



Universidad de San Carlos de Guatemala  
Escuela de Ciencias Físicas y Matemáticas  
Departamento de Física

## **DESARROLLO DE RED NEURONAL CONVOLUCIONAL PARA ANÁLISIS DE SEÑALES SÍSMICAS**

**Julio Antonio Medina Chután**

Asesorado por Lic. Robin Yani

Guatemala, septiembre de 2021



UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



ESCUELA DE CIENCIAS FÍSICAS Y MATEMÁTICAS

**DESARROLLO DE RED NEURONAL  
CONVOLUCIONAL PARA ANÁLISIS DE SEÑALES  
SÍSMICAS**

TRABAJO DE GRADUACIÓN  
PRESENTADO A LA JEFATURA DEL  
DEPARTAMENTO DE FÍSICA  
POR

**JULIO ANTONIO MEDINA CHUTÁN**  
ASESORADO POR LIC. ROBIN YANI

AL CONFERÍRSELE EL TÍTULO DE  
**LICENCIADO EN FÍSICA APLICADA**

GUATEMALA, SEPTIEMBRE DE 2021



UNIVERSIDAD DE SAN CARLOS DE GUATEMALA  
ESCUELA DE CIENCIAS FÍSICAS Y MATEMÁTICAS



## CONSEJO DIRECTIVO

DIRECTOR M.Sc. Edgar Aníbal Cifuentes Anléu  
SECRETARIO ACADÉMICO Ing. José Rodolfo Samayoa Dardón

## TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

EXAMINADOR Dr. Enrique Pazos  
EXAMINADOR Lic. Hector Perez  
EXAMINADOR Dr. Rodrigo Sacahui



# ÍNDICE GENERAL

<b>ÍNDICE DE FIGURAS</b>	<b>IX</b>
<b>ÍNDICE DE TABLAS</b>	<b>XII</b>
<b>OBJETIVOS</b>	<b>XIII</b>
<b>INTRODUCCIÓN</b>	<b>xv</b>
<b>1. Mecánica Estadística de las Redes Neuronales</b>	<b>1</b>
1.1. Estudios primordiales de redes neuronales en Física . . . . .	1
1.2. Principios generales de Mecánica Estadística . . . . .	2
1.3. Modelos de sistemas magnéticos con spin Ising . . . . .	6
1.4. Aproximación de campo medio . . . . .	7
1.5. Modelo de Ising con interacciones débiles de largo alcance . . . . .	11
1.6. Física del Estado del Vidrio de Spin . . . . .	14
1.6.1. Promedios apagados y recocidos . . . . .	14
1.7. El modelo SK . . . . .	16
1.7.1. Método de Réplicas . . . . .	18
1.7.2. Cálculo de la Energía Libre . . . . .	19
1.7.3. Solución simétrica de réplicas . . . . .	21
1.7.4. Ruptura de la simetría de réplicas de Parisi . . . . .	23
1.7.5. Ruptura de la simetría de replicas de escala completa . . . . .	27
1.8. Modelos Estadísticos de Redes Neuronales . . . . .	31
1.8.1. El modelo de Hopfield . . . . .	36
<b>2. Desarrollo de Red Neuronal</b>	<b>43</b>
2.1. Adquisición de Datos . . . . .	43
2.1.1. Datos sismológicos puros . . . . .	44
2.1.1.1. Formas de Onda . . . . .	44
2.1.1.2. Archivos de análisis sísmico . . . . .	46

2.1.2.	Método de Adquisición . . . . .	46
2.1.2.1.	Limpieza de datos . . . . .	47
2.1.2.2.	Re-muestreo . . . . .	47
2.1.2.3.	Normalización . . . . .	48
2.1.2.4.	Creación y recolección de etiquetas . . . . .	48
2.1.2.5.	Aumentación de datos . . . . .	51
2.2.	Diseño de arquitectura de red neuronal . . . . .	52
2.2.1.	Arquitecturas investigadas . . . . .	53
2.2.2.	Capas de procesamiento . . . . .	54
2.2.3.	Capa de salida . . . . .	55
2.2.4.	Función de error . . . . .	56
2.2.5.	Arquitectura de SeismicNet . . . . .	56
2.3.	Implementación de SeismicNet . . . . .	57
2.3.1.	Herramientas de desarrollo . . . . .	57
2.3.2.	Costo computacional . . . . .	59
2.3.3.	Entrenamiento . . . . .	60
<b>3. Resultados</b>		<b>61</b>
3.1.	Resultados tabulados . . . . .	62
3.1.1.	TestNet . . . . .	62
3.1.1.1.	Enero 2021 . . . . .	62
3.1.1.2.	Febrero 2021 . . . . .	63
3.1.2.	TestNet2 . . . . .	63
3.1.2.1.	Enero 2021 . . . . .	63
3.1.2.2.	Febrero 2021 . . . . .	64
3.1.3.	TestNet3 . . . . .	65
3.1.3.1.	Enero 2021 . . . . .	65
3.1.3.2.	Febrero 2021 . . . . .	65
3.1.4.	TestNet4 . . . . .	66
3.1.4.1.	Enero 2021 . . . . .	66
3.1.4.2.	Febrero 2021 . . . . .	67
3.1.5.	SeismicNet . . . . .	67
3.1.5.1.	Enero 2021 . . . . .	67
3.1.5.2.	Febrero 2021 . . . . .	68
3.2.	Comparaciones . . . . .	68
3.2.1.	Enero 2021 . . . . .	69

3.2.2. Febrero 2021 . . . . .	71
3.3. Predicciones de SeismicNet . . . . .	73
3.3.1. Sin Aciertos . . . . .	73
3.3.2. Un acierto . . . . .	77
3.3.3. Dos aciertos . . . . .	80
3.4. Discusión de resultados . . . . .	83
<b>CONCLUSIONES</b>	<b>85</b>
<b>RECOMENDACIONES</b>	<b>87</b>
<b>BIBLIOGRAFÍA</b>	<b>89</b>
<b>A. Apéndice</b>	<b>95</b>
A.1. Método de Réplicas . . . . .	95



# ÍNDICE DE FIGURAS

1.1.	Retículo de spins . . . . .	7
1.2.	Vecinos mas cercanos . . . . .	7
1.3.	Energía libre del imán ferromagnético de Ising. Se puede ver en color azul para $T > T_c$ y en naranja para $T < T_c$ . . . . .	11
1.4.	Clasificación de replicas en bloques para la ruptura de la simetría de replicas de dos pasos. . . . .	28
1.5.	Esquema de definición de los elementos matriciales $Q_{ab}$ para la ruptura de simetría de replicas de dos pasos. . . . .	28
2.1.	Ejemplo de archivo mseed, se pueden varias formas de onda, todas estas formas de onda sirven para caracterizar un evento sísmico. . . .	45
2.2.	Ejemplo de sismograma en estación STG2 . . . . .	45
2.3.	Ejemplo de un S-file . . . . .	46
2.4.	Ejemplo de forma de onda re-muestreada . . . . .	48
2.5.	Ejemplo de creación de etiquetas, en el gráfico superior se ve el sismograma y la picada de fase P se aprecia en color azul, la picada de fase S se ve en color rojo. En el segundo gráfico se ve la distribución de probabilidad asignada al arribo de fase P. En la tercera gráfica se ve la dist. de probabilidad para la fase S. Por ultimo en el gráfico inferior se ve la distribución de probabilidad para el ruido de la traza.	50
2.6.	Ejemplo de creación de etiquetas, en el gráfico superior se ve el sismograma y la picada de fase P se aprecia en color azul, la picada de fase S se ve en color rojo. En el segundo gráfico se ve la distribución de probabilidad asignada al arribo de fase P. En la tercera gráfica se ve la dist. de probabilidad para la fase S. Por ultimo en el gráfico inferior se ve la distribución de probabilidad para el ruido de la traza.	50

2.7. Ejemplos de sub-ajuste y sobre-ajuste. En la trazas superiores izquierda y derecha se puede ver el fenómeno del sub-ajuste, en este caso se quiere ajustar los puntos a la traza que se aprecia en color verde. En el grafico inferior izquierdo se aprecia un buen ajuste al gráfico verde. En la traza inferior derecha se puede ver claramente el sobre-ajuste.	52
2.8. Función ReLU . . . . .	55
2.9. Arquitectura de SeimicNet, se pueden ver la forma que toman los datos en cada capa de procesamiento. . . . .	57
 3.1. Gráfico de barras en el que se comparan los resultados de los distintos modelos evaluados. En este caso se han contabilizado el numero de predicciones de fase P con $\Delta t < 1s$ entre la predicción y la etiqueta de entrenamiento. . . . .	69
3.2. Gráfico de barras en el que se comparan los resultados de los distintos modelos evaluados. En este caso se han contabilizado el número de predicciones de fase S con $\Delta t < 1s$ entre la predicción y la etiqueta de entrenamiento. . . . .	69
3.3. En este gráfico de barras se aprecian los resultados en los que se han contabilizado las predicciones en cualquiera de las 2 fases sísmicas S o P. . . . .	70
3.4. El criterio mas riguroso en el que se han contabilizado predicciones correctas del modelo en las 2 fases sísmicas, se tuvo una localización de fases sísmicas completamente acertada en las muestras. . . . .	70
3.5. Gráfico de barras en el que se comparan los resultados de los distintos modelos evaluados. En este caso se han contabilizado el numero de predicciones de fase P con $\Delta t < 1s$ entre la predicción y la etiqueta de entrenamiento. . . . .	71
3.6. Gráfico de barras en el que se comparan los resultados de los distintos modelos evaluados. En este caso se han contabilizado el número de predicciones de fase S con $\Delta t < 1s$ entre la predicción y la etiqueta de entrenamiento. . . . .	71
3.7. En este gráfico de barras se aprecian los resultados en los que se han contabilizado las predicciones en cualquiera de las 2 fases sísmicas S o P. . . . .	72



3.14. Muestra sin coincidencias. En este caso el evento sísmico no es muy notorio en la traza, hay varios picos en la dist. de prob., los máximos están cercanos a las etiquetas pero sin la suficiente precisión. . . . .	76
3.15. Muestra sin coincidencias. Se tienen picos de las dist. de prob para los dos eventos en la traza, se tienen máximos en el segundo evento, para esta muestra se ha etiquetado el primer evento resultando en una predicción errónea. . . . .	76
3.16. Muestra en la que se tiene coincidencia para la picada de fase P, la predicción de la fase S esta cercana a la etiqueta pero hay una diferencia de mas de 3 segundos, se puede notar que el pico de la fase P es bastante elevado, por lo regular se tiene este comportamiento para buenas predicciones. . . . .	77
3.17. En las predicciones de esta muestra se puede ver que el modelo acertó en la fase S y se equivocó en la fase P, se puede ver varios picos en la dist. de prob. de la fase S, es normal que la fase S tenga varios picos de este tipo, en general las fases S plantean un reto ya que vienen superpuestas con las ondas P por lo que su correcta caracterización es mas complicada. . . . .	77
3.18. En este caso se tiene una predicción correcta para la fase P y para la fase S se tiene una predicción totalmente desacertada, como se aprecia la predicción esta muy lejos de la etiqueta, la diferencia es de mas de 19 segundos, resultados como este son los que se pueden mejorar con la utilización de mas datos en el entrenamiento del modelo. . . . .	78
3.19. Se tiene una picada de fase S predicha exitosamente a pesar de que tienen 2 eventos en la traza. Los picos de prob. son bajos para las picadas predichas. . . . .	78
3.20. Aunque se tiene una predicción perfecta para la fase P, el modelo identificó en la traza un arribo posterior de fase S como la predicción para la traza, de nuevo se tiene el caso con varias señales sísmicas en una sola traza. . . . .	79
3.21. Para esta muestra no se pudo identificar a la fase S correctamente, el único pico de probabilidad en la dist. de fase S está en la posición equivocada. Hay varios picos en para la fase P, el máximo corresponde a una identificación correcta. . . . .	79

3.22. Muestra en la que se identificaron correctamente a las 2 fases sísmicas, para las 2 fases se tiene $\Delta t < 0$ , en el caso de la onda P no hay diferencia entre la etiqueta y la predicción. . . . .	80
3.23. Se tienen dos aciertos, se puede notar que para la distribución de fase S hay varios picos de probabilidad, el máximo corresponde bastante bien con la posición de la etiqueta recolectada. . . . .	80
3.24. En este caso se pueden ver varios picos en las 2 distribuciones de probabilidad los picos mas altos coinciden con las etiquetas. . . . .	81
3.25. En esta muestra el evento sísmico no se ve a simple vista, para hacer un análisis de picadas de fase seria necesario aplicar acercamientos y filtros, SeismicNet reconoce exitosamente a las 2 fases. . . . .	81
3.26. En esta predicciones correctas de SeimicNet la muestra es visiblemente ruidosa, i.e. las amplitudes de la señal sísmicas son comparables o menores que el ruido, esto significa que el modelo aprende a reconocer otros parámetros como los cambios en el contenido de frecuencias de las ondas. . . . .	82
3.27. Se tiene varios picos cercanos a los máximos de las 2 distribuciones, las predicciones son acertadas. . . . .	82



# ÍNDICE DE TABLAS

3.1. Medidas de tendencia central para modelo de prueba TestNet, datos de enero 2021 . . . . .	62
3.2. Contabilización de predicciones acertadas del modelo TestNet en los datos de enero 2021 . . . . .	62
3.3. Medidas de tendencia central para modelo de prueba TestNet, datos de febrero 2021 . . . . .	63
3.4. Contabilización de predicciones acertadas del modelo TestNet en los datos de febrero 2021 . . . . .	63
3.5. Medidas de tendencia central para modelo de prueba TestNet2, datos de enero 2021 . . . . .	63
3.6. Contabilización de predicciones acertadas del modelo TestNet2 en los datos de enero 2021 . . . . .	64
3.7. Medidas de tendencia central para modelo de prueba TestNet2, datos de febrero 2021 . . . . .	64
3.8. Contabilización de predicciones acertadas del modelo TestNet2 en los datos de febrero 2021 . . . . .	64
3.9. Medidas de tendencia central para modelo de prueba TestNet3, datos de enero 2021 . . . . .	65
3.10. Contabilización de predicciones acertadas del modelo TestNet3 en los datos de enero 2021 . . . . .	65
3.11. Medidas de tendencia central para modelo de prueba TestNet3, datos de febrero 2021 . . . . .	65
3.12. Contabilización de predicciones acertadas del modelo TestNet3 en los datos de febrero 2021 . . . . .	66
3.13. Medidas de tendencia central para modelo de prueba TestNet3, datos de enero 2021 . . . . .	66
3.14. Contabilización de predicciones acertadas del modelo TestNet4 en los datos de enero 2021 . . . . .	66

3.15. Medidas de tendencia central para modelo de prueba TestNet4, datos de febrero 2021. . . . .	67
3.16. Contabilización de predicciones acertadas del modelo TestNet4 en los datos de febrero 2021 . . . . .	67
3.17. Medidas de tendencia central para modelo SeismicNet, datos de enero 2021. . . . .	67
3.18. Contabilización de predicciones acertadas del modelo SeismicNet en los datos de enero 2021 . . . . .	68
3.19. Medidas de tendencia central para modelo de prueba SeismicNet, datos de febrero 2021. . . . .	68
3.20. Contabilización de predicciones acertadas del modelo SeismicNet en los datos de febrero 2021 . . . . .	68

# **OBJETIVOS**

## **Generales**

Investigar las redes neuronales desde el contexto de la física y la mecánica estadística. Estudiar las nuevas tendencias en el uso de modelos predictivos basados en inteligencia artificial para su aplicación en el análisis de señales de origen natural como lo son los sismogramas. Desarrollar una red neuronal para detección de fases sísmicas.

## **Específicos**

1. Desarrollar arquitecturas de redes convolucionales para detección de fases sísmicas que mejoren el desempeño de los modelos realizados durante prácticas finales.
2. Comparar el desempeño de las redes neuronales creadas con el fin de optimizar hiper-parámetros.
3. Determinar mas problemas en los que las herramientas de las redes neuronales se puedan aplicar al procesamiento de datos en geofísica.



## INTRODUCCIÓN

El uso de la inteligencia artificial(IA) es cada vez mas prevalente y su aplicación a diversas áreas ha tenido un auge en años recientes. Su uso se ha hecho ubicuo en todas las áreas de investigación científica y ha sido una herramienta fundamental para el desarrollo de nuevas tecnologías que la mayoría de la sociedad usa con regularidad. El éxito actual de la IA se debe al evidente incremento que se ha tenido en la disponibilidad de conjuntos masivos de datos (Big Data), el poder de procesamiento de las computadoras actuales, la ubicuidad de los sistemas de computo que se ha dado con los teléfonos móviles y el rápido desarrollo que se ha tenido de parte de la comunidad técnica y científica; vale mencionar que la comunidad técnica y académica dedicada al desarrollo de nuevos conceptos y técnicas en IA es muy abierta y colaborativa. Para poner en contexto el incremento de la disponibilidad de datos en la actualidad se tienen ejemplos, como el hecho que hay mas de un trillón de paginas web<sup>1</sup>, mas de una hora de video es subida a YouTube cada segundo<sup>2</sup>, los genomas de miles de personas los cuales tienen una longitud de  $3.8 \times 10^9$  de pares base han sido secuenciados en laboratorios alrededor del mundo, el archivo de formas de onda sísmicas públicamente disponible en IRIS(Incorporated Research Institutions for Seismology<sup>3</sup>) crece exponencialmente [22], los datos de sismogramas identificados en el INSIVUMEH crece rápidamente[29]. Esta evidente disponibilidad de datos de toda índole y los avances en el hardware han hecho que los conceptos de redes neuronales creados en los 60s, 70s y 80s hayan despegado en la actualidad, creando nuevas perspectivas de investigación y nuevos conceptos. Aunque el consenso científico con respecto al uso de la IA en la investigación es variado, su proliferación en prácticamente toda la ciencia es innegable.

---

<sup>1</sup><http://googleblog.blogspot.com/2008/07/we-knew-web-was-big.html>

<sup>2</sup>[http://www.youtube.com/t/press\\_statistics](http://www.youtube.com/t/press_statistics)

<sup>3</sup><https://www.iris.edu>

Las aplicaciones de la inteligencia artificial en desarrollo de tecnologías son muy variadas, por mencionar algunas que han tenido un éxito sustancial se tienen: los sistemas de reconocimiento de imágenes(visión computarizada), sistemas de reconocimiento de voz, procesamiento de lenguaje natural, sistemas de recomendación automática, navegadores de Internet, desarrollo de videojuegos, robótica, Big Data y minado de datos y muchas otras aplicaciones en generación de vídeo. En física como en otras ciencias se ha visto una tendencia en los últimos años de la aplicación de técnicas de inteligencia artificial a varios problemas en distintas ramas, por mencionar algunas aplicaciones interesantes en ramas muy diversas de la física se tienen los siguiente ejemplos:

- Óptica. [10]
- Astrofísica. [38]
- Física de Partículas. [6]
- Física del estado sólido. [40]
- Física computacional. [21], [43]
- Mecánica Cuántica. [37]
- Mecánica Estadística. [3]

El interés de la física en las redes neuronales no se da sólo en su aplicación y sus características pragmáticas sino que también hay un interés en ayudar a dar intuiciones profundas y mejorar el entendimiento teórico que se tiene actualmente [3], la física se interesa particularmente en la mecánica estadística del aprendizaje automático, siendo este un campo abierto de investigación donde hay una gran colaboración interdisciplinaria. El estudio del procesamiento de la información y la mecánica estadística quizá tiene su formulación primordial con la teoría de la información de Claude Shannon.

