal{N}}_{ij}(dt, du, d\xi) \quad (5.27)$$

donde  $(\tilde{\mathcal{N}}_{jk}(dt, du, d\xi))_{j \neq k}$  es una colección de copias de  $\mathcal{N}_{12}$  independientes. Notar que en este caso  $\tilde{\mathcal{N}}_{jk}(dt, du, d\xi)$  y  $\tilde{\mathcal{N}}_{kj}(dt, du, d\xi)$  están definidos como dos procesos independientes, a diferencia de  $\mathcal{N}_{jk}$  y  $\mathcal{N}_{kj}$  que son iguales por definición. Observar que de esta manera efectivamente no hay proceso que implique saltos simultáneos entre las neuronas. Por ejemplo, suponiendo  $n = 5$ , el evento donde la neurona  $k = 4$  emite un firing y selecciona a las neuronas  $i = 2$  y  $j = 3$  en  $\mathbf{Z}_t$  involucra un átomo del proceso de Poisson  $\mathcal{N}_{23}(dt, du, (3, 4])$ , presente en  $\mathcal{M}^2$ ,  $\mathcal{M}^3$  y  $\mathcal{N}^4$ . En cambio para el proceso  $\mathbf{V}_t$  dicho evento implica átomos en los procesos  $\tilde{\mathcal{N}}_{23}(dt, du, (3, 4])$ ,  $\tilde{\mathcal{N}}_{32}(dt, du, (3, 4])$  y  $\mathcal{N}_{23}(dt, du, (3, 4])$ , presentes en  $\mathcal{R}^2$ ,  $\mathcal{R}^3$  y  $\mathcal{N}^4$  respectivamente.

Notar entonces que  $\mathcal{R}^i$  es un proceso de Poisson con la misma distribución que  $\mathcal{M}^i$ , y más aun,  $\mathcal{N}^1, \dots, \mathcal{N}^n$  y  $\mathcal{R}^1, \dots, \mathcal{R}^n$  son independientes, ya que ningún bloque de  $(\mathcal{N}_{jk})_{j < k}$  o de  $(\tilde{\mathcal{N}}_{jk})_{j \neq k}$  se repite.

De esta manera, definimos  $V_0^i = \mathbf{Z}_0^i$  y  $(V_t^i)_{t \geq 0}$  como

$$\begin{aligned} dV_t^i &= -V_t^i dt - V_t^i \int_{\mathbb{R}_+} \mathbf{1}(u \leq \gamma V_t^i) \mathcal{N}^i(dt, du) \\ &\quad + \int_{\mathbb{R}_+} \int_{(0, N]} \mathbf{1}(u \leq \gamma F_t^i(\mathbf{Z}_t, \xi)) \mathcal{R}^i(dt, du, d\xi) \end{aligned} \quad (5.28)$$

Observemos que la construcción de esta SDE es esencialmente la misma que define a  $Z_t^i$  (5.23), incluso con la variable  $F_t^i(\mathbf{Z}_t, \xi)$  utilizada para samplear  $f_t$ , pero con el proceso  $\mathcal{R}^i$  en lugar de  $\mathcal{M}^i$ . Por lo tanto  $V_t^1, \dots, V_t^n$  son procesos no lineales, y más aún, son independientes. Ésta última afirmación no es trivial: si bien los procesos de Poisson que definen las SDEs son efectivamente independientes, la variable  $F_t^i(\mathbf{Z}_t, \xi)$  podría a priori inducir correlaciones. Nos referimos a [CF16, Lema 6] para una prueba de que éste no es el caso.

*Observación 5.6.1.* La construcción de estos acoplamientos con  $\kappa \in \{1, \dots, N\}$  general son análogos, dejamos las construcciones explícitas disponibles en el apéndice, ver A.

Verifiquemos ahora que la ley de  $\mathbf{V}_t$  (que es  $f_t^{\otimes n}$ ) permanece cerca de  $\mathcal{L}^n(\mathbf{Z}_t)$ . Observemos que

$$W_1(f_t^{\otimes n}, \mathcal{L}^n(\mathbf{Z}_t)) \leq \mathbb{E} \left[ \frac{1}{n} \sum_{i=1}^n |Z_t^i - V_t^i| \right] = \mathbb{E}[|Z_t^1 - V_t^1|] := h_t,$$

donde hemos usado intercambiabilidad en la última igualdad. Por lo tanto, basta con acotar  $h_t$ . Utilizando (5.23) y (5.28), tenemos para  $0 \leq s \leq t$ :

$$h_t - h_s = - \int_s^t h_r dr + \mathbf{F} + \mathbf{E} + \mathbf{E}^Z + \mathbf{E}^V, \quad (5.29)$$

donde  $\mathbf{F}$  es el término asociado a los firing (es decir, a los átomos de  $\mathcal{N}^1$ ),  $\mathbf{E}$  corresponde a excitaciones conjuntas (primer término de (5.26), que es el mismo que el primer término de (5.27)),  $\mathbf{E}^Z$  corresponde a las excitaciones exclusivas de  $Z^1$  (segundo término de (5.26)), y finalmente  $\mathbf{E}^V$  es el término correspondiente a las excitaciones exclusivas de  $V^1$  (segundo término de (5.27)). La mayor parte de los saltos se encuentran en  $\mathbf{F}$  y  $\mathbf{E}$ , y la idea es mostrar que estos términos son no positivos, mientras que los términos  $\mathbf{E}^Z$  y  $\mathbf{E}^V$  son de orden  $(t-s)\frac{n}{N}$ . En consecuencia, de (5.29), obtendremos la desigualdad

$$\partial_t h_t \leq -h_t + C \frac{n}{N}, \quad (5.30)$$

donde  $C$  es una constante que puede cambiar de línea en línea. El resultado quedará entonces demostrado por el lema de Grönwall.

Procedemos entonces a escribir  $\mathbf{F}$ ,  $\mathbf{E}$ ,  $\mathbf{E}^Z$  y  $\mathbf{E}^V$  explícitamente y mostrar que satisfacen las cotas correspondientes. Sea  $\Delta_r := |Z_r^1 - V_r^1| - |Z_{r^-}^1 - V_{r^-}^1|$ . Entonces:

- De (5.26) y (5.27), se implica inmediatamente que

$$\mathbf{F} = \mathbb{E} \int_s^t \int_0^\infty \Delta_r \mathcal{N}^1(dr, du).$$

Y con un argumento análogo al de (5.13), se puede ver fácilmente que

$$\mathbf{F} = -\gamma \mathbb{E} \int_s^t (Z_r^1 - V_r^1)^2 dr \leq 0.$$

- También a partir de (5.26) y (5.27), se deduce que

$$\mathbf{E} = \sum_{j>n} \mathbb{E} \int_s^t \int_0^\infty \int_n^N \Delta_r \mathcal{N}_{1j}(dr, du, d\xi).$$

Como dijimos, este término codifica los saltos conjuntos, y por lo tanto deberíamos tener  $\mathbf{E} = 0$ . En efecto, como la misma indicadora  $\mathbf{1}(u \leq \gamma F_r^1(\mathbf{Z}_{r^-}, \xi))$  aparece tanto en (5.23) como en (5.28), vemos que en este caso  $Z^1$  y  $V^1$  reciben el impulso excitatorio de manera conjunta o bien no reciben impulso. Por lo tanto  $\Delta_r = 0$  en esta integral, y en consecuencia  $\mathbf{E} = 0$ .

- De nuevo, a partir de (5.26) y (5.27), deducimos que

$$\mathbf{E}^Z = \sum_{j \neq 1} \mathbb{E} \int_s^t \int_0^\infty \int_0^N \Delta_r \mathbf{1}((j \leq n) \vee (\xi \leq n)) \mathcal{N}_{1j}(dr, du, d\xi)$$

En este caso, dado que sólo  $Z^1$  recibe el impulso excitatorio, tenemos:

$$\begin{aligned} \Delta_r &= \mathbf{1}\left(u \leq \gamma F_r^1(\mathbf{Z}_{r^-}, \xi)\right) \left[|Z_r^1 + 1 - V_r^1| - |Z_{r^-}^1 - V_{r^-}^1|\right] \\ &\leq \mathbf{1}\left(u \leq \gamma F_r^1(\mathbf{Z}_{r^-}, \xi)\right). \end{aligned}$$

Recordemos que la intensidad de  $\mathcal{N}_{1j}(dr, du, d\xi)$  es  $\frac{2 \mathbf{1}([\xi] \notin \{1, j\})}{(N-1)(N-2)} dr du d\xi$ , ver (5.17). Por lo tanto, tenemos que:

$$\mathbf{E}^Z \leq \sum_{j \neq 1} \mathbb{E} \int_s^t \int_0^N \gamma F_r^1(\mathbf{Z}_{r^-}, \xi) \mathbf{1}((j \leq n) \vee (\xi \leq n)) \frac{2 \mathbf{1}([\xi] \notin \{1, j\})}{(N-1)(N-2)} d\xi dr.$$

Ahora, reescribiendo la indicadora  $\mathbf{1}((j \leq n) \vee (\xi \leq n))$  como la suma de  $\mathbf{1}(j \leq n, \xi > n)$ ,  $\mathbf{1}(j > n, \xi \leq n)$  y  $\mathbf{1}(j \leq n, \xi \leq n)$ , obtenemos tres términos, que llamamos respectivamente  $\mathbf{E}_1^Z$ ,  $\mathbf{E}_2^Z$  and  $\mathbf{E}_3^Z$ . Por ejemplo,  $\mathbf{E}_1^Z$  está dada por

$$\mathbf{E}_1^Z = \frac{2\gamma}{(N-1)(N-2)} \sum_{j=2}^n \sum_{k=n+1}^N \int_s^t \mathbb{E} \int_{k-1}^k F_r^1(\mathbf{Z}_r, \xi) d\xi dr.$$

Pero gracias a (5.22), tenemos que  $\mathbb{E} \int_{k-1}^k F_r^1(\mathbf{Z}_r, \xi) d\xi = m_1(f_r)$ , que está acotado uniformemente en el tiempo gracias al Lema 5.6.4 más abajo. Observando que hay  $(n-1)(N-n)$  términos en la doble sumatoria, obtenemos que:

$$\mathbf{E}_1^Z = \frac{2\gamma(n-1)(N-n)}{(N-1)(N-2)} \int_s^t m_1(f_r) dr \leq \frac{n}{N} C(t-s).$$

Con el mismo argumento se prueba que  $\mathbf{E}_2^Z$  y  $\mathbf{E}_3^Z$  satisfacen la misma cota. Por lo tanto,  $\mathbf{E}^Z \leq \frac{n}{N} C(t-s)$ , como se había afirmado.

