

TESIS DE MAESTRÍA EN CIENCIAS FÍSICAS

**INTERACCIÓN FUNCIONAL ENTRE NEURONAS:
APLICACIÓN EN LA RETINA**

Gabriel Matías Lorenz
Maestrando

Dr. Damián Hernández Lahme
Director

Miembros del Jurado
Dr. Ariel Curiale (CONICET-Instituto Balseiro)
Dr. Sebastián Risau (CONICET)
Dr. Eugenio Urdapilleta (CONICET-Instituto Balseiro)

Neurociencia Computacional – Dept. Física Médica – Centro
Atómico Bariloche

Instituto Balseiro
Universidad Nacional de Cuyo
Comisión Nacional de Energía Atómica
Argentina

Índice de contenidos

Índice de contenidos	iii
Índice de figuras	v
Resumen	vii
Abstract	ix
1. Introducción	1
1.1. La retina	3
2. Modelo de interacción neuronal	5
2.1. Notación	5
2.2. Modelo marginal para una neurona	6
2.3. Modelo logarítmico-lineal para dos neuronas	8
2.4. Relación entre modelos	10
3. Estimadores de $P(\sigma \mathbf{s})$	11
3.1. Estimadores para el modelo marginal	11
3.2. Estimadores para el modelo de interacción	14
3.3. Estimadores de $\alpha(\mathbf{s})$	15
3.3.1. Estimador bayesiano	16
3.4. Rendimiento del estimador bayesiano	18
3.5. Comparación de estimadores	19
4. Resultados y discusión	21
4.1. Datos usados	21
4.2. Procesamiento	22
4.3. Resultados	23
4.3.1. Consistencia con el modelo LN	27
4.3.2. Validación	28
4.4. Discusión y propuestas a futuro	30

5. Conclusiones	35
Bibliografía	37

Índice de figuras

1.1. Esquema de una retina	3
2.1. Ejemplos de estímulos	6
2.2. Ejemplos de filtros	7
2.3. Ejemplos de funciones no lineales	7
2.4. Esquema del modelo LN en dos neuronas	8
2.5. Esquema del modelo de interacción en dos neuronas	9
3.1. Ejemplo de STA	12
3.2. Ejemplo de STC	13
3.3. Función ψ	17
3.4. Rendimiento del estimador bayesiano	18
3.5. Rendimiento del estimador	19
4.1. Filtros espaciales en dos células OFF	24
4.2. Filtros temporales en dos células OFF	24
4.3. Funciones no lineales en dos células OFF	25
4.4. Resultados para células ON	25
4.5. Resultados para células ON-OFF	26
4.6. Resultados para células mixtas	26
4.7. Equivalencias de filtros	27
4.8. Equivalencia de funciones no lineales	28
4.9. Crossvalidation	29
4.10. Propuesta de redes neuronales	31

Resumen

La función principal de las neuronas es la de codificar información: pueden codificar la información de un estímulo, transmitir órdenes a otras regiones o servir de enlace para otras neuronas, entre otras cosas. Al estudiar cómo una población de neuronas codifica información externa, es común que se analice la dependencia de esta información con la actividad individual de cada neurona, ignorando las correlaciones dentro de la población. Si bien esto contribuye en gran parte a caracterizar neuronas y poblaciones de ellas, se ha encontrado que las neuronas realizan cálculos complejos, aún en capas de bajo nivel, gracias a las interacciones entre neuronas de una misma capa. Por lo tanto, es necesario estudiar el comportamiento *conjunto* de las neuronas en la codificación. Existen modelos que incorporan las interacciones entre pares de neuronas y que logran describir mejor el comportamiento de algunas poblaciones de neuronas. No obstante, no proveen una perspectiva completa a la hora de entender el origen de las correlaciones o la función que cumplen en la codificación del estímulos. En este trabajo proponemos una forma de entender esta codificación al buscar no sólo las características del estímulo que inducen una respuesta de una determinada neurona sensorial, sino también aquellas que modulan la interacción entre ésta y una neurona vecina de su mismo nivel. Para ello, planteamos una distribución de probabilidad conjunta para el disparo de dos neuronas considerando que la correlación entre ellas es guiada por campos receptivos en el estímulo. A fin de poder calcular los parámetros de este modelo en datos experimentales construimos un estimador que discrimina los campos receptivos individuales de los de interacción, aún en casos de pocas muestras. Aplicamos el modelo a mediciones de células ganglionares de la retina de ratón. Encontramos evidencia de campos receptivos de la interacción inhibitorios y los resultados fueron consistentes con los obtenidos por el modelo marginal. Además, las pruebas de validación hechas sobre el modelo muestran una mejora en la capacidad de predicción de datos en comparación al modelo de probabilidad marginal en tres de cuatro casos estudiados.

Palabras clave: NEUROCIENCIA COMPUTACIONAL, CORRELACIONES NEURONALES, INFERENCIA BAYESIANA

Abstract

Coding information is the principal function of neurons: they can code information from stimuli, give orders to other regions in the body and be messengers for other neurons, among other things. When one studies how a neuronal population codes external information, it is common to analyze the dependence of this information with every individual neural activity, disregarding correlations inside the population. Although this contributes greatly to characterize neurons and populations, it has been found that neurons can make complex calculations even in low level layers thanks to interactions between neurons in the same layer. Because of this, it is necessary to search how to study the joint behavior of neurons in coding. Currently, there are models that incorporate pairwise interactions to describe population behavior. Nevertheless, they do not provide foundations necessary to understand the origin of these correlations or its function in stimulus coding. In this work we propose a probability model for a pair of neurons, incorporating explicitly a correlation modulated by the stimulus. We constructed a bayesian estimator to fit the model to experimental data. This estimator is capable of discriminate between individual and coupling receptive fields, even in low-count measurements scenarios. We applied the model in retinal ganglion cells of mice. We found evidence of inhibitory coupling receptive fields and its features were consistent with ones from the marginal model. We run a cross-validation test and found positive results for three of four studied pairs.

Keywords: COMPUTATIONAL NEUROSCIENCE, NEURAL CORRELATIONS, BAYESIAN INFERENCE

Capítulo 1

Introducción

Las neuronas se comunican entre sí mediante potenciales eléctricos y neurotransmisores. Entender cómo las neuronas codifican la información mediante estas señales es una pregunta fundamental que guía las investigaciones en neurociencia.

Dentro del cerebro hay un gran número de diferentes tipos de neuronas que cumplen distintas funciones. Entre estas funciones se encuentran la de expresar características del exterior (neuronas sensoriales), transmitir órdenes a otras regiones del cuerpo (neuronas motoras) y servir de enlace para otras neuronas (interneuronas). En el caso de las neuronas sensoriales, se ha encontrado que no responden a cambios en cualquier parte del estímulo, sino que sólo reaccionan a un conjunto específico de sus propiedades. Se encontraron dependencias entre la orientación del estímulo visual con la actividad de las neuronas en la corteza visual primaria V_1 [1] y neuronas en la región inferotemporal que reaccionan a la presencia de manos y caras en una imagen[2]. Los hallazgos realizados no se limitan a la visión: se han encontrado células de lugar en el hipocampo que se activan cuando el sujeto se encuentra en determinadas posiciones de su ambiente[3].

Estas observaciones junto con otras llevaron a modelar los campos receptivos en la neurona: regiones específicas del estímulo a las que la neurona presta atención, descartando la información fuera de ellas [4]. Estos campos pueden variar entre las neuronas de una misma población y son relativamente constantes, por lo que sirven de herramienta para distinguir neuronas según su función[5].

El estudio de los campos receptivos proporcionó herramientas para investigar el camino que realiza la información de un estímulo por el cerebro. En términos generales, la codificación de un estímulo puede modelarse por capas donde las neuronas reciben la información de una capa anterior y después de procesarla la transmiten a la siguiente. A partir de la información que le llega a su campo receptivo, la neurona puede procesarla de distintas maneras a fin de lograr una codificación óptima[6]. En términos generales, las capas van codificando informa-

ción cada vez más compleja: los campos receptivos de las neuronas más cercanas al estímulo codifican rasgos básicos y cada capa siguiente recibe información cada vez más procesada y compleja[7]. En la realidad este proceso no es tan lineal y simple como lo describimos[7], pero es útil para comenzar a entender cómo codifica una neurona.

Dado que la actividad de una neurona está fuertemente condicionada por la información proveniente de la capa más básica, al analizarla es común marginar el comportamiento de las células vecinas. Este enfoque permite un análisis menos complejo de la función de una neurona, pero limita la comprensión de cómo las poblaciones de neuronas codifican en conjunto. Si bien se ha observado que las correlaciones entre pares de neuronas y correlaciones de mayor nivel pueden ser débiles ($p(\sigma_1, \dots, \sigma_2) \approx \prod p(\sigma_i)$)[8], influyen fuertemente el comportamiento de muchos sistemas neuronales[8, 9]. Por ende, tener en cuenta estas correlaciones puede contribuir en gran medida a las predicciones que uno realiza sobre el comportamiento de una población de neuronas.

Los métodos para determinar estos campos receptivos individuales no pueden capturar la información de estas interacciones ya que no discriminan entre las contribuciones del estímulo exclusivas a una célula y las sinérgicas. Se han propuesto métodos y modelos para estudiar las correlaciones de a pares en una población de neuronas. Schneidman y Granot-Atedgi proponen modelos de máxima entropía en los que se ajustan coeficientes de interacción fijos para cada par de neuronas de una medición[9, 10]. En el caso de la visión, se ha propuesto usar redes neuronales convolucionales para reproducir el comportamiento de las células ganglionares[11].

El principal inconveniente de estos métodos es que sus resultados no son lo suficientemente interpretables: si bien consiguen mejorar las predicciones sobre la actividad del sistema de neuronas que se estudia, no proveen una estructura interna que permita entender qué causa estas interacciones ni qué función tienen en la codificación del estímulo.

En este trabajo proponemos un modelo de disparo de las neuronas en el que la correlación entre neuronas del mismo nivel es modulada por el estímulo. Este modelo busca las regiones del estímulo que afectan la correlación entre neuronas. Buscamos con esto dar información interpretable para investigar el papel de las correlaciones en la codificación de la información.

Si bien el trabajo presentado es usado en neuronas sensoriales, esto puede ser generalizado para cualquier variable externa al sistema de neuronas. En particular, este modelo puede aplicarse para relacionar la actividad neuronal con el comportamiento del sujeto.

1.1. La retina

En este trabajo el modelo que proponemos es aplicado sobre datos de la retina de ratón, por lo que brindamos algunos datos sobre ésta.

La retina de los vertebrados (Fig. 1.1) es una capa fina de tejido neuronal que se encuentra en la parte posterior del ojo. Contiene cuatro clases básicas de neuronas, además de fotorreceptores, organizados en tres capas celulares (nucleares). Estas capas se encuentran separadas por dos capas sinápticas (plexiformes), donde se producen prácticamente todas las conexiones entre las neuronas de la retina.

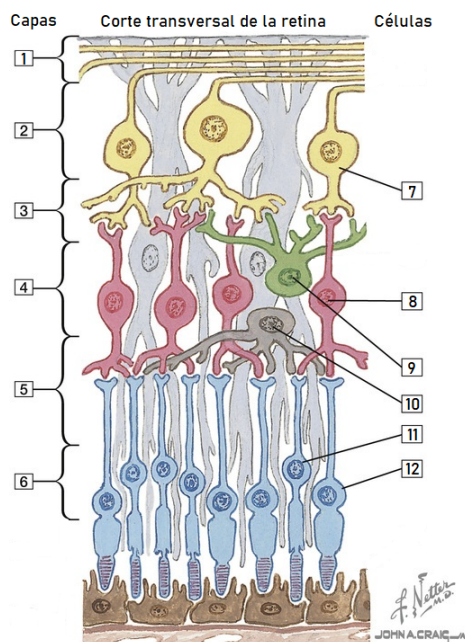


Figura 1.1: Esquema de una retina obtenida de *Netter's Neuroscience Flash Cards*. 1: nervios ópticos que van al cerebro. 2: capa celular ganglionar. 3: capa sináptica interna. 4: capa celular interna. 5: capa sináptica externa. 6: capa celular externa. 7: célula ganglionar. 8: célula bipolar. 9: célula amacrina. 10: célula horizontal. 11: fotorreceptor tipo bastón. 12: fotorreceptor tipo cono.

El cuerpo de los fotorreceptores, quienes son los que se encuentran más alejados del frente del ojo, componen la capa nuclear externa, mientras que la capa nuclear interna está compuesta por los cuerpos de las células horizontales, bipolares y amacrinas. Las células ganglionares forman la última capa nuclear que es la más cercana al frente del ojo. Cuando la luz entra al ojo, atraviesa la retina transparente y es capturada por los fotorreceptores. La información de éstos atraviesa la capa plexiforme externa y llega a las células bipolares y horizontales. Las células bipolares reciben toda la información que entregan los fotorreceptores y emiten sus respuestas hacia la capa plexiforme interna donde se activarán las células amacrinas y ganglionares. En cambio, las acciones de las células horizontales se encuentran confinadas en la capa plexiforme externa, mediando interacciones

laterales en esa capa. Lo mismo ocurre para las células amacrinas en la capa plexiforme interna. Por último, las células ganglionares son las neuronas de salida de la retina y sus conexiones se reúnen para formar el nervio óptico y transmitir la información visual a centros visuales del cerebro[5].