Este trabajo de graduación consta de tres de partes principales, en la primera parte se han investigado teóricamente a las redes neuronales en el contexto de la mecánica estadística. Se hace una revisión rápida de los conceptos de mecánica estadística concernientes, se introducen los sistemas de ferromagnéticos de Ising y la aproximación de campo medio para resolver el modelo de Ising en situaciones interesantes. También se presentan a los vidrios de spin modelados a partir del modelo de Ising esto da lugar al método de replicas, la solución de Parisi y la ruptura de simetría de escala completa. Se termina la parte de investigación con modelos estadísticos de redes neuronales, en específico el modelo de Hopfield.

En la segunda parte de este trabajo se exponen los detalles del desarrollo de un modelo de red neuronal convolucional con el fin de detección de arribos de fases en señales sísmicas(sismogramas), dicha red se ha llamado SeismicNet. Se utilizaron datos de la red sísmica del departamento de Geofísica del INSIVUMEH, de los años 2019 y 2020 para entrenar al modelo. Los modelos investigados mejoraron el desempeño obtenido previamente [29] y muestran potencial para aplicaciones reales como la localización automática de eventos y como una herramienta de apoyo en el análisis rutinario de sismogramas.

En la tercera parte se presentan los resultados de SeismicNet en conjuntos de datos que el modelo no ha visto previamente, para probar las predicciones de la red neuronal se usaron datos de enero y febrero 2021.



# 1. Mecánica Estadística de las Redes Neuronales

La analogía entre sistemas ferromagnéticos y redes neuronales artificiales puede entenderse intuitivamente de la siguiente manera, las unidades de activación o neuronas de una red neuronal corresponden a los spins en un retículo que modela a un material ferromagnético y los acoplos o interacciones entre spins corresponden a las conexiones entre neuronas en una red artificial, con esto en mente se pueden estudiar a estos sistemas en el contexto de la mecánica estadística.

## 1.1. Estudios primordiales de redes neuronales en Física

Estudios de sistemas desordenados han generado nuevas ideas del comportamiento cooperativo y de las propiedades computacionales emergentes de redes de procesadores simples con alta conectividad que tienen semejanza a las neuronas biológicas. Una red neuronal artificial es un ensamblaje altamente interconectado de elementos simples. Estos elementos conocidos en la literatura de IA como neuronas o activaciones pueden ser funciones o dispositivos con dos estados, que pueden pasar de un estado a otro o pueden prenderse y apagarse, dependiendo si el valor de sus entradas excede un umbral preestablecido. En este sentido las neuronas o nodos de la red artificial tienen similitud con neuronas de origen biológico, las cuales se disparan i.e. envían un pulso de voltaje a través de sus axones cuando la suma de las entradas en sus sinapsis excede un umbral de activación o de disparo. Las redes neuronales artificiales son sobre simplificaciones de la biología, al principio se pretendía modelar a sistemas biológicos a través del uso de este tipo de redes, sin embargo el uso actual es mucho más amplio y general. Intentos por modelar el funcionamiento del cerebro con redes de neuronas simples se puede trazar a 1943 cuando Walter Pitts y Warren McCulloch [28] propusieron redes de elementos de umbral de dos estados que son capaces de realizar operaciones lógicas arbitrarias y fueron de los primeros en abordar el problema de reconocimiento de patrones. En los 60s Rosenblatt [39]

desarrolló una red de propagación directa llamada perceptron, se le conocía también como red adaptativa ya que realizaba tareas de reconocimiento simple. Durante los 60s Bernard Widrow [45] desarrollo una máquina adaptativa lineal, llamada Adaline. Una variedad de modelos de redes para desarrollo de memoria asociativa y reconocimiento de patrones fueron desarrollados en los 80s por diversos equipos, Shun-ichi Amari [2], Grossberg [13] y otros.

El interés inicial de la física en las redes neuronales surge de la analogía entre dichas redes y sistemas magnéticos simples. Esta analogía fue sugerida inicialmente por William Little [26], subsecuentemente se tuvo el trabajo de John Hopfield [18] quien señaló la equivalencia entre el comportamiento a largo plazo de redes con conexiones simétricas y las propiedades de equilibrio de sistemas magnéticos como el vidrio de espín(spin glass), en particular Hopfield demostró como se puede explotar esta equivalencia para diseñar circuitos neuronales para memoria asociativa y otras tareas de computación.

## 1.2. Principios generales de Mecánica Estadística

En términos simples los enunciados básicos de la mecánica estadística pueden ser introducidos de la siguiente manera. Se tienen sistemas en los que se permite que los estados microscópicos de un sistema macroscópico que tiene determinado numero de grados de libertad(un numero grande de grados de libertad) sean descritos por las configuraciones de  $N$  variables microscópicas  $\{s_i\}$ , ( $i = 1, 2, \dots, N$ ). La cantidad básica que caracteriza a los estados microscópicos es la energía  $H$  y está definida como una función de las variables microscópicas  $\{s_i\}$ , es decir  $H = H(s_1, s_2, \dots, s_N) \equiv H[s]$ .

El comportamiento dinámico microscópico del sistema esta definido por un conjunto ecuaciones diferenciales dinámicas, tal que en general la energía del sistema tiende a un mínimo. Ademas se asume que ningún sistema observable puede ser perfectamente aislado del entorno y el efecto de la interacción con el entorno conocida como baño térmico tiene como resultado único el llamado ruido térmico en las ecuaciones exactas que rigen su dinámica. El ruido térmico(ruido blanco) actúa como fluctuaciones aleatorias no correlacionadas que producen la aleatorización y la mezcla de las trayectorias dinámicas exactas del sistema.

Si se tiene que  $A[s]$  es una cantidad observable, las cantidades de interés en mecánica

estadística, son los valores promedio de las cantidades observables, en otras palabras en vez de estudiar el cambio exacto en el tiempo de los valores del observable  $A[s(t)]$ , se introducen la cantidad promedio

$$\langle A \rangle = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t dt' A[s(t')] \quad (1.1)$$

que puede ser formalmente obtenida después de haber obtenido observaciones durante un periodo de tiempo infinito(en la practica un tiempo largo en el que se han tomado un numero grande de medidas es suficiente).

La hipótesis fundamental de equilibrio en mecánica estadística se da de la siguiente manera, asumiendo que debido al proceso de mezcla de las trayectorias dinámicas después de un periodo infinito de observación el sistema en general visitara todos sus posibles microestados muchas veces, entonces es posible encontrar la cantidad promedio en 1.1 al promediar sobre el ensamble de microestados en lugar de promediar en el tiempo es decir

$$\langle A \rangle = \int ds_1 ds_2, \dots, ds_N A[s] P(s_1, s_2, \dots, s_N) \quad (1.2)$$

Donde  $P[s]$  es la función de distribución de probabilidad de los microestados del sistema. Con esto se tiene que debido a la mezcla de las trayectorias dinámicas después de un tiempo infinito de observación, en vez de resolver las ecuaciones dinámicas de movimiento de una manera exacta y analítica, el sistema puede ser descrito estadísticamente en términos de las probabilidades de los microestados dados por la función  $P[s]$ . Sin importar la forma de la función de distribución de probabilidad, esta debe estar apropiadamente normalizada

$$\int ds_1 ds_2 \dots ds_N P(s_1, s_2, \dots, s_N) = 1 \quad (1.3)$$

La cantidad fundamental en mecánica estadística que caracteriza la distribución de probabilidad es la entropía que está definida como el promedio del logaritmo de la función de distribución:

$$S = -\langle \log(P[s]) \rangle \equiv - \int ds_1 ds_2 ds_N P[s] \log(P[s]) \quad (1.4)$$

Se puede notar de la definición anterior que debido a la normalización en 1.3 la entropía es positiva. En general el valor de la entropía puede dar información de hasta que extensión la distribución de probabilidad del sistema esta “ordenada”.

Para iluminar la definición de entropía se puede tener el siguiente ejemplo ilustrativo, si se tienen un sistema con microestados discretos que se pueden identificar con el indice  $\alpha$  y se asume que la distribución de probabilidad es tal que solo  $L$  numero de estados(entre todos los microestados) tienen la misma probabilidad distinta de cero, es decir los  $L$  estados son equiprobables. Entonces debido a la condición de normalización 1.3, la probabilidad de cualquiera de estos  $L$  estados debe ser igual a  $\frac{1}{L}$ , entonces de la definición de entropía 1.4 se obtiene que

$$S = - \sum_{\alpha}^L P_{\alpha} \log P_{\alpha} = \log L.$$

Del ultimo resultado se obtiene que para una distribución que está mas esparcida en el espacio de estados( $L$  aumenta) se tiene que la entropía crece. Por el otro lado, si se tiene una distribución mas concentrada en el espacio de estados( $L$  disminuye) mas pequeña es la entropía. En el caso extremo en el que solo hay un microestado que puede ser ocupado por el sistema la entropía se desvanece por completo,  $S = 0$ . En general, el valor  $\exp(S)$  puede interpretarse como el promedio del numero de estados ocupados por el sistema con probabilidades no despreciables.

Si ahora se considera la forma que debería de tener la función de distribución de probabilidad de manera general, siguiendo la hipótesis básica el valor promedio de la energía del sistema es

$$E \equiv \langle H \rangle \equiv \sum_{\alpha} P_{\alpha} H_{\alpha} \quad (1.5)$$

La entropía del sistema está dada por

$$S = - \sum_{\alpha} P_{\alpha} \log(P_{\alpha}) \quad (1.6)$$

Por conveniencia se ha asumido un sistema con estados discretos. La interacción del sistema con su entorno produce los siguientes efectos fundamentales: Primero, en equilibrio térmico el valor promedio de su energía se conserva. Segundo, la naturaleza está estructurada de tal manera que cualquiera que sea la estructura interna del sistema, el valor de la entropía en el estado de equilibrio tiende a un máximo posible(limitado por la condición que el valor promedio de la energía permanezca constante), hasta cierto sentido esto es natural: el ruido aleatorio hace que el sistema

esta tan desordenado como sea posible. Ahora se analiza la forma que debe tener la función de probabilidad para maximizar la entropía. Para considerar a las restricciones de la conservación del promedio de la energía 1.5 y la normalización 1.3, se puede usar el método de multiplicadores de Lagrange para la expresión 1.6 con las restricciones:

$$\sum_{\alpha} P_{\alpha} H_{\alpha} - E = 0$$

$$\sum_{\alpha} P_{\alpha} - 1 = 0$$

obteniendo la siguiente expresión a maximizar para todos los posibles valores de la distribución  $P_{\alpha}$ :

$$S_{\beta, \gamma}[P] = - \sum_{\alpha} P_{\alpha} \log(P_{\alpha}) - \beta \left( \sum_{\alpha} P_{\alpha} H_{\alpha} - E \right) - \gamma \left( \sum_{\alpha} P_{\alpha} - 1 \right) \quad (1.7)$$

Donde  $\beta$  y  $\gamma$  son los multiplicadores de Lagrange. Derivando con respecto a  $P_{\alpha}$  para encontrar el máximo da como resultado

$$P_{\alpha} = \frac{1}{Z} \exp(-\beta H_{\alpha}) \quad (1.8)$$

Donde  $Z = \exp(\gamma + 1)$ , de la condición de normalización  $\sum_{\alpha} P_{\alpha} = 1$ , aplicada a 1.8 se obtiene

$$Z = \sum_{\alpha} \exp(\beta H_{\alpha}) \quad (1.9)$$

conocida como función de partición y el parámetro  $\beta$  definido por la condición:

$$\frac{1}{Z} \sum_{\alpha} H_{\alpha} \exp(\beta H_{\alpha}) = E \quad (1.10)$$

Normalmente se define a  $\beta = \frac{1}{K_B T}$ , donde  $K_B$  es el conocido factor de Boltzmann, aquí se toma  $K_B = 1$ . En la práctica la temperatura  $T$  es el parámetro independiente y la energía promedio se obtiene como función de la temperatura por medio de 1.10. La otra cantidad de interés para esta breve introducción de mecánica estadística es la energía libre definida como

$$F = E - TS \quad (1.11)$$

Usando la ecuación 1.8 como la distribución de probabilidad se pueden derivar la siguiente relaciones entre la energía libre, la función de partición y la energía pro-

medio.

$$F = -T \log(Z) \quad (1.12)$$

$$S = \beta^2 \frac{\partial F}{\partial \beta} \quad (1.13)$$

$$E = -\frac{\partial}{\partial \beta} \log(Z) = F + \beta \frac{\partial F}{\partial \beta}. \quad (1.14)$$

Se puede hacer la observación, que según la definición 1.11, el principio de máximo de la entropía es equivalente al mínimo de la energía libre. Se puede verificar que al tomar a la energía libre como cantidad fundamental(en lugar de la entropía) y minimizarla con respecto a las distribución de probabilidad se obtiene la misma forma de función de distribución de probabilidad de la expresión 1.8.

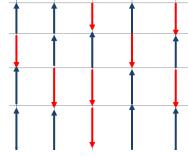
### 1.3. Modelos de sistemas magnéticos con spin Ising

En los materiales magnéticos se supone que el estado microscópico del sistema está definido por los valores de las magnetizaciones de spin locales. En varios materiales magnéticos los electrones responsables por el comportamiento magnético están localizados cerca de los átomos en el retículo cristalino y la fuerza que tiende a orientar a los spins es una interacción de intercambio de corto alcance.

Unos de los modelos mas populares que cualitativamente describen la situación descrita anteriormente son los llamados modelos de Ising. El modelo de Ising fue formulado por Lenz [25] en 1920 y el modelo unidimensional fue resuelto por el estudiante de Lenz, Ernst Ising en su tesis doctoral [19] de 1924. Las variables microscópicas de este modelo son los denominados spin de Ising  $\sigma_i$  que por definición solo toman los valores  $+1$  y  $-1$ , esto consiste en asignar un valor discreto a la orientación relativa del momento dipolar magnético(spin). Se asigna  $+1$  a una orientación paralela al eje  $y$  en el retículo formado por las posiciones discretizadas de las moléculas. ver fig. 1.1 y se le asigna  $-1$  a la orientación anti-paralela.

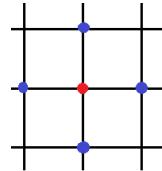
La forma tradicional en los modelos Ising para la energía microscópica que también se conoce como Hamiltoniano del sistema, es la siguiente:

$$\mathcal{H} = - \sum_{\langle i,j \rangle} J_{ij} \sigma_i \sigma_j - h \sum_i \sigma_i \quad (1.15)$$



**Figura 1.1.** Retículo de spins.

El indice  $< i, j >$  significa la suma de las interacciones de los vecinos mas cercanos, estos son por ejemplo si se toma a la posición  $(x, y)$  sus vecinos serian  $(x, y + 1), (x, y - 1), (x - 1, y), (x + 1, y)$ , los vecinos se pueden ver en color azul en la fig 1.2. y la posición  $(x, y)$  se aprecia en color rojo;  $J_{ij}$  son los valores de las interacciones spin-spin(acoplos),  $h$  es el valor del campo magnético externo. Si todos los valores de los  $J_{ij}$  son iguales a un constante positiva entonces se obtiene lo que se conoce como el modelo ferromagnético de Ising, si todos los  $J_{ij}$  son iguales a una constante negativa entonces se obtiene el modelo antiferromagnético de Ising. El primer termino de 1.15 representa a la energías de interacción introducidas para llevar a un estado ferromagnético ordenado, el segundo termino representa la interacción entre el campo magnético aplicado  $h$  y los spines del sistema esta interacción es de carácter puramente paramagnético.



**Figura 1.2.** Vecinos mas cercanos

## 1.4. Aproximación de campo medio

A pesar de la aparente simplicidad del modelo de Ising, solo existen unas cuantos casos para los que existe una solución exacta. Como es de esperarse para problemas de varios cuerpos en física, en los que soluciones exactas son extremadamente raras. Algunos casos que se han resuelto son los casos unidimensional y bidimensional en los que no hay campo magnético externo aplicado,  $h = 0$ . En la mayoría de los otros casos se necesitan de métodos de aproximación. Vale la pena mencionar que resolver modelos con spin de Ising es de interés pragmático ya que tiene uso para modelar varios fenómenos termodinámicos y sirve para probar muchos modelos

teóricos. Uno de los métodos mas simples para aproximar la solución es el llamado método de aproximación de campo medio, en muchos casos esta aproximación da resultados que no están lejos de los resultados exactos y muy frecuentemente hace posible que se tenga algún entendimiento cualitativo del comportamiento del sistema bajo consideración.

El punto inicial para la aproximación de campo medio es la asunción que se hace en la forma de la función de distribución de probabilidad  $P[\sigma] = \frac{1}{Z} \exp(-\beta H[\sigma])$ . Se asume que se puede aproximar como

$$P[\sigma] = \frac{1}{Z} \exp(-\beta H[\sigma]) \simeq \prod_i P_i(\sigma_i) \quad (1.16)$$

Donde  $P[\sigma]$  es el producto de las distribuciones independientes de los sitios del retículo del sistema, las distribuciones independientes de las posiciones en el retículo tiene la siguiente forma

$$P_i(\sigma_i) = \frac{1+m_i}{2} \delta(\sigma_i - 1) + \frac{1-m_i}{2} \delta(\sigma_i + 1) \quad (1.17)$$

donde  $m_i$  son parámetros que se tienen que obtener. La factorización en la expresión 1.16, significa que el promedio de cualquier producto de funciones de las magnetizaciones de los sitios se puede escribir como un producto de los promedios independientes, es decir

$$\langle f(\sigma_i)g(\sigma_i) \rangle = \langle f(\sigma_i) \rangle \langle g(\sigma_i) \rangle \quad (1.18)$$

donde  $\langle f(\sigma_i) \rangle = \sum_i P_i(\sigma_i) f(\sigma_i)$ , utilizando el ansatz 1.17, se obtiene

$$\langle f(\sigma_i) \rangle = \frac{1+m_i}{2} f(1) + \frac{1-m_i}{2} f(-1) \quad (1.19)$$

Para las magnetizaciones promedio de cada sitio se obtiene fácilmente

$$\langle \sigma_i \rangle = m_i \quad (1.20)$$

Entonces se puede interpretar físicamente a los parámetros  $\{m_i\}$ , como el promedio de las magnetizaciones de spin en cada sitio. De acuerdo a los principios generales de la mecánica estadística 1.2, estos parámetros deben de ser tales que minimicen la energía libre del sistema. Utilizando la ecuación 1.16 en 1.4 se obtiene la siguiente

expresión para la entropía

$$\begin{aligned} S &= -\langle \log(P[\sigma]) \rangle \simeq -\sum_i \langle \log(P_i(\sigma_i)) \rangle \\ S &\simeq -\sum_i \left[ \frac{1+m_i}{2} \log\left(\frac{1+m_i}{2}\right) + \frac{1-m_i}{2} \log\left(\frac{1-m_i}{2}\right) \right] \end{aligned} \quad (1.21)$$

Para la energía promedio se utilizan las expresiones 1.15 y 1.17 en 1.5, haciendo uso de los resultados 1.18 y 1.20, se obtiene el resultado

$$E = -\sum_{\langle i,j \rangle} J_{ij} m_i m_j - h \sum_i m_i \quad (1.22)$$

Para la energía libre se utiliza la ecuación 1.11 obteniendo

$$F = -\sum_{\langle i,j \rangle} J_{ij} m_i m_j - h \sum_i m_i + T \sum_i \left[ \frac{1+m_i}{2} \log\left(\frac{1+m_i}{2}\right) + \frac{1-m_i}{2} \log\left(\frac{1-m_i}{2}\right) \right] \quad (1.23)$$

Como se discutió anteriormente para encontrar las condiciones que maximicen la entropía se puede optar equivalentemente por minimizar la energía libre  $F$ . Las condiciones para encontrar el mínimo de  $F$  son  $\frac{\partial F}{\partial m_i} = 0$ , las cuales resultan en

$$-2 \sum_i J_{ij} m_j - h + T \operatorname{arctanh}(m_i) = 0 \quad (1.24)$$

Las ecuaciones resultantes que definen a los parámetros  $m_i$  son:

$$m_i = \tanh \left[ \beta \left( 2 \sum_j J_{ij} m_j + h \right) \right] \quad (1.25)$$

(se ha tomado a  $\beta = \frac{1}{T}$  i.e.  $K_B = 1$ ). Para tener un ejemplo concreto, se considera el sistema ferromagnético que reside en un retículo tridimensional cubico con campo magnético aplicado igual a cero. En este caso todos los enlaces spin-spin son iguales  $J_{ij} = J > 0$  y cada sitio tiene 6 vecinos cercanos, es decir todos equidistan de dicho sitio. Ya que el sistema es homogéneo, es natural esperar que todos los promedios de las magnetizaciones  $m_i$  deban de ser igual a una constante  $m$ , con esta información se tiene que la ecuación 1.25 se reduce a

$$m = \tanh [\beta(12Jm)] \quad (1.26)$$

Se sabe de cálculo que el mínimo de la energía libre está condicionado por  $\frac{\partial^2 F}{\partial^2 m} > 0$ ,

esta condición se puede reducir a

$$\frac{1}{1-m^2} > 12\beta J \quad (1.27)$$

De este ultimo resultado es fácil notar que si  $T > T_c = \frac{J}{12}$ , la única solución de la ecuación 1.26 es  $m = 0$ , esta solución también satisface 1.27, entonces para temperaturas mayores a la temperatura critica  $T_c$ , el mínimo de la energía libre se alcanza en el estado en el que todos los spins de magnetización son cero.