- Para probar la cota para  $\mathbf{E}^V$ , podemos repetir el argumento anterior, donde ahora  $V^1$  es quien recibe los impulsos excitatorios en lugar de  $Z^1$ , y con el proceso de Poisson  $\tilde{\mathcal{N}}_{1j}$  en lugar de  $\mathcal{N}_{1j}$ . De esta manera obtenemos también la cota  $\mathbf{E}^V \leq \frac{n}{N} C(t-s)$ .

Estas cotas implican la desigualdad (5.30), lo cual concluye la prueba.  $\square$

**Lema 5.6.4.** *Notemos  $m_r(f_t) = \int z^r f_t(dz) = \mathbb{E} Z_t^r$  para  $f_t = \mathcal{L}(Z_t)$ . Fijamos un entero  $1 \leq r \leq 3$ . Si  $m_r(f_0) < \infty$ , entonces  $\sup_{t \geq 0} m_r(f_t) < \infty$ .*

*Demostración.* Observemos que  $\mathbb{E} \left( (Z_t + \rho)^r - Z_t^r \right) = \sum_{q=0}^{r-1} \binom{r}{q} \rho^{r-q} \mathbb{E} Z_t^q$ . La fórmula de Dynkin aplicada a (4.6) implica que

$$\frac{d}{dt} m_r(f_t) = -\mu m_r(f_t) - \gamma m_{r+1}(f_t) + \kappa \gamma m_1(f_t) \sum_{q=0}^{r-1} \binom{r}{q} \rho^{r-q} m_q(f_t). \quad (5.31)$$

Para  $r = 1$ , (5.31) resulta en  $\frac{d}{dt} m_1(f_t) = (\kappa \rho \gamma - \mu) m_1(f_t) - \gamma m_2(f_t)$ . Consecuentemente,  $m_2(f_t) \leq m_1(f_t)^2$  implica que

$$\frac{d}{dt} m_1(f_t) \leq (\kappa \rho \gamma - \mu) m_1(f_t) - \gamma m_1(f_t)^2.$$

Esta desigualdad muestra que  $\frac{d}{dt} m_1(f_t) \leq 0$  siempre que  $m_1(f_t) \geq \kappa \rho - \mu/\gamma$ . Por lo tanto,

$$c_1 := \sup_{t \geq 0} m_1(f_t) \leq \max\{m_1(f_0), \kappa \rho - \mu/\gamma\}.$$

Ahora, (5.31) para  $r = 2$  implica que

$$\begin{aligned} \frac{d}{dt} m_2(f_t) &= -2\mu m_2(f_t) - \gamma m_3(f_t) + \kappa \gamma m_1(f_t) (\rho^2 + 2\rho m_1(f_t)) \\ &\leq -2\mu m_2(f_t) + \kappa \gamma (c_1 + \rho)^3, \end{aligned}$$

que muestra que  $\frac{d}{dt} m_2(f_t) \leq 0$  siempre que  $m_2(t) \geq (2\mu)^{-1} \kappa \gamma (c_1 + \rho)^3$ . Entonces,

$$c_2 := \sup_{t \geq 0} m_2(f_t) \leq \max\{m_2(f_0), (2\mu)^{-1} \kappa \gamma (c_1 + \rho)^3\}.$$

Análogamente, (5.31) para  $r = 3$  implica que

$$\begin{aligned} \frac{d}{dt}m_3(f_t) &= -3\mu m_3(f_t) - \gamma m_4(f_t) + \kappa\gamma m_1(f_t)\left(\rho^3 + 3\rho^2 m_1(f_t) + 3\rho m_2(f_t)\right) \\ &\leq -3\mu m_3(f_t) + \kappa\gamma(c_1 \vee c_2 + \rho)^4. \end{aligned}$$

El mismo argumento utilizado para  $r = 2$  ahora muestra que

$$c_3 := \sup_{t \geq 0} m_3(f_t) \leq \max\{m_3(f_0), (3\mu)^{-1}\kappa\gamma(c_1 \vee c_2 + \rho)^4\}. \quad \square$$

*Prueba del Teorema 5.3.1.* Como  $\frac{1}{N} \sum_{i=1}^N \delta_{(Z_t^i, X_t^i)}$  constituye un acoplamiento entre  $\bar{\mathbf{Z}}_t$  y  $\bar{\mathbf{X}}_t$ , vemos que

$$W_1(f_t, \bar{\mathbf{X}}_t) \leq W_1(f_t, \bar{\mathbf{Z}}_t) + W_1(\bar{\mathbf{Z}}_t, \bar{\mathbf{X}}_t) \leq W_1(f_t, \bar{\mathbf{Z}}_t) + \frac{1}{N} \sum_{i=1}^N |Z_t^i - X_t^i|,$$

por lo que, por intercambiabilidad, vemos que

$$\mathbb{E} \left[ W_1(f_t, \bar{\mathbf{X}}_t) \right] \leq \mathbb{E} \left[ W_1(f_t, \bar{\mathbf{Z}}_t) \right] + h_t,$$

donde  $h_t := \mathbb{E} [|Z_t^1 - X_t^1|]$ . Por el Lema 5.6.2, basta con estimar  $h_t$ .

A partir de (5.20) y (5.23) vemos que

$$\partial_t h_t \leq -h_t - \gamma \mathbb{E} \left[ (Z_t^1 - X_t^1)^2 \right] + 2\gamma \mathbb{E} \int_1^N \left| F_t^1(\mathbf{Z}_t, \xi) - X_t^{[\xi]} \right| \frac{d\xi}{N-1}.$$

Simplemente descartamos el segundo término, mientras que en el tercer término sumamos y restamos  $Z_t^{[\xi]}$ . Usando (5.21) e intercambiabilidad de las coordenadas, obtenemos

$$\begin{aligned} \partial_t h_t &\leq -h_t + 2\gamma \mathbb{E} \left[ W_1(f_t, \bar{\mathbf{Z}}_t^1) \right] + 2\gamma \mathbb{E} \left[ \frac{1}{N-1} \sum_{i=2}^N |Z_t^i - X_t^i| \right] \\ &\leq (\theta_c - 1)h_t + \frac{C}{N^{1/3}}, \end{aligned}$$

gracias al Lema 5.6.2. Por lo tanto, usando el lema de Grönwall se obtiene el resultado.  $\square$

### 5.6.1. Persistencia (Teorema 5.4.1)

Nuevamente, recordamos la suposición  $\kappa = 2$ . También asumimos  $\rho = \mu = 1$ .

*Prueba del teorema 5.4.1.* Primero tratamos el caso  $\theta > 1$ , que implica  $\theta_c > 1$ . Llamamos  $\epsilon := \frac{1}{2} \inf_{t \geq 0} \mathbb{E}[Z_t]$ ; gracias al Teorema 5.2.2, sabemos que  $\epsilon > 0$ . Entonces, para todo  $t \geq 0$ , tenemos

$$2\epsilon \leq \mathbb{E}[Z_t] = W_1(f_t, \delta_0) \leq \mathbb{E}[W_1(f_t, \bar{\mathbf{X}}_t)] + \mathbb{E}[W_1(\bar{\mathbf{X}}_t, \delta_0)] \leq \frac{C e^{(\theta_c - 1)t}}{N^{1/3}} + \mathbb{E}[X_t^1],$$

donde en el último paso usamos el Teorema 5.3.1 e intercambiabilidad. Claramente,  $C e^{(\theta_c - 1)t} N^{-1/3} \leq \epsilon$  si y sólo si  $t \leq \tilde{c} + c \log N$  para  $c = 1/[3(\theta_c - 1)] > 0$  y  $\tilde{c} = [1/(\theta_c - 1)] \log(\epsilon/C) \in \mathbb{R}$ , de donde se deduce el resultado.