Capítulo 2

Modelo de interacción neuronal

En este capítulo desarrollaremos el modelo que usamos para representar la interacción funcional. Comenzaremos por introducir la notación empleada tanto para representar la actividad eléctrica de las neuronas como de los estímulos que se les presentan. También se explicará el modelo lineal - no lineal de la actividad neuronal, que servirá para proveer herramientas en la construcción de nuestro propio modelo y para comparar el rendimiento entre ambos.

2.1. Notación

Para poder trabajar con modelos matemáticos del disparo neuronal, es necesario introducir algún tipo de notación para poder representar tanto el estímulo de la neurona como su respuesta.

Una neurona es capaz de generar distintos patrones de disparo [12] y puede transmitir información en distintos atributos de éstos (ej.: frecuencia, tiempo, etc.). Para simplificar nuestro trabajo, caracterizamos la condición de una neurona según englobadas por un estado inactivo y un estado excitado. Cuando en la ventana temporal en la que observamos la neurona ocurre un potencial de acción, decimos que la neurona se encuentra excitada y lo representamos con $\sigma_i = 1$. En caso de que no haya habido una activación en ese tiempo, la neurona se encuentra inactiva con $\sigma = 0$. Esto es lo que se conoce como el tren de disparos de la neurona.

Los vectores son la notación usada para describir no sólo estímulos relacionados a distintos sentidos (ej.: visual y auditivos), sino que tengan en cuenta las distintas características que conforman a cada tipo de estímulo (ej.: evolución temporal, intensidad y color para estímulos visuales; timbre, volumen y frecuencia para estímulos auditivos). La forma de un vector \mathbf{s} puede elegirse acorde para que sus componentes puedan representar distintas características del estímulo según el interés del investigador. En el caso de estímulos visuales, las componentes s_i pue-

den representar intensidades en función del tiempo t_i en el caso de querer estudiar campos receptivos temporales (Fig. 2.1, caso a), posiciones en el espacio (x_i, y_i) a un tiempo específico (Fig. 2.1, caso b) o una combinación de ambas en el caso de campos espacio-temporales (Fig. 2.1, caso c).

Los estímulos que pretendemos analizar en este trabajo tienen una forma establecida. En el experimento [13] la retina es estimulada con una sucesión de matrices de dimensión 15×20 , donde cada componente puede valer 0 o 1 con probabilidad uniforme en cada ventana temporal e independientemente de las otras componentes de la matriz. Para esta situación entonces las componentes de \mathbf{s} representarán la posición en cada una de estas grillas. Si bien en el experimento, y por lo tanto en nuestro análisis, se considera la progresión temporal entre estas grillas (Fig. 2.1, caso c), para la explicación de distintos modelos teóricos y estimadores estadísticos sólo analizaremos estímulos espaciales (como en el segundo caso de la Figura 2.1) a fin de simplificar los ejemplos.

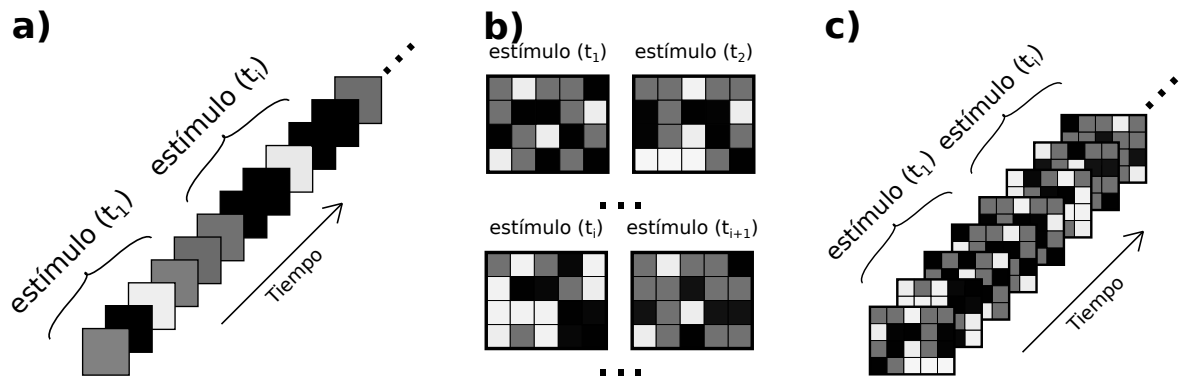


Figura 2.1: Ejemplos de estímulos visuales según la representación elegida. Las componentes pueden representar valores en el tiempo (izquierda), posiciones en el espacio (centro) o una combinación de ambas (derecha). Imagen inspirada de [14]

2.2. Modelo marginal para una neurona

Una neurona puede procesar hasta cierta cantidad de información entrante debido a los límites en la variabilidad de sus respuestas. Es inevitable entonces que la información que llega a codificar la neurona sea menor que la información completa del estímulo mismo. Hallazgos en distintas regiones del cerebro [1–3] dan pie a modelar campos receptivos en la neurona, regiones del estímulo a la que la neurona presta atención, ignorando el resto.

Uno puede modelar los campos receptivos como vectores relevantes dentro del espacio de estímulos. De esta manera, la información del estímulo que procesa la neurona se consigue proyectándolo sobre estas direcciones relevantes, también

llamadas filtros en la literatura. En la Figura 2.2 se pueden ver ejemplos de vectores relevantes del espacio de los estímulos espaciales de la Figura 2.1.

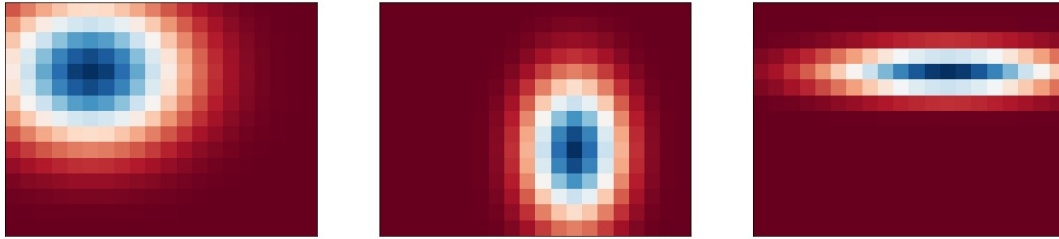


Figura 2.2: Ejemplos de direcciones relevantes para el caso del espacio de estímulos visuales de tamaño 15×20 .

Entonces, uno representa la probabilidad de disparo condicionada por el estímulo $P(\sigma = 1|\mathbf{s})$ como una función no necesariamente lineal que tiene como argumentos las proyecciones en los subespacios relevantes:

$$P(\sigma = 1|\mathbf{s}) = \psi(\mathbf{k}_1 \cdot \mathbf{s}, \dots, \mathbf{k}_m \cdot \mathbf{s}), \quad (2.1)$$

donde \mathbf{k}_i es el filtro correspondiente a cada uno de dichos subespacios y ψ es un función no lineal con imagen en el intervalo $[0, 1]$. En la Figura 2.3 se pueden observar ejemplos de funciones de probabilidad comunes para una o dos direcciones relevantes.

la proyección del estímulo sobre direcciones relevantes y el cómputo de la probabilidad de disparo a través de ellas componen el modelo lineal-no lineal o modelo LN, que logra describir el comportamiento de muchos tipos de neuronas sensoriales [15]. En la Figura 2.4 se muestra un esquema que resume el procedimiento del modelo. En el caso de que la región relevante para una neurona no sea un subespacio dentro del espacio de estímulos, sino una superficie arbitraria dentro de él, el modelo toma como subespacio relevante al menor subespacio que contenga esta superficie.

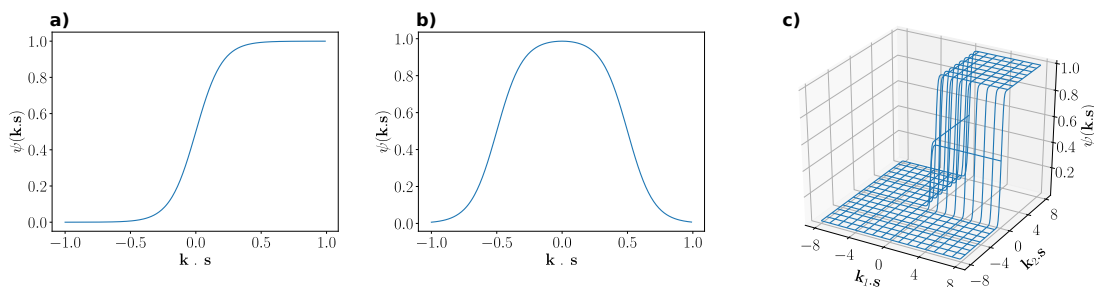


Figura 2.3: Ejemplos de funciones de probabilidad $P(\sigma|\mathbf{s})$ para uno o dos argumentos.

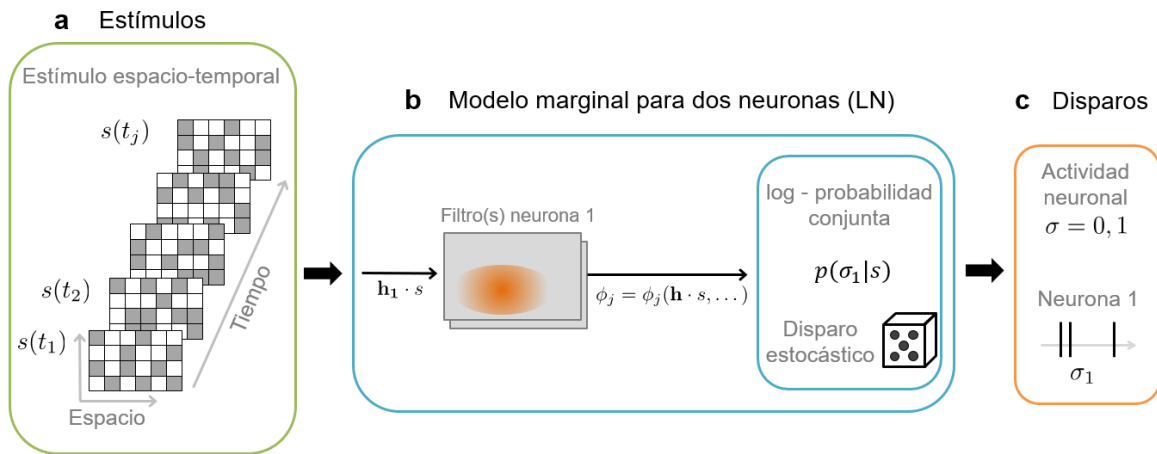


Figura 2.4: Esquema del modelo LN.

La siguiente sección presentará nuestro modelo que busca representar la actividad de pares de neuronas teniendo en cuenta una interacción dependiente del estímulo.

2.3. Modelo logarítmico-lineal para dos neuronas

Para diseñar el modelo para pares de neuronas se buscó que no pierda generalidad, evitando incorporar suposiciones extra al modelo. La distribución de probabilidad restringida a determinados valores medios que introduce menos sesgo es aquella que tiene máxima entropía dadas las restricciones. En una población de neuronas la distribución de máxima entropía fijado el estímulo y restringiendo los valores medios de las actividades individuales $\langle \sigma_i \rangle$ y de la correlación $\langle \sigma_i \sigma_j \rangle$, tiene la forma

$$\ln P(\{\sigma\}|\mathbf{s}) = \sum_i \sigma_i \cdot \alpha_i(\mathbf{s}) + \sum_{i,j>i} \sigma_i \sigma_j \cdot \alpha_{ij}(\mathbf{s}) - \ln Z, \quad (2.2)$$

donde Z es la constante de normalización.

La distribución elegida es perteneciente a la familia de las distribuciones logarítmico-lineales. Las funciones α_i representan la contribución individual para la probabilidad de disparo de cada neurona. Las funciones α_{ij} modifican la probabilidad del disparo simultáneo de las neuronas i y j y está relacionada con la información mutua entre ellas. En particular, la información mutua condicionada al estímulo \mathbf{s} entre las dos neuronas es cero cuando $\alpha_{12} = 0$ [16].

En el caso de un sistema de dos neuronas, la expresión 2.2 define la única distribución de probabilidad posible dadas las restricciones pedidas. La probabilidad de cada estado se encuentra definida de la siguiente forma:

$$P(\sigma_1 = 0, \sigma_2 = 0 | \mathbf{s}) = q_{00}(\mathbf{s}) = \frac{1}{Z}$$

$$P(\sigma_1 = 1, \sigma_2 = 0 | \mathbf{s}) = q_{10}(\mathbf{s}) = \frac{e^{\alpha_1(\mathbf{s})}}{Z}$$

$$P(\sigma_1 = 0, \sigma_2 = 1 | \mathbf{s}) = q_{01}(\mathbf{s}) = \frac{e^{\alpha_2(\mathbf{s})}}{Z}$$

$$P(\sigma_1 = 1, \sigma_2 = 1 | \mathbf{s}) = q_{11}(\mathbf{s}) = \frac{e^{\alpha_1(\mathbf{s}) + \alpha_2(\mathbf{s}) + \alpha_{12}(\mathbf{s})}}{Z}$$

Para poder interpretar los resultados del modelo en función del estímulo, escribimos las funciones $\alpha(\mathbf{s})$, así como la $P(\sigma | \mathbf{s})$ en el modelo LN, como dependientes de proyecciones en vectores relevantes. Las funciones α quedan expresadas de la forma

$$\alpha_1(\mathbf{s}) = \zeta_1(\mathbf{h}_1^1 \cdot \mathbf{s}, \dots, \mathbf{h}_m^1 \cdot \mathbf{s}), \quad (2.3)$$

$$\alpha_2(\mathbf{s}) = \zeta_2(\mathbf{h}_1^2 \cdot \mathbf{s}, \dots, \mathbf{h}_r^2 \cdot \mathbf{s}), \quad (2.4)$$

$$\alpha_{12}(\mathbf{s}) = \zeta_{12}(\mathbf{h}_1^{12} \cdot \mathbf{s}, \dots, \mathbf{h}_l^{12} \cdot \mathbf{s}), \quad (2.5)$$

donde $\{\mathbf{h}_i^1\}$ y $\{\mathbf{h}_i^2\}$ son los filtros individuales de cada neurona y $\{\mathbf{h}_i^{12}\}$ es el conjunto de filtros asociado a la interacción de las neuronas. Las funciones ζ son funciones arbitrarias con imagen en todo el espacio real. En la Figura 2.5 se muestra un esquema que resume lo descrito anteriormente.