Sin embargo si la temperatura es menor a la temperatura critica,  $T < T_c$ , entonces adicionalmente a la solución  $m = 0$ , la ecuación 1.26 tiene como soluciones no triviales a  $m = \pm m(T) \neq 0$ . Se puede verificar fácilmente que en esta región de la temperatura la solución  $m = 0$  se convierte en máximo de la energía libre, mientras que los mínimos están en  $\pm m(T)$ . Entonces para la región de temperaturas bajas i.e.  $T < T_c$ , la energía libre tiene dos mínimos que están caracterizados por magnetización de sitio de signos opuestos.

Cerca a  $T_c$  la magnetización  $m(T)$  es pequeña, para este caso la expansión en serie de potencias de  $m(T)$  en la ecuación 1.26 es posible. El termino dominante en la expansión es  $\tau \equiv (1 - T/T_c) \ll 1$ , con esto se obtiene la expresión:

$$\begin{aligned} m(T) &= C\tau^{1/2} \\ m(T) &= C\left(1 - \frac{T}{T_c}\right)^{1/2} \end{aligned} \quad (1.28)$$

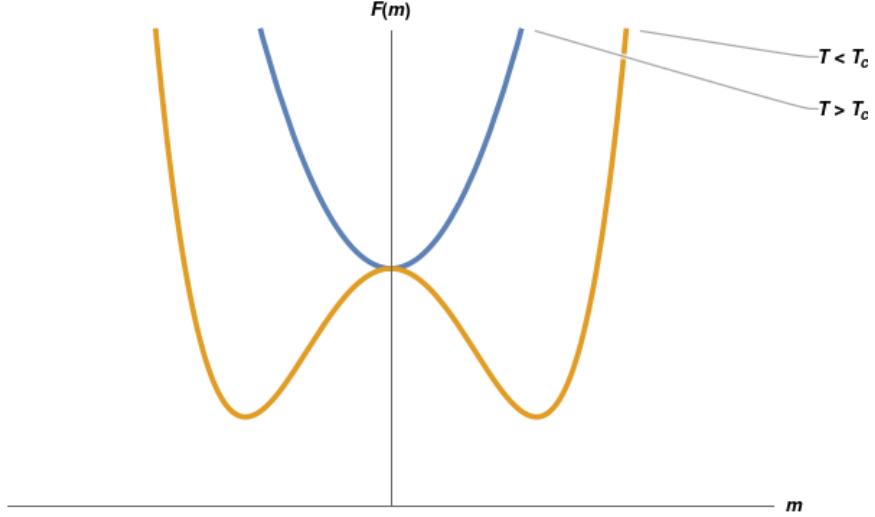
donde C es una constante de proporcionalidad, cuando  $T \rightarrow T_c$  desde abajo,  $m(T) \rightarrow 0$ . Utilizando la ecuación 1.23 para la energía libre como función de  $m$  se obtiene:

$$F(m) = NT \left[ -6Jm^2 + \frac{1+m}{2} \log\left(\frac{1+m}{2}\right) + \frac{1-m}{2} \log\left(\frac{1-m}{2}\right) \right] \quad (1.29)$$

donde  $N$  es el numero total de sitios en el sistema. La forma cualitativa de la función  $F(m)$  para  $T > T_c$  y  $T < T_c$  se puede ver en la figura 1.3.

Debido a que la energía libre es proporcional al volumen del sistema, que es esencialmente  $N$ , el valor de la barrera en la energía libre que separa a los estados  $\pm m(T)$  cuando  $T < T_c$  también es proporcional al volumen del sistema  $N$ . Por lo tanto en el limite termodinámico  $N \rightarrow \infty$ (que corresponde a las consideraciones del sistema macroscópico) la barrera que separa a los dos estados tiende al infinito.

Las simples consideraciones anteriores demuestran a nivel cualitativo el fenómeno fundamental que se conoce como ruptura espontanea de simetría. A la temperatura



**Figura 1.3.** Energía libre del imán ferromagnético de Ising. Se puede ver en color azul para  $T > T_c$  y en naranja para  $T < T_c$ .

$T = T_c$  la transición de fase de segundo orden ocurre, de tal manera que en la región de bajas temperaturas  $T < T_c$  la simetría con respecto al cambio global de los signos de los spins se rompe y aparecen dos estados base en lugar de uno. Estos estados base difieren por el signo de la magnetización de spins promedio y están separados por la barrera macroscópica de la energía libre.

## 1.5. Modelo de Ising con interacciones débiles de largo alcance

Este es el sistema ferromagnético de Ising descrito por el Hamiltoniano 1.15 en el que se asume que los valores de las interacciones spin-spin los  $J_{ij}$  son los mismos para todos los pares de spins y no solo para los vecinos mas cercanos. Es obvio que este sistema es uno que dista mucho de un sistema físico, ya que las interacciones entre spin-spin son independientes de la distancia espacial entre ellos, sin embargo este modelo simplificado permite una solución exacta y esta solución es una solución simple no trivial.

Debido a que la energía del sistema debe de ser proporcional al volumen  $N$ , mientras que el total de interacciones entre pares de spins es proporcional a  $N^2$ , entonces se tiene que el valor de las interacciones spin-spin en este modelo deben de ser proporcionales a  $N^{-1}$ . Si se toma a  $J_{ij} = J/N$ , el Hamiltoniano de este sistema se

puede escribir como:

$$\begin{aligned} H &= -\frac{J}{2N} \sum_{ij}^N \sigma_i \sigma_j - h \sum_i^N \sigma_i \\ &= -\frac{J}{2N} \left( \sum_i^N \sigma_i \right)^2 - h \sum_i^N \sigma_i \end{aligned} \quad (1.30)$$

Para la función de partición se obtiene fácilmente

$$\begin{aligned} Z &= \sum_{\sigma} \exp \left[ \frac{\beta J}{2N} \left( \sum_i^N \sigma_i \right)^2 + \beta h \sum_i^N \sigma_i \right] = \\ &\left( \frac{\beta N J}{2\pi} \right)^{1/2} \int_{-\infty}^{+\infty} dm \sum_{\sigma} \exp \left[ -\frac{N\beta J}{2} m^2 + \beta J m \left( \sum_i^N \sigma_i \right) + \beta h \sum_i^N \sigma_i \right] \end{aligned} \quad (1.31)$$

El parámetro Gaussiano  $m$  introducido en la ecuación anterior para linealizar la suma sobre spins tiene una interpretación física que se expondrá en los siguientes argumentos. En la expresión 1.31 para la función de partición se puede asumir que esta corresponde al sistema descrito por el Hamiltoniano

$$H[m, \sigma] = \frac{NJ}{2} m^2 - Jm \left( \sum_i^N \sigma_i \right) - h \sum_i^N \sigma_i \quad (1.32)$$

La energía libre correspondiente como función del parámetro  $m$  sería

$$F = -T \log \left[ \sum_{\sigma} \exp (-\beta H[m, \sigma]) \right] \quad (1.33)$$

El mínimo de  $F$  se obtiene de nuevo con  $\partial F / \partial m = 0$ , computaciones sencillas usando 1.32 y 1.33 dan como resultado la solución para  $m$

$$m = \frac{1}{N} \sum_i^N \langle \sigma_i \rangle \quad (1.34)$$

donde

$$\langle \sigma_i \rangle \equiv \frac{1}{Z} \sum_{\sigma} \exp (\beta H[m, \sigma]) \quad (1.35)$$

es el valor de la magnetización de spins promediado “termalmente”. Entonces el parámetro  $m$  describe el promedio de la magnetización del sistema. Los cálculos para

la función de partición son relativamente sencillos, de la ecuación 1.31 se obtiene

$$\begin{aligned}
Z &= \left( \frac{\beta N J}{2\pi} \right)^{1/2} \int_{-\infty}^{+\infty} dm \exp \left( \frac{-N\beta J}{2} m^2 \right) \prod_i^N \left[ \sum_{\sigma_i=\pm 1} \exp(\beta(Jm + h)\sigma_i) \right] = \\
&= \left( \frac{\beta N J}{2\pi} \right)^{1/2} \int_{-\infty}^{+\infty} dm \exp \left( \frac{-N\beta J}{2} m^2 \right) [2 \cosh \beta(Jm + h)]^N = \\
&= \left( \frac{\beta N J}{2\pi} \right)^{1/2} \int_{-\infty}^{+\infty} dm \exp \left( \frac{-N\beta J}{2} m^2 + N \ln [2 \cosh \beta(Jm + h)] \right) = \\
&= \left( \frac{\beta N J}{2\pi} \right)^{1/2} \int_{-\infty}^{+\infty} dm \exp (-\beta J N f(m))
\end{aligned} \tag{1.36}$$

con  $f(m)$  definido por

$$f(m) = \frac{1}{2} m^2 - \frac{1}{\beta J} \ln (2 \cosh \beta(Jm + h)) \tag{1.37}$$

Del hecho que el exponente en la ultima integral de la expresión 1.36 sea proporcional al volumen del sistema  $N$  en el limite macroscópico  $N \rightarrow \infty$  la contribución dominante en la función de partición(y en el energía libre) viene de la vecindad con el mínimo de la función  $f(m)$ . El mínimo de obtiene se la manera estándar  $\partial f / \partial m = 0$ , obtenido el resultado para m:

$$m = \tanh \beta(Jm + h) \tag{1.38}$$

que coincide con la ecuación para la magnetización correspondiente a la aproximación de campo medio 1.26(se obtuvo para el caso  $h = 0$ ). Esto significa que para el modelo en consideración la aproximación de campo medio parece ser exacta. Para la energía libre del sistema en el termino dominante en  $N$  se tiene

$$F = NJ \left[ \frac{1}{2} m^2 - \frac{1}{\beta J} \ln (2 \cosh \beta(Jm + h)) \right] \tag{1.39}$$

donde el valor de  $m$  se obtiene de la ecuación 1.38.

## 1.6. Física del Estado del Vidrio de Spin

Cuando las interacciones entre spins no son uniformes en el espacio, el análisis del modelo de Ising no funciona. Particularmente cuando se tienen interacciones ferromagnéticas para algunos acoplos y se tienen interacciones anti-ferromagnéticas para otros la orientación de los spins no puede ser uniforme en el espacio a diferencia de sistemas ferromagnéticos, incluso a bajas temperaturas. Se tiene el fenómeno que bajo las circunstancias descritas a veces los spins se “congelan” aleatoriamente, aleatorios en el espacio pero congelados en el tiempo, esta es la visión intuitiva de la fase del vidrio de spin.

Hay varios modelos en mecánica estadística para los vidrios de spin, uno de los mas simples y en cierto sentido de los mas generales es el que se da en términos de spins de Ising clásicos, descritos por el Hamiltoniano:

$$H = -\frac{1}{2} \sum_{i \neq j}^N J_{ij} \sigma_i \sigma_j \quad (1.40)$$

El sistema consiste de  $N$  spins Ising  $\{\sigma_i\}$  ( $i = 1, 2, \dots, N$ ), que toman valores  $\pm 1$ , posicionados en los vértices de un retículo como se ve en la fig. 1.1, numerados por el indice  $i$ . Las interacciones spin-spin  $J_{ij}$  son aleatorias en sus valores y sus signos. El modelo en si, se define por la elección de la función de distribución  $P(J_{ij})$  para las interacciones entre spins. La elección mas simple y suficientemente realista desde el punto de vista experimental para la distribución de probabilidad de las interacciones, es aquella con valores de la interacciones distintos de cero solo para las interacciones de vecinos mas cercanos y su distribución de probabilidad siendo una Gaussiana e independiente para todos los pares de spin con interactuantes:

$$P(J_{ij}) = \frac{1}{\sqrt{2\pi J^2}} \exp\left(-\frac{J_{ij}^2}{2J^2}\right) \quad (1.41)$$

El parámetro  $J$  es el valor característico de interacciones spin-spin.

### 1.6.1. Promedios apagados y recocidos

En los modelos de spin en los que se dice que el desorden esta “apagado” (*quenched*), se entiende que las variables “desordenadas” permanecen fijas mientras los spins fluctúan. Desde un punto de vista experimental esto corresponde a una situación en la

que la escala dinámica del tiempo del desorden (e.g. los acoplos de spin en un vidrio de spin) es mucho mayor que la escala de tiempo dinámica de las fluctuaciones de spin. En una muestra experimental las variables desordenadas asumen un valor bien definido(aunque desconocido) del tiempo que es independiente. La descripción en términos de variables aleatorias se debe interpretar de la siguiente manera: cada realización dada de las variables aleatorias corresponde a una muestra específica del sistema, mientras que la distribución de la cual fueron tomadas describe a las fluctuaciones muestra a muestra. Se tiene una situación distinta cuando el desorden esta “recocido”(*annealed*), esto significa que la escala de tiempo del desorden y la escala de tiempo de las fluctuaciones de spin tienen magnitudes comparables. En este caso, dada una muestra experimental se tiene que las variables del desorden varían en el tiempo y cuyas estadísticas son descritas por la distribución correspondiente.

Para el desorden apagado hay una jerarquía entre las variables de spin(rápidas) y las variables del desorden(lentas) esto es crucial en muchos aspectos. Por ejemplo, para determinar el comportamiento termodinámico, idealmente para una muestra dada se quieren computar los promedios sobre la medida de Boltzmann 1.8 y obtener las propiedades de equilibrio de dicha muestra. Sin embargo debido al desorden, solo se pueden calcular cantidades promediadas sobre la distribución del desorden. Una pregunta natural que surge, es entender hasta que punto estas cantidades promediadas describen la física de una sola muestra.

Es importante notar que para observables extensivos como la energía libre, se tienen buenos comportamientos, esto se debe a que sus densidades asociadas se auto-promedian en el límite termodinámico, esto significa que asumen el mismo valor para cada realización del desorden con una probabilidad finita. En este caso las fluctuaciones entre muestras se van desvaneciendo cuando el volumen del sistema se hace infinito y el valor del promedio apagado coincide con el promedio normal. Por el contrario para variables que no se auto promedian se tiene que las fluctuaciones son grandes entre muestras y cuando se computa el promedio del desorden solo algunas muestras raras con probabilidad que se desvanece resultan en una contribución finita. La propiedad de auto-promediamiento de las variables extensivas significa que son estas las variables necesarias para calcular correctamente y apropiadamente el comportamiento de un sistema físico. Desde un punto de vista técnico esto hace que los cálculos sean mas complicados que lo usual. Si se considera por ejemplo el vidrio de spin donde el desorden aparece como acoplos aleatorios (e.g.  $J_{ij}$  en el modelo de Ising), para describir la termodinámica de este sistema se considera a la energía libre del sistema ya que esta es una variable extensiva y por lo tanto auto-promediable.

La densidad de energía libre  $f_J$  para una realización desordenada  $J$  se puede definir como

$$f_J = -\frac{1}{\beta N} \ln Z_J = -\frac{1}{\beta N} \ln \text{Tr}_{\{S_i\}} \exp(-\beta H_J[S]) \quad (1.42)$$

donde  $Z_J$  es la función de partición del modelo. El valor promedio sobre la distribución del desorden es entonces

$$f = \int dJP(J)f_J = \overline{f_J} = -\frac{1}{\beta N} \overline{\ln Z_J} \quad (1.43)$$

se ha indicado con una barra el promedio sobre la distribución desordenada  $P(J)$ . En la expresión anterior se tiene que calcular el promedio de un logaritmo lo cual no es simple y no es común de realizar. Esto es consecuencia de la naturaleza “apagada” del desorden que requiere utilizar observables extensivos para promediar en lugar de usar por ejemplo la función de partición, por esta razón el promedio 1.43 es conocido en mecánica estadística como promedio “apagado”. Es importante notar que hay dos promedios siendo calculados en secuencia para la expresión 1.43, el primero se trata del promedio termodinámico que se obtienen al promediar sobre la medida de Boltzmann 1.8 y el segundo promedio se hace sobre el desorden. Un computo mucho mas fácil es obtenido al promediar directamente a la función de partición sobre el desorden y después aplicar el logaritmo:

$$f_a n = -\frac{1}{\beta N} \ln \overline{Z_j} = -\frac{1}{\beta N} \ln \int dJP(J) \text{Tr}_{\{S_i\}} \exp(-\beta H_J[S_i]) \quad (1.44)$$

En este caso la medida de Boltzmann y la distribución del desorden aparecen con la misma jerarquía y los dos promedios se realizan simultáneamente. Este procedimiento es adecuado en el caso que se tiene un desorden recocido, es por esta razón que la expresión 1.44 se conoce como promedio recocido.

## 1.7. El modelo SK

El modelo de réplicas o modelo de Sherrington y Kirkpatrick(SK) del vidrio de spins está descrito por el Hamiltoniano

$$H = - \sum_{i < j}^N J_{ij} \sigma_i \sigma_j \quad (1.45)$$

. Las interacciones spin-spin  $J_{ij}$  en este sistema son las variables aleatorias apagadas que son independientes para cualquier par de sitios  $(i, j)$  y que están descritos por la distribución Gaussiana:

$$P[J_{ij}] = \prod_{i < j} \left[ \sqrt{\frac{N}{2\pi}} \exp \left( -\frac{J_{ij}^2 N}{2} \right) \right] \quad (1.46)$$

De acuerdo a las definiciones anteriores, cada spin puede interaccionar con cualquier otro spin del sistema, por esta razón en este modelo no existe el concepto de estructura del espacio: dimensionalidad, tipo de retículo, vecinos cercanos, etc. El espacio en este modelo es solo el conjunto de los  $N$  sitios en donde se han situado a los spins de Ising. Se puede interpretar al modelo SK en el límite en el que  $N \rightarrow \infty$  como el sistema de retículo de spins en un espacio dimensional infinito.