Ahora, trabajamos el caso  $\theta < 1$ . La convergencia en la métrica  $\omega(\cdot, \cdot)$  es exactamente (5.2), por lo tanto no hay nada que probar. En el caso  $2\gamma = \theta_c < 1$ , de (5.20) es fácil mostrar que

$$\frac{d}{dt} \mathbb{E}[X_t^1] = -\mathbb{E}[X_t^1] - \gamma \mathbb{E}[(X_t^1)^2] + \frac{2\gamma}{N-1} \sum_{i=2}^N \mathbb{E}[X_t^i] \leq -(1 - \theta_c) \mathbb{E}[X_t^1],$$

donde en el último paso, simplemente hemos descartado el término  $-\gamma \mathbb{E}[(X_t^1)^2] \leq 0$ , y hemos usado intercambiabilidad. El resultado entonces se sigue del lema de Grönwall.  $\square$



## Capítulo 6

# Simulaciones

En esta sección desarrollamos algunas simulaciones que reflejan los resultados obtenidos en la sección anterior, así como también aportan intuición para conjeturar otras propiedades del sistema.

En todas las simulaciones, salvo que se indique otra cosa, se realizaron experimentos para una población de  $N = 1000$  neuronas, y con parámetros de firing, decaimiento y magnitud de impulso  $\gamma = \mu = \rho = 1$ .

En el teorema 5.1.1 vimos que, independientemente del estado inicial o de los parámetros, el proceso  $\mathbf{X}_t$  converge casi seguramente a  $\mathbf{0}$ . Sin embargo, en este mismo resultado vemos cómo el *parámetro de reproducción*  $\theta$  (ver (5.1)) resulta crítico en la velocidad de convergencia al único estado absorbente (el trivial), dependiendo de si es mayor o menor a 1. Más aún, en los subsiguientes resultados este mismo parámetro cobra relevancia fundamental: en el teorema 5.2.2 establece una transición de fase que determina si el proceso límite  $(Z_t)_{t \geq 0}$  converge a cero o bien permanece (posiblemente bajo una ley estacionaria), mientras que el teorema 5.4.1, es definitorio en el fenómeno de persistencia de una neurona típica del sistema finito; ambos bajo la misma condición que en 5.1.1. Por lo tanto, el parámetro  $\theta$  tendrá un rol central en la construcción de las simulaciones.

### 6.1. El tiempo de extinción

Como dijimos en los párrafos previos, el teorema 5.1.1 indica que para una población finita de neuronas, el proceso indefectiblemente converge a  $\mathbf{0}$ . Por lo tanto, tiene sentido de hablar de un *tiempo de extinción* del sistema. Para definirlo de manera formal, observar que casi seguramente, existen finitos interacciones para el sistema (ver demostración del teorema 5.1.1), y por lo tanto podemos definir  $T_{\text{ext}}$  como el supremo (máximo) de los tiempos de interacción, es decir (con la notación de 4.1.3):

$$T_{\text{ext}} := \sup_{k \in \mathbb{N}_0} \{\hat{T}_k : \hat{T}_k < \infty\} \quad (6.1)$$

Notar que éste no es un tiempo de parada.

### 6.1.1. Evolución exponencial del tiempo de extinción:

En las siguientes simulaciones 6.1 se plasma la evolución del tiempo de extinción promedio en función de  $\theta$ , para  $\kappa = 2$  y  $\gamma = \gamma(\theta)$ :

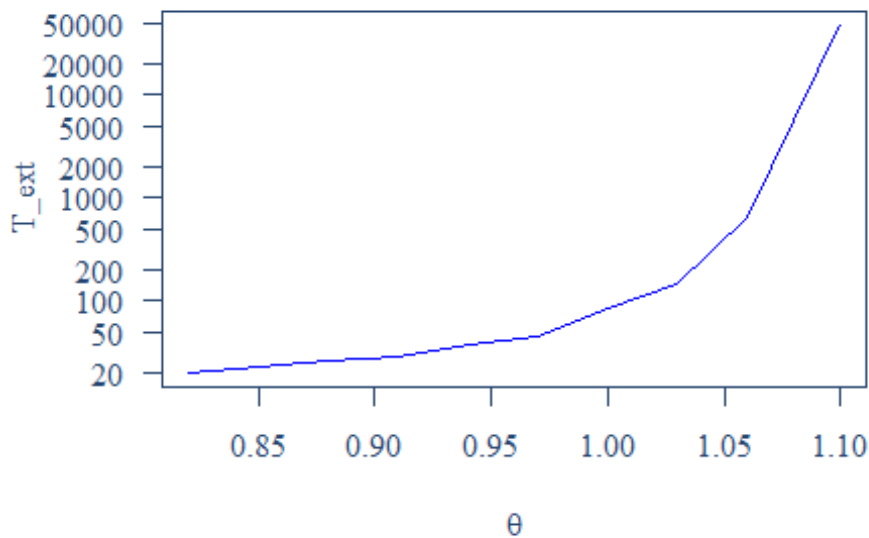


Figura 6.1: Tiempo de extinción promedio en función de  $\theta$

El crecimiento exponencial del tiempo de extinción para valores de  $\theta$  mayores a 1 es acorde a lo esperado a partir del teorema 5.2.2, correspondiente al proceso límite  $(Z_t)_{t \geq 0}$  para estos valores de  $\theta$ . Notar que el gráfico está presentado en escala logarítmica en el eje de ordenadas, por lo que se puede deducir que el tiempo de extinción explota rápidamente tras cruzar el umbral de criticalidad de  $\theta$ .

### 6.1.2. Distribución del tiempo de extinción:

Para distintos valores de  $\theta$  y  $\kappa = 2$  con  $N = 500$  neuronas, realizamos además simulaciones para analizar la distribución del tiempo de extinción, correspondientes a las figuras 6.2 a 6.8, donde presentamos histogramas de frecuencia del tiempo de extinción para los distintos valores de  $\theta$ .

Como puede apreciarse en los casos subcríticos (figuras 6.2 y 6.3, con  $\theta = 0,6$  y  $\theta = 0,8$  respectivamente), el tiempo de extinción se mantiene acotado. Luego, las figuras 6.4 a 6.6, correspondientes a los valores “críticos”  $\theta = 0,99$ ,  $\theta = 1,01$  y  $\theta = 1,03$ , ilustran cómo el tiempo de extinción va creciendo conforme la constante de reproducción aumenta y se vuelve supercrítico. Finalmente, las figuras 6.7 y 6.8 (correspondientes a los valores supercríticos  $\theta = 1,05$  y  $\theta = 1,1$  respectivamente), muestran valores del tiempo de extinción mucho mayor,



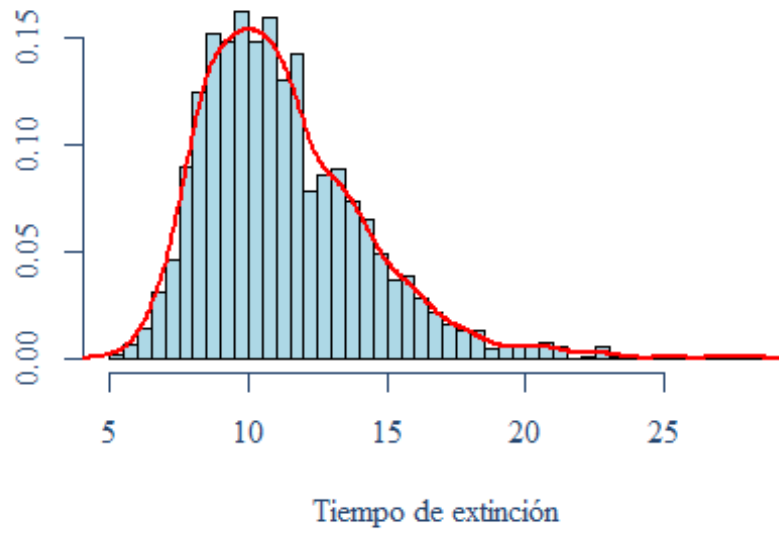


Figura 6.2: Distribución estimada del tiempo de extinción para  $\theta = 0,6$ . El tiempo promedio estimado es de 11,2

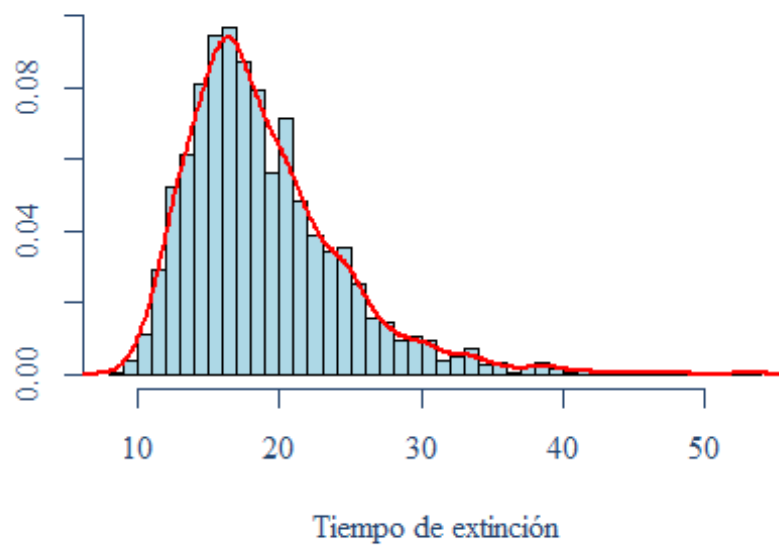


Figura 6.3: Distribución estimada del tiempo de extinción para  $\theta = 0,8$ . El tiempo promedio estimado es de 18,9