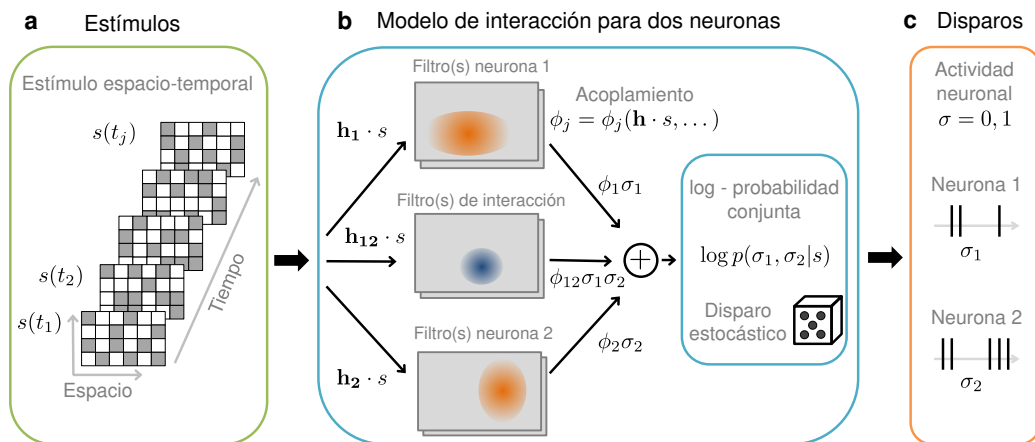


Figura 2.5: Esquema del modelo de interacción para un sistema de dos neuronas.

2.4. Relación entre modelos

Como nuestro modelo es un caso más general construido a partir del modelo marginal de una neurona, debería ser posible volver a este último si marginamos la respuesta de una de las neuronas en nuestro modelo. Para facilitar esta comparación, podemos reescribir 2.1 de manera de que también sea de la forma logarítmica-lineal:

$$P(\sigma = 1|\mathbf{s}) = \frac{1}{e^{-\phi(\mathbf{k}_1 \cdot \mathbf{s}, \dots, \mathbf{k}_m \cdot \mathbf{s})} + 1}, \quad (2.6)$$

donde ϕ es una función con dominio en \mathbb{R} . Si bien esta forma de escribir la probabilidad no permite trabajar con distribuciones donde la probabilidad sea 1 o 0 para algún estímulo, sigue siendo lo suficientemente general para utilizarlo como reemplazo de la ec. 2.1 teniendo en cuenta esta limitación.

En un sistema de dos neuronas, si marginamos la respuesta de la segunda neurona, obtenemos la siguiente expresión:

$$\begin{aligned} P(\sigma_1 = 1|\mathbf{s}) &= P(\sigma_1 = 1; \sigma_2 = 0|\mathbf{s}) + P(\sigma_1 = 1; \sigma_2 = 1|\mathbf{s}) \\ &= \frac{e^{\alpha_1(\mathbf{s})}}{Z} (1 + e^{\alpha_2(\mathbf{s}) + \alpha_{12}(\mathbf{s})}). \end{aligned}$$

Si usamos la definición de Z y usamos factor común obtenemos

$$P(\sigma_1 = 1|\mathbf{s}) = \frac{e^{\alpha_1} (1 + e^{\alpha_2 + \alpha_{12}})}{1 + e^{\alpha_2} + e^{\alpha_1} (1 + e^{\alpha_2 + \alpha_{12}})},$$

donde vemos que parte del denominador tiene la misma forma que su numerador. Usando esto e igualando esta probabilidad a la forma que plantea el modelo LN resulta que

$$\frac{1}{e^{-\phi_1(\mathbf{k}_1 \cdot \mathbf{s}, \dots, \mathbf{k}_m \cdot \mathbf{s})} + 1} = \frac{1}{e^{-\alpha_1} \left[\frac{1 + e^{\alpha_2}}{1 + e^{\alpha_2 + \alpha_{12}}} \right] + 1}.$$

Despejando esta igualdad obtenemos

$$\phi_1(\mathbf{k}_1 \cdot \mathbf{s}, \dots, \mathbf{k}_m \cdot \mathbf{s}) = \alpha_1(\mathbf{s}) - \ln \left[\frac{1 + e^{\alpha_2(\mathbf{s})}}{1 + e^{\alpha_2(\mathbf{s}) + \alpha_{12}(\mathbf{s})}} \right]. \quad (2.7)$$

Cabe notar que si no hubiese una interacción entre las neuronas de este sistema $\alpha_{12}(\mathbf{s})$ sería nulo y tendríamos una correspondencia entre la función ϕ_1 , asociada a la actividad marginal de la neurona, y la contribución individual α_1 de nuestro modelo.

Capítulo 3

Estimadores de $P(\sigma|\mathbf{s})$

En el capítulo anterior desarrollamos nuestro modelo de actividad para un par de neuronas que incorpora la interacción entre ellas, describiendo esta última en función del estímulo. Además, la relacionamos con el modelo LN que describe la dependencia de la neurona con el estímulo, marginando la actividad neuronal vecina. Pero, para que estos modelos sean útiles, es necesario encontrar los métodos que puedan extraer sus parámetros de las mediciones reales sobre las neuronas.

En este capítulo presentaremos los estimadores usados para ambos modelos. Describiremos cómo se derivan del modelo y qué supuestos necesitan. Se comparará su rendimiento usando datos simulados y se detallarán sus limitaciones.

3.1. Estimadores para el modelo marginal

Para poder describir la respuesta de una neurona en función del modelo LN se necesitan encontrar las direcciones relevantes de las neuronas ($\{\mathbf{k}_i\}$) y la dependencia entre las proyecciones sobre éstas dada por $\phi(\{\mathbf{k}_i \cdot \mathbf{s}\})$. Para poder encontrar los campos receptivos bajo los que una neurona opera hay dos técnicas estadísticas principales: el *spike-triggered average* (STA), dado por el promedio de los estímulos que excitaron la neurona, y el *spike-triggered covariance* (STC), promedia la covarianza de estos mismos estímulos[14, 17].

Para usar estos métodos, es necesario que los estímulos presentados posean una distribución de probabilidad esférica[14]. En el caso de nuestro estímulo, esto se cumple dado que los valores de las celdas en la grilla siguen una distribución uniforme y son independientes entre sí espacial- y temporalmente. Además, si uno procesa previamente los estímulos de manera que cumplan las condiciones

$$\langle \mathbf{s} \rangle = \sum_{\mathbf{s}} \mathbf{s} = 0 \quad \text{y} \quad (3.1)$$

$$C = \sum_{\mathbf{s}} \mathbf{s}\mathbf{s}^T = \mathbf{1}_{N \times N}, \quad (3.2)$$

las expresiones de los estimadores STA y STC se simplifican. Usando la notación definida en el capítulo anterior, el estimador STA se escribe entonces de la siguiente manera:

$$\mathbf{k} = \frac{\sum_{\mathbf{s}} \sigma(\mathbf{s})\mathbf{s}}{\sum_{\mathbf{s}} \sigma(\mathbf{s})}. \quad (3.3)$$

El vector \mathbf{k} que estima este método se encuentra dentro del subespacio relevante. En el caso de que la neurona tenga más de una dirección relevante, este método captura una combinación lineal entre ellos. Para ilustrar esto, en la Figura 3.1 se puede observar el filtro conseguido usando los datos simulados de una neurona que posee los primeros dos filtros de la Figura 2.2 y su función de probabilidad es la tercera de las mostrada en la Figura 2.3.

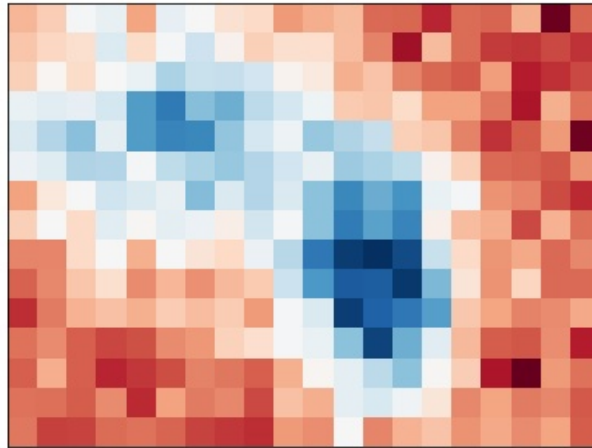


Figura 3.1: El resultado de un *spike-triggered average* usando las respuestas simuladas ante 10000 estímulos de una neuronas con los primeros dos filtros de la Figura 2.2 y la función de la derecha de la Figura 2.3.

Para obtener obtener los vectores que generan un subespacio relevante de más de una dimensión, uno tiene que usar el método de *spike-triggered covariance*. Esto se hace calculando la matriz

$$\langle \hat{\mathbf{C}}_s \rangle = \frac{1}{\sum_{\mathbf{s}} \sigma(\mathbf{s}) - 1} \sum_{\mathbf{s}} (\sigma(\mathbf{s})\mathbf{s} - \mathbf{k})(\sigma(\mathbf{s})\mathbf{s} - \mathbf{k})^T. \quad (3.4)$$

Al diagonalizar 3.4, aquellos autovectores que corresponden a direcciones irrelevantes del estímulo poseen sus autovalores cercanos a 1 (Ec. 3.2). Los autovalores más alejados del resto son aquellos cuyos autovectores pertenecen al subespacio

relevante. En la Figura 3.2 se muestran los autovalores encontrados usando el caso simulado para la Figura 3.1 y los filtros encontrados a partir de los autovalores marcados.

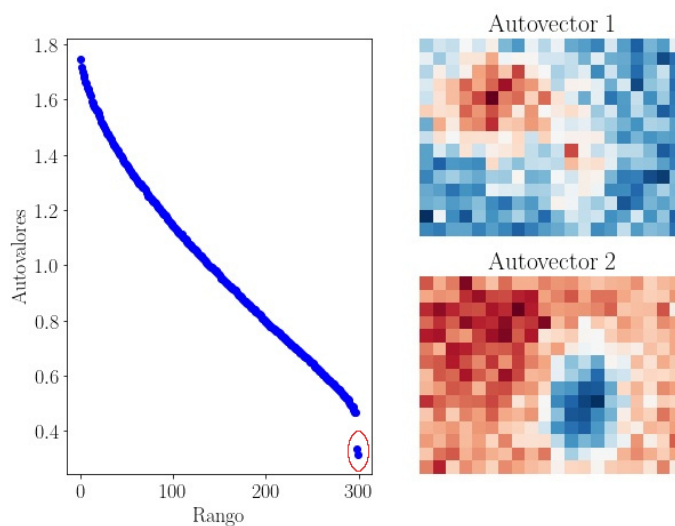


Figura 3.2: A la izquierda, los autovalores de la matriz $\langle \hat{C}_s \rangle$ conseguidos con *spike-triggered covariance* para el caso simulado de la Figura 3.1. A la izquierda, se muestran los autovalores provenientes de la diagonalización de esta matriz. Los autovalores más alejados de la unidad son aquellos cuyos autovectores son generadores del espacio relevante (figuras de la derecha).

Una vez conseguidas las direcciones relevantes para la neurona, uno puede aproximar la probabilidad de disparo $P(\sigma = 1|\mathbf{s})$, realizando un histograma de las proyecciones de todos los estímulos usados. Si uno normaliza el histograma, se obtiene una aproximación de la probabilidad de disparo en el rango.