De acuerdo con la distribución de probabilidad 1.46 se obtiene

$$\ll J_{ij} \gg = 0 ; \quad \ll J_{ij}^2 \gg = \frac{1}{N} \quad (1.47)$$

donde  $\ll \dots \gg$  significa que se promedia sobre las realizaciones de las variables aleatorias  $J_{ij}$ 's y esta definido por

$$\ll (\dots) \gg \equiv \int D\mathbf{J} P[\mathbf{J}] (\dots) = \prod_{i < j} \left[ \sqrt{\frac{N}{2\pi}} \int_{-\infty}^{+\infty} dJ_{ij} \exp \left\{ -\frac{J_{ij}^2 N}{2} \right\} \right] (\dots) \quad (1.48)$$

La expresión anterior es equivalente a la expresión en 1.43. Se puede verificar que debido a la normalización del orden  $1/N$  para el valor promedio del cuadrado de los  $J_{ij}$  en 1.47, el promedio de la energía del sistema estimada usando 1.45 es del orden  $N$ , como debería de serlo para un sistema físico definido apropiadamente. Esta claro que la estructura microscópica definida en el modelo SK es una concepción teórica altamente no física, sin embargo este modelo tiene dos grandes ventajas: tiene una solución exacta y los resultados obtenidos de la solución exacta son no triviales y pueden ser generalizadas para sistemas físicos aleatorios “normales”. Una característica importante del modelo SK es el hecho que tiene una solución exacta y la razón de esto no debería asombrar ya que en un espacio con dimensionalidad infinita la de campo medio es exacta.

### 1.7.1. Método de Rélicas

Se puede introducir el método de replicas(ver Apéndice A.1 ) como sigue, la energía libre(auto-promediable) del sistema apagado se obtiene al promediar sobre la aleatoriedad del logaritmo de la función de partición:

$$F \equiv \ll F_J \gg = -\frac{1}{\beta} \ll \ln Z_J \gg \quad (1.49)$$

donde la función de partición

$$Z_J = \sum_{\sigma} \exp\{-\beta H[\sigma]\} \quad (1.50)$$

debe de ser calculada para las variables aleatorias fijas  $J_{ij}$ . Para realizar el procedimiento de promediar se inventó la siguiente técnica o método([9], [32], [36]). Si se considera una potencia entera  $n$  de la función de partición 1.50, se tendría la función de partición de  $n$  replicas idénticas del sistema original sin interacción, todas las replicas tienen valor idéntico de interacción spin-spin  $J_{ij}$ :

$$\begin{aligned} Z_J^n &= \left( \sum_{\sigma^1} \exp\{-\beta H[\sigma^1]\} \right) \left( \sum_{\sigma^2} \exp\{-\beta H[\sigma^2]\} \right) \cdots \left( \sum_{\sigma^n} \exp\{-\beta H[\sigma^n]\} \right) \\ Z_J^n &= \sum_{\sigma^1} \sum_{\sigma^2} \cdots \sum_{\sigma^n} \exp \left\{ -\beta \sum_{a=1}^n H[\sigma^a] \right\} \\ Z_J^n &= \left( \prod_{a=1}^n \sum_{\sigma^a} \right) \exp \left\{ -\beta \sum_{a=1}^n H[\sigma^a] \right\} \end{aligned} \quad (1.51)$$

Utilizando 1.51 y 1.45 se obtiene

$$Z_J^n = \left( \prod_{a=1}^n \sum_{\sigma^a} \right) \exp \left\{ \beta \sum_{a=1}^n \sum_{i < j}^N J_{ij} \sigma_i^a \sigma_j^a \right\} \quad (1.52)$$

donde el superíndice  $a$  etiqueta a las replicas. Es útil introducir a la siguiente cantidad:

$$F_n = -\frac{1}{\beta n} \ln Z_n \quad (1.53)$$

donde

$$Z_n \equiv \ll Z_J^n \gg \quad (1.54)$$

Si se toma el límite  $n \rightarrow \infty$  de la expresión 1.53 se recobra la expresión original para la energía libre 1.49, ver A.8

$$\lim_{n \rightarrow 0} F_n = -\frac{1}{\beta} \ll \ln Z_J \gg = F \quad (1.55)$$

De tal manera el esquema del método de replicas es el siguiente, primero se calcula la cantidad  $F_n$  para el entero  $n$ , después se hace la continuación analítica de la función obtenida del parámetro  $n$ , para un  $n$  arbitrario no entero. Después de esto se halla el límite cuando  $n \rightarrow 0$ .

### 1.7.2. Cálculo de la Energía Libre

Para calcular la energía libre de las replicas  $F_n$ , de acuerdo a las expresiones 1.53 y 1.54, se tiene que calcular el promedio “recocido” (*annealed average*) de la enésima función de partición:

$$Z_n = \sum_{\sigma_i^a} \int D J_{ij} \exp \left\{ \beta \sum_{a=1}^n \sum_{i < j}^N J_{ij} \sigma_i^a \sigma_j^a - \frac{1}{2} \sum_{i < j} J_{ij}^2 N \right\} \quad (1.56)$$

integrando con respecto a  $J_{ij}$  utilizando

$$\int_{-\infty}^{+\infty} e^{-\gamma x^2 + \alpha x} dx = \sqrt{\frac{\pi}{\gamma}} e^{\frac{\alpha^2}{4\gamma}} \quad (1.57)$$

se obtiene

$$Z_n = \sum_{\sigma_i^a} \exp \left\{ \frac{\beta^2}{2N} \sum_{i < j}^N \left( \sum_{a=1}^n \sigma_i^a \sigma_j^a \right)^2 \right\} \quad (1.58)$$

o equivalentemente se tiene la siguiente expresión donde se han intercambiado el orden de las sumatoria en el exponente

$$Z_n = \sum_{\sigma_i^a} \exp \left\{ \frac{1}{4} \beta^2 N n + \frac{\beta^2 N}{2} \sum_{a < b}^n \left( \frac{1}{N} \sum_{i=1}^N \sigma_i^a \sigma_i^b \right)^2 \right\} \quad (1.59)$$

La sumatoria sobre los sitios en la expresión anterior puede ser linealizada utilizando la expresión 1.57 en sentido opuesto del que se utilizó al introducir una matriz de

replicas  $\hat{Q}$  y sus elementos  $Q_{ab}$  de la siguiente manera:

$$Z_n = \prod_{a < b}^n \left( \int dQ_{ab} \right) \sum_{\sigma_i^a} \exp \left\{ \frac{1}{4} \beta^2 N n - \frac{\beta^2 N}{2} \sum_{a < b}^n Q_{ab}^2 + \beta^2 \sum_{a < b}^n \sum_i^N Q_{ab} \sigma_i^a \sigma_i^b \right\} \quad (1.60)$$

En donde se ha usado 1.57. Se puede hallar una interpretación física para los parámetros  $Q_{ab}$ , de acuerdo con la ecuación anterior los valores de equilibrio de los elementos matriciales  $Q_{ab}$  están definidos por la ecuación  $\delta Z_n / \delta Q_{ab} = 0$ , que resulta en:

$$Q_{ab} = \frac{1}{N} \sum_i^N \langle \sigma_i^a \sigma_i^b \rangle \quad (1.61)$$

Debido a que la expresión en el exponente de 1.60 es lineal en la sumatoria con respecto a  $\sigma_i^a$ , la función de partición se puede escribir utilizando un producto en donde cada factor es la función de partición de cada sitio independiente:

$$Z_n = \prod_{a < b}^n \left( \int dQ_{ab} \right) \exp \left\{ \frac{1}{4} \beta^2 N n - \frac{\beta^2 N}{2} \sum_{a < b}^n Q_{ab}^2 \right\} \prod_i^N \left[ \sum_{\sigma_i^a} \exp \left\{ \beta^2 \sum_{a < b}^n Q_{ab} \sigma_i^a \sigma_i^b \right\} \right] \quad (1.62)$$

usando las propiedades exponenciales y logarítmicas se puede expresar la  $Z_n$  como

$$Z_n = \prod_{a < b}^n \left( \int dQ_{ab} \right) \exp \left\{ \frac{1}{4} \beta^2 N n - \frac{\beta^2 N}{2} \sum_{a < b}^n Q_{ab}^2 + N \ln \left[ \sum_{\sigma_i^a} \exp \left\{ \beta^2 \sum_{a < b}^n Q_{ab} \sigma_i^a \sigma_i^b \right\} \right] \right\} \quad (1.63)$$

Esta ecuación puede representarse como

$$Z_n = \int D\hat{Q} \exp(-\beta n N f[\hat{Q}]) \quad (1.64)$$

donde se tiene la definición

$$f[\hat{Q}] = -\frac{1}{4} \beta + \frac{\beta}{2n} \sum_{a < b}^n Q_{ab}^2 - \frac{1}{\beta n} \ln \left[ \sum_{\sigma_i^a} \exp \left\{ \beta^2 \sum_{a < b}^n Q_{ab} \sigma_i^a \sigma_i^b \right\} \right] \quad (1.65)$$

Con esto se ha transferido el problema al espacio de replicas.

En el límite termodinámico en el término dominante para  $N$ , la integral de en la función de partición 1.64 se puede aproximar por el punto de ensilladura de la función  $f[\hat{Q}]$ :

$$Z_n \simeq \left[ \det \frac{\delta^2 f}{\delta \hat{Q}^2} \right]^{(-1/2)} \exp(-\beta n N f[\hat{Q}^*]) \quad (1.66)$$

Donde  $\hat{Q}^*$  es la matriz que corresponde al mínimo de la función  $f$  y que esta definido por la ecuación de punto de ensilladura

$$\frac{\delta f}{\delta Q_{ab}} = 0 \quad (1.67)$$

Al revisar el esquema general del método de replicas se puede notar que la cantidad  $f[\hat{Q}]$  representaría la densidad de la energía libre del sistema. Del desarrollo anterior se tiene entonces que el procedimiento para resolver el problema consiste en encontrar la expresión para la energía libre 1.65 como función de  $\hat{Q}$ , después se tiene que resolver la ecuación de punto de ensilladura 1.67 para obtener  $\hat{Q}^*$  y el valor correspondiente de  $f[\hat{Q}^*]$  finalmente se toma el limite cuando  $n \rightarrow \infty$ , sin embargo se tiene el inconveniente que en general el procedimiento anterior no se puede completar, ya que para una matriz arbitraria  $\hat{Q}$  la expresión 1.65 no se puede calcular. También hay algunas consideración a tomar en el limite cuando  $n \rightarrow \infty$  con respecto a que soluciones de 1.67 deben de tomarse en cuenta. Para poder solucionar entonces el problema se han propuesto procedimientos en los que simplemente se “conjetura” la forma correcta de la matriz  $\hat{Q}$  que dependería de un numero finito de parámetros y después de esto estos parámetros se obtienen a través de ecuaciones de ensilladura como 1.67, con esto es posible hallar el punto los extremos dentro de un sub-espacio de espacio de las matrices  $\hat{Q}$ . Si ademas fuera posible probar que la matriz Hessiana  $\frac{\delta^2 f}{\delta \hat{Q}^2}$  está definida positivamente entonces se tendría la certeza que se ha hallado el punto de interés, se estudian algunos casos en las siguientes secciones.

### 1.7.3. Solución simétrica de réplicas

Debido a que todas la replicas del sistema son equivalentes, se puede asumir inocentemente que la forma adecuada para la matriz  $\hat{Q}^*$  es tal que todos sus elementos son iguales.

$$Q_{ab} = q; \quad \forall a \neq b \quad (1.68)$$

La matriz de replicas  $\hat{Q}$  es una matriz simétrica( $Q_{ab} = Q_{ba}$ ) de dimensiones  $n \times n$ , con ceros en la diagonal, este ansatz es conocido como la aproximación de replicas simétrica. De hecho esta es solo una hipótesis que dice que solo existe un estado base en el sistema. Todos los cálculos para la aproximación de replicas simétricas

son relativamente directos, para la energía libre 1.65 se obtiene

$$f(q) = -\frac{1}{4}\beta + \frac{\beta}{2n} \frac{n(n-1)}{2} q^2 - \frac{1}{\beta n} \ln \left[ \sum_{\sigma_a} \exp \left\{ \frac{1}{2}\beta^2 (\sum_a^n \sigma_a)^2 q - \frac{1}{2}\beta^2 nq \right\} \right] \quad (1.69)$$

o equivalentemente utilizando resultados de la integral Gaussiana se obtiene

$$f(q) = -\frac{1}{4}\beta + \frac{1}{2}\beta q + \frac{1}{4}(n-1)\beta q^2 - \frac{1}{\beta n} \ln \left[ \int_{-\infty}^{+\infty} \frac{dz}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \prod_{a=1}^n \left( \sum_{\sigma_a=\pm 1} \exp\{\beta\sigma_a\sqrt{q}z\} \right) \right] \quad (1.70)$$

tomando  $\lim_{n \rightarrow 0} f(q)$  se obtiene

$$f(q) = -\frac{1}{4}\beta(1-q)^2 - \frac{1}{\beta} \int_{\infty}^{+\infty} \frac{dz}{\sqrt{(2\pi)}} \exp\left\{-\frac{1}{2}z^2\right\} \ln(2 \cosh(\beta\sqrt{q}z)) \quad (1.71)$$

La ecuación de ensilladura para la función  $f(q)$  con respecto a  $q$  da como resultado

$$q = \int_{-\infty}^{+\infty} \frac{dz}{2\pi} \exp\left(-\frac{1}{2}z^2 \tanh^2(\beta\sqrt{q}z)\right) \quad (1.72)$$

Recordando  $\beta = 1/T$ , se tiene de 1.72 que cuando  $T > 1$  la única solución es  $q = 0$ . Para  $T > 1$  hay soluciones no triviales  $q \neq 0$ . Cuando  $(1-T) \equiv \tau \ll 1$  entonces  $q(\tau) \simeq \tau$ . Cuando  $T \rightarrow 0$  entonces  $q \rightarrow 1$ .

De acuerdo con 1.61, la solución  $q(T)$  es el parámetro de orden físico:

$$q = \frac{1}{N} \sum_i^N \langle \sigma_i \rangle^2 \quad (1.73)$$

El hecho que  $q$  no sea igual a cero en la fase de temperaturas bajas significa que los spins del sistema están “congelados” en un estado aleatorio. Ya que solo una solución para  $q(T)$  para el régimen de temperaturas bajas, se tiene que ese estado desordenado es el estado base y es único. De estos resultados se pueden obtener todas las cantidades termodinámicas correspondientes, como el calor específico, la entropía, la susceptibilidad etc. Esto aparenta que el ansatz considerado para el sistema de replicas produce una solución completa al problema, sin embargo no es una solución consistente. Unos de los problemas mas fáciles de ver con la solución obtenida es calcular la entropía del sistema. Se puede hallar relativamente fácil que la entropía del sistema a temperaturas bajas se hace negativa, en  $T = 0$  la entropía es  $S = -\frac{1}{2\pi} \simeq -0.17$ . La razón para esta paradoja reside en el hecho que el

determinante de la matriz Hessiana es negativa  $\det(\delta^2 f / \delta \hat{Q}^2)$ , esto nos dice que esta solución es inestable en el régimen de temperaturas bajas. Detalles de estos cálculos pueden encontrarse en [9]. Esto sugiere que la verdadera solución está mas allá del subespacio de replicas simétricas de las matrices  $\hat{Q}$ .

#### 1.7.4. Ruptura de la simetría de réplicas de Parisi

La estrategia para encontrar la verdadera solución a la matriz  $\hat{Q}$  en el modelo de replicas en el límite cuando  $n \rightarrow \infty$  es conocido como el esquema de ruptura en la simetría de replicas de Parisi [36]. El primer paso del esquema de Parisi consiste en hallar otra estructura de prueba para la matriz  $\hat{Q}$  pero sin simetría de replicas, esto formaría un nuevo subespacio para una posible solución, dentro de dicho subespacio se busca calcular el extremo de la energía libre de replicas  $f$ . Después se debe calcular a la matriz Hessiana  $\frac{\delta^2 f}{\delta \hat{Q}^2}$  para verificar la estabilidad de la solución obtenida. Debido al fracaso del modelo de replicas expuesto anteriormente la estrategia de Parisi consiste en hacer un ansatz con una estructura matricial que contenga mas parámetros. La situación es mas sofisticada de lo que parece ya que como se explicara brevemente, no hay ansatz con un numero finito de parámetros que den una solución estable. Sin embargo el esquema tiene como fin probar distintas estructuras de prueba para  $\hat{Q}$  y calcular los distintos eigenvalores de la matriz Hessiana para determinar que estructuras son mas estables, este proceso iterativo resulta como una guía para determinar que “dirección” tomar en el espacio de las matrices  $\hat{Q}$  con el propósito de hallar la verdadera solución u obtener una aproximación con precision arbitraria. El modelo de Parisi a veces conocido como RSB, se resume en la secuencia infinita de ansatz que aproximan a la verdadera solución con cada búsqueda de mejores estructuras matriciales, con esto se puede formular a la solución verdadera con una descripción de funciones continuas que es obtenida en el límite de la secuencia infinita en los subespacios de  $\hat{Q}$ . Como ganancias del esquema de Parisi se tiene que se puede verificar a la estabilidad de la solución obtenida y ademas es posible definir cantidades físicas simples que hacen mas explicito demostrar la física detrás de este modelo iterativo. Después de haber expuesto paso a paso el proceso de Parisi se entiende mejor el porque de este tipo de acercamiento para la solución del problema.