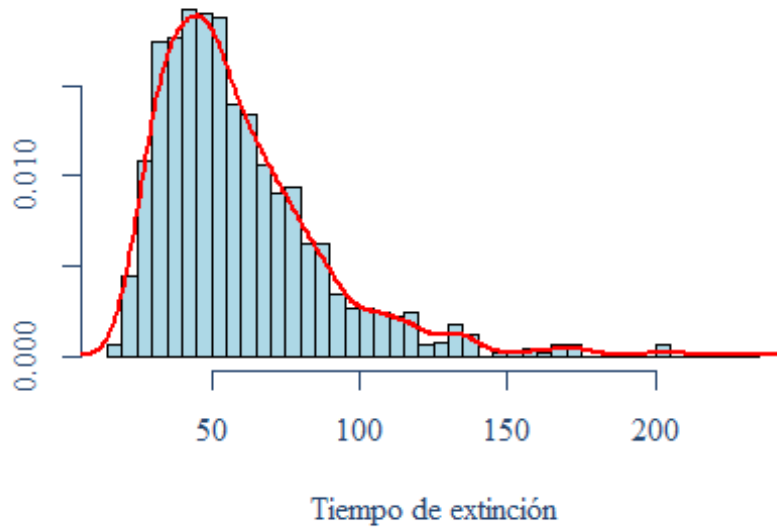


Figura 6.4: Distribución estimada del tiempo de extinción para  $\theta = 0,99$ . El tiempo promedio estimado es de 59,7

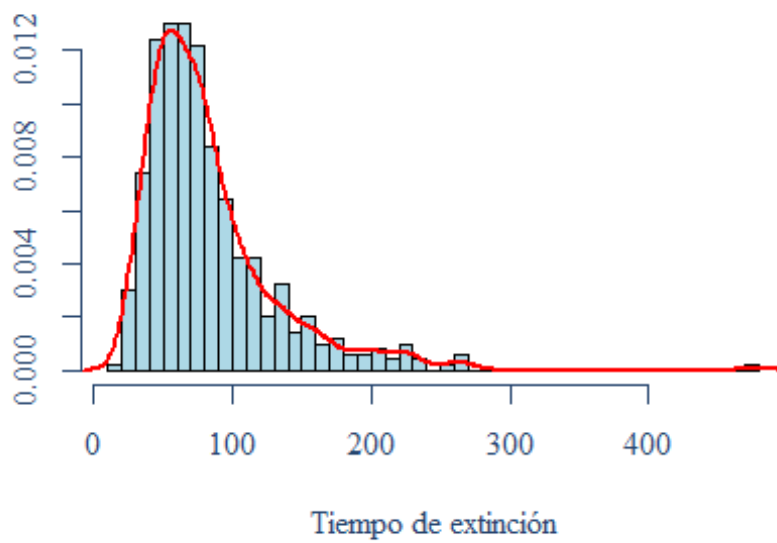


Figura 6.5: Distribución estimada del tiempo de extinción para  $\theta = 1,01$ . El tiempo promedio estimado es de 82,8

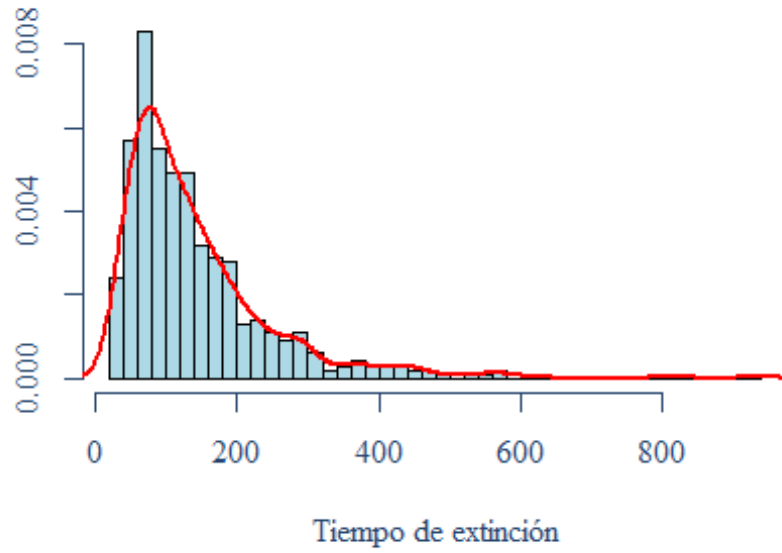


Figura 6.6: Distribución estimada del tiempo de extinción para  $\theta = 1,03$ . El tiempo promedio estimado es de 142,2

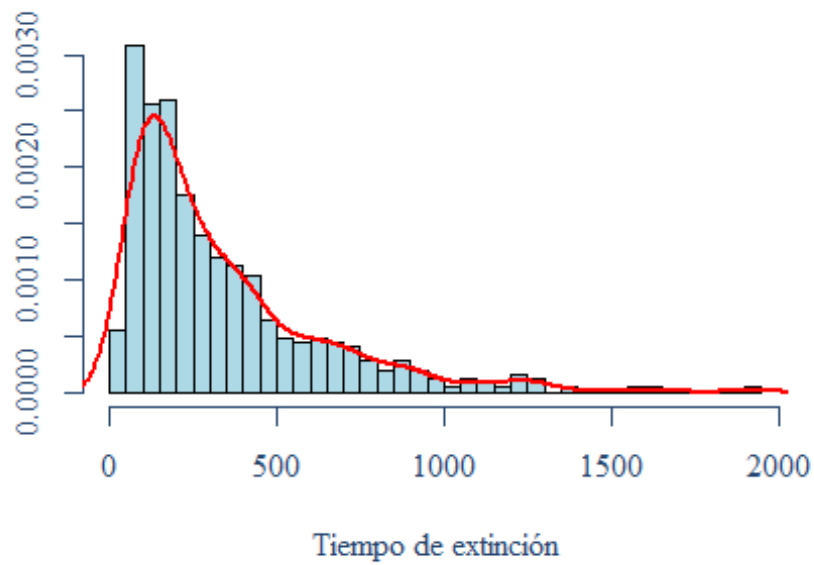


Figura 6.7: Distribución estimada del tiempo de extinción para  $\theta = 1,05$ . El tiempo promedio estimado es de 330,2

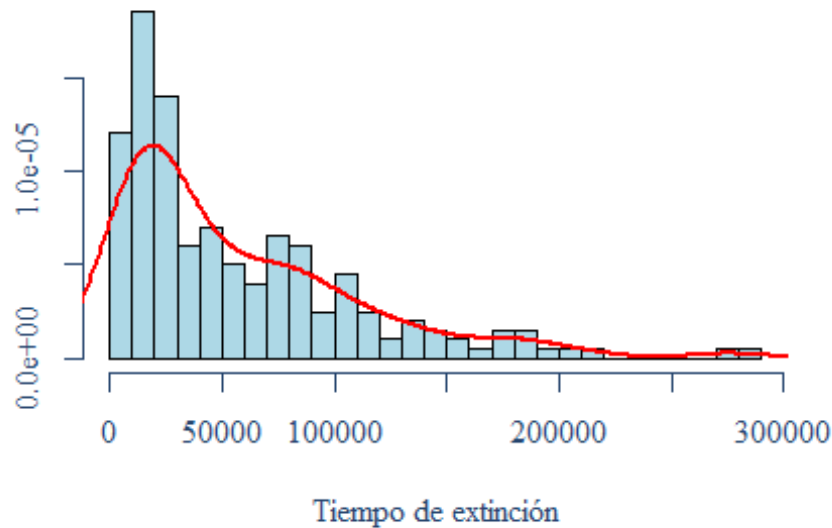


Figura 6.8: Distribución estimada del tiempo de extinción para  $\theta = 1, 1$ . El tiempo promedio estimado es de 56896,6

que además sugieren un fenómeno de persistencia acorde al teorema 5.4.1, y más aún, sugieren una escala superior a la de orden  $\log(N)$  obtenida en dicho teorema.

## 6.2. Comportamiento de la media empírica:

Aquí simulamos la trayectoria de la media empírica. Se puede observar (figura 6.9) que cuando  $\theta < 1$ , la actividad se extingue, mientras que en las figuras 6.10 a 6.12, donde  $\theta > 1$ , la actividad neuronal persiste, de acuerdo a lo esperado según los teoremas 5.2.2 y 5.4.1. Pero más aún, de estos últimos gráficos también se puede intuir un fenómeno de metaestabilidad en torno de cierto valor para cada  $\theta$  fijo, lo que podría sugerir la existencia de una distribución cuasiestacionaria, que capturaría este fenómeno de estabilidad que se observa, teniendo en cuenta que eventualmente la actividad se extingue. Al respecto, cabe mencionar que estamos en vías de probar teóricamente la existencia de cuasiestacionariedad, aunque aún no hemos llegado a dicho resultado.

## 6.3. Actividad de una partícula marcada:

Finalmente, En las figuras 6.13 a 6.15 simulamos el valor medio una partícula marcada en una población de neuronas con  $\theta > 1$ <sup>1</sup> a tiempo  $\log(N)$ . Observamos que los resultados

---

<sup>1</sup>La elección de los  $\theta$  en esta sección está realizada para coincidir con los valores de  $\theta$  correspondientes a la simulación de la media empírica de la sección anterior.