3.2. Estimadores para el modelo de interacción

Se pueden conocer más propiedades de α analizando sus momentos. Para una de las funciones α_l , las componentes i del primer momento están dadas por

$$A_{li}^{(1)} = \langle s_i \alpha_l(\mathbf{s}) \rangle_{P(\mathbf{s})} \quad (3.5)$$

El teorema de Novikov [18] relaciona para una distribución gaussiana de los estímulos con media nula, el primer momento de una función con sus derivadas en la forma

$$A_{li}^{(1)} = \sum_{j=1}^l C_{ij} \left\langle \frac{\partial \alpha_l(\mathbf{s})}{\partial s_j} \right\rangle. \quad (3.6)$$

Como definimos los estímulos tales que $C_{ij} = \delta_{ij}$, obtenemos que

$$A_{li}^{(1)} = \left\langle \frac{\partial \alpha_l(s)}{\partial s_i} \right\rangle. \quad (3.7)$$

Entonces, aplicando la derivada en la expresión en 2.3, uno consigue

$$A_{li}^{(1)} = \langle s_i \alpha_l(\mathbf{s}) \rangle = \sum_{\mu=1}^l \left\langle \frac{\partial \zeta_{\mu}(\mathbf{s})}{\partial (\mathbf{s} \cdot \mathbf{k}_{\mu})} \right\rangle \mathbf{k}_{\mu}(i), \quad (3.8)$$

lo que significa que el primer momento de la función α está dentro del subespacio relevante. Experimentalmente, uno muchas veces no posee la distribución de probabilidad de los estímulos necesaria para calcularlo. Aún así, es posible estimar el valor del primer momento para conjuntos grandes de datos con la expresión

$$\mathbf{h}_l = \hat{\mathbf{A}}_l^{(1)} = \frac{1}{|S|} \sum_S \alpha_l(\mathbf{s}) \mathbf{s}, \quad (3.9)$$

realizando el promedio sobre los estímulos del experimento. La ley de los grandes números nos asegura que esta aproximación tiende a $\mathbf{A}^{(1)}$.

Entonces, podemos inferir los campos receptivos tanto individuales como de interacción de las respuestas de un par de neuronas ante varios estímulos.

Analizando de forma similar el segundo momento

$$\mathbf{A}_{lij}^{(2)} = \delta_{ij} \langle \alpha(\mathbf{s}) \rangle + \left\langle \frac{\partial \alpha_l(s)}{\partial s_i \partial s_j} \right\rangle \quad (3.10)$$

uno llega a que la matriz $A_{ij}^{(2)} - \delta_{ij} \langle \alpha(\mathbf{s}) \rangle$ puede tener tantos autovalores no nulos como filtros a los que está asociado α y sus autovectores generan el subespacio relevante. Esto se deberá tomar en cuenta cuando las funciones α dependan de

más de un filtro[17].

3.3. Estimadores de $\alpha(\mathbf{s})$

Por lo visto anteriormente, uno puede conseguir una dirección relevante para cada contribución sabiendo el valor de la función α para cada estímulo medido. Uno puede escribir estas funciones a partir de las probabilidades de los cuatro estados para cada estímulo. Escribiendo las probabilidades de los cuatro estados posibles como $P(\sigma_1\sigma_2|\mathbf{s}) \equiv q_{\sigma_1\sigma_2}(\mathbf{s})$, las funciones α pueden expresarse como

$$\alpha_1(\mathbf{s}) = \ln \left(\frac{q_{10}(\mathbf{s})}{q_{00}(\mathbf{s})} \right), \quad (3.11)$$

$$\alpha_2(\mathbf{s}) = \ln \left(\frac{q_{01}(\mathbf{s})}{q_{00}(\mathbf{s})} \right). \quad (3.12)$$

$$\alpha_{12}(\mathbf{s}) = \ln \left(\frac{q_{11}(\mathbf{s}) \cdot q_{00}(\mathbf{s})}{q_{10}(\mathbf{s}) \cdot q_{01}(\mathbf{s})} \right). \quad (3.13)$$

Podemos entonces realizar estimaciones del valor de estas funciones ante cierto estímulo a partir de estimaciones de las probabilidades $\mathbf{q} = (q_{00}, q_{10}, q_{01}, q_{11})$.

Si uno maximiza la probabilidad de los datos dados los parámetros (en este caso, las probabilidades \mathbf{q}), se obtiene el estimador de máxima verosimilitud (ML, por su siglas en ingles). Esto es equivalente a maximizar la probabilidad a posteriori suponiendo una prior constante en \mathbf{q} . En este caso se aproxima la probabilidad de que ocurra uno de los estados como la frecuencia relativa de las mediciones:

$$\hat{q}_{\sigma_1\sigma_2 ML}(\mathbf{s}) = \frac{N_{\sigma_1\sigma_2}(\mathbf{s})}{N(\mathbf{s})} \quad (3.14)$$

Una propiedad de este estimador es que la estimación del valor medio de una función es igual a la función evaluada en la estimación de sus argumentos:

$$\hat{f}(x, y, z)_{ML} = f(\hat{x}_{ML}, \hat{y}_{ML}, \hat{z}_{ML}).$$

Entonces, el valor medio de α_{12} queda de la forma:

$$\hat{\alpha}_{12 ML}(\mathbf{s}) = \ln \left(\frac{N_{11}(\mathbf{s}) \cdot N_{00}(\mathbf{s})}{N_{10}(\mathbf{s}) \cdot N_{01}(\mathbf{s})} \right). \quad (3.15)$$

El problema con este método surge cuando hay pocas mediciones por estímulo y no hay ocurrencias de alguno de los cuatro estados posibles. Esto genera una divergencia en el estimador inevitable en la mayoría de las mediciones.

Si en vez de encontrar la distribución \mathbf{q} más probable se busca aquella con

menor error cuadrático medio, la estimación es de la forma

$$\langle q_{\sigma_1\sigma_2}(\mathbf{s}) \rangle_2 = \frac{N_{\sigma_1\sigma_2} + 1}{N + 4} \quad (3.16)$$

Si bien esto evita las divergencias en caso de que $N_{\sigma_1\sigma_2}$ sea cero, cuando $N = 1$ se genera un sesgo fuerte en las probabilidades marginales de disparo (se aproximan a una distribución uniforme).

3.3.1. Estimador bayesiano

Dado que el valor de $\alpha_l(\mathbf{s})$ ante un estímulo puede ser calculado con las probabilidades de disparo $\mathbf{q} = (q_{\sigma_1\sigma_2}) = (q_{00}, q_{10}, q_{01}, q_{11})$ (Ec. 3.13), podemos estimarlo como el valor medio de $\alpha_l(\mathbf{q}(\mathbf{s}))$ ante una distribución de probabilidad $P(\mathbf{q}(\mathbf{s}))$. Si medimos N veces la respuesta del sistema ante un estímulo \mathbf{s} y armamos el vector $\mathbf{N} = (N_{\sigma_1\sigma_2})$, podemos escribir la estimación como

$$\langle \alpha_{12}(\mathbf{N}(\mathbf{s})) \rangle = \int d\mathbf{q} \alpha_{12}(\mathbf{q}(\mathbf{s})) P(\mathbf{q}(\mathbf{s})|\mathbf{N}(\mathbf{s})) \quad (3.17)$$

Usando la regla de Bayes podemos invertir la probabilidad condicional como

$$\langle \alpha_{12}(\mathbf{N}(\mathbf{s})) \rangle = \int d\mathbf{q} \alpha_{12}(\mathbf{q}(\mathbf{s})) \frac{P(\mathbf{N}(\mathbf{s})|\mathbf{q}(\mathbf{s})) P(\mathbf{q}(\mathbf{s}))}{P(\mathbf{N}(\mathbf{s}))} \quad (3.18)$$

La probabilidad $P(\mathbf{N}(\mathbf{s})|\mathbf{q}(\mathbf{s}))$ sigue una distribución multinomial, por lo que reemplazando esta expresión y la de α_{12} obtenemos que

$$\langle \alpha_{12}(\mathbf{N}(\mathbf{s})) \rangle = \int d\mathbf{q} \frac{P(\mathbf{q})}{P(\mathbf{N}(\mathbf{s}))} \ln \left(\frac{q_{11} \cdot q_{00}}{q_{10} \cdot q_{01}} \right) \prod q_i^{N_i} \quad (3.19)$$

Es necesario detenerse aquí para considerar las condiciones del sistema que queremos analizar y elegir una distribución prior $P(\mathbf{q})$ adecuada. En general, hay un tiempo limitado para medir la actividad de una neurona, por lo que no se suele medir su respuesta bajo muchas repeticiones. Tampoco es extraño que en algunos experimentos no se repita ningún estímulo, por lo que uno sólo tiene una medición para cada estímulo. Debido a eso, buscamos una distribución que limite el sesgo causada por el el muestreo finito. Esto significa, por ejemplo, no suponer probabilidad nula para estados no medidos. Una distribución de la misma familia que la distribución multinomial que evita esto es la distribución de Dirichlet:

$$P(\mathbf{q}) = \frac{\prod q_i^{\beta-1}}{B}, \quad \beta \geq 0 \quad (3.20)$$

donde B es un factor de normalización. Incluyendo esto, la probabilidad de α_{12} es

$$\langle \alpha_{12}(\mathbf{N}(\mathbf{s})) \rangle \propto \int \mathbf{d}\mathbf{q} \ln \left(\frac{q_{11} \cdot q_{00}}{q_{10} \cdot q_{01}} \right) \left(\prod q_i^{N_i + \beta - 1} \right) \quad (3.21)$$

Integrando esto sobre el espacio de probabilidad, el estimador queda de la forma[19]

$$\langle \alpha_{12}(\mathbf{N}(\mathbf{s})) \rangle = \psi(N_{11} + \beta) + \psi(N_{00} + \beta) - \psi(N_{10} + \beta) - \psi(N_{01} + \beta) \quad (3.22)$$

La función ψ , llamada función digamma, es una función especial obtenida mediante la derivación del logaritmo de la función gamma Γ . Esta función tiende a la función $-x^{-1}$ cuando el argumento es muy pequeño y tiende a $\ln(x)$ para valores muy grandes (ver Fig. 3.3). Uno puede observar entonces que para casos de muchas mediciones y con valores en \mathbf{n} grandes el valor de $\langle \alpha_{12}(\mathbf{q}) \rangle$ tiende a ser el mismo que el de máxima verosimilitud.

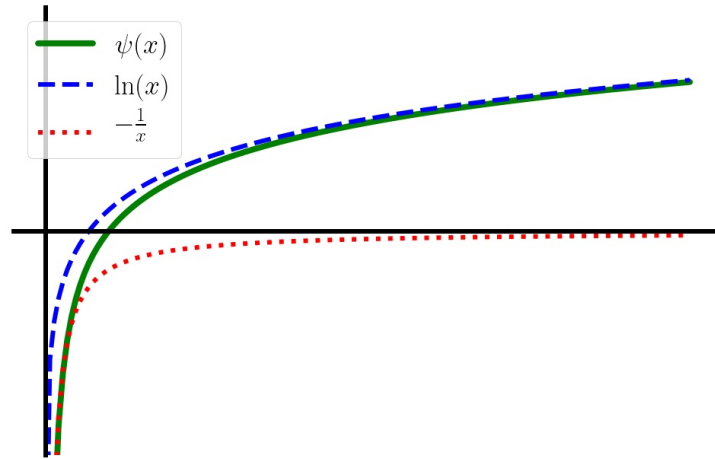


Figura 3.3: Esquema del comportamiento de la función ψ .

Repitiendo el cálculo anterior para α_1 y α_2 tenemos

$$\begin{cases} \langle \alpha_1(\mathbf{s}) \rangle = \psi(n_{10} + \beta) - \psi(n_{00} + \beta) \\ \langle \alpha_2(\mathbf{s}) \rangle = \psi(n_{01} + \beta) - \psi(n_{00} + \beta) \end{cases} \quad (3.23)$$

La función de β en el estimador puede ser interpretada como la de pseudo-cuentas que se agregan a las cuentas medidas para cada estado. Mientras más alto sea el valor de β que se elige, las cuentas \mathbf{n} tienen menos relevancia y \mathbf{q} tiende a una distribución plana. En cambio, valores muy pequeños de β permiten más variabilidad en los resultados y pueden llegar a propagar las variaciones debidas al muestreo finito. Según [20], un valor óptimo para el estimador que genera un bajo

sesgo en la entropía de $\mathbf{q}(\mathbf{s})$ es el inverso del número de estados posible ($\beta = 1/4$ para dos neuronas).

3.4. Rendimiento del estimador bayesiano

Para determinar si el estimador es útil para extraer la información específica de la interacción entre neuronas, lo probamos con datos simulados a partir de nuestro modelo (Ec. 2.2).

Los filtros usados para simular los resultados fueron los que se observan en la Figura 3.4. En estos datos simulados, las funciones ζ sólo dividen su argumento (la proyección del estímulo sobre el campo receptivo) por la norma del filtro y el valor de β usado es de $1/4$.

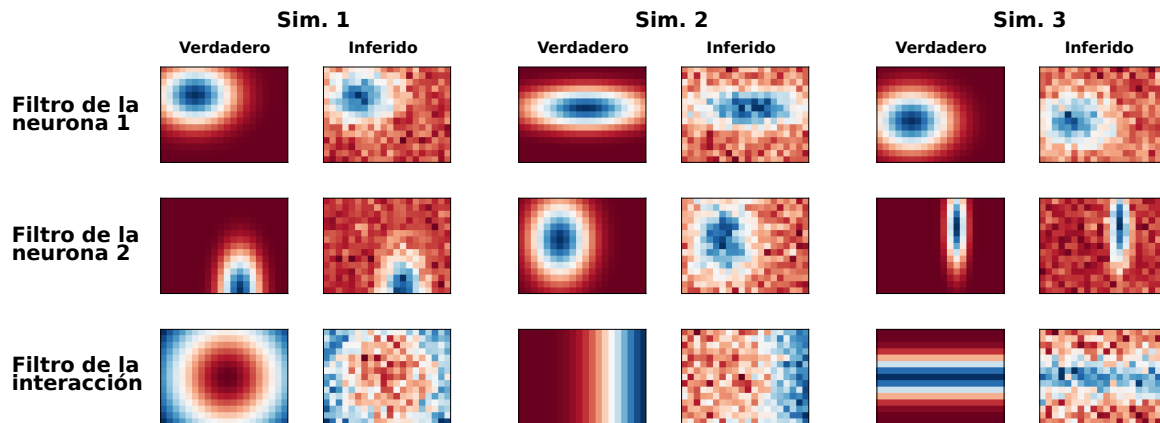


Figura 3.4: Comparación de los filtros verdaderos usados en la simulación con los filtros inferidos por el estimador, usando 5 mediciones de 5000 estímulos distintos.