El primer paso consiste en dividir  $n$  replicas en  $n/m$  grupos con  $m$  replicas en cada grupo donde hasta el momento se asume que  $n$  y  $m$  son enteros. Después se procede a definir a la matriz  $\hat{Q}$  como  $Q_{ab} = q_1$  si las replicas  $a$  y  $b$  pertenecen al mismo

grupo, y  $Q_{ab} = q_0$  si las replicas  $a$  y  $b$  pertenecen a distintos grupos, los elementos diagonales son de nuevo cero 1.68 de manera compacta se puede escribir como:

$$Q_{ab} = \begin{cases} q_0, & I\left(\frac{a}{m}\right) \neq I\left(\frac{b}{m}\right) \\ q_1, & I\left(\frac{a}{m}\right) = I\left(\frac{b}{m}\right) \end{cases} \quad (1.74)$$

Donde  $I(x)$  es la función que devuelve al entero mas pequeño que sea mayor o igual a  $x$ , en el lenguaje python por ejemplo seria la función `int()`. La estructura de la matriz de la matriz  $Q_{ab}$  para el caso  $n = 8, m = 4$  es la siguiente

$$\left( \begin{array}{cccc|c} 0 & q_1 & q_1 & q_1 & q_0 \\ q_1 & 0 & q_1 & q_1 & \\ q_1 & q_1 & 0 & q_1 & \\ q_1 & q_1 & q_1 & 0 & \\ \hline & & & & \begin{array}{cccc} 0 & q_1 & q_1 & q_1 \\ q_1 & 0 & q_1 & q_1 \\ q_1 & q_1 & 0 & q_1 \\ q_1 & q_1 & q_1 & 0 \end{array} \\ q_0 & & & & \end{array} \right) \quad (1.75)$$

Para  $n = 16, m = 4$  se tiene la siguiente forma matricial:

$$\left( \begin{array}{c|ccccc}
0 & q_1 & q_1 & q_1 & & \\
q_1 & 0 & q_1 & q_1 & q_0 & \\
q_1 & q_1 & 0 & q_1 & & \\
q_1 & q_1 & q_1 & 0 & & \\ \hline
& 0 & q_1 & q_1 & q_1 & q_0 \\
& q_0 & q_1 & 0 & q_1 & \\
& q_1 & q_1 & 0 & q_1 & \\
& q_1 & q_1 & q_1 & 0 & \\ \hline
& 0 & q_1 & q_1 & q_1 & q_0 \\
& q_0 & q_1 & 0 & q_1 & \\
& q_1 & q_1 & 0 & q_1 & \\
& q_1 & q_1 & q_1 & 0 & \\ \hline
& 0 & q_1 & q_1 & q_1 & q_0 \\
& q_0 & q_1 & 0 & q_1 & \\
& q_1 & q_1 & 0 & q_1 & \\
& q_1 & q_1 & q_1 & 0 & 
\end{array} \right) \quad (1.76)$$

Para este ansatz se tienen 3 parámetros  $q_0$ ,  $q_1$  y  $m$  que son definidos a través de las ecuaciones de ensilladura correspondientes. Utilizando la forma explícita de  $\hat{Q}$  en 1.65 se obtiene

$$f[\hat{Q}] = -\frac{1}{4}\beta + \frac{\beta}{2n} \sum_{a<b}^n Q_{ab}^2 - \frac{1}{\beta n} \ln Z([\hat{Q}]) \quad (1.77)$$

donde

$$Z([\hat{Q}]) = \sum_{\sigma_a} \exp \left( \beta^2 \sum_{a<b}^n Q_{ab} \sigma_a \sigma_b \right) \quad (1.78)$$

Se pueden utilizar cálculos directos para simplificar

$$\sum_{a<b}^n Q_{ab} \sigma_a \sigma_b = \frac{1}{2} \left[ q_0 \left( \sum_{a=1}^n \sigma_a \right)^2 + (q_1 - q_0) \sum_{k=1}^{n/m} \left( \sum_{c_k=1}^m \sigma_{c_k} \right)^2 - nq_1 \right] \quad (1.79)$$

Donde  $k$  numera a cada bloque de replicas y  $c_k$  numera a cada replica dentro de cada bloque. Se utilizan transformaciones con la conocida integral Gaussiana para la expresión 1.78 en los términos cuadrados y se obtiene

$$Z[q_1, q_0, m] = \int \frac{dz}{\sqrt{2\pi q_0}} \exp\left(-\frac{z^2}{2q_0}\right) \prod_{k=0}^{n/m} \left( \int \frac{dy_k}{\sqrt{2\pi(q_1 - q_0)}} \exp\left(-\frac{y_k^2}{2(q_1 - q_0)}\right) \times \sum_{\sigma_a} \exp\left\{ \beta \left[ z \sum_a^n \sigma_a + \sum_{k=0}^{n/m} y_k \left( \sum_{c_k=1}^m \sigma_{c_k} \right) \right] - \frac{1}{2} \beta^2 n q_1 \right\} \right) \quad (1.80)$$

La suma sobre spins resulta en

$$Z[q_1, q_0, m] = \int \frac{dz}{\sqrt{2\pi q_0}} \exp\left(-\frac{z^2}{2q_0}\right) \left[ \int \frac{dy}{\sqrt{2\pi(q_1 - q_0)}} \exp\left(-\frac{y^2}{2(q_1 - q_0)(2 \cosh \beta(z + y))^m}\right) \times \exp\left(-\frac{1}{2} \beta^2 n q_1\right) \right]^{n/m} \quad (1.81)$$

Para el segundo término en 1.77 se tiene que

$$\frac{\beta}{2n} \sum_{a < b}^n Q_{ab}^2 = \frac{\beta}{4} [q_1^2(m-1) + q_0^2(n-m)] \quad (1.82)$$

Después de esto se tiene que tomar el límite cuando  $n \rightarrow \infty$  para las ecuaciones 1.81 y 1.82 y sustituirlas en la ecuación 1.77 dando como resultado:

$$f(q_1, q_0, m) = -\frac{1}{4} \beta (1 + mq_0^2 + (1-m)q_1^2 - 2q_1) - \frac{1}{m\beta} \int \frac{dz}{\sqrt{2\pi q_0}} \exp\left(-\frac{z^2}{2q_0}\right) \ln \left[ \frac{dy}{\sqrt{2\pi(q_1 - q_0)}} \exp\left(-\frac{y^2}{2(q_1 - q_0)}\right) (\cosh \beta(z + y)^m) \right] \quad (1.83)$$

Del último resultado es fácil verificar que se obtiene de vuelta el modelo de replicas simétricas cuando  $m = 0$  y  $m = 1$  con  $q = q_0$  y  $q = q_1$  respectivamente. Algo importante a notar aquí es el hecho que en este caso se busca el máximo y no el mínimo de la energía libre, la razón formal reside en el hecho que en el límite cuando  $n \rightarrow \infty$  el numero de componentes de del parámetro de orden  $\hat{Q}$  se hace negativo, por ejemplo en el caso de la ruptura de simetría de un paso cada linea de la matriz  $\hat{Q}$  contiene  $(m-1) < 0$  componentes que son igual a  $q_1$  y  $(n-m) \rightarrow -m < 0$  componentes iguales a  $q_0$ . Este fenómeno puede ejemplificarse también en el caso en el que la energía libre contiene solo al termino  $\frac{\beta}{2n} \sum_{a < b}^n Q_{ab}^2$  sin ningún termino de interacción que es el caso para el límite de altas temperaturas, entonces se puede

ver de

$$\lim_{n \rightarrow \infty} \left[ \frac{\beta}{n} \sum_{a < b} Q_{ab}^2 \right] = -\frac{\beta}{2} [(1-m)q_1^2 + mq_0^2] \quad (1.84)$$

que el “extremo correcto” en el que la matriz Hessiana esta definida positivamente es el máximo y no el mínimo con respecto a  $q_0$  y  $q_1$ .

Para obtener los puntos de ensilladura se tienen que seguir los pasos estándar, encontrar las derivadas con respecto a  $q_0$ ,  $q_1$  y  $m$  igualar a cero y resolver, estos resultados se resumen en la siguiente lista

1. En la fase de bajas temperaturas  $T < 1$  la función  $f$  tiene de hecho un máximo en donde  $0 \leq m(T) \leq 1$ ;  $0 \leq q_0(T) \leq 1$ ;  $0 \leq q_1(T) \leq 1$ , tanto para  $T \rightarrow 1$  y para  $T \rightarrow 0$  se obtiene que  $m(T) \rightarrow 0$ .
2. Aunque la entropía a bajas temperaturas se hace negativa de nuevo,  $S(T = 0) \simeq -0.01$ , en este caso el valor es mucho mas pequeño que para el caso de la solución de replicas simétricas  $S(T = 0) \simeq -0.17$ .
3. El eigenvalor mas negativo de la matriz Hessiana cerca de  $T_c$  es igual a  $-\frac{c(T-T_c)^2}{9}$ , mientras que para la solución de replicas simétricas se tiene que el eigenvalor mas negativo cercano a  $T_c$  es  $-c(T - T_c)^2$ , por lo que en cierto sentido la inestabilidad se ha reducido por un factor de 9.

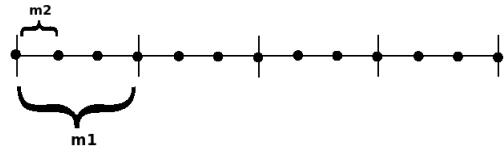
Resumiendo los resultados de la ruptura de simetría de replicas se tiene que aunque no se obtuvo una solución perfecta, se obtuvo una mejor aproximación de la solución que para las replicas simétricas, la estrategia es entonces moverse mas lejos cada vez en la “dirección” escogida en los subespacios de replicas.

### 1.7.5. Ruptura de la simetría de replicas de escala completa

Se continua el camino de la sección anterior para tratar de generalizar la estructura de la matriz  $\hat{Q}$  para mas pasos de la ruptura de simetría de replicas. Se inicia por introducir una serie de numero enteros  $m_i$  ( $i = 1, 2, \dots, k+1$ ) tal que  $m_0 = n$ ,  $m_{k+1} = 1$ , con todas razones  $m_i/m_{i+1}$  enteras. Después se dividen a  $n$  replicas en  $\frac{n}{m_1}$  grupos de tal manera que cada grupo consiste de  $m_1$  replicas, despues se divide nuevamente cada grupo de  $m_1$  replicas en  $\frac{m_1}{m_2}$  subgrupos con  $m_2$  replicas en cada subgroupo y así sucesivamente ver fig. 1.4.

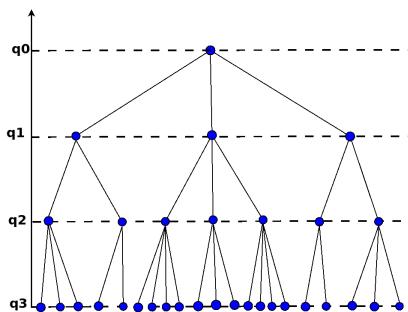
Todos los elementos no diagonales de la matriz  $\hat{Q}$  se definen de la siguiente manera:

$$Q_{ab} = q_i, \text{ para } I\left(\frac{a}{m_i + 1}\right) = I\left(\frac{b}{m_i + 1}\right) \quad (1.85)$$



**Figura 1.4.** Clasificación de replicas en bloques para la ruptura de la simetría de replicas de dos pasos.

donde  $\{q_i\}$  es un conjunto de  $(k + 1)$  parámetros ( $k = 1$  corresponde a la ruptura de simetría de replicas de una paso tratada anteriormente). Esta definición de elementos matriciales puede ser representada también de manera gráfica utilizando un árbol jerárquico como se ve en la figura 1.5, este árbol jerárquico tiene una arquitectura similar a la que se utiliza en las redes neuronales profundas también conocidos como perceptrones de varias capas. La forma explícita de la matriz  $\hat{Q}$  para el caso  $k = 2$



**Figura 1.5.** Esquema de definición de los elementos matriciales  $Q_{ab}$  para la ruptura de simetría de replicas de dos pasos.

es:

$$\left( \begin{array}{cc|c|c}
 0 & q_2 & q_1 & q_1 \\
 q_2 & 0 & & \\
 \hline
 q_1 & 0 & q_2 & q_1 \\
 & q_2 & 0 & \\
 \hline
 q_1 & q_1 & 0 & q_2 \\
 & & q_2 & 0
 \end{array} \right) \xrightarrow{q_0} \left( \begin{array}{cc|c|c}
 0 & q_2 & q_1 & q_1 \\
 q_2 & 0 & & \\
 \hline
 q_1 & 0 & q_2 & q_1 \\
 & q_2 & 0 & \\
 \hline
 q_1 & q_1 & 0 & q_2 \\
 & & q_2 & 0
 \end{array} \right) \xrightarrow{q_0} \left( \begin{array}{cc|c|c}
 0 & q_2 & q_1 & q_1 \\
 q_2 & 0 & & \\
 \hline
 q_1 & 0 & q_2 & q_1 \\
 & q_2 & 0 & \\
 \hline
 q_1 & q_1 & 0 & q_2 \\
 & & q_2 & 0
 \end{array} \right) \xrightarrow{q_0} \left( \begin{array}{cc|c|c}
 0 & q_2 & q_1 & q_1 \\
 q_2 & 0 & & \\
 \hline
 q_1 & 0 & q_2 & q_1 \\
 & q_2 & 0 & \\
 \hline
 q_1 & q_1 & 0 & q_2 \\
 & & q_2 & 0
 \end{array} \right) \xrightarrow{q_0} \left( \begin{array}{cc|c|c}
 0 & q_2 & q_1 & q_1 \\
 q_2 & 0 & & \\
 \hline
 q_1 & 0 & q_2 & q_1 \\
 & q_2 & 0 & \\
 \hline
 q_1 & q_1 & 0 & q_2 \\
 & & q_2 & 0
 \end{array} \right) \xrightarrow{(1.86)}$$

El siguiente paso para la ruptura de escala completa es calcular la energía libre de las ecuaciones 1.77 y 1.77 que dependen de  $(k+1)$  parámetros  $q_i$  y  $k$  parámetros  $m_i$ . Después de esto se toma el límite  $n \rightarrow \infty$ . Los parámetros  $q_i$  están definidos de tal manera que  $1 \leq m_{i+1} \leq m_i \leq n$ , de nuevo se tiene estas condiciones se convierten a  $0 \leq m_i \leq m_{i+1} \leq 1$  en el caso  $n \rightarrow \infty$ . Como ultimo paso se tiene que hallar a los parámetros  $q_i$  y  $m_i$  por medio de las ecuaciones de ensilladura correspondientes.

$$\frac{\partial f}{\partial q_i} = 0; \quad \frac{\partial f}{\partial m_i} = 0 \tag{1.87}$$

El cálculo explícito de la energía libre es análogo al caso tratado anteriormente en la sub-sección 1.7.4. El resultado es el siguiente:

$$\begin{aligned}
f[q_0, q_1, \dots, q_k; m_1, m_2, \dots, m_k] = & \\
& -\frac{1}{4}\beta\left[1 + \sum_{i=1}^k(m_{i+1} - m_i)q_i^2 - 2q_k\right] - \\
& \frac{1}{m_1} \int \frac{dz_0}{\sqrt{2\pi q_0}} \exp\left(-\frac{z_0^2}{2q_0}\right) \ln \left\{ \int \frac{dz_1}{\sqrt{2\pi(q_1 - q_0)}} \exp\left(-\frac{z_1^2}{2(q_1 - q_0)}\right) \times \right. \\
& \left[ \int \frac{dz_2}{\sqrt{2\pi(q_2 - q_1)}} \exp\left(-\frac{z_2^2}{2(q_2 - q_1)}\right) \right] \cdots \left[ \int \frac{dz_k}{\sqrt{2\pi(q_k - q_{k-1})}} \exp\left(-\frac{z_k^2}{2(q_k - q_{k-1})}\right) \right] \times \\
& \left. (2 \cosh \beta (\sum_{i=0}^k z_k))^{m_k} \right]^{m_{k-1}/m_k} \cdots \left. \right]^{m_2/m_3} \left. \right]^{m_1/m_2} \quad (1.88)
\end{aligned}$$

Lamentablemente no es posible obtener soluciones a las ecuaciones de ensilladura correspondiente para la ecuación anterior para un  $k$  arbitrario, para  $k = 3$  la solución numérica para la entropía a temperatura cero  $S(T = 0) \simeq -0.003$ . Para obtener la solución entonces se debe tomar el límite  $k \rightarrow \infty$ , de esta manera en el esquema del método de ruptura de simetría de replicas se consigue  $n \rightarrow 0$  a través de  $n \rightarrow \infty$ , esto puede parecer contradictoria y extraordinario, sin embargo este truco es algo normal y se considera natural en términos de números p-adic [9], que es quizás el lenguaje natural para tratar las estructuras consideradas en este esquema. Para el límite  $k \rightarrow \infty$  en vez de utilizar un conjunto infinito de parámetros  $q_i$  es conveniente introducir a la función  $q(x)$  definida como:

$$q(x) = q_i, \text{ para } m_i \leq x \leq m_{i+1} \quad (1.89)$$

con esta definición la energía libre se convierte en un funcional de la función  $q(x)$  y el problema se traslada a calcular el máximo de este funcional con respecto de la función  $q(x)$

$$\frac{\partial f}{\partial q(x)} \quad (1.90)$$

Esta la ecuación de ensilladura para el parámetro de orden  $q(x)$ . Cerca a  $T_c$  todos los cálculos pueden llevarse a cabo analíticamente, esto hace posible obtener la forma de función  $q(x)$  explícitamente para obtener una entendimiento cualitativo de los efectos en los cambios de temperatura.

## 1.8. Modelos Estadísticos de Redes Neuronales

Inicialmente el interés de la mecánica estadística en las redes neuronales se da por el potencial de estas redes para modelar el funcionamiento del cerebro humano y el emulamiento de algunas capacidades cognitivas, este es el enfoque tratado en esta sección. Las redes neuronales artificiales con propósitos puramente pragmáticos que no tienen nada que ver con el modelamiento de sistemas biológicos se tratan más adelante. La rama del estudio de las redes neuronales dentro de la teoría del vidrio de spin es relativamente nueva. Sorprendentemente, la teoría simple del vidrio de spin basada principalmente en el modelo SK, que es obviamente una concepción artificial, ha adquirido nueva significancia debido a que parece ser capaz de proveer una descripción cualitativa de la memoria. En los años 90s se demostró convincentemente que los modelos estadísticos de redes neuronales podían imitar bastante bien algunas de las características más generales de la memoria asociativa, la experiencia ganada con dichos modelos puede dar intuiciones del funcionamiento del cerebro. Adicionalmente el proceso de reconstrucción de patrones en estos modelos está relacionado con la computación en paralelo por lo que son de gran interés práctico y científico.

Estos modelos están inspirados en parte por algunos hechos biológicos, tanto experimentales como hipotéticos:

1. El elemento de estructura mínima es la “neurona”, cuyo estado puede ser determinado por una variable real  $\sigma$  que corresponde a la tasa de disparo de una neurona real. Experimentalmente se han identificado dos estados neuronales, inactivo y de disparo, con una transición continua entre dichos estados. Esto se modela regularmente al introducir una función de energía apropiada con dos mínimos, que asegura que para la mayoría del tiempo la neurona está en la vecindad de dichos estados. Por esta razón el modelo simple de Ising es utilizado,  $\sigma = \pm 1$ . A veces se utiliza un representación alternativa, con el cambio de variables  $V = (1 + \sigma)/2$ , que remapea los estados a  $V = 0$ ,  $V = 1$  que está más acorde a la convención en neurobiología en la que  $V = 0$  corresponde a el estado inactivo y  $V = 1$  al estado de disparo.
2. El proceso de reconocimiento de patrones ocurre a través de la dinámica paralela de las neuronas, la configuración inicial de las neuronas corresponde a la imagen visual “actual”, mientras que la configuración estable final de las neuronas corresponde al patrón de memoria “ideal”.

3. La memoria de los patrones ideales memorizados se guarda a través de las propiedades de las interacciones inter-neuronales. En el cerebro real las interacciones inter-neuronales son el resultado de las conexiones neurona-neurona, cuyas propiedades son gobernadas principalmente por la propiedades de las uniones entre las dendritas de una neurona y los axones de otra neurona. Existen al menos dos tipos de uniones, excitatorias e inhibitorias, y existen muchas conexiones, alrededor de  $10^{14}$  en el cerebro de los mamíferos, por lo que la cantidad de información que puede guardarse en ellas es muy grande, probablemente unos  $10^{14}$  bits de información, se asume que las uniones son modificadas en el proceso de aprendizaje de tal manera que guarden la información requerida.

De consideraciones generales, es improbable que el tipo de uniones sinápticas sea muy grande, debido a que se sabe que para una un tipo de unión sináptica dada se tiene un comportamiento bastante estable con respecto a influencias externas. En la mayoría de modelos se inició modelando al tipo de unión por una variable continua  $J$ (de tal manera que hay un numero infinito de uniones), pero estudio numéricos y analíticos de los modelos modificados en los que se hace el cambio  $J \rightarrow \text{sign } J$ , que entonces corresponde a solo dos estados el excitatorio y el inhibitorio, muestran que solo hay pequeñas diferencias entre sus propiedades funcionales.

4. La memoria es de contenido direccionable(*content-addressable*), es decir que el patrón completo se puede recuperar de una pequeña parte del contenido.

Antes de discutir los detalles concretos de los modelos estadísticos de redes neuronales es importante hacer notar que la asunción principal acerca de estos modelos es la simetría en las interacciones neuronales. Esta asunción es totalmente irrealista en sistemas neurales biológicos, en los que las interacciones neuronales son esencialmente asimétricas, sin embargo se ha probado que para modelos como el de Hopfield y sus análogos asimétricos tienen propiedades con diferencias pequeñas, ver [9], [36], [32].