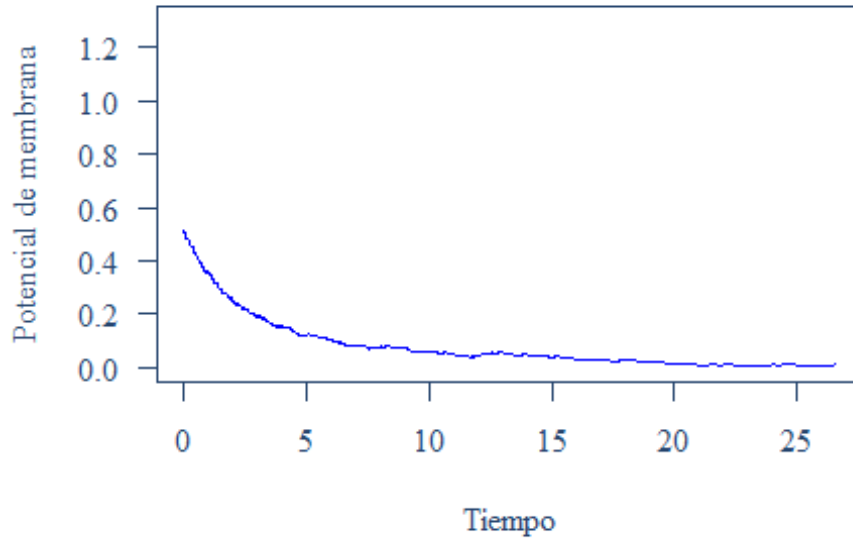


Figura 6.9: Evolución temporal de la media empírica para  $N = 1000$ ,  $\kappa = 2$  y  $\theta \sim 0,9$  (estableciendo  $\gamma = 0,6$ ).

muestran actividad neuronal en el tiempo observado, con un potencial que coincide con el valor en el que se estabiliza la media empírica para ese  $\theta$ . Esto reafirma nuevamente el teorema 5.4.1.

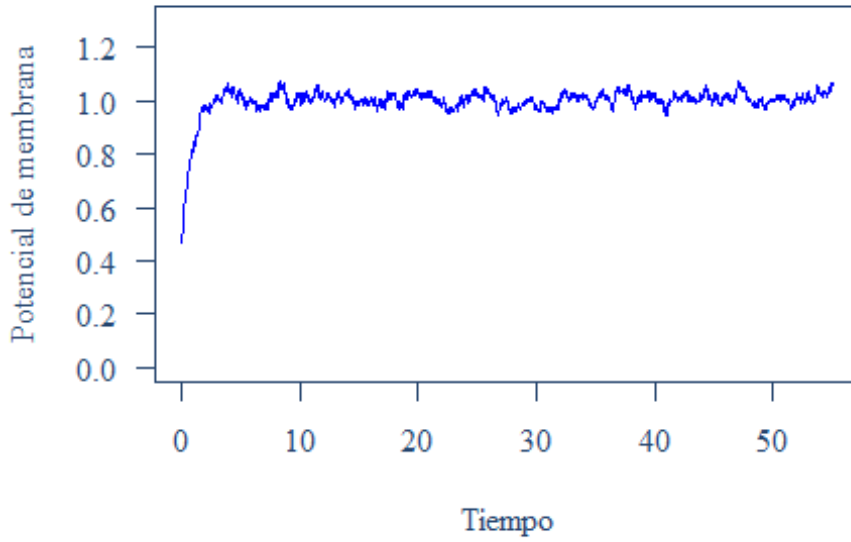


Figura 6.10: Evolución temporal de la media empírica para  $N = 1000$ ,  $\kappa = 3$  y  $\theta \sim 1,89$ . La media empírica se estabiliza en torno a 1,02

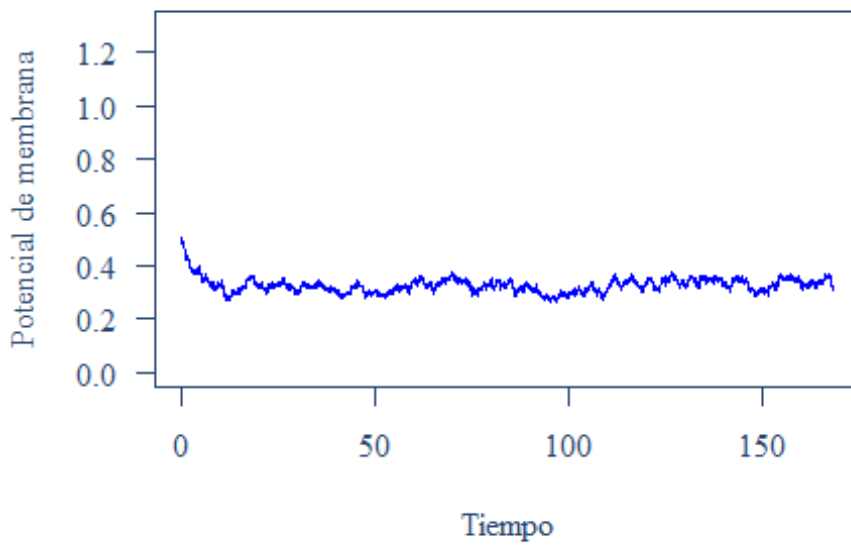


Figura 6.11: Evolución temporal de la media empírica para  $N = 1000$ ,  $\kappa = 2$  y  $\theta \sim 1,26$ . La media empírica se estabiliza en torno a 0,33

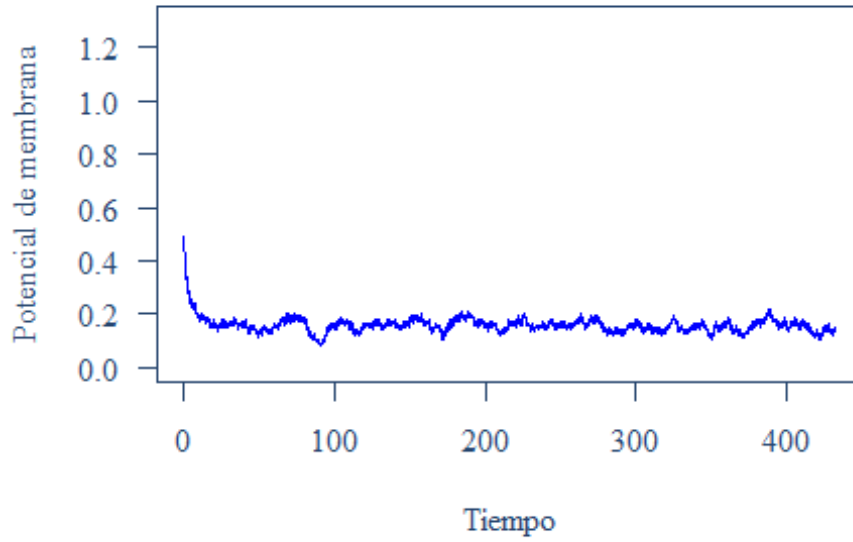


Figura 6.12: Evolución temporal de la media empírica para  $N = 1000, \kappa = 2$  y  $\theta \sim 1,11$  (estableciendo  $\gamma = 0,8$ ). La media empírica se estabiliza en torno a 0,15

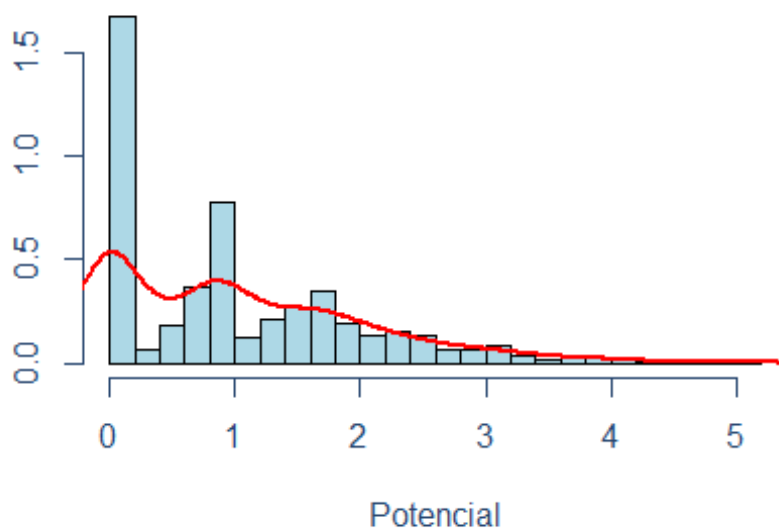


Figura 6.13: Histograma del potencial de una partícula “típica” a tiempo  $\log(N)$  para  $\theta = 1,89$ . El valor medio es  $\sim 1,03$

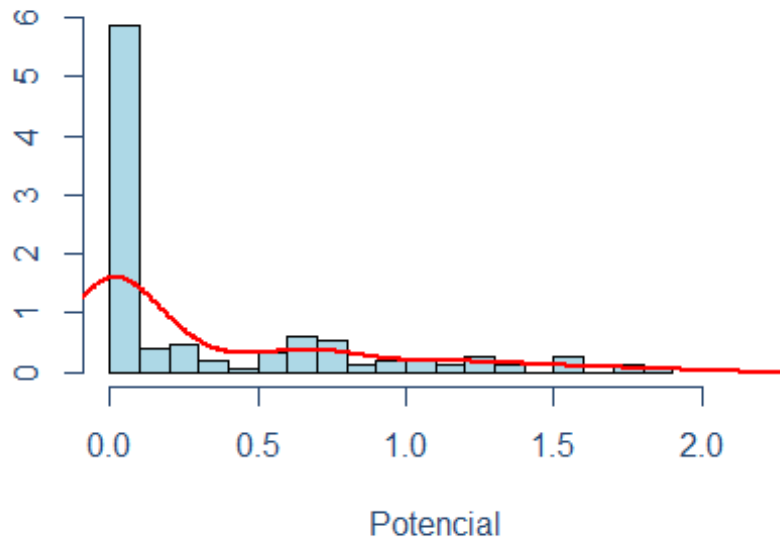


Figura 6.14: Histograma del potencial de una partícula “típica” a tiempo  $\log(N)$  para  $\theta = 1,26$ . El valor medio es  $\sim 0,34$