En la Figura 3.4 se puede notar que el estimador infiere tanto los filtros individuales como los de interacción dado un número bajo de repeticiones. Por ende, de aquí podemos constatar de que este estimador es apropiado para caracterizar las correlaciones en las respuestas de un par de neuronas. Cuando simulamos un sistema con filtro de interacción nulo, inferimos ruido al calcular \mathbf{h}_{12} aún si los filtros individuales se superponen. Esto demuestra que en las mediciones podemos diferenciar las neuronas que están correlacionadas por tener campo receptivo en común de las que lo están por tener un filtro de interacción.

Para comparar los filtros estimados con los usados en la simulación, definimos una distancia como la distancia euclídea entre los filtros normalizados. En el gráfico de la izquierda, se observa cómo el rendimiento de nuestro modelo aumenta con el número de estímulos distintos. En comparación con el modelo marginal, su rendimiento es por que él cuando no se usan repeticiones pero lo sobrepasa con usando 5 repeticiones por estímulo. En el gráfico de la derecha en la Figura 3.5 se puede

observar cómo disminuye la distancia entre el filtro individual de la simulación 1 (Fig. 3.4) y los filtros inferidos mientras se aumenta la cantidad de repeticiones por estímulo fijando el número de estímulos distintos.

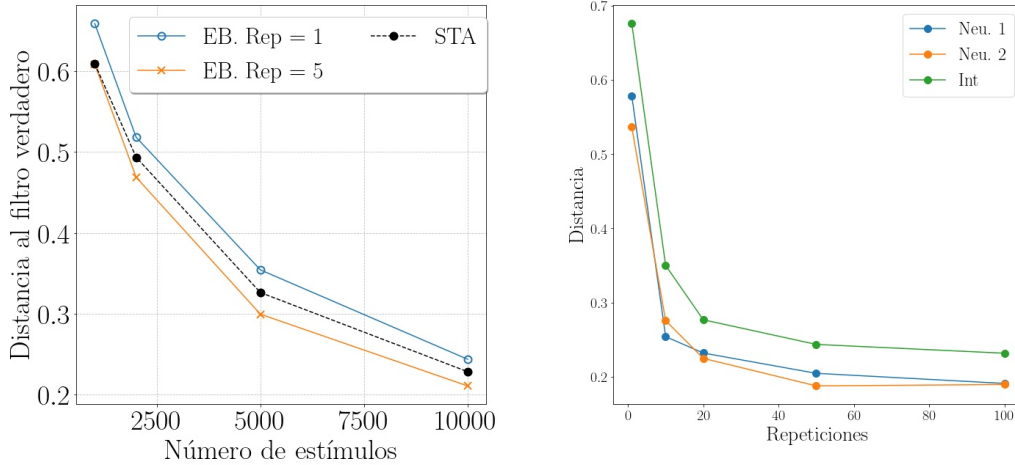


Figura 3.5: A la izquierda, se muestra la distancia entre el filtro original y los inferidos por cada modelo con los datos simulados de una neurona única. La figura de la derecha muestra la distancia entre filtros originales y los filtros con los datos simulados del primer caso de la Figura 3.4.

3.5. Comparación de estimadores

Como vimos en el capítulo anterior, existe una relación entre las funciones usadas para el modelo marginal y el conjunto. Uno puede esperar entonces que haya algún tipo de relación entre los filtros obtenidos por cada estimador.

En el caso en que sólo se presenta una vez cada estímulo a las neuronas ($N(\mathbf{s}) = 1$), uno puede reescribir las ecuaciones 3.22 y 3.23 usando la siguiente propiedad de la función digamma:

$$\psi(z + 1) - \psi(z) = \frac{1}{z}.$$

Usando esto, uno puede escribir el valor estimado de las funciones α como

$$\begin{cases} \langle \alpha_1(\mathbf{s}) \rangle &= \beta^{-1}(2\sigma_1(\mathbf{s}) - 1)(1 - \sigma_2(\mathbf{s})) \\ \langle \alpha_2(\mathbf{s}) \rangle &= \beta^{-1}(2\sigma_2(\mathbf{s}) - 1)(1 - \sigma_1(\mathbf{s})) \\ \langle \alpha_{12}(\mathbf{s}) \rangle &= \beta^{-1}(4\sigma_1(\mathbf{s})\sigma_2(\mathbf{s}) - 2(\sigma_1(\mathbf{s}) + \sigma_2(\mathbf{s})) - 1) \end{cases} \quad (3.24)$$

Observando los casos posibles, los posibles valores de estas funciones son limi-

tados:

$$\langle \alpha_1(\mathbf{s}) \rangle = \begin{cases} \beta^{-1}, & \text{si } \sigma_1 = 1 \text{ y } \sigma_2 = 0 \\ 0, & \text{si } \sigma_2 = 1 \\ -\beta^{-1}, & \text{si } \sigma_1 = 0 \text{ y } \sigma_2 = 0 \end{cases} \quad (3.25)$$

$$\langle \alpha_{12}(\mathbf{s}) \rangle = \begin{cases} \beta^{-1}, & \text{si } (\sigma_1, \sigma_2) = (1, 1) \text{ o } (0, 0) \\ -\beta^{-1}, & \text{si no} \end{cases} \quad (3.26)$$

A fin de facilitar las cuentas, trabajaremos con una transformación de las variables de la forma

$$\begin{aligned} \tilde{\mathbf{k}}_i &= N_i \mathbf{k}_i = \sum_{\mathbf{s}} \sigma_i(\mathbf{s}) \mathbf{s}, \quad i = 1, 2, 12 \\ \tilde{\mathbf{h}}_j &= N\beta \mathbf{h}_j = \beta \sum \alpha_j(\mathbf{s}) \mathbf{s}, \quad j = 1, 2, 12 \end{aligned}$$

donde \mathbf{k}_{12} se refiere al STA realizado sobre el disparo sincrónico de ambas neuronas.

Usando las ecuaciones en 3.24 uno obtiene las siguientes relaciones entre ambos grupos de filtros:

$$\begin{cases} \tilde{\mathbf{k}}_1 &= \tilde{\mathbf{h}}_1 + \frac{1}{2}\tilde{\mathbf{h}}_{12} \\ \tilde{\mathbf{k}}_2 &= \tilde{\mathbf{h}}_2 + \frac{1}{2}\tilde{\mathbf{h}}_{12} \\ \tilde{\mathbf{k}}_{12} &= \frac{1}{2}(\tilde{\mathbf{h}}_1 + \tilde{\mathbf{h}}_2) + \frac{3}{4}\tilde{\mathbf{h}}_{12} \end{cases} \quad (3.27)$$

$$\begin{cases} \tilde{\mathbf{h}}_1 &= \tilde{\mathbf{k}}_1 - (2\tilde{\mathbf{k}}_{12} - \tilde{\mathbf{k}}_1 - \tilde{\mathbf{k}}_2) \\ \tilde{\mathbf{h}}_2 &= \tilde{\mathbf{k}}_2 - (2\tilde{\mathbf{k}}_{12} - \tilde{\mathbf{k}}_1 - \tilde{\mathbf{k}}_2) \\ \tilde{\mathbf{h}}_{12} &= 2(2\tilde{\mathbf{k}}_{12} - \tilde{\mathbf{k}}_1 - \tilde{\mathbf{k}}_2) \end{cases} \quad (3.28)$$

Cabe observar que si las dos neuronas actúan independientemente, el filtro de interacción \mathbf{h}_{12} tiende al vector nulo. En este caso, los filtros individuales coinciden y el filtro STA del disparo sincrónico \mathbf{k}_{12} es igual al promedio de los filtros individuales.

Capítulo 4

Resultados y discusión

En los capítulos anteriores propusimos un modelo de disparo de dos neuronas con la correlación entre ellas modulada por el estímulo. También propusimos un estimador para los campos receptivos de cada contribución que evita las divergencias producidas por muestreos pequeños y que probamos con datos simulados. En este capítulo aplicaremos el estimador en datos experimentales para inferir los parámetros del modelo en el caso de neuronas reales.

Los datos elegidos son mediciones hechas en células ganglionares de la retina. Elegimos este tipo de células debido a que hay numerosos trabajos que prueban que la inclusión de las correlaciones entre pares de neuronas en los modelos de disparo mejora apreciablemente su rendimiento[9, 10]. En [21] se observa que al calcular los campos receptivos del disparo sincrónicos de dos neuronas, el resultado difiere de aquel esperado de dos neuronas independientes. Por estas razones es interesante observar si estas correlaciones codifican regiones del estímulo.

4.1. Datos usados

Se usaron las mediciones realizadas en la investigación de [13] sobre la diversidad funcional de las células ganglionares de la retina. El trabajo es una categorización no supervisada sobre más de 11000 células extraídas de retinas de ratón. Esta categorización es de índole funcional, por lo que se realiza observando las diferencias en las respuestas a distintos tipos de estímulos. Dada la naturaleza del trabajo, éste se focaliza en analizar la actividad *marginal* de las células ganglionares.

Lo que nos interesa a nosotros es estudiar la codificación *conjunta*. Dado que uno de los estímulos sigue una distribución de probabilidad esférica, esto nos permite usar los métodos de estimación propuestos en el capítulo anterior.

Las mediciones se realizan con microscopía de dos fotones para detectar el nivel Ca^{2+} , el cual sirve de indicador de la excitación de la neurona. La retina

se separa del ojo y se coloca sobre una placa transparente con una solución para alargar su vida útil. Mientras un proyector estimula la retina, el microscopio de dos fotones mide el nivel de Ca^{2+} que liberan las células ganglionares como respuesta. El instrumento registra la señal de Ca^{2+} en forma de su variación respecto a la media ($\Delta F/F$). En los datos medidos se incorpora un marcador de los tiempos del estímulos para poder alinearlos con una precisión de 2ms.

4.2. Procesamiento

La señal de Ca^{2+} atraviesa un filtrado paso alto sobre 0,1 Hz para eliminar tendencias globales a lo largo de la medición (*detrend*). Para determinar los disparos que realiza la neurona se analiza una interpolación de la señal de Ca^{2+} que tiene una frecuencia diez veces más grande que la señal original. A partir de ahí se determinan los disparos de la neurona donde la derivada de la señal realiza un pico. La altura mínima a la que identificamos un pico como correspondiente a un potencial de acción lo definimos como $\bar{\sigma} = \frac{\langle \dot{r}(t) \rangle}{0,6745}$, donde $\dot{r}(t)$ es la derivada de la señal de Ca^{2+} que se analiza, en línea con el preprocesamiento realizado por [13].

La retina es estimulada con un video de una matriz con dimensiones 15×20 con valores binarios de intensidad. Los valores de cada píxel cambian con probabilidad uniforme e independiente entre sí con una frecuencia de 5 Hz.

El campo receptivo de estas células es espacio-temporal, por lo que definimos como el estímulo a analizar a los cuadros mostrados en un intervalo de tiempo. Tomamos como estímulo los distintos segmentos de 1,7 s dentro del video entero, el mismo que el usado por el laboratorio que realizó las mediciones.

En el caso del STA, el campo se calcula de la forma

$$\mathbf{F}_{STA}(x, y, \tau) = \frac{1}{M} \sum_{i=1}^M \mathbf{s}(x, y, t_i + \tau), \quad (4.1)$$

donde $\mathbf{s}(x, y, \tau)$ es el estímulo, t_i es un retraso que varía desde -320 a 1380 ms y M es el número de picos en la señal.

En el caso de nuestro estimador de interacción el cálculo se realiza de manera similar:

$$\mathbf{F}_\alpha^{(k)}(x, y, \tau) = \frac{1}{T} \sum_{i=1}^T \alpha_k(\mathbf{s}(x, y, t_i)) \mathbf{s}(x, y, t_i + \tau), \quad (4.2)$$

donde T representa el número total de estímulos.

En ambos casos el campo conseguido se suaviza usando un filtro gaussiano para cada instancia de tiempo por separado. Si suponemos una independencia en-

tre las contribuciones temporales y espaciales, podemos extraer campos receptivos separados. En primer lugar, se calcula la desviación estándar de F a lo largo de las instancias de tiempo τ ($s.d.[F(x, y, \tau)]_\tau$). Sobre esto se ajusta una distribución gaussiana. A partir de esto, el campo receptivo temporal se calcula realizando un promedio temporal en $F(x, y, \tau)$ de los píxeles con valores con el 20% más alto en $s.d.[F(x, y, \tau)]_\tau$ y ponderados con los valores de dicha distribución gaussiana. Esta descomposición la usamos para poder analizar los filtros conseguidos, pero las proyecciones se calculan con el filtro F completo.