En términos de mecánica estadística el problema de los sistemas de redes neuronales se puede formular al considerar un sistema de  $N$  unidades binarias de Ising  $\{\sigma_i\}$  ( $i = 1, 2, \dots, N$ ) que toman valores  $\pm 1$ . También es posible introducir una

dinámica discreta que tiene la forma

$$\sigma_i(t+1) = \text{sign} \left( \sum_{i \neq j}^N J_{ij} \sigma_j(t) \right) \quad (1.91)$$

donde sign es la función signo y  $(\sum_{i \neq j}^N J_{ij} \sigma_j(t)) \equiv h_i(t)$  son los campos locales que actúan sobre las neuronas y las interacciones  $J_{ij}$  son los parámetros. Si se asume que los acoplamientos  $J_{ij}$  son simétricos se tiene un tipo de dinámica que corresponde al relajamiento de la energía mínima:

$$H = -\frac{1}{2} \sum_{i \neq j}^N J_{ij} \sigma_i \sigma_j \quad (1.92)$$

que ocurre a una temperatura exactamente cero, por lo que no se tiene ruido termal presente. El proceso correspondiente de relajamiento para temperaturas distintas de cero puede ser descrito por el siguiente algoritmo probabilístico discretizado en el tiempo:

$$\sigma_i(t+1) = \begin{cases} \sigma_i(t), & \text{con probabilidad } p_+ = \frac{\exp(\beta \sigma_i(t) h_i(t))}{\exp(\beta \sigma_i(t) h_i(t)) + \exp(-\beta \sigma_i(t) h_i(t))} \\ -\sigma_i(t), & \text{con probabilidad } p_- = 1 - p_+ \end{cases} \quad (1.93)$$

El parámetro  $\beta = 1/T$  determina la intensidad del ruido termal. A diferencia del proceso de relajamiento al introducir la dinámica regida por 1.91 que conduce al mínimo de la energía 1.92, el proceso de relajamiento probabilístico de 1.93 con poco ruido termal conduce a una solución que fluctúa en la vecindad del mínimo. Este tipo de dinámica describe el relajamiento del mínimo de la energía definido de manera usual:

$$F = -\frac{1}{\beta} \ln \left[ \sum_{\sigma} \exp(-\beta H[\sigma]) \right] \quad (1.94)$$

El problema entonces se puede formular como sigue, se fijan  $P$  configuraciones de spin aleatorios, que usualmente se denotan como  $\xi_i^{(\mu)}$  ( $\mu = 1, 2, \dots, P$ ), y se busca definir a los acoplos  $J_{ij}$  de tal manera que el mínimo mas “profundo” de la energía libre sea “paralelo” a los patrones pre-asignados  $\xi_i^{(\mu)}$ . En términos mas estrictos, dados los patrones  $\xi_i^{(\mu)}$  se tienen que definir a los acoplos  $J_{ij}$  de tal manera que para el parámetro de orden

$$m_{\mu} = \frac{1}{N} \sum_i^N \langle \sigma_i \rangle \xi_i^{\mu} \quad (1.95)$$

se obtengan estados en los que uno de los componentes del vector  $m_\mu$  es finito y todos los demás componentes son cero en el límite termodinámico  $N \rightarrow \infty$ .

Si se considera una definición tradicional(en la teoría del vidrio de spin) para los acoplos  $J_{ij}$  que se conoce como regla de aprendizaje Hebbiana:

$$J_{ij} = \frac{1}{N} \sum_i^P \xi_i^\mu \xi_j^\mu \quad (1.96)$$

Al usar la regla de aprendizaje Hebbiana en 1.92, se obtiene un sistema descrito por el Hamiltoniano

$$H = -\frac{1}{2} N \sum_\mu^P \left( \frac{1}{N} \sum_i^N \sigma_i \xi_i^\mu \right)^2 \quad (1.97)$$

El caso en el que  $P = 1$  es relativamente sencillo, de acuerdo al Hamiltoniano 1.97 el sistema tiene solo 2 estados base(ferromagnéticos),  $\sigma_i = \pm \xi_i^\mu$ , por lo tanto si la temperatura es lo suficientemente baja i.e.  $T < T_c$  se tiene que para cualquier estado inicial del sistema el proceso de relajamiento 1.93 recobraría rápidamente al sistema a cualquiera de los estados base(mínimos) de energía.

Para valores arbitrarios de  $P$  se tiene una situación mucho mas complicada, antes de hacer los cálculos termodinámicos para el modelo descrito por el Hamiltoniano 1.97 completos es conveniente estimar hasta que valores del parámetro  $P$  se tienen spin paralelos a los patrones guardados, es decir  $\sigma_i = \xi_i^\mu$ , que sean por lo menos estable para la dinámica mas simple de temperatura cero, 1.91. La estabilidad del patrón  $\mu$  significa que se tiene que satisfacer la siguiente expresión para todos los sitios del sistema

$$\xi_i^\mu = \text{sign} \left( \sum_{j \neq i}^N J_{ij} \xi_j^\mu \right) \quad (1.98)$$

De acuerdo con 1.96 para el campo local  $h_i^\mu \equiv \sum_{i \neq j}^N J_{ij} \sigma_j$  con  $\sigma_j = \xi_j^\mu$  se obtiene

$$h_i^\mu \equiv \sum_{j \neq i}^N J_{ij} \xi_j^\mu = \xi_i^\mu + \frac{1}{N} \sum_{\nu \neq \mu}^P \xi_i^\nu \sum_{j \neq i}^N \xi_j^\nu \xi_j^\mu \quad (1.99)$$

estos campos se pueden representar como:

$$h_i^\mu = \xi_i^\mu + \zeta_i^\mu \quad (1.100)$$

al primer termino se le conoce como la *señal* ya que provee la estabilidad del patrón mientras que al segundo termino se le conoce *ruido* y es debido a este que se produce

la inestabilidad. Si el ruido es pequeño entonces las ecuaciones 1.98 se satisfacen directamente ya que  $h_i^\mu \approx \xi_i^\mu$ . Estimando el valor característico del ruido se tiene que

$$\langle\langle \zeta^2 \rangle\rangle = \frac{1}{N^2} \sum_{\nu,\gamma}^P \sum_{j,l}^N \langle\langle \xi_i^\nu \xi_i^\gamma \xi_j^\nu \xi_j^\mu \xi_l^\gamma \xi_l^\mu \rangle\rangle \quad (1.101)$$

Ya que se asumen que todos los patrones son independientes i.e. no están correlacionados, la única contribución distinta de cero viene de los términos  $\nu = \gamma$  y  $j = l$ , por lo que el resultado de la expresión anterior es

$$\langle\langle \zeta^2 \rangle\rangle = \frac{P}{N} \equiv \alpha \quad (1.102)$$

por lo tanto se tienen que si el numero de patrones se acerca al orden  $N$ , el ruido debido a la interferencia de los patrones puede producir la inestabilidad de los estados con spin paralelo a los patrones. De hecho se sabe que cuando se incrementa el parámetro  $P$  se espera que el “daño” local de los patrones guardados ocurra antes que el numero de patrones se haga de orden  $N$ . El punto importante a notar es que los patrones guardados  $\xi_i^{(\mu)}$  no son las únicas soluciones estacionarias a la ecuación 1.98. Muchas de las soluciones llamadas “espurias” o estados mezclados como  $\xi_i = \text{sign}(\xi_i^{(\mu_1)} + \xi_i^{(\mu_2)} + \xi_i^{(\mu_3)} + \dots)$  pueden parecer ser soluciones estacionarias. De hecho según la ecuación 1.102 si  $1 \ll P \ll N$  el campo de ruido  $\zeta$  se convierte en una variable aleatoria Gaussiana con media igual a cero y varianza igual a  $\sqrt{P/N}$ , con esto se tiene que la probabilidad de hallar un sitio en el que el término de ruido en 1.99-1.100 sea mas grande que la señal es

$$p \approx \sqrt{\frac{N}{2\pi P}} \int_1^\infty d\zeta \exp \left[ -\frac{N}{2P} \zeta^2 \right] \approx \sqrt{\frac{P}{2\pi N}} \exp \left[ -\frac{N}{2P} \right] \quad (1.103)$$

Entonces el numero de sitios donde en el estado estacionario de configuraciones de spin puede ser diferente al patrón guardado es

$$pN \approx \sqrt{\frac{pN}{2\pi}} \exp \left[ -\frac{N}{2P} \right] \quad (1.104)$$

Por lo tanto la probabilidad de encontrar un patrón guardado sin errores es del orden 1 solo cuando el parámetro  $P$  es menor a  $\frac{N}{\ln N}$ .

Para  $P$  más grandes el patrón recuperado difiere del patrón guardado en un número finito de sitios, es más, se hallan un numero grande de patrones distintos y estables en la vecindad del patrón guardado. En términos de la “superficie de energía” se puede

decir que un numero pequeño(locales) de mínimos espurios de energía aparecen en el “fondo” de un mínimo de considerable tamaño que corresponde a un patrón guardado y el sistema siempre se ve atrapado en uno de ellos, en esta situación incluso una pequeña cantidad de ruido térmico es capaz de suavizar los mínimos pequeños dejando intactos los mínimos de mayor tamaño, aunque no es posible recobrar el patrón guardado exacto. Para estos casos la calidad del patrón recobrado es tal que una pequeña fracción de errores está presente en el patrón de salida.

Finalmente se tiene que cuando el número de patrones guardados se acerca al orden  $N$ , de acuerdo con la estimación 1.102, el nivel del ruido se acerca al orden de la señal en todos los sitios del sistema y en este caso no es posible recobrar ningún patrón.

### 1.8.1. El modelo de Hopfield

El modelo de Hopfield se puede definir de acuerdo a la regla de aprendizaje Hebbiana 1.96 y por el Hamiltoniano 1.97, este modelo también se conoce como modelo de Hopfield de redes neuronales. Los cálculos termodinámicos son análogos a los cálculos llevados acabo para el modelo SK(sección 1.7) con la diferencia que el desorden apagado del modelo de Hopfield es debido a los patrones apagados  $\xi_i^\mu$  y no a los acoplos de los spins. La otra diferencia importante a notar es que no se está interesado en el estado del vidrio de spin(también presente en este modelo) sino que se quiere entender el estado con parámetro de orden “ferromagnético” 1.95 distinto de cero, que regularmente se conoce como fase de memoria, o fase de recuperación [9].

Para la función de partición del esquema de replicas promediada sobre los patrones  $\xi_i^\mu$  se obtiene:

$$\langle\langle Z^n \rangle\rangle = \sum_{\xi=\pm 1} \sum_{\sigma=\pm 1} \exp \left\{ \frac{1}{2} \beta N \sum_a^n \sum_\mu^P \left( \frac{1}{N} \sum_i^N \sigma_i^a \xi_i^\mu \right)^2 \right\} \quad (1.105)$$

Donde el subíndice  $a$  se utiliza para etiquetar a las replicas, después de utilizar las propiedades de la integral Gaussiana en 1.105 se obtiene

$$\langle\langle Z^n \rangle\rangle = \sum_{\xi=\pm 1} \sum_{\sigma=\pm 1} \int Dm_\mu^a \exp \left\{ -\frac{1}{2} \beta N \sum_a^n \sum_\mu^P (m_\mu^a)^2 + \beta \sum_a^n \sum_\mu^P \sum_i^N m_\mu^a \sigma_i^a \xi_i^\mu \right\} \quad (1.106)$$

De acuerdo con la ecuación anterior el parámetro  $m_\mu^a$  corresponde al parámetro de orden 1.95. Se tiene interés en el estado en el que uno de los componentes de  $m_\mu^a$  es finito y todos los demás componentes son pequeños(ruido) que de acuerdo con 1.95 debe de ser del orden  $1/\sqrt{N}$ , tomando esto en cuenta se tiene que debido a que el numero específico del componente finito de interés no es de importancia entonces se puede tomar  $\mu = 1$  y obviar el subíndice  $\mu = 1$  de las formulas i.e.  $m_{\mu=1}^a \equiv m_a$ ,  $\xi_i^{\mu=1} \equiv \xi_i$ . Para los traslapes del “ruido” restante se hace un re-escalamiento definido por  $m_\mu^a \rightarrow \frac{m_\mu^a}{\sqrt{N}}$  ( $\mu \neq 1$ ), con estos cambios se obtiene que

$$\langle\langle Z^n \rangle\rangle = \sum_{\xi=\pm 1} \sum_{\sigma=\pm 1} \int Dm_a \exp \left\{ -\frac{1}{2}\beta N \sum_a^n (m_a)^2 + \beta \sum_a^n \sum_i^N m_a \sigma_i^a \xi_i \right\} \times Z_1 \quad (1.107)$$

con  $Z_1$  definido por:

$$Z_1 = \sum_{\xi^{\mu>1}=\pm 1} \int Dm_{\mu>1}^a \exp \left\{ -\frac{1}{2}\beta \sum_a^n \sum_{\mu=2}^P (m_\mu^a)^2 + \frac{\beta}{\sqrt{N}} \sum_a^n \sum_{\mu=2}^P \sum_i^N m_\mu^a \sigma_i^a \xi_i^\mu \right\} \quad (1.108)$$

Después de promediar sobre  $\xi$  se obtiene

$$Z_1 = \int Dm_{\mu>1}^a \exp \left\{ -\frac{1}{2}\beta \sum_a^n \sum_{\mu=2}^P (m_\mu^a)^2 \right\} \times \prod_i^N \prod_{\mu=2}^P \left[ \sum_{\xi_i^\mu=\pm 1} \exp \left( \sum_a^n m_\mu^a \sigma_i^a \xi_i^\mu \right) \right] = \\ \int Dm_{\mu>1}^a \exp \left\{ -\frac{1}{2}\beta \sum_a^n \sum_{\mu=2}^P (m_\mu^a)^2 + \sum_{\mu=2}^P \sum_i^N \ln \cosh \left( \frac{\beta}{\sqrt{N}} \sum_a^n m_\mu^a \sigma_i^a \right) \right\} \quad (1.109)$$

donde se ha usado el hecho que  $\exp \left( \sum_a f(x_a) \right) = \prod_a \exp \left( f(x_a) \right)$  y la definición del coseno hiperbólico. Usando la expansión de Taylor para  $\log \cosh(\dots)$  en 1.109 en potencias de  $1/\sqrt{N}$  se obtiene

$$Z_1 = \int D_{\mu>1}^a \exp \left\{ -\frac{1}{2}\beta \sum_a^n \sum_{\mu=2}^P (m_\mu^a)^2 + \frac{1}{2} \sum_{\mu=2}^P \sum_a^n \sum_b^n m_\mu^a m_\mu^b \left( \frac{1}{N} \sum_i^N \sigma_i^a \sigma_i^b \right) \right\} = \\ \prod_{\mu=2}^P \left[ \int Dm_\mu^a \exp \left\{ -\frac{1}{2}\beta \sum_a^n \sum_b^n m_\mu^a m_\mu^b (\delta_{ab} - \beta Q_{ab}) \right\} \right] \times \prod_{a \neq b} \delta(Q_{ab} - \frac{1}{N} \sum_i^N \sigma_i^a \sigma_i^b) \quad (1.110)$$

donde  $Q_{aa} \equiv 1$ , integrando sobre  $m_\mu^a$  se obtiene(ver apéndice PENDIENTE)

$$Z_1 = \left( \exp \left\{ -\frac{1}{2} \text{Tr} \ln(\hat{1} - \beta \hat{Q}) \right\} \right)^P \prod_{a \neq b} \delta(Q_{ab} - \frac{1}{N} \sum_i^N \sigma_i^a \sigma_i^b) = \\ \exp \left\{ -\frac{\alpha N}{2} \text{Tr} \ln(\hat{1} - \beta \hat{Q}) \right\} \prod_{a \neq b} \delta(Q_{ab} - \frac{1}{N} \sum_i^N \sigma_i^a \sigma_i^b) \quad (1.111)$$

al hacer una redefinición  $r_{ab} \rightarrow i \frac{\alpha \beta^2}{2} r_{ab}$  para la función de partición del esquema de replicas 1.107 se obtiene:

$$\langle \langle Z^n \rangle \rangle = \sum_{\xi=\pm 1} \sum_{\sigma=\pm 1} \int Dm_a \int D\hat{Q} \int D\hat{r} \exp \left\{ -\frac{1}{2} \beta N \sum_a^n (m_a)^2 - \frac{N}{2} \alpha \beta^2 \sum_{a \neq b} r_{ab} Q_{ab} \right. \\ \left. - \frac{\alpha N}{2} \text{Tr} \ln(\hat{1} - \beta \hat{Q}) + \beta \sum_a^n \sum_i^N m_a \sigma_i^a \xi_i + \frac{1}{2} \alpha \beta^2 \sum_i^N \sum_{a \neq b} r_{ab} \sigma_i^a \sigma_i^b \right\} \quad (1.112)$$

debido a que la en la expresión anterior la suma espacial es lineal se puede representar como

$$\exp \left\{ N \ln \left[ \sum_{\alpha} \exp \left( \beta \sum_a^n m_a \sigma^a \xi + \frac{1}{2} \alpha \beta^2 \sum_{a \neq b} r_{ab} \sigma^a \sigma^b \right) \right] \right\} \quad (1.113)$$

redefiniendo  $m_a \rightarrow \xi m_a$ , se obtiene la siguiente expresión para la función de partición del método de replicas:

$$\langle \langle Z^n \rangle \rangle = \int Dm_a \int D\hat{Q} \int D\hat{r} \exp \left\{ -\beta n N F[m_a, \hat{Q}, \hat{r}] \right\} \quad (1.114)$$

donde la energía libre de la replicas  $F[m_a, \hat{Q}, \hat{r}]$  es

$$F[m_a, \hat{Q}, \hat{r}] = \frac{1}{2n} \sum_a^n (m_a)^2 + \frac{1}{2n} \alpha \beta \sum_{a \neq b} r_{ab} Q_{ab} + \frac{\alpha}{2\beta n} \text{Tr} \ln(\hat{1} - \beta \hat{Q}) - \\ \frac{1}{\beta n} \ln \left[ \sum_{\sigma} \exp \left( \beta \sum_a^n m_a \sigma^a + \frac{1}{2} \alpha \beta^2 \sum_{a \neq b} r_{ab} \sigma^a \sigma^b \right) \right] \quad (1.115)$$

siguiendo el esquema de replicas el siguiente paso para hallar la energía libre física se tiene que hallar las soluciones a las ecuaciones de ensilladura para  $m_a$ ,  $\hat{Q}$  y  $\hat{r}$  y hallar en límite cuando  $n \rightarrow 0$ . El siguiente paso es hallar la energía libre para la aproximación de replicas simétrica, asumiendo que para las replicas las siguientes

definiciones

$$\begin{aligned} Q_{ab} &= q, \text{ para } a \neq b \\ r_{ab} &= r, \text{ para } a \neq b \\ m_a &= m, \text{ para cualquier } a \end{aligned} \quad (1.116)$$

con los elementos diagonales  $Q_{aa} \equiv 1$ , con este ansatz para la matriz  $Q_{ab}$  se puede hallar que la matriz  $(\hat{1} - \beta \hat{Q})$  tiene un eigenvalor igual a  $1 - \beta + (1 - n)\beta q$  y  $(n - 1)$  eigenvalores iguales a  $1 - \beta + \beta q$ , con esto se puede hallar la traza en la expresión 1.115 y después el límite respectivo:

$$\begin{aligned} \lim_{n \rightarrow 0} \frac{1}{n} \text{Tr} \ln(\hat{1} - \beta \hat{Q}) &= \lim_{n \rightarrow 0} \frac{1}{n} \left[ \ln(1 - \beta + (1 - n)\beta q) + (n - 1) \ln(1 - \beta + \beta q) \right] = \\ &\ln(1 - \beta + \beta q) - \frac{\beta q}{1 - \beta + \beta q} \quad (1.117) \end{aligned}$$