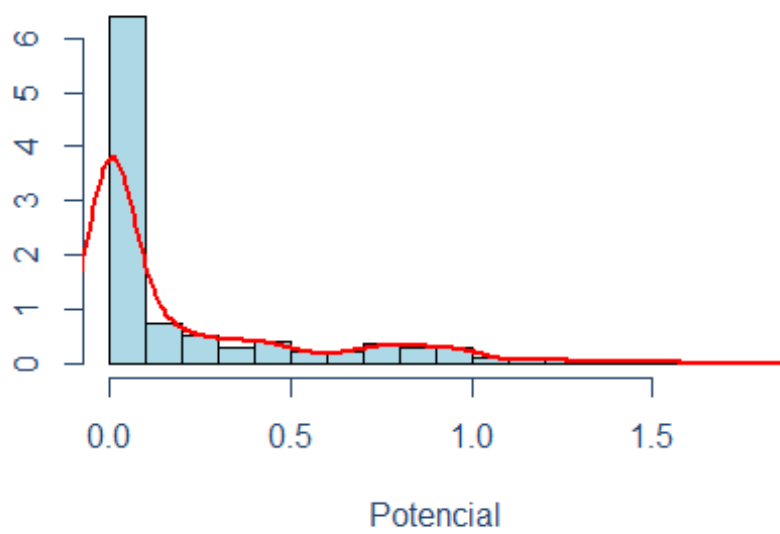


Figura 6.15: Histograma del potencial de una partícula “típica” a tiempo  $\log(N)$  para  $\theta \sim 1,11$  (estableciendo  $\gamma = 0,8$ ). El valor medio es  $\sim 0,18$



# Capítulo 7

## Conclusiones

En este trabajo hemos presentado un modelo estocástico para una red neuronal que tiene características distintivas y que resulta relevante tanto para la investigación biológica como para la investigación matemática, como argumentamos en 1.4. Demostramos la convergencia del sistema finito a un proceso no lineal, para el que mostramos una transición de fase que luego se traduce en un fenómeno de persistencia para la población finita. Varias líneas de investigación pueden derivarse de estos resultados. En los siguientes párrafos, enumeramos algunas de ellas.

### **Algunas generalizaciones y resultados naturales:**

En primer lugar, como fue observado 5.2.2, creemos fuertemente que el proceso no lineal  $Z_t$  tiene un único límite no trivial  $Z_\infty$  en el caso  $\theta > 1$ . Esta conjetura está apoyada por las simulaciones numéricas de la red finita para  $N$  relativamente grande, como puede apreciarse en el capítulo anterior 6: la medida empírica del proceso parecería estabilizarse cuando el proceso corre por un tiempo largo (incluso considerando el hecho de que el Teorema 5.1.1 implica que  $\mathbf{X}_t$  eventualmente decaerá, ver Observación 5.4.1). Es posible que esta distribución no trivial pueda ser dilucidada mediante un procedimiento de punto fijo. Sería muy deseable tener una prueba rigurosa de este fenómeno y esto resulta una línea de trabajo interesante para trabajo futuro.

En relación a este fenómeno, parecería natural investigar la emergencia de *distribuciones cuasi-estacionarias*, que capturan precisamente fenómenos de metaestabilidad en procesos que se extinguen casi seguramente. Hemos comenzado a trabajar en esta línea, con la esperanza de obtener resultados relevantes al respecto.

Tal como se observó (5.4.1), la estimación de orden  $\log(N)$  en el teorema de persistencia 5.4.1 no es óptima. Por lo tanto, también resultaría interesante explorar métodos alternativos para obtener una estimación más fina de la cota para el resultado de persistencia, ya que tanto la intuición como las simulaciones indicarían que una cota inferior óptima para el tiempo de actividad del sistema es de orden mayor al obtenido. Sin embargo, no parece que tal resultado pueda obtenerse de manera directa a partir de los resultados aquí expuestos, y

por lo tanto, deben considerarse estrategias distintas a las presentadas.

Es pertinente mencionar también que una extensión natural sería considerar como magnitudes aleatorias tanto el alcance local  $\kappa$  así como la magnitud del impulso excitatorio  $\rho$ . Creemos que un modelo con estas características debería derivar en esencialmente los mismos resultados (reemplazando  $\kappa$  y  $\rho$  con  $\mathbb{E}[\kappa]$  y  $\mathbb{E}[\rho]$ , respectivamente) para  $\kappa$  y  $\rho$  en condiciones suficientemente “buenas”, ver Observación 5.3.1.

### Un modelo generalizado sobre un grafo aleatorio:

Finalmente, una generalización importante sería la introducción de una estructura espacial en el modelo, lo cual se traduce esencialmente en establecer una estructura *dinámica* de grafo aleatorio subyacente, representando la interconexión entre las neuronas de la población.

El carácter dinámico de la estructura de grafo representaría así la escala de tiempo a la que la población neuronal ve modificadas sus interconexiones. Éste es un fenómeno biológico real, pero que sucede a escala de tiempos significativamente menores que la escala de tiempo correspondiente a las interacciones aquí estudiadas.

Ésta es por lo tanto una generalización que añadiría verosimilitud al modelo acá presentado (que representa una primera aproximación a este fenómeno), aunque resulta bastante más compleja de manipular matemáticamente. En este contexto, en el otro extremo del espectro se encontraría la consideración de un grafo aleatorio fijado, es decir, uno donde las conexiones sinápticas no se ven modificadas una vez establecidas. Este caso sería también de gran interés como una buena aproximación al fenómeno real, y en particular, sería muy interesante obtener resultados análogos a los de este trabajo y observar si los parámetros críticos cambian de manera significativa o qué paralelismos podrían trazarse. Podría ser muy informativo para poder establecer conjeturas sobre lo que se espera del caso “intermedio”, es decir, con un grafo aleatorio dinámico pero “lento”.

En este sentido, las técnicas y los resultados en los trabajos [LRW19] y [ORS20] podrían ser un buen punto de partida que permita obtener un sistema de ecuaciones diferenciales asociadas a un árbol de Galton-Watson, en lugar de un resultado de campo medio como en este trabajo. Para ello, sería necesario encuadrar el modelo dentro del contexto de *convergencia débil local* para grafos aleatorios *sparse* (es decir, de grado finito, al menos en esperanza), que es una modalidad de convergencia natural para este tipo de grafos, ver [LRW19, Sección 2.2]. La estructura espacial en este marco cobra suma importancia a la hora de obtener resultados asintóticos, que ya no tienen que ver con el fenómeno de propagación del caos.

En este caso, también sería de gran relevancia obtener resultados de persistencia y poder compararlos con los obtenidos en esta tesis.

## Apéndice A

# Construcción de los acoplamientos para $\kappa$ general:

Fijamos nuevamente un entero  $2 \leq n \leq N$ . El procedimiento para la construcción de  $\mathbf{Z}_t$  será idéntico, mientras que la construcción de  $\mathbf{V}_t$  será ligeramente distinta pero análoga a la desarrollada en 5.6.3. Aquí  $K$  siempre notará un subconjunto de  $\{1, \dots, N\}$  de  $\kappa$  elementos (al conjunto de todos los subconjuntos  $K$  lo notamos  $\mathbb{K}$ ) y dado  $i \in K$  notamos  $K^i := K \setminus i$ . Notamos también  $m(K^i)$  al mínimo del conjunto  $K^i$ . De manera análoga a (5.18) y (5.19), dado  $i \in \{1, \dots, N\}$ , definimos

$$\mathcal{N}^i(dt, du) = \sum_{K:K \ni i} \mathcal{N}_K(dt, du, (i-1, i]) \quad (\text{A.1})$$

$$\mathcal{M}^i(dt, du, d\xi) = \sum_{K:K \ni i} \mathcal{N}_K(dt, du, d\xi). \quad (\text{A.2})$$

Observar que en este caso obtenemos la misma expresión (5.20) para la ecuación (4.1) que define a  $X_t^i$ . Por lo tanto, en este caso la definición de nuestro acoplamiento  $\mathbf{Z}_t = (Z_t^1, \dots, Z_t^N)$  se define de manera idéntica a (5.23) como la solución para

$$\begin{aligned} dZ_t^i &= -Z_t^i dt - Z_t^i \int_0^\infty \mathbf{1}(u \leq \gamma Z_t^i) \mathcal{N}^i(dt, du) \\ &\quad + \int_0^\infty \int_{(0, N]} \mathbf{1}(u \leq \gamma F_t^i(\mathbf{Z}_t, \xi)) \mathcal{M}^i(dt, du, d\xi), \end{aligned} \quad (\text{A.3})$$

con  $Z_0^i = X_0^i$ .