Para obtener la función que actúa sobre las proyecciones a los campos receptivos, ya sea ϕ para el STA o α para el estimador de interacción, se ajusta una función exponencial sobre la proyección de los estímulos y los datos correspondientes (disparos en el caso de STA, valor de α en el caso de interacción). Para reducir la variabilidad de los datos y lograr un mejor ajuste, hicimos grupos de 1000 puntos en los que realizamos un promedio en ambos ejes.

Los resultados de nuestro modelo se comparan con aquellos obtenidos del modelo LN. Hacemos esto no sólo para ayudar en la interpretación de los datos obtenidos, sino también para validar su capacidad de describir las células.

4.3. Resultados

Los datos medidos por [13], además de asignar a cada neurona un grupo funcional, incluyen un puntaje que mide la calidad del filtro encontrado. En esta sección presentamos los resultados de un conjunto de pares específicos, los cuales se eligieron entre las mediciones que poseen un mayor puntaje de calidad dado. En la Tabla 4.1 se muestran los tipos funcionales de neuronas que integran cada par según la categorización de [13]. Estas categorías apuntan al tipo de campo receptivo de la neurona: por ejemplo, una célula ON tiende a excitarse cuando una determinada región del estímulo aumenta y luego disminuye su intensidad por un cierto periodo de tiempo.

Tabla 4.1: Tipos de neurona en cada par estudiado

	Neurona 1	Neurona 2
1	OFF	OFF
2	ON	ON
3	ON-OFF	ON-OFF
4	ON-OFF	OFF

En la Figura 4.1 se muestran los campos receptivos espaciales para el primer par de neuronas. Los filtros espaciales individuales conseguidos por nuestro método poseen una forma muy similar a los de LN y sus puntos máximos se encuentran

en la misma posición en la matriz. En el modelo de interacción, los filtros tienen una distribución que es más ancha en la dirección al campo receptivo de la otra neurona. Por otra parte, el campo de la interacción tiene el mismo tipo de perfil que los campos individuales y su centro se encuentra en un punto medio de éstos.

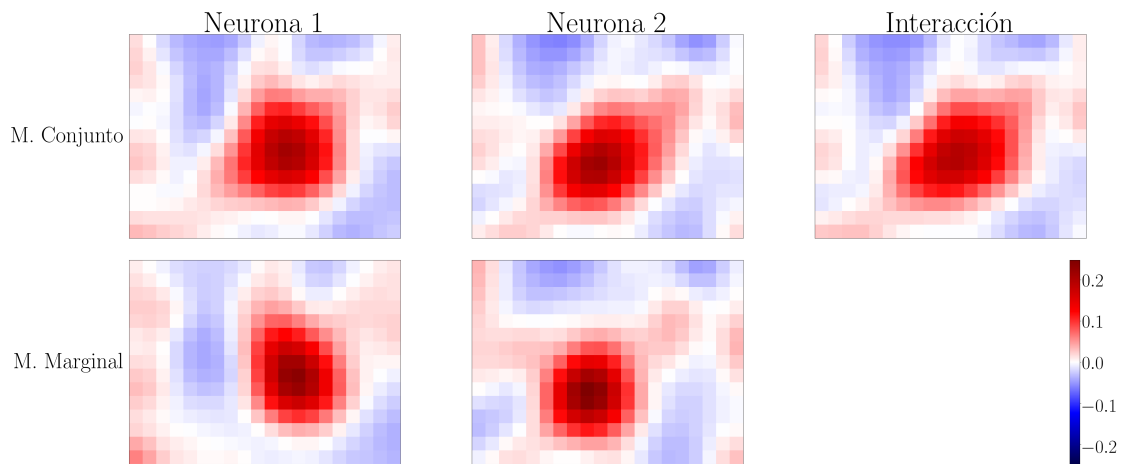


Figura 4.1: Filtros espaciales obtenidos con el modelo de interacción (conjunto) y el modelo LN (marginal) del primer par de neuronas (células OFF).

Los filtros temporales de nuestro modelo también son parecidos a aquellos del modelo LN (Fig. 4.2). Si bien los filtros individuales no son iguales entre cada método, poseen la misma forma funcional. De forma similar al caso espacial, el filtro temporal de la interacción se encuentra aproximadamente entre los dos filtros individuales.

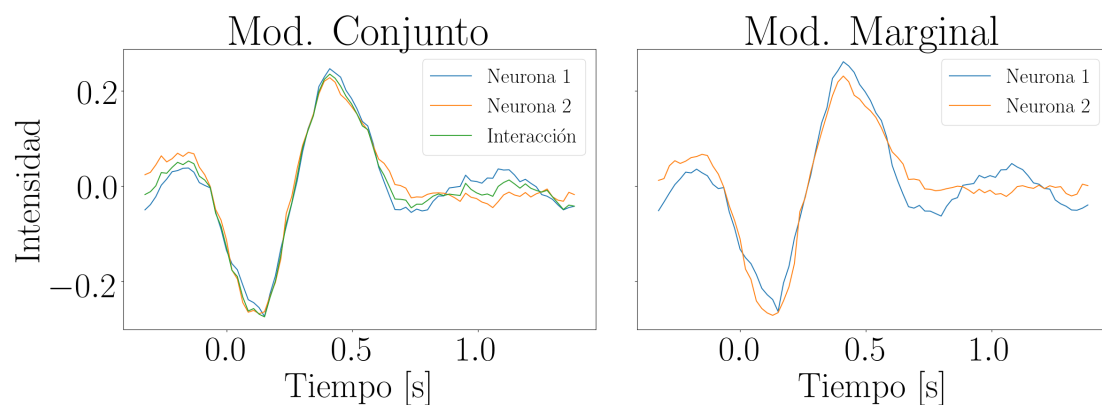


Figura 4.2: Filtros temporales obtenidos con el modelo de interacción y el modelo LN del primer par de neuronas (células OFF).

El hecho de que los perfiles de ambos modelos tengan una misma forma funcional a pesar de sus diferencias indica que los filtros individuales de nuestro método puede servir para categorizar funcionalmente a las neuronas.

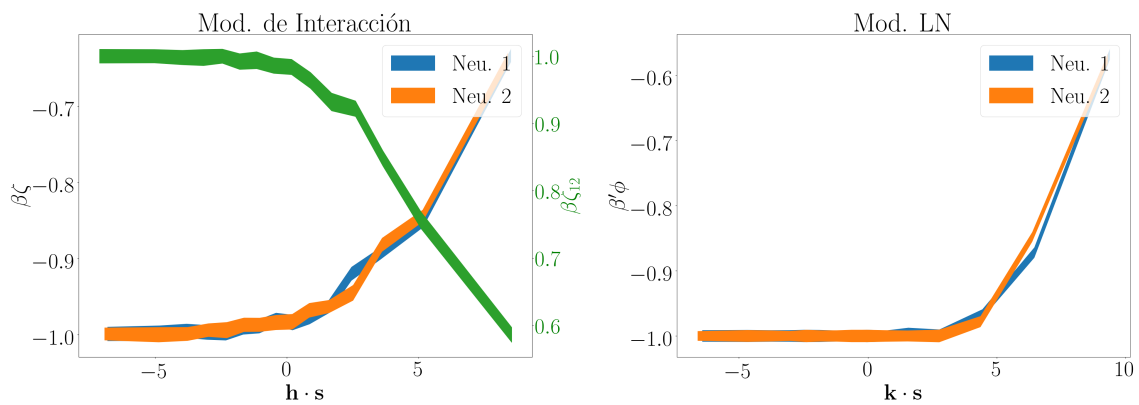


Figura 4.3: Funciones no lineales de dos células OFF obtenidos usando cada modelo. En la Figura de la izquierda hay un segundo eje para incluir la función de la interacción.

En la Figura 4.3 se observan las funciones no lineales obtenidas de cada modelo. Para obtener las funciones no lineales del modelo marginal se usó nuestro estimador bayesiano para el caso de dos estados. Se usó como valor de β uno tal que reproduzca la frecuencia de disparo total de la medición.

Se puede observar que la función asociada a la interacción tiene una tendencia contraria a la de las contribuciones individuales. Como su filtro es cercano a los individuales, esto significa que al ser estimulada esta región, las contribuciones individuales aumentan la probabilidad de que dispare cada uno, pero la contribución de la interacción disminuye esta probabilidad. Esto se interpreta como que el rol del filtro de interacción es disminuir la correlación cuando se recibe un estímulo en esta zona intermedia.

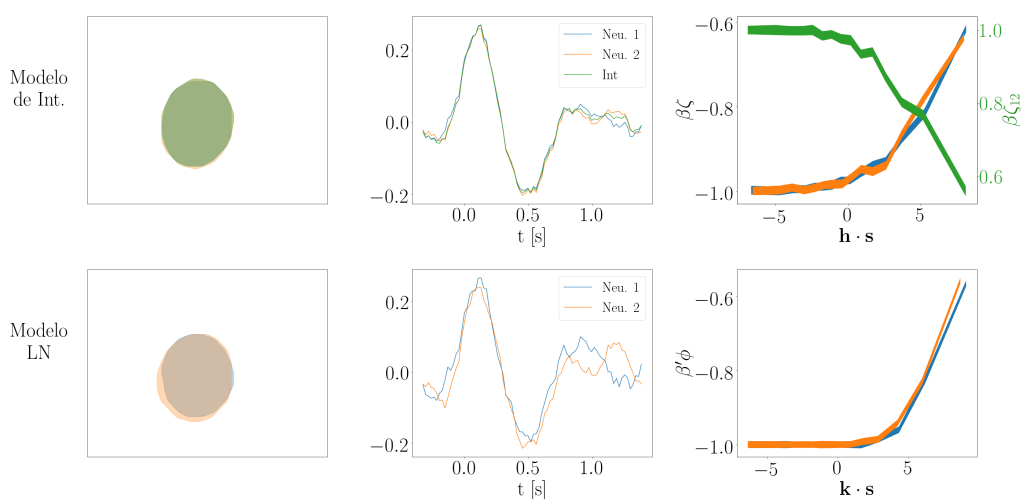


Figura 4.4: Resultados para un par de células ON. A la izquierda se muestran las regiones del estímulo donde los filtros son mayores al 60% del valor máximo. En la segunda columna se encuentran los filtros temporales y a la derecha mostramos las funciones marginales ϕ y las funciones ζ de nuestro modelo.

En las Figuras 4.4 y 4.5 se muestra la comparación entre los modelos para un par de células ON y un par de células ON-OFF, respectivamente. Se puede ver que tanto las contribuciones individuales como de interacción comparten el mismo filtro. Aún así, la función no lineal de la interacción es tal que disminuye el disparo simultáneo de las neuronas.

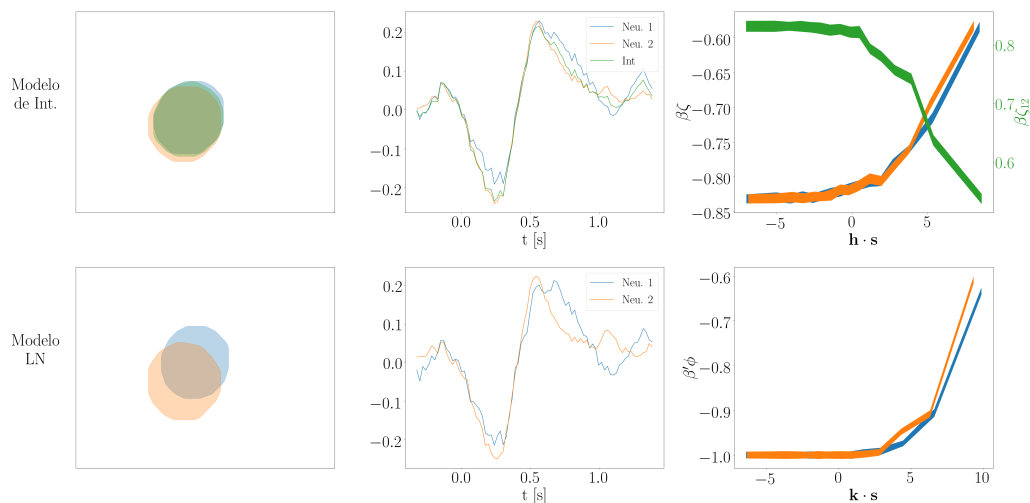


Figura 4.5: Resultados para un par de células ON-OFF.

En la Figura 4.6 se muestra la comparación los parámetros anteriores en un sistema con una célula ON-OFF y una célula OFF. El filtro de interacción espacial no se limita a la intersección entre los campos sino que ocupa la mayor parte de la unión de ellos. Se eligió el signo del filtro temporal de interacción para que tenga una forma similar a los filtros individuales. Al igual que en los casos anteriores, la función no lineal de la interacción es tal que disminuye el disparo simultáneo de las neuronas.