Para el primer término de 1.115 se obtiene:

$$\lim_{n \rightarrow 0} \frac{1}{n} n(n - 1)qr = -qr \quad (1.118)$$

y para el segundo se obtiene fácilmente que

$$\lim_{n \rightarrow 0} \sum_a (m_a)^2 = m^2 \quad (1.119)$$

Y para el último término de 1.115 se tiene que calcula el límite como sigue

$$\begin{aligned} \lim_{n \rightarrow 0} \frac{1}{n} \ln \left[ \sum_{\sigma} \exp \left( \beta \sum_a m_a \sigma^a + \frac{1}{2} \alpha \beta^2 \sum_{a \neq b} r_{ab} \sigma^a \sigma^b \right) \right] &= \\ \lim_{n \rightarrow 0} \frac{1}{n} \ln \left[ \sum_{\sigma} \exp \left( \beta m \sum_a \sigma^a + \frac{1}{2} \alpha \beta^2 r (\sum_a \sigma^a)^2 - \frac{n}{2} \alpha \beta^2 r \right) \right] &= \\ -\frac{1}{2} \alpha \beta^2 r + \lim_{n \rightarrow 0} \frac{1}{n} \ln \left[ \int \frac{dz}{\sqrt{2\pi}} \exp \left( -\frac{z^2}{2} \right) \sum_{\sigma=\pm 1} \exp(\beta(m + \sqrt{\alpha r} z) \sum_a \sigma^a) \right] &= \\ -\frac{1}{2} \alpha \beta^2 r + \lim_{n \rightarrow 0} \frac{1}{n} \ln \left\{ \int \frac{dz}{\sqrt{2\pi}} \exp \left( -\frac{z^2}{2} \right) [2 \cosh(\beta(m + \sqrt{\alpha r} z))^n] \right\} &= \\ -\frac{1}{2} \alpha \beta^2 r + \int \frac{dz}{\sqrt{2\pi}} \exp \left( -\frac{z^2}{2} \right) \ln [2 \cosh(\beta(m + \sqrt{\alpha r} z))] & \quad (1.120) \end{aligned}$$

usando los resultados anteriores se puede escribir finalmente la expresión para la energía libre del esquema de replicas simétricas.

$$F[m, q, r] = \frac{1}{2}m^2 + \frac{1}{2}\alpha\beta r(1 - q) + \frac{\alpha}{2\beta} \left[ \ln(1 - \beta + \beta q) - \frac{\beta q}{1 - \beta + \beta q} \right] - \frac{1}{\beta} \langle \langle \ln[2 \cosh(\beta(m + \sqrt{\alpha r}z))] \rangle \rangle \quad (1.121)$$

donde  $\langle \langle \dots \rangle \rangle$  en este caso significa la integral Gaussiana sobre  $z$  o equivalente el promedio Gaussiano sobre  $z$ . Las soluciones a la ecuaciones de ensilladura  $\frac{\partial F}{\partial m} = 0$ ,  $\frac{\partial F}{\partial q} = 0$ ,  $\frac{\partial F}{\partial r} = 0$  son respectivamente:

$$m = \langle \langle \tanh[\beta(m + \sqrt{\alpha r}z)] \rangle \rangle \quad (1.122)$$

$$r = \frac{q}{(1 - \beta + \beta q)^2} \quad (1.123)$$

$$q = \langle \langle \tanh^2[\beta(m + \sqrt{\alpha r}z)] \rangle \rangle \quad (1.124)$$

Es importante revisar los casos limitantes para obtener un entendimiento del comportamiento en este modelo:

1. Caso  $\alpha \rightarrow 0$ : Este caso corresponde a la situación en la que el numero de patrones guardados  $P$  es mucho menor a  $N$ (el caso  $\alpha = 0$  significa que en el límite termodinámico  $N \rightarrow \infty$  el numero de patrones  $P$  permanece finito). Cuando  $\alpha \rightarrow 0$  se obtiene una ecuación de campo medio ordinaria para la transición de fase ferromagnética

$$m = \tanh(\beta m) \quad (1.125)$$

y se tiene que  $q = m^2$ , para este caso el sistema tiene la transición de fase de la fase paramagnética  $T > T_c = 1$  con  $m = 0$  a la fase ferromagnética o fase de “memoria”  $T < 1$  con un traslape finito  $m = m(T) \neq 0/$

2. Límite de baja temperatura  $T \rightarrow 0$ : Cuando el parámetro  $\alpha$  es finito la situación se hace mas compleja, sin embargo, en la situación que  $T \rightarrow 0$  se pueden

simplificar de gran manera a la expresiones 1.122, 1.123 y 1.124, por ejemplo:

$$\begin{aligned}
m = \lim_{T \rightarrow 0} \langle \langle \tanh[\beta(m + \sqrt{\alpha r}z)] \rangle \rangle &= \lim_{\beta \rightarrow \infty} \langle \langle \tanh[\beta(m + \sqrt{\alpha r}z)] \rangle \rangle = \\
&\int \frac{dz}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right) \text{sign}(m + \sqrt{\alpha r}z) = \\
&- \int_{-\infty}^{z=-\frac{m}{\sqrt{\alpha r}}} \frac{dz}{\sqrt{2\pi}} \exp\left\{\frac{-z^2}{2}\right\} + \int_{z=-\frac{m}{\sqrt{\alpha r}}}^{+\infty} \frac{dz}{\sqrt{2\pi}} \exp\left\{\frac{-z^2}{2}\right\} = \\
&- \int_{-\infty}^{-\frac{m}{\sqrt{\alpha r}}} \frac{dz}{\sqrt{2\pi}} \exp\left\{\frac{-z^2}{2}\right\} + \int_{-\frac{m}{\sqrt{\alpha r}}}^{+\frac{m}{\sqrt{\alpha r}}} \frac{dz}{\sqrt{2\pi}} \exp\left\{\frac{-z^2}{2}\right\} + \int_{+\frac{m}{\sqrt{\alpha r}}}^{+\infty} \frac{dz}{\sqrt{2\pi}} \exp\left\{\frac{-z^2}{2}\right\} = \\
&\frac{2}{\sqrt{2\pi}} \int_0^{\frac{m}{\sqrt{\alpha r}}} dz \exp\left\{-\frac{z^2}{2}\right\} = \text{erf}\left(\frac{m}{\sqrt{2\alpha r}}\right) \quad (1.126)
\end{aligned}$$

en el cálculo anterior se uso el hecho que  $\lim_{a \rightarrow \infty} \tanh(ax) = \text{sign}(x)$ , la función de error esta definida por

$$\text{erf}(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x dz \exp(-z^2) \quad (1.127)$$

Es conveniente introducir a la cantidad  $C = \beta(1 - q)$  en vez del parámetro q

$$\begin{aligned}
C &= \lim_{\beta \rightarrow \infty} \left[ \beta(1 - \langle \langle \tanh^2[\beta(m + \sqrt{\alpha r}z)] \rangle \rangle) \right] \\
&= \lim_{\beta \rightarrow \infty} \left[ \beta \left( \int \frac{dz}{\sqrt{2\pi}} \exp\left\{\frac{-z^2}{2}\right\} [1 - \tanh^2(\beta(m + \sqrt{\alpha r}z))] \right) \right] = \\
&= \lim_{\beta \rightarrow \infty} \left[ \beta \left( \int \frac{dz}{\sqrt{2\pi}} \exp\left\{\frac{-z^2}{2}\right\} \cosh^{-2}(\beta(m + \sqrt{\alpha r}z)) \right) \right] = \\
&= \sqrt{\frac{2}{\pi \alpha r}} \exp\left(-\frac{m^2}{2\alpha r}\right)
\end{aligned} \quad (1.128)$$

de tal manera que en el límite  $T \rightarrow 0$  se obtiene  $q \rightarrow 1$  y el parámetro  $C = \frac{1}{T}(1 - q)$  permanece finito. Para  $r$  se obtiene :

$$r = \frac{1}{(1 - C)^2} \quad (1.129)$$

Se introduce  $y = \frac{m}{\sqrt{2\alpha r}}$  en lugar de  $m$  obteniendo

$$\begin{aligned} y &= \frac{1}{\sqrt{2\alpha r}} \operatorname{erf}(y) \\ C &= \sqrt{\frac{2}{\pi\alpha r}} \exp(-y^2) \\ \frac{1}{\sqrt{r}} &= 1 - C \end{aligned} \tag{1.130}$$

Eliminando a  $r$  y  $C$  de las ecuaciones 1.130 se obtiene

$$y = \frac{\operatorname{erf}(y)}{\sqrt{2\alpha} + \frac{2}{\sqrt{\pi}} \exp(-y^2)} \tag{1.131}$$

La ecuación anterior tiene a la solución  $y = 0$  que corresponde a la fase de vidrio de spin en la que no hay traslape entre los patrones guardados, que significa que  $m = 0$  y

$$\begin{aligned} \sqrt{r} &= 1 + \sqrt{\frac{2}{\pi\alpha}} \\ C &= \frac{\sqrt{\frac{2}{\pi\alpha}}}{1 + \sqrt{\frac{2}{\pi\alpha}}} \end{aligned} \tag{1.132}$$

Otro punto importante es que si  $\alpha$  no es muy grande se tiene una solución no trivial con  $y \neq 0$  que corresponde a la fase de “memoria” en la que  $m \neq 0$ . Si  $\alpha$  es pequeño entonces

$$y \approx \frac{1}{\sqrt{2\alpha}} \tag{1.133}$$

y para el traslape de  $m$  se obtiene

$$m \approx 1 - \frac{2}{\sqrt{\pi}} \exp\left(-\frac{1}{2\alpha}\right) \tag{1.134}$$

## 2. Desarrollo de Red Neuronal

En este capítulo se detalla el desarrollo de una red neuronal convolucional(CNN) que se ha llamado SeismicNet siguiendo la convención en la comunidad del aprendizaje automático de nombrar a las redes neuronales con nombres similares. En este caso se ha omitido expandir en el desarrollo teórico de la redes neuronales profunda(DNN), para ver detalles del desarrollo de una DNN y una CNN se puede ver [29], en donde se inició la investigación de este trabajo de tesis, uno de los objetivos de este trabajo es mejorar los resultados obtenidos previamente. En el trabajo de prácticas citado anteriormente se investigó la posibilidad del desarrollo de una CNN y se desarrollo un modelo “inocente” de red neuronal para la detección de fases sísmicas, ahí se utilizó una arquitectura que tiene similitudes con los modelos utilizados en la detección de objetos a través de la visión computarizada, una aplicación popular de estos modelos se da en el desarrollo de automóviles autónomos, ver por ejemplo [1]. Para referenciar el modelo anterior se utilizara el alias TestNet. Para mejorar el desempeño obtenido en TestNet se aumentó la complejidad del modelo de varias maneras. Se utilizó un numero mas grande de muestras para entrenamiento, se incrementaron las capas de convolución del modelo ademas de usar capas adicionales, se manipularon los datos de tal manera de darle mas expresividad al modelo, se exploró un mayor numero de posibles arquitecturas, se optimizaron los algoritmos de adquisición de datos. Todos estos detalles se expanden a continuación.

### 2.1. Adquisición de Datos

Hay varios motivos por los que uno de los pasos fundamentales en el desarrollo de una CNN o DNN es la adquisición de datos. En esta etapa se pueden evaluar muchos aspectos del proyecto(recursos computacionales, herramientas tecnológicas, posibles complicaciones etc.) y se puede tener un análisis temprano de que tipo de problema se va abordar. Uno de los aspectos claves de la adquisición de datos es que en esta etapa se puede empezar a diseñar la arquitectura del modelo ya

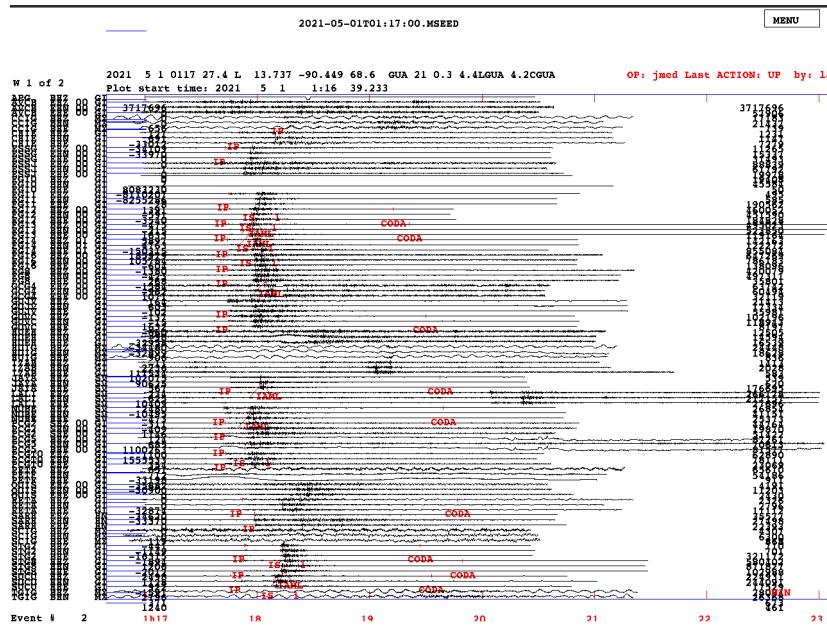
que son los datos en su forma cruda los que determinan últimamente cual es la forma de la primer capa en un modelo de DNN o sirve para determinar la primer capa de convolución en una CNN. Uno de los aspectos de importancia para este proyecto es el uso de una mayor cantidad de datos para la creación de conjuntos de entrenamiento, con esto en mente se utilizaron datos del los años 2019 y 2020, para probar las predicciones del modelo se utilizaron datos de los meses de enero y febrero del año 2021. Esto contrasta con TestNet en el que se usaron datos de 5 meses para entrenar al modelo.

### **2.1.1. Datos sismológicos puros**

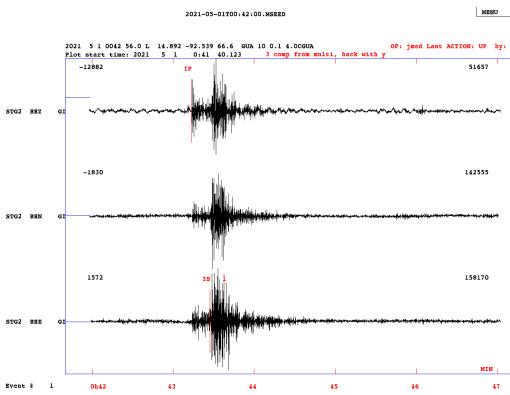
Los datos en su forma original consisten en los datos sismológicos generados en el departamento de Geofísica del INSIVUMEH, el proceso de recolección de estos datos consiste en una infraestructura compleja. Se tiene instrumentación de medición alrededor de Guatemala, esto consiste en sensores de movimiento de varias tecnologías llamados sismógrafos, estos son configurados dependiendo de la tecnología para transmitir datos en tiempo real. Estos datos son transformados a través de varias capas y resultan en un sismograma que observadores analizan en tiempo real. Estos sismogramas tienen una segunda o a veces varias revisiones con el fin de asegurar que el catalogo final del INSIVUMEH sea lo mas preciso posible. El proceso de generación de estos datos asegura que se tienen datos representativos de los eventos sísmicos detectados en Guatemala.

#### **2.1.1.1. Formas de Onda**

Los sismogramas o formas de onda de los eventos sísmicos contienen información de las llegadas de ondas mecánicas a los sensores, estas ondas se conocen como fases sísmicas. En la mayoría de casos para cada estación sísmica alrededor del país hay 3 canales de detección: Norte-Sur, Este-Oeste, Vertical. Las formas de onda para cada estación y cada canal son recolectadas a través de varios sistemas que funcionan en armonía: SeedLink, SEISAN, SeisComp, y muchos otros. El resultado final de esta recolección es un archivo de formato estándar en sismología llamado MiniSeed. Un ejemplo de estos archivos se puede ver en la figura 2.1, como se puede ver son varias formas de onda para cada evento sísmico. Las únicas formas de onda que pueden utilizarse en SeismicNet son aquellas en las que exista información de la llegada de fases sísmicas, que puede apreciarse como indicadores de color rojo en 2.1.



**Figura 2.1.** Ejemplo de archivo mseed, se pueden varias formas de onda, todas estas formas de onda sirven para caracterizar un evento sísmico.



**Figura 2.2.** Ejemplo de sismograma en estación STG2

Estas formas de onda son los datos con los que se alimentara a la red neuronal, en este caso se utilizan las amplitudes normalizadas para caracterizar a cada muestra. En la figura 2.2, se puede ver 3 formas de onda en este caso se ha separado a una solo estación para visualizar mejor la naturaleza de un sismograma en el que se han hecho análisis de arribos de fases sísmicas, cada forma de onda corresponde a una canal que determina la dirección del movimiento, se puede apreciar también las

picadas de fase asignadas después del análisis.

### 2.1.1.2. Archivos de análisis sísmico

El software estándar para realizar observación sísmica en el INSIVUMEH es el software libre SEISAN(ver [34], [35]). SEISAN es un conjunto de herramientas que facilitan el análisis sismológico, una de las herramientas importantes en la localización de eventos sísmicos es el localizador de fases sísmicas como se puede ver en la figura 2.2, este modulo provee de una interfaz de usuario gráfica(GUI) que permite a los analistas tener una representación gráfica y versátil de los sismogramas registrados. El resultado del análisis en esta interfaz gráfica resulta en un archivo con un formato conocido como Nordic, estos archivos se conocen como S-files y son la unidad básica de almacenamiento del software SEISAN, en si los S-files son archivos de texto con una definición que se puede consultar en el apéndice A de [34]. En los S-files reside información de cada evento sísmico, esta información incluye: magnitudes, tiempo de arribo de fases sísmicas en cada estación y canal, profundidades, localización de archivo de forma de onda(.mseed), un ejemplo de estos archivos puede verse en la figura 2.3. Estos archivos se utilizan para obtener las etiquetas de cada muestra de entrenamiento.

```
2019 11 4 0613 44.7 L 14.047 -89.798 2.8 GUA 5 0.7 2.3LGUA 3.2CGUA      1
GAP=205          1.62        12.1     30.9 33.8 0.2277E+03 -0.8256E+03 -0.3804E+03E
2019 11 4 0613 44.695 14.04737 -89.79838    2.830  0.674                      H
ACTION:UPD 19-12-16 17:02 OP:laa STATUS:           ID:20191104061260 L   I
OLDACT:ARG 19-11-12 20:52 OP:seis STATUS:          ID:20191104061260 L   3
2019-11-04T06:13:00.00000Z.MSEED                  6
STAT SP IPHASW D HRMM SECON CODA AMPLIT PERI AZIMU VELO AIN AR TRES W  DIS CAZ7
NUBE HZ IP      61348.605  61                   97  -0.0810 16.2 173
NUBE HE ES     1       61351.380               97  -0.42 7 16.2 173
JUAM EZ EP      61351.437               56  -0.1610 33.4 17
JUAM EN IS     1       61356.531               56  -0.45 7 33.4 17
JUAM EN IAML    61356.791    404.2 0.22            33.4 17
MT03 HZ EP      61356.369  29                   56  -0.4510 61.2 50
CHIE HZ IP      61358.730               47  0.1110 75.1 41
CHIE HN ES     1       614 8.707                47  -0.78 7 75.1 41
SARH EZ EP      614 9.256                 47  1.5210 138 54
```

**Figura 2.3.** Ejemplo de un S-file

### 2.1.2. Método de Adquisición

Para generar conjuntos de datos de entrenamiento y de prueba se pre-procesaron los datos descritos en la sección anterior a través de varias etapas, estas etapas for-

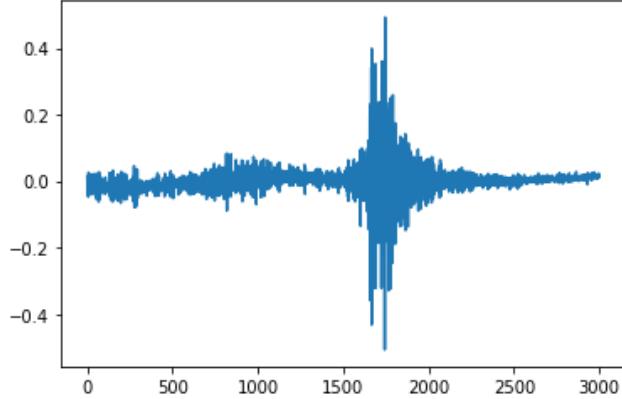
man la base del algoritmo de adquisición de datos.