Ahora especificamos la construcción de  $\mathbf{V}_t$ . Con exactamente los mismos argumentos que en 5.6.3, fijada  $i \in \{1, \dots, N\}$  construimos  $V_t^i$  mediante una SDE análoga, dirigida por el proceso de poisson  $\mathcal{N}^i$  de (A.1) y por un proceso  $\mathcal{R}^i$  cuidadosamente elegido a partir de  $\mathcal{M}^i$  de (A.2). Al igual que en (5.26), y notando  $\mathcal{N}_K = \mathcal{N}_K(dt, du, d\xi)$  escribimos

$$\mathcal{M}^i(dt, du, d\xi) = \sum_{\substack{K \ni i \\ m(K^i) > n}} \mathbf{1}(\xi > n) \mathcal{N}_K + \sum_{K \ni i} \mathbf{1}\left((\xi \leq n) \vee (m(K^i) \leq n)\right) \mathcal{N}_K \quad (\text{A.4})$$

Observar que la condición  $m(K^i) \leq n$  representa que alguna neurona  $j \in K^i$  (o sea, **aparte** de  $i$ ) se encuentra entre las primeras  $n$  neuronas, evento que implica los saltos simultáneos que

queremos desacoplar. Para lograr esto vamos a definir una colección de copias independientes que reemplacen a los procesos de Poisson del segundo término, de manera análoga a los procesos  $(\tilde{\mathcal{N}}_{jk}(dt, du, d\xi))_{j \neq k}$  de (5.27). Sea entonces  $K_0 \subseteq \{1, \dots, N\}$  un subconjunto fijo de  $\kappa$  elementos, y definimos una colección  $(\tilde{\mathcal{N}}_{(j, K^j)})_{K \in \mathbb{K}, j \in K}$  independiente de copias de  $\mathcal{N}_{K_0}$ . Consecuentemente, definimos  $\mathcal{R}^i = \mathcal{R}^i(dt, du, d\xi)$  como

$$\mathcal{R}^i = \sum_{\substack{K \ni i \\ m(K^i) > n}} \mathbf{1}(\xi > n) \mathcal{N}_K + \sum_{K \ni i} \mathbf{1}\left((\xi \leq n) \vee (m(K^i) \leq n)\right) \tilde{\mathcal{N}}_{(i, K^i)} \quad (\text{A.5})$$

Se puede ver con idénticos razonamientos a los expuestos en 5.6.3 por qué esta definición implica que  $\mathcal{R}^i$  es un proceso de Poisson con la misma distribución que  $\mathcal{M}^i$ , y que  $\mathcal{N}^1, \dots, \mathcal{N}^n$  y  $\mathcal{R}^1, \dots, \mathcal{R}^n$  son independientes, ya que ningún “bloque” perteneciente a  $(\mathcal{N}_K)_{K \in \mathbb{K}}$  o a  $(\tilde{\mathcal{N}}_{(j, K^j)})_{K \in \mathbb{K}, j \in K}$  se repite y por lo tanto no hay saltos simultáneos. Por ejemplo, si  $n = 5$  y  $\kappa = 3$ , sea el evento donde la neurona 1 emite un firing y excita a las neuronas 2, 3 y 4. Entonces definiendo  $K := \{1, 2, 3, 4\}$ , el evento implica para  $\mathbf{Z}_t$  un átomo en  $\mathcal{N}_{K^1}(dt, du, (0, 1])$ , presente en  $\mathcal{M}^2, \mathcal{M}^3, \mathcal{M}^4$  y  $\mathcal{N}^1$ ; mientras que para el proceso  $\mathbf{V}_t$  dicho evento implica átomos en los procesos  $\tilde{\mathcal{N}}_{(2, K^2)}(dt, du, (0, 1])$ ,  $\tilde{\mathcal{N}}_{(3, K^3)}(dt, du, (0, 1])$ ,  $\tilde{\mathcal{N}}_{(4, K^4)}(dt, du, (0, 1])$  y  $\mathcal{N}_{K^1}(dt, du, (0, 1])$ , presentes en  $\mathcal{R}^2, \mathcal{R}^3, \mathcal{R}^4$  y  $\mathcal{N}^1$  respectivamente.

Finalmente  $(V_t^i)_{t \geq 0}$  queda definida por  $V_0^i = \mathbf{Z}_0^i$  y

$$\begin{aligned} dV_t^i &= -V_t^i dt - V_t^i \int_{\mathbb{R}_+} \mathbf{1}(u \leq \gamma V_t^i) \mathcal{N}^i(dt, du) \\ &\quad + \int_{\mathbb{R}_+} \int_{(0, N]} \mathbf{1}(u \leq \gamma F_t^i(\mathbf{Z}_t, \xi)) \mathcal{R}^i(dt, du, d\xi) \end{aligned} \quad (\text{A.6})$$

que es la misma definición que en (5.28), con los procesos de Poisson respectivos.

# Bibliografía

- [ABG<sup>+</sup>14] R. Azais, 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.
- [ACJL23] Maximiliano Altamirano, Roberto Cortez, Matthieu Jonckheere, and Lasse Leskelä. Persistence in a large network of sparsely interacting neurons. *Journal of Mathematical Biology*, 86(1):16, 2023.
- [Adr14] E. Adrian. The all-or-none principle in nerve. *The Journal of physiology*, 47(6):460, 1914.
- [AFG11] Amine Asselah, Pablo A Ferrari, and Pablo Groisman. Quasistationary distributions and fleming-viot processes in finite spaces. *Journal of Applied Probability*, 48(2):322–332, 2011.
- [BFFT12] Javier Baladron, Diego Fasoli, Olivier Faugeras, and Jonathan 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.
- [BG90] Carol Bezuidenhout and Geoffrey Grimmett. The critical contact process dies out. *The Annals of Probability*, 18(4):1462–1482, 1990.
- [BGSR17] Thierry Bodineau, Isabelle Gallagher, and Laure Saint-Raymond. From hard sphere dynamics to the stokes–fourier equations: An analysis of the boltzmann–grad limit. *Annals of PDE*, 3(1):1–118, 2017.
- [Bic02] Klaus Bichteler. *Stochastic integration with jumps*. Number 89. Cambridge University Press, 2002.
- [BKK<sup>+</sup>92] 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.
- [Bog18] Vladimir Igorevich Bogachev. *Weak convergence of measures*. American Mathematical Society Providence, 2018.
- [BRC<sup>+</sup>19] Michael Beyeler, Emily L Rounds, Kristofor D Carlson, Nikil Dutt, and Jeffrey L Krichmar. Neural correlates of sparse coding and dimensionality reduction. *PLoS computational biology*, 15(6):e1006908, 2019.

- [Bré81] Pierre Brémaud. *Point processes and queues: martingale dynamics*, volume 50. Springer, 1981.
- [BRTV98] S. Benachour, B. Roynette, D. Talay, and P. Vallois. Nonlinear self-stabilizing processes – i existence, invariant probability, propagation of chaos. *Stochastic Processes and their Applications*, 75(2):173 – 201, 1998.
- [BRV01] A. Barabási, E. Ravasz, and T. Vicsek. Deterministic scale-free networks. *Physica A: Statistical Mechanics and its Applications*, 299(3):559–564, 2001.
- [Bur06a] Anthony N Burkitt. A review of the integrate-and-fire neuron model: I. homogeneous synaptic input. *Biological cybernetics*, 95(1):1–19, 2006.
- [Bur06b] Anthony N Burkitt. A review of the integrate-and-fire neuron model: II. inhomogeneous synaptic input and network properties. *Biological cybernetics*, 95(2):97–112, 2006.
- [CCP11] María J Cáceres, José A Carrillo, and Benoît Perthame. Analysis of nonlinear noisy integrate & fire neuron models: blow-up and steady states. *The Journal of Mathematical Neuroscience*, 1(1):7, 2011.
- [CD09] Shirshendu Chatterjee and Rick Durrett. Contact processes on random graphs with power law degree distributions have critical value 0. *The Annals of Probability*, 37(6):2332–2356, 2009.
- [CD21] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: a review of models, methods and applications. *arXiv preprint arXiv:2106.14812*, 2021.
- [CDP20] J. A. Carrillo, M. G. Delgadino, and G. A. Pavliotis. A  $\lambda$ -convexity based proof for the propagation of chaos for weakly interacting stochastic particles. *Journal of Functional Analysis*, 279(10):108734, 2020.
- [CF16] Roberto Cortez and Joaquin Fontbona. Quantitative propagation of chaos for generalized Kac particle systems. *The Annals of Applied Probability*, 26(2):892 – 916, 2016.
- [Chi01] E.J. Chichilnisky. A simple white noise analysis of neuronal light responses. *Network: Computation in Neural Systems*, 12(2):199–213, 2001.
- [Dav84] Mark H. A. Davis. Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society B*, 46(3):353–388, 1984.
- [Dav18] Mark H. A. Davis. *Markov models and optimization*. Routledge, 2018.
- [DIR<sup>+</sup>15] François Delarue, James Inglis, Sylvain Rubenthaler, Etienne Tanré, et al. Global solvability of a networked integrate-and-fire model of mckean–vlasov type. *The Annals of Applied Probability*, 25(4):2096–2133, 2015.
- [DMGLP15] A. De Masi, A. Galves, E. Löcherbach, and E. Presutti. Hydrodynamic limit for interacting neurons. *Journal of Statistical Physics*, 158(4):866–902, 2015.