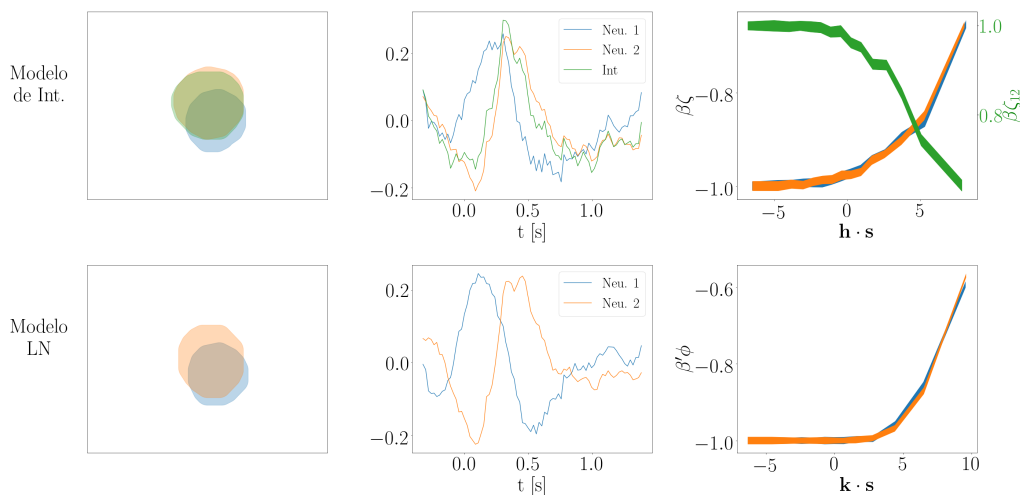


Figura 4.6: Resultados para dos neuronas de tipo ON-OFF (Neurona 1) y OFF (Neurona 2).

4.3.1. Consistencia con el modelo LN

Como argumentamos previamente, nuestro modelo para la actividad conjunta tiene que ser capaz de reproducir la actividad marginal. Esta correspondencia se ve reflejada en la relación entre los filtros mostrada en la sección 3.5. En la Figura 4.7 comparamos los filtros conseguidos del estimador STA con su combinación equivalente de los filtros de nuestro modelo de interacción dada por la Ec. 3.28. Se puede observar que los filtros inferidos por nuestro método reproducen exactamente tanto las características espaciales como el perfil temporal del filtro de STA.

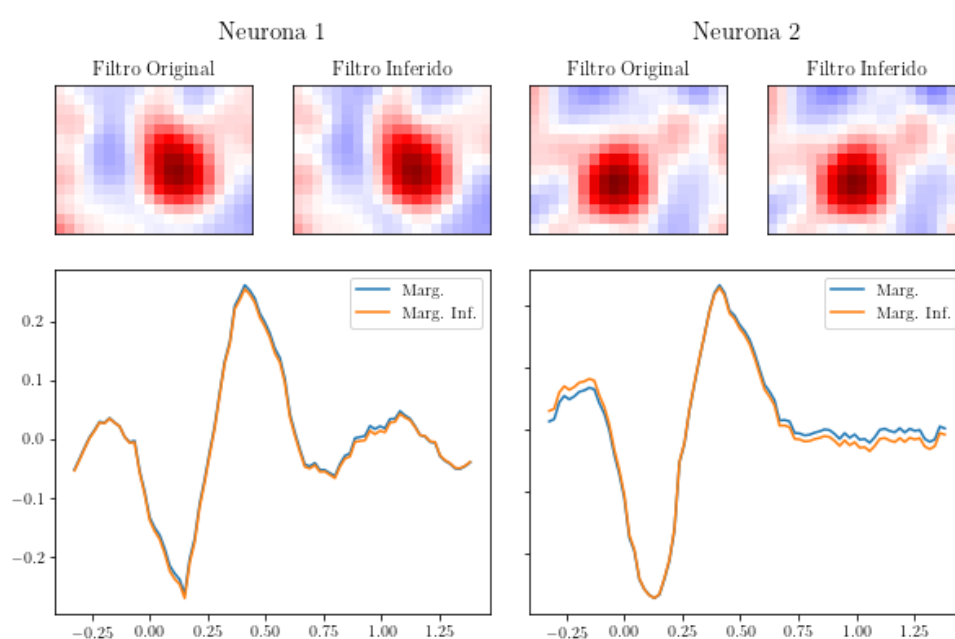


Figura 4.7: Equivalencia en un par de células OFF entre los filtros marginales y los inferidos mediante la transformación $\tilde{\mathbf{k}}_i = \hat{\mathbf{h}}_i + \frac{1}{2}\tilde{\mathbf{h}}_{12}$. A la izquierda se muestran los resultados de la neurona 1 y a la derecha, los resultados de la neurona dos. En la parte superior, el filtro de la izquierda muestra el filtro espacial conseguido mediante STA y el de la derecha, el inferido a partir de los filtros del modelo conjunto mediante la transformación. En la parte inferior, se muestran los filtros temporales superpuestos.

Para las funciones no lineales ϕ y α que transforman las proyecciones sobre estos filtros, se observó que los valores elegidos para β afectaban la tasa de disparo promedio estimada. Elegimos valores distintos a los sugeridos por [20] para los estimadores marginal y conjunto de manera de acercarse a la tasa de disparo medida.

La comparación entre la función ϕ proveniente del modelo LN y la reconstrucción a partir de las funciones ζ propuesta en 2.7 puede verse en la Figura 4.8. Es evidente que se consigue una gran coincidencia al igual que el caso de los campos receptivos.

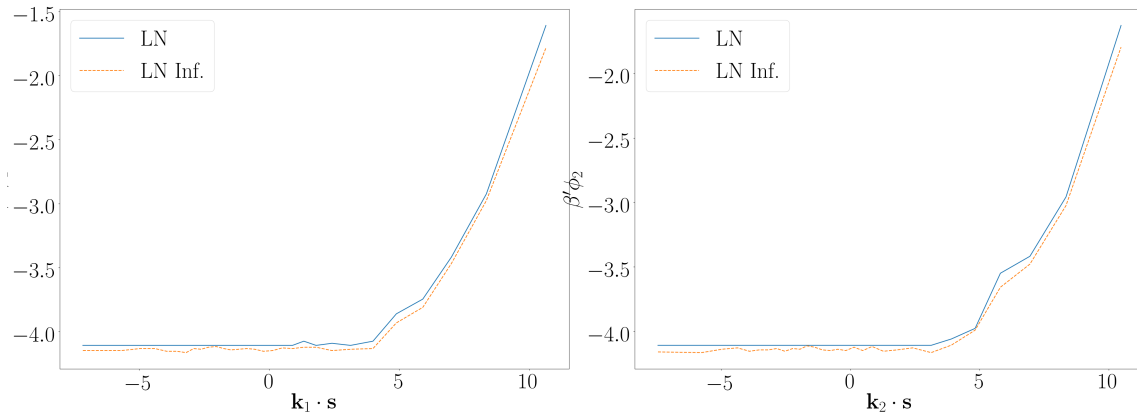


Figura 4.8: Funciones no lineales asociadas a cada modelo.

Los resultados de las Figuras 4.7 y 4.8 fueron equivalentes en los otros pares de neuronas, por lo que confirmamos para los casos estudiados la compatibilidad predicha entre ambos modelos.

4.3.2. Validación

En la sección anterior vimos que los resultados que arroja nuestro método sobre un par de neuronas puede usarse para reproducir los producidos por STA. Como nuestro modelo necesita ajustar más parámetros con los mismos datos, es esperable que sea capaz de describir mejor la actividad neuronal.

Para validar un modelo de este tipo se usan dos conjuntos de datos, uno de entrenamiento y uno de prueba en el que ambos poseen los estímulos y las respuestas de las neuronas a éstos. Los parámetros θ del modelo (los filtros y las funciones ζ) se ajustan con las respuestas a los estímulos de entrenamiento y, a partir de ahí, se intenta reproducir los resultados del grupo de prueba, aplicando el modelo con los parámetros conseguidos a los estímulos de prueba.

Una manera directa de hacer esto es presentar a las neuronas los estímulos de ambos grupos varias veces. En nuestro caso, al no tener mediciones repetidas usaremos un intervalo del estímulo total del sistema y lo usaremos como grupo de prueba. Este método se llama validación cruzada (*cross validation*).

Para cuantificar el error que tiene el modelo prediciendo la actividad en el grupo de prueba, calculamos el valor de entropía cruzada (*cross entropy*) de las predicciones. Esta es una medida en bits que indica la separación entre una distribución de probabilidad verdadera p y una distribución inferida q que tiene el mismo soporte \mathcal{X} :

$$CE = - \sum_{x \in \mathcal{X}} p(x) \log q(x) \quad (4.3)$$

Mientras más alejadas se encuentren dos distribuciones, el valor de CE aumentará. Estimamos esta cantidad calculando para cada estímulo \mathbf{s}_i la probabilidad que había inferido el modelo para la respuesta verdadera σ_i y lo promediamos para todos los estímulos:

$$CE = -\frac{1}{N} \sum_{i=1}^N \log q(\sigma_i | \theta, \mathbf{s}_i). \quad (4.4)$$

En este análisis hacemos cinco mediciones usando grupos de prueba distintos. Antes de realizar el ajuste del modelo a cada grupo de entrenamiento, es necesario asegurar las condiciones estadísticas por separado. Para realizar una validación no sólo de nuestro modelo sino también del modelo marginal, comparamos ambos con un modelo más básico: el de estimar la probabilidad de uno de los cuatro estados con una probabilidad λ uniforme en el tiempo. Este modelo no tiene en cuenta el tiempo de los disparos ni el estímulo, por lo que es esperable que los otros dos lo superen a la hora de predecir actividad nueva. Al compararlo, fue superado por ambos modelos dependientes del estímulo ($p < 0,01$).

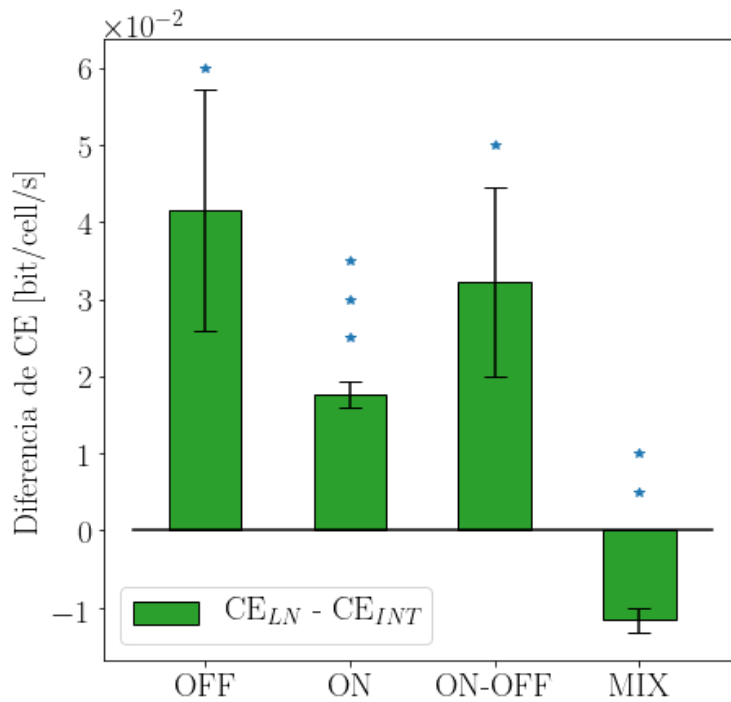


Figura 4.9: Diferencia de la entropías cruzadas del modelo marginal y el de interacción. La presencia de una estrella (*) sobre una medición quiere decir que tal medición es estadísticamente significativa con $p < 0,05$. De manera similar, dos estrellas (**) significan que $p < 0,01$ y tres estrellas (***) significan que $p < 0,001$.

Como el número de disparos de las neuronas es escaso y distribuido irregularmente, los valores de CE varían de forma apreciable para distintos grupos de prueba y no es adecuado promediar sus valores. Debido a esto, para comparar los rendimientos entre los dos modelos analizamos la diferencia que tienen entre ellos. En la Figura 4.9 podemos ver la diferencia de entropía cruzada que hay entre ambos modelos dependientes del estímulo. La variación que hay entre nuestro modelo y el modelo LN no es grande, por lo que realizamos una prueba t de Student con la hipótesis nula de que el rendimiento de nuestro modelo es igual o peor que el del modelo LN. Realizando una prueba de t de Student corroboramos que estas diferencias son estadísticamente significativas. En los tres casos podemos rechazar la hipótesis nula dado que obtenemos un p-valor menor a 0,04. También se observa que el rendimiento es mejor en el caso del par de neuronas OFF, lo cual va de la mano con la observación de [21] de que este tipo de células presentaba mayor comportamiento correlacionado.

En el caso del par de células de distinto tipo vemos que el rendimiento del modelo de interacción es menor al del modelo LN. Esto puede deberse a que las células no posean una correlación grande entre ellas por lo que la descripción marginal de ellas es suficiente.

4.4. Discusión y propuestas a futuro

En todos los casos estudiados se pudo obtener un campo receptivo para la correlación entre neuronas. En los casos observados el filtro es cercano a ambos filtros individuales y, junto con ellos, consistente con los filtros marginales. En todos los casos $\alpha_{12}(\mathbf{h} \cdot \mathbf{s})$ es una función decreciente en función de las proyecciones de su filtro, al contrario que las funciones α individuales. Esto indica que la estimulación del estímulo en la región relevante para la interacción disminuye la probabilidad de un disparo conjunto, es decir, disminuyen su correlación a lo largo de esa dirección. Como en las simulaciones del Cap. 3 se observó que el estimador distingue la presencia de un filtro de interacción aún cuando los filtros individuales se superponen entre sí, podemos suponer que esta correlación no es producto de la intersección de sus campos receptivos sino que existe un mecanismo extra que los modula. En [21] se descartó la posibilidad de que las uniones gap entre las células ganglionares sea la causa de estas correlaciones. Esto, junto con el hecho de que la correlación no está explicada por un estímulo compartido de una célula bipolar, sugiere que la correlación se produce por las células amacrinas que sean intermediarias en estos pares. Estas células se extienden lateralmente conectando células ganglionares y su rol es inhibitorio, tal como el tipo de interacciones que observamos.