- 
- [Dud18] Richard M Dudley. *Real analysis and probability*. CRC Press, 2018.
- [dVCD13] Oswaldo Luiz do Valle Costa and François Dufour. *Continuous average control of piecewise deterministic Markov processes*. Springer, 2013.
- [FG00] Pablo A Ferrari and Antonio Galves. Construction of stochastic processes, coupling and regeneration. *XIII Escuela Venezolana de Matemática*, 2000.
- [FG15] Nicolas Fournier and Arnaud Guillin. On the rate of convergence in wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3-4):707–738, 2015.
- [FGM08] J Fontbona, H Guérin, and S Méléard. Measurability of optimal transportation and convergence rate for landau type interacting particle systems, to appear in *probab. Theory Related Fields*, 2008.
- [FM01] Nicolas Fournier and Sylvie Méléard. A markov process associated with a boltzmann equation without cutoff and for non-maxwell molecules. *Journal of Statistical Physics*, 104(1):359–385, 2001.
- [FM07] Pablo Ferrari and Nevena Maric. Quasi stationary distributions and fleming- viot processes in countable spaces. *Electronic Journal of Probability*, 12:684–702, 2007.
- [FM16] Nicolas Fournier and Stéphane Mischler. Rate of convergence of the nanbu particle system for hard potentials and maxwell molecules. *The Annals of Probability*, 44(1):589–627, 2016.
- [Gär88] Jürgen Gärtner. On the mckean-vlasov limit for interacting diffusions. *Mathematische Nachrichten*, 137(1):197–248, 1988.
- [GM97] Carl Graham and Sylvie Méléard. Stochastic particle approximations for generalized boltzmann models and convergence estimates. *The Annals of probability*, 25(1):115–132, 1997.
- [Grü71] F Alberto Grünbaum. Propagation of chaos for the boltzmann equation. *Archive for Rational Mechanics and Analysis*, 42(5):323–345, 1971.
- [GWL08] Evan M Galloway, Newton H Woo, and Bai Lu. Persistent neural activity in the prefrontal cortex: a mechanism by which bdnf regulates working memory? *Progress in brain research*, 169:251–266, 2008.
- [HH52] Alan L Hodgkin and Andrew F Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500, 1952.
- [HL78] Richard Holley and Thomas M Liggett. The survival of contact processes. *The Annals of Probability*, 6(2):198–206, 1978.
- [HM14] Maxime Hauray and Stéphane Mischler. On kac’s chaos and related problems. *Journal of Functional Analysis*, 266(10):6055–6157, 2014.

- [JW17] Pierre-Emmanuel Jabin and Zhenfu Wang. Mean field limit for stochastic particle systems. In *Active Particles, Volume 1*, pages 379–402. Springer, 2017.
- [Kac56] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III*, pages 171–197, Berkeley and Los Angeles, 1956. University of California Press.
- [Kad09] L. Kadanoff. More is the same; phase transitions and mean field theories. *Journal of Statistical Physics*, 137(5-6):777–797, 2009.
- [Kan39] Leonid V Kantorovich. The mathematical method of production planning and organization. *Management Science*, 6(4):363–422, 1939.
- [Kin92] John Frank Charles Kingman. *Poisson processes*, volume 3. Clarendon Press, 1992.
- [KPW13] M. Kochmański, T. Paszkiewicz, and S. Wolski. Curie–weiss magnet—a simple model of phase transition. *European Journal of Physics*, 34(6):1555, 2013.
- [KS12] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 2012.
- [Lap07] Louis Lapicque. Recherches quantitatives sur l’excitation électrique des nerfs traitée comme une polarisation. *Journal de Physiologie et de Pathologie Generalej*, 9:620–635, 1907.
- [LBMM07] 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.
- [Lin92] T Lindvall. Lectures on the coupling method (mineola, ny, 1992).
- [LRW19] Daniel Lacker, Kavita Ramanan, and Ruoyu Wu. Local weak convergence for sparse networks of interacting processes. *arXiv preprint arXiv:1904.02585*, 2019.
- [Mal01] F. Malrieu. Logarithmic sobolev inequalities for some nonlinear pde’s. *Stochastic Processes and their Applications*, 95(1):109 – 132, 2001.
- [McK69] HP McKean. Propagation of chaos for a class of nonlinear parabolic equations, lecture in differential equations, vol. 2, aziz, ak, 1969.
- [Mél96] Sylvie Méléard. Asymptotic behaviour of some interacting particle systems; mckean-vlasov and boltzmann models. *Probabilistic models for nonlinear partial differential equations*, pages 42–95, 1996.
- [Mie06] Grégory Miermont. Advanced probability. *University of Cambridge, Cambridge*, 2006.
- [MJ67] HP McKean Jr. An exponential formula for solving boltzmann’s equation for a maxwellian gas. *Journal of Combinatorial Theory*, 2(3):358–382, 1967.



- 
- [MM13] Stéphane Mischler and Clément Mouhot. Kac’s program in kinetic theory. *Invent. Math.*, 193(1):1–147, 2013.
- [MMW15] Stéphane Mischler, Clément Mouhot, and Bernt Wennberg. A new approach to quantitative propagation of chaos for drift, diffusion and jump processes. *Probability Theory and Related Fields*, 161(1):1–59, 2015.
- [MT04] Guy Major and David Tank. Persistent neural activity: prevalence and mechanisms. *Current opinion in neurobiology*, 14(6):675–684, 2004.
- [Oel84] Karl Oelschläger. A martingale approach to the law of large numbers for weakly interacting stochastic processes. *The Annals of Probability*, pages 458–479, 1984.
- [Oks13] Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.
- [ORS20] Roberto I Oliveira, Guilherme H Reis, and Lucas M Stolerman. Interacting diffusions on sparse graphs: hydrodynamics from local weak limits. *Electronic Journal of Probability*, 25:1–35, 2020.
- [PPSP08] Dorte P Pelvig, Henning Pakkenberg, Anette Kirstine Stark, and Bente Pakkenberg. Neocortical glial cell numbers in human brains. *Neurobiology of aging*, 29(11):1754–1762, 2008.
- [PPU<sup>+</sup>05] Jonathan W Pillow, Liam Paninski, Valerie J Uzzell, Eero P Simoncelli, and EJ Chichilnisky. Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. *Journal of Neuroscience*, 25(47):11003–11013, 2005.
- [PSP<sup>+</sup>08] Jonathan W Pillow, Jonathon Shlens, Liam Paninski, Alexander Sher, Alan M Litke, EJ Chichilnisky, and Eero P Simoncelli. Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, 454(7207):995–999, 2008.
- [RD10] Edmund T Rolls and Gustavo Deco. *The noisy brain: stochastic dynamics as a principle of brain function*, volume 34. Oxford university press Oxford, 2010.
- [RT16] Philippe Robert and Jonathan Touboul. On the dynamics of random neuronal networks. *Journal of Statistical Physics*, 165(3):545–584, 2016.
- [RTB<sup>+</sup>18] Bastian Rieck, Matteo Togninalli, Christian Bock, Michael Moor, Max Horn, Thomas Gumbsch, and Karsten Borgwardt. Neural persistence: A complexity measure for deep neural networks using algebraic topology. *arXiv preprint arXiv:1812.09764*, 2018.
- [RTK15] 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.
- [SM12] Donald L Snyder and Michael I Miller. *Random point processes in time and space*. Springer Science & Business Media, 2012.

- [Spo12] Herbert Spohn. *Large scale dynamics of interacting particles*. Springer Science & Business Media, 2012.
- [Szn84] Alain-Sol Sznitman. Équations de type de boltzmann, spatialement homogenes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 66(4):559–592, 1984.
- [Szn86] Alain-Sol Sznitman. A propagation of chaos result for burgers’ equation. *Probability theory and related fields*, 71(4):581–613, 1986.
- [Szn91] Alain-Sol Sznitman. Topics in propagation of chaos. In *Ecole d’été de probabilités de Saint-Flour XIX—1989*, pages 165–251. Springer, 1991.
- [T<sup>+</sup>14] Jonathan Touboul et al. Propagation of chaos in neural fields. *The Annals of Applied Probability*, 24(3):1298–1328, 2014.
- [Tan83] Hiroshi Tanaka. Some probabilistic problems in the spatially homogeneous boltzmann equation. In *Theory and application of random fields*, pages 258–267. Springer, 1983.
- [Tou12] Jonathan 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.
- [Tou14] Jonathan Touboul. Spatially extended networks with singular multi-scale connectivity patterns. *Journal of Statistical Physics*, 156(3):546–573, 2014.
- [Tug14] Julian Tugaut. Phase transitions of mckean–vlasov processes in double-wells landscape. *Stochastics*, 86(2):257–284, 2014.
- [Vas69] Leonid Nisonovich Vaserstein. Markov processes over denumerable products of spaces, describing large systems of automata. *Problemy Peredachi Informatsii*, 5(3):64–72, 1969.
- [Vil09] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer, 2009.
- [Wag96] Wolfgang Wagner. A functional law of large numbers for boltzmann type stochastic particle systems. *Stochastic Analysis and Applications*, 14(5):591–636, 1996.
- [Wei07] P. Weiss. L’hypothèse du champ moléculaire et la propriété ferromagnétique. *Journal de Physique*, 6(661):36, 1907.
- [ZS17] Joel Zylberberg and Ben W Strowbridge. Mechanisms of persistent activity in cortical circuits: possible neural substrates for working memory. *Annual review of neuroscience*, 40:603, 2017.