Para mejorar el rendimiento que tiene nuestro modelo en relación al modelo marginal proponemos dos estrategias posibles:

1. el valor de β tiene un efecto grande sobre la estimación de ζ cuando tenemos una única medición por estímulo. Esto es porque elegir un valor de β constante para los cuatro estados es suponer una equivalencia entre los cuatro estados. Una mejora sería introducir valores β_i distintos para los cuatro casos a fin de mejorar el rendimiento. Esta elección no corresponde a ninguna distribución de probabilidad a priori, por lo que sólo se muestra como una aproximación a los valores que se deberían usar en los cálculos. Otra posible estrategia para validar el modelo es usar redes neuronales para acoplar al estimador analítico. Entrenando una red densa con las proyecciones de los estímulos con los filtros obtenidos por el estimador se podría conseguir una estimación de ζ menos sesgada. Si uno construye la red de forma de tener una última capa de cuatro unidades que representen las probabilidades y sus valores se calculen usando el mismo algoritmo de nuestro modelo de exponenciar sus pesos y normalizarlo (esto se llama activación *softmax* en el marco de redes neuronales), uno puede alterar la capa anterior y sus conexiones de tal manera de que uno consiga los valores de α acordes a cada modelo. En la Figura 4.10 se muestra cómo serían las arquitecturas propuestas para cada modelo. El número de unidades de la primera capa debería decidirse tomando en cuenta el tamaño y complejidad de los datos con los que se entrena. El método de cálculo de error sería usando el algoritmo de *categorical cross entropy*, que es el mismo que usado en este trabajo para el caso analítico.

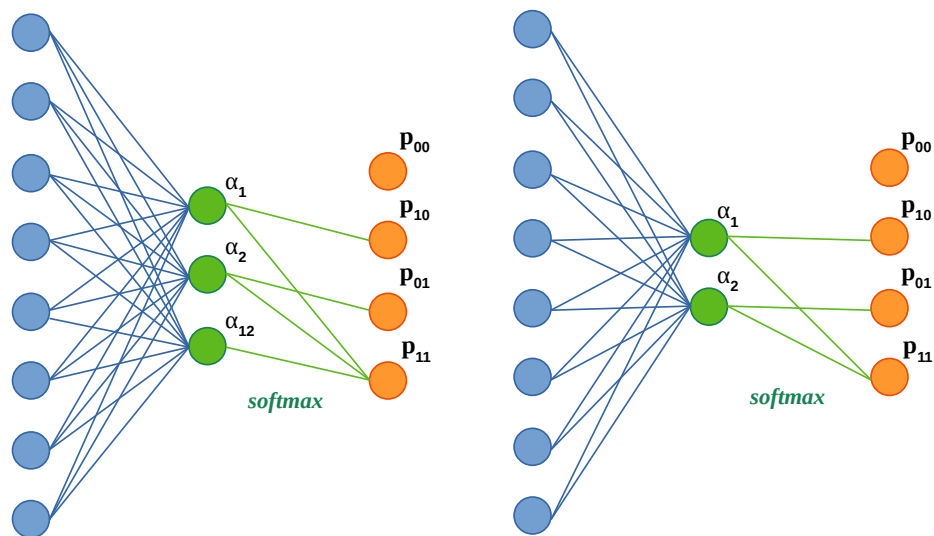


Figura 4.10: Propuesta de redes neuronales para acoplar al estimador analítico. La red de la izquierda está diseñada para calcular la probabilidad de los disparos usando el modelo de interacción, mientras que el de la derecha usa el modelo marginal.

2. Los pares de neuronas tienen distintos grados de correlación entre sí. Si bien se eligieron neuronas que posean un filtro de interacción definido, la correlación entre ellas puede ser lo suficientemente débil para que no haya una diferencia grande entre los rendimientos de ambos modelos. Con esta idea en mente, uno debería aplicar el modelo a una población entera de neuronas y observar si la validación en este caso es más significativa.

Las propuestas mencionadas anteriormente fueron para mejorar nuestra capacidad de estimar los parámetros del modelo en datos. En lo relativo a las posibles aplicaciones de nuestro método, gracias a su flexibilidad podemos expandirlo y modificarlo para analizar distintos aspectos de la codificación conjunta de las neuronas. En primer lugar, el modelo puede expandirse para tener en cuenta correlaciones entre tres neuronas o más. Esto nos permitiría contestar cómo interactúan grupos de tres neuronas y cómo sus correlaciones codifican información del estímulo. Por otro lado, en nuestro trabajo estudiamos siempre el caso de una dirección relevante para cada contribución al sistema. Esto lo hicimos ya sea en las formulaciones teóricas como en nuestra elección de usar los estimadores para encontrar filtros únicos en el subespacio relevante. Un camino a recorrer es el de buscar más direcciones relevantes usando el segundo momento $\mathbf{A}_{ij}^{(2)}$ (Eq. 3.10) para estudiar si se encuentra una modulación más compleja de la correlación y ver cómo se relaciona con los filtros del modelo LN.

Si bien nuestro modelo es completamente genérico, nuestro estimador está restringido a estímulos distribuidos esféricamente. Debido a esto, no puede ser usado en estímulos naturales. Una estrategia para superar esta limitación es el uso de redes neuronales. En resultados anteriores[22] mostramos que estas redes pueden inferir los coeficientes del modelo a partir de los estímulos. Lo que queda entonces es encontrar cómo extraer datos interpretables de la red entrenada.

Otra expansión posible para nuestro método es incluir correlaciones temporales entre las dos neuronas. En [8] se postuló un modelo de neuronas acopladas en los que la actividad de cada una puede influir en la probabilidad de disparo futura de la otra neurona o de sí misma. Este tratamiento permite no sólo mejorar los filtros obtenidos, sino que permite reproducir distintos patrones de disparo que tienen las neuronas. Es por eso que es atrayente la posibilidad de integrar estas correlaciones temporales en nuestro modelo para observar si hay características del estímulo que poseen efectos sobre las neuronas a una escala de tiempo más larga.

Para finalizar, nuestro modelo puede usarse para analizar preguntas específicas de la codificación en poblaciones. En la retina se han encontrado grupos de neuronas que reaccionan colectivamente al estímulo para codificar información no incluida en los campos receptivos de sus integrantes (modos colectivos)[23]. Nues-

tro método puede usarse para ver los filtros asociados a las correlaciones dentro de un clúster e intentar reconstruir la información del estímulo que codifica el conjunto a partir de la información que codifican las neuronas y sus correlaciones.

Capítulo 5

Conclusiones

En este trabajo apuntamos a resolver el interrogante de cómo la información se codifica en la actividad *conjunta* de las neuronas. Para ello, planteamos un modelo de probabilidad de máxima entropía para un sistema de dos neuronas a fin de estudiar cómo el estímulo puede afectar las correlaciones entre ellas. Usando un estimador bayesiano para pocas cuentas, aplicamos el modelo sobre pares de células ganglionares en retina de ratón y encontramos regiones relevantes del estímulo que modifican la correlación de estas neuronas. Verificamos también la relación esperada entre estas regiones y aquellas obtenidas mediante el modelo marginal de una neurona. No sólo mostramos que se puede extraer información del estímulo a partir de las correlaciones, si no también que dicha información es útil para predecir actividad neuronal en datos no vistos, con resultados positivos en tres de cuatro pares analizados de células ganglionares.

Como principal limitación, debemos notar que para la estimación –no así para el modelo– es necesaria una distribución esférica de estímulos de entrada. Si bien esto todavía permite cierta libertad en la clase de estímulos que uno considera, es una barrera a superar. Mediante pruebas preliminares y en resultados no mostrados acá, hemos comprobado que la estimación también podría realizarse por medio de redes neuronales que mapeen la relación entre la entrada y los coeficientes del modelo. En principio, esto permitiría incluir estímulos más naturalistas y relevantes biológicamente. Por otro lado, si bien el estimador bayesiano propuesto es capaz de extraer los filtros, todavía es necesario un estudio más exhaustivo del sesgo que éste introduce cuando hay pocas muestras.

En definitiva, nuestro enfoque provee un método directo para extraer la información del estímulo que se codifica en las correlaciones entre neuronas y, como mostramos para células ganglionares, abre una dimensión extra para interpretar datos de neuronas sensoriales. El método puede ampliarse para incorporar más neuronas y mayor cantidad de regiones relevantes en el estímulo. Debido a la ge-

neralidad de su construcción, este enfoque podría aplicarse no sólo en neuronas sensoriales sino también en neuronas motoras o de otro tipo, buscando cómo las correlaciones codifican la información, ya sea de entrada o de salida, asociada a la población estudiada.

Bibliografía

- [1] Hubel, D. H., Wiesel, T. N. Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology*, **148** (3), 574–591, 1959. [1](#), [6](#)
- [2] Gross, C. G., Rocha-Miranda, C. E., Bender, D. B. Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology*, **35** (1), 96–111, 1972. URL <https://doi.org/10.1152/jn.1972.35.1.96>, pMID: 4621506. [1](#)
- [3] O'Keefe, J., Dostrovsky, J. The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, **34** (1), 171 – 175, 1971. URL <http://www.sciencedirect.com/science/article/pii/0006899371903581>. [1](#), [6](#)
- [4] Dayan, P., Abbott, L. Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems. Computational Neuroscience Series. Massachusetts Institute of Technology Press, 2001. URL <https://books.google.com.ar/books?id=5GSKQgAACAAJ>. [1](#)
- [5] Squire, L. R., Dronkers, N., Baldo, J. Encyclopedia of neuroscience. Elsevier, 2009. [1](#), [4](#)
- [6] Schwartz, O., Simoncelli, E. P. Natural signal statistics and sensory gain control. *Nature neuroscience*, **4** (8), 819, 2001. [1](#)
- [7] Kandel, E. R., Schwartz, J. H., Jessell, T. M., of Biochemistry, D., Jessell, M. B. T., Siegelbaum, S., *et al.* Principles of neural science, tomo 4. McGraw-hill New York, 2000. [2](#)
- [8] Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E., *et al.* Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, **454** (7207), 995, 2008. [2](#), [32](#)
- [9] Schneidman, E., Berry II, M. J., Segev, R., Bialek, W. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*, **440** (7087), 1007, 2006. [2](#), [21](#)

- [10] Granot-Atedgi, E., Tkačik, G., Segev, R., Schneidman, E. Stimulus-dependent maximum entropy models of neural population codes. *PLoS computational biology*, **9** (3), e1002922, 2013. [2](#), [21](#)
- [11] McIntosh, L., Maheswaranathan, N., Nayebi, A., Ganguli, S., Baccus, S. Deep learning models of the retinal response to natural scenes. *En: Advances in neural information processing systems*, págs. 1369–1377. 2016. [2](#)
- [12] Shinomoto, S., Shima, K., Tanji, J. Differences in spiking patterns among cortical neurons. *Neural computation*, **15** (12), 2823–2842, 2003. [5](#)
- [13] Baden, T., Berens, P., Franke, K., Rosón, M. R., Bethge, M., Euler, T. The functional diversity of retinal ganglion cells in the mouse. *Nature*, **529** (7586), 345, 2016. [6](#), [21](#), [22](#), [23](#)
- [14] Samengo, I., Gollisch, T. Spike-triggered covariance: geometric proof, symmetry properties, and extension beyond gaussian stimuli. *Journal of computational neuroscience*, **34** (1), 137–161, 2013. [6](#), [11](#)
- [15] Gerstner, W., Kistler, W., Naud, R., Paninski, L. *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press, 2014. URL <https://books.google.com.ar/books?id=D4j2AwAAQBAJ>. [7](#)
- [16] Amari, S.-I. Information geometry on hierarchy of probability distributions. *IEEE transactions on information theory*, **47** (5), 1701–1711, 2001. [8](#)
- [17] Bialek, W., van Steveninck, R. R. Features and dimensions: Motion estimation in fly vision. *arXiv preprint q-bio/0505003*, 2005. [11](#), [15](#)
- [18] Cáceres, M. O. *Non-equilibrium statistical physics with application to disordered systems*. Springer, 2017. [14](#)
- [19] Wolpert, D. H., Wolf, D. R. Estimating functions of probability distributions from a finite set of samples. *Phys. Rev. E*, **52**, 6841–6854, Dec 1995. URL <https://link.aps.org/doi/10.1103/PhysRevE.52.6841>. [17](#)
- [20] Schürmann, T., Grassberger, P. Entropy estimation of symbol sequences. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, **6** (3), 414–427, 1996. [17](#), [27](#)
- [21] Meister, M., Lagnado, L., Baylor, D. A. Concerted signaling by retinal ganglion cells. *Science*, **270** (5239), 1207–1210, 1995. [21](#), [30](#)
- [22] Lorenz, M. Interacción funcional entre neuronas de la: descripción con estimadores Bayesianos y redes neuronales. Proyecto Fin de Carrera, Instituto Balseiro, 2019. [32](#)

-
- [23] Prentice, J. S., Marre, O., Ioffe, M. L., Loback, A. R., Tkačik, G., Berry II, M. J. Error-robust modes of the retinal population code. *PLoS computational biology*, **12** (11), e1005148, 2016. [32](#)