### **2.1.2.1. Limpieza de datos**

El hecho de utilizar mas datos significa que hay que usar algoritmos de limpieza mas robustos que permitan utilizar todos los datos útiles, en aprendizaje automático un tema recurrente es la necesidad de usar mas y mas datos con el fin de mejorar los modelos. Para desarrollar modelos de predicción exitosos es necesario conocer bien los datos con los que se esta tratando. En este análisis previo de datos se pueden desarrollar herramientas que den un perspectiva cualitativa de la naturaleza de las muestras en su forma original. Este proceso se usa también para determinar que factores usar para adquirir muestras útiles. En este caso se depuran formas de onda para las que no existe picadas de fase sísmica o estaciones en las que solo se tiene una picada se fase sísmica, también se excluyen formas de onda que vienen en demasiadas secciones. Otro factor importante es el numero de amplitudes en una muestra, en esta etapa se asegura tener un numero de puntos necesario para el proceso de re-muestreo.

### **2.1.2.2. Re-muestreo**

Para crear una arquitectura de red neuronal es fundamental tener un conjunto de datos consistente, en muchos casos los datos con los que se pretende entrenar a un modelo vienen en una variedad de formatos y tamaños. Por esta razón es casi siempre obligatorio pasar los datos a través de un re-muestreo que depure y genere datos consistentes. En el caso de las formas de onda a utilizar, se tiene el caso que no todas vienen con el mismo numero de amplitudes en un sismograma específico, algunas tienen tan solo 2000 puntos en la serie de tiempo y otras tienen hasta 100000. Por esta razón se re-muestrearon todas la formas de onda para tener un vector de amplitudes consistente, se utilizaron dimensiones de este vector de entre 2000 – 3000 puntos, la elección de este parámetro estuvo inspirada en algunos artículos([48], [46], [22] ), así como también una búsqueda de hiper-parámetros afinada cuidadosamente. Para realizar este proceso se utilizaron herramientas de la biblioteca ObsPy de python, ver [33]. Un ejemplo de una forma de onda re-muestreada se puede ver en figura 2.4.



**Figura 2.4.** Ejemplo de forma de onda re-muestreada

#### 2.1.2.3. Normalización

Una práctica común en el aprendizaje automático es normalizar todas las muestras para el entrenamiento, esto se hace con fines de convergencia de los algoritmos de minimización de la función de error. Esto tiene que ver con la precisión interna de las computadoras y el paso que se da en cada iteración de algoritmos como el clásico descenso del gradiente y variantes más eficientes y complejas. Otra razón específica del problema a resolver para normalizar es el hecho que no se están distinguiendo las diferentes magnitudes de cada sismo.

Aquí se ha utilizado la convención de normalización de media:

$$x'_i = \frac{x_i - \mu(\vec{x})}{\max(\vec{x}) - \min(\vec{x})} \quad (2.1)$$

donde  $x'_i$  es el componente transformado,  $x_i$  es el componente a transformar,  $\mu(\vec{x})$  es el promedio de los componentes  $x_i$  del vector  $\vec{x}$ ,  $\max(\vec{x})$  es el valor máximo de los componentes de  $\vec{x}$  y  $\min(\vec{x})$  es el mínimo.

#### 2.1.2.4. Creación y recolección de etiquetas

Las etiquetas de entrenamiento se crearon a través de la información contenida en los archivos S-files, donde la información vital son los tiempos de arribo de las fases sísmicas. La creación de etiquetas de entrenamiento es un proceso que va amarrado al diseño de la arquitectura a implementar. Debido que se investigaron varias arquitecturas de CNN y DNN se tuvieron que recolectar dos tipos de etiquetas para las variaciones en las arquitecturas. El primer tipo de etiquetas consiste en

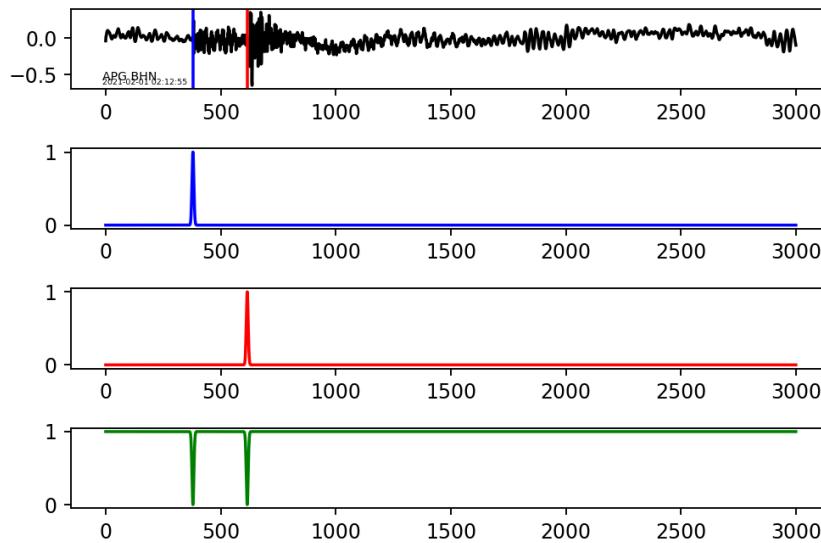
transformar en números reales los tiempos de arribo, esto se hace teniendo en cuenta el lapso del evento, con esto se puede identificar la posición relativa de las fases sísmicas en un espacio arbitrario, en este caso el espacio arbitrario corresponde a la dimensión del vector de entrada de la CNN o el numero de neuronas en la primer capa de procesamiento. El otro tipo de etiquetas recolectado consiste en distribuciones de probabilidad, esto se hizo con la finalidad de construir arquitecturas de CNN de mejor desempeño. Las distribuciones de probabilidad son generadas para caracterizar la fase P, la fase S y el ruido en una muestra, esto ayuda a la convergencia del modelo como se expondrá mas adelante. Al final de la creación de las distribuciones de probabilidad se puede obtener una matriz con la siguiente forma:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 0.9 & 0.88 & 0.81 & 0.78 & \dots & 1 & 1 & 1 & 1 & \dots & 0.9 & 0.85 & 0.73 & 0.7 & \dots \\ 0 & 0 & 0 & 0 & \dots & 0.1 & 0.12 & 0.19 & 0.22 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0.1 & 0.15 & 0.27 & 0.3 & \dots \end{bmatrix} \quad (2.2)$$

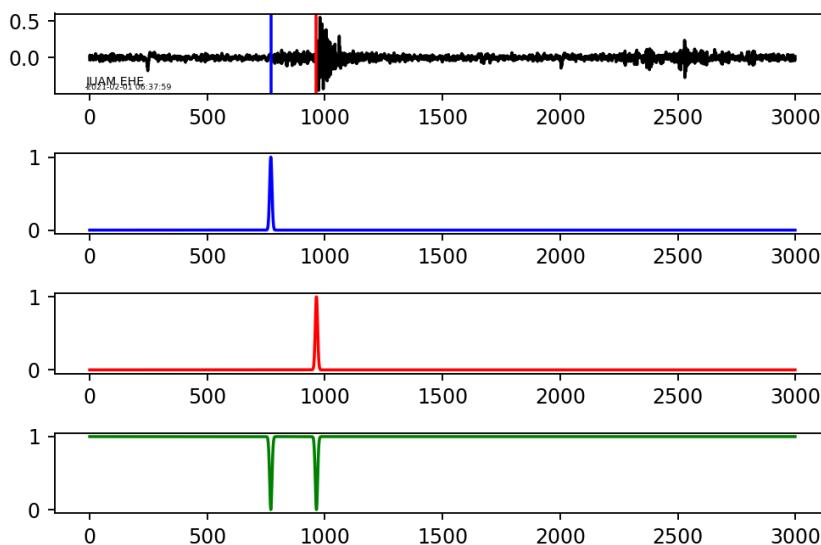
esta matriz representaría la etiqueta de una muestra y tiene dimensiones de  $3 \times n$  donde  $n$  es la dimensión del vector de entrada(2000-3000). Ejemplos de la recolección de etiquetas pueden verse en las figuras 2.5, 2.7. Alternativamente se pueden pensar en 2.2 como 3000 etiquetas con la siguiente forma:

$$\begin{bmatrix} 0.8 \\ 0.2 \\ 0 \end{bmatrix} \quad (2.3)$$

esto parece trivial pero es parte del diseño de una red neuronal que clasifica 3000 puntos como ruido, fase P o fase S, este punto se hará mas claro en las siguientes secciones, como se puede notar todas las columnas de las matriz 2.1 suman 1 al igual que los elementos del vector-columna 2.2.



**Figura 2.5.** Ejemplo de creación de etiquetas, en el gráfico superior se ve el sismograma y la picada de fase P se aprecia en color azul, la picada de fase S se ve en color rojo. En el segundo gráfico se ve la distribución de probabilidad asignada al arribo de fase P. En la tercera gráfica se ve la dist. de probabilidad para la fase S. Por ultimo en el gráfico inferior se ve la distribución de probabilidad para el ruido de la traza.



**Figura 2.6.** Ejemplo de creación de etiquetas, en el gráfico superior se ve el sismograma y la picada de fase P se aprecia en color azul, la picada de fase S se ve en color rojo. En el segundo gráfico se ve la distribución de probabilidad asignada al arribo de fase P. En la tercera gráfica se ve la dist. de probabilidad para la fase S. Por ultimo en el gráfico inferior se ve la distribución de probabilidad para el ruido de la traza.

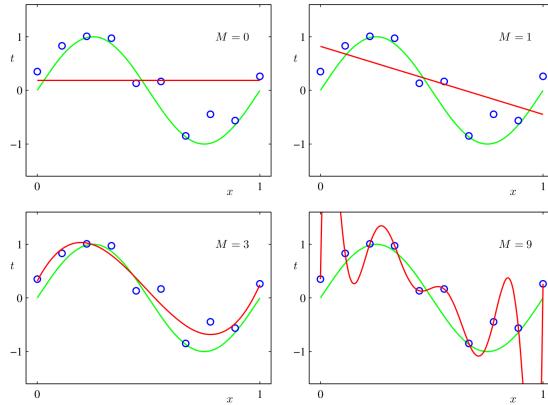
#### **2.1.2.5. Aumentación de datos**

Para aumentar el numero de muestras a añadir al conjunto de entrenamiento se consideró que las picadas de fase en una estación se hacen en canales específicos, las picadas de fase P se hacen por lo general en el canal vertical(eje z) y las picadas de fase S se hacen en los canales horizontales, con esto en mente se decidió asignar las 2 picadas de fase para todos los canales de una misma estación siempre y cuando existan estos datos(hay estaciones con una sola picada de fase). Hay otras técnicas para aumentar datos como la aplicación de transformaciones de varios tipos a los datos, aquí se quería trabajar con datos puros y se considerara usar datos sintéticos en proyectos posteriores. Los conjuntos de datos aumentados ayudan en el proceso de entrenamiento, en general se observó que los criterios utilizados para aumentar los datos son consistentes con el análisis sísmico.

Otro factor que ayudó a aumentar el número de datos fue la optimización que se hizo a los algoritmos de adquisición de datos, se aplicaron procedimientos de ObsPy como la función merge(), que corrige el seccionamiento en las formas de onda en su formato original, se corrigieron bugs en el código, todos estos detalles se pueden ver reflejados en el código del proyecto.

## 2.2. Diseño de arquitectura de red neuronal

Con el propósito de mejorar los modelos creados previamente se abordaron varias arquitecturas que tuviesen mas expresividad y que fuesen mas complejas. Una de las elecciones mas fáciles fue elegir una arquitectura de CNN sobre una DNN, este punto ya se había establecido previamente. Las capas de convolución en una CNN dan una ventaja significativa sobre una DNN(profunda). Las capas de convolución son capas de pre-procesamiento que sirven como filtros que extraen características de los datos, de hecho las capas convolutivas se conocen en la literatura de AI como capas de extracción de características(*feature extraction layers*), estas son optimizadas usando variantes mas sofisticadas del descenso del gradiente. Estas capas son diseñadas de tal manera que el proceso de optimización asegure que al menos se ha tratado de minimizar el error en el espacio de filtros posibles. Otro punto clave para el éxito del desarrollo de SeismicNet como se ha mencionado anteriormente es el uso de mas datos de entrenamiento, esto es algo que siempre se busca en el desarrollo de una CNN exitosa. Al incrementar el tamaño del conjunto de entrenamiento se evitan problemas en la modelación como el fenómeno de sobre-ajuste. Al evitar el sobre-ajuste se permite al modelo generalizar la solución del problema y no hacer un ajuste “perfecto” a los datos de entrenamiento. Este fenómeno puede apreciarse en la la traza inferior derecha de la imagen 2.7



**Figura 2.7.** Ejemplos de sub-ajuste y sobre-ajuste. En la trazas superiores izquierda y derecha se puede ver el fenómeno del sub-ajuste, en este caso se quiere ajustar los puntos a la traza que se aprecia en color verde. En el grafico inferior izquierdo se aprecia un buen ajuste al gráfico verde. En la traza inferior derecha se puede ver claramente el sobre-ajuste.

Algunos modelos que sirvieron de inspiración para el diseño de SeismicNet son el Zhu & Berroza [48] y el de Woollam & De Angelis [46]. En los artículos citados

anteriormente se desarrollaron modelos de análisis sísmico que abordan el problema de detección de fases en un sismograma.

### 2.2.1. Arquitecturas investigadas

El proceso de diseño de una red neuronal involucra crear varios modelos con el fin de encontrar un modelo óptimo que capture los patrones deseados en los datos. La naturaleza de los datos es clave para determinar qué tipo de arquitectura puede tener éxito, sin embargo esto no es el fin de la historia ya que el espacio de arquitecturas posibles es infinito y hay que sopesar varios factores. Con la referencia del proyecto anterior que se ha llamado TestNet, se pudo determinar que las capas de convolución son fundamentales para mejorar el desempeño. Por lo que fue clave en el desarrollo de este proyecto determinar si una arquitectura en la que se utilizaran convoluciones bidimensionales podría tener éxito en la localización de fases sísmicas, esto se hizo inspirado en el éxito que tienen los algoritmos de visión computarizada ([14], [24], [30]), en los que el motor principal de los modelos son las capas de convolución bidimensionales, estas ayudan a extraer características propias de cada imagen. El modelo creado que utiliza capas convolucionales bidimensionales en su arquitectura se ha denominado TestNet2 para futura referencia. Después de haber investigado las capas de convolución bidimensionales con imágenes de baja calidad para ahorrar espacio en el disco y en la memoria RAM a la hora de entrenar el modelo, de hecho este es un tema que se ampliará en la sección 2.3.2, se decidió utilizar convoluciones unidimensionales en los datos preprocesados y se encontraron muy buenos resultados. Este tipo de exploración en las posibilidades de la arquitectura de una CNN es fundamental, en este caso determinó la primera capa del modelo de SeismicNet.

Otro punto a determinar fue el tipo de etiqueta a recolectar como se expuso en 2.1.2.4, las primeras arquitecturas creadas se basaron en 2 etiquetas para caracterizar los arribos relativos de las fases sísmicas, esto tiene similitud a arquitecturas utilizadas para detección de objetos en tiempo real como se hace en el desarrollo de vehículos autónomos [1]. De las arquitecturas creadas se pudo determinar que era mejor explorar un tipo alternativo de etiquetas que proveyeran más información al modelo con tal de utilizar la información auxiliar implícita de los sismogramas, en este caso se utilizó como información auxiliar al ruido en la traza ya que se tienen muchos puntos que son considerados ruido. Esto fue clave para cambiar al modelo de un modelo de regresión en el que devolvían 2 números reales a un modelo

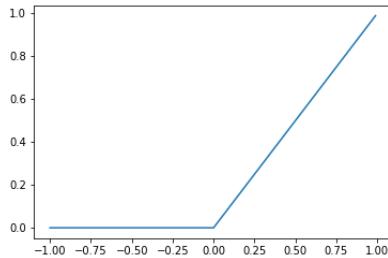
de clasificación de clases múltiples independientes. En lugar de dar como salidas 2 números reales se da como salida 3 distribuciones de probabilidad(9000 números) que corresponden a las fase P, fase S y el ruido. En otras palabras se tienen 3000 etiquetas del tipo que tiene en 2.6. De esta manera se tiene que hacer 3000 decisiones por la red neuronal es decir determinar para cada etiqueta del tipo 2.6 cual es la probabilidad para cada clase(P, S, ruido). Este cambio le da al modelo un grado superior de expresividad para poder distinguir las fases del ruido. Esto se expande en la sección 2.2.4. Se crearon 2 arquitectas de prueba con el esquema descrito para las etiquetas de salida en forma de distribuciones de probabilidad: TestNet3 y TestNet4. TestNet3 explora el espacio de las funciones de activación, y TestNet4 explora el impacto que tiene disminuir la resolución en el vector de entrada de la red.

## 2.2.2. Capas de procesamiento

Como se hizo ver con anterioridad las capas de convolución fueron fundamentales en la optimización del desempeño de SeismicNet. Adicionalmente se utilizaron capas de agrupamiento en las que se reduce el numero de neuronas, su funcionamiento es bastante sencillo, se agrupan los datos y se escoge el máximo, alternativamente se genera el promedio del grupo, este simple procedimiento reduce considerablemente(dependiendo del tamaño del grupo) la cantidad de neuronas, en las herramientas de implementación de CNN y DNN se les conoce a estas capas como *Pooling Layers*. Otra capa que fue de utilidad fueron las capas de deconvolución estas funcionan de manera inversa a las convoluciones, estas se utilizan para decodificar las características encontradas por las convoluciones, su uso de esta manera no es común, se utilizó este truco ya que se quería aumentar la dimensión de las convoluciones con el fin de incrementar las neuronas en las capas finales de procesamiento, esto fue fundamental para encontrar una salida con la dimension deseada, que en este caso es una vector de dimension 9000 que se re-dimensiona a una matriz de dimensiones  $3 \times 3000$  como 2.2. La función de activación escogida para las neuronas y para las capas de convolución es la la función rectificador o unidad lineal rectificada(ReLU) definida como

$$f(x) = x^+ = \max(0, x) \quad (2.4)$$

esta función devuelve 0 si  $x < 0$  de lo contrario devuelve  $x$ , para referencia ver la figura 2.8, esta elección tiene que ver con una convergencia mas rápida del descenso del gradiente esto se ve reflejado a la hora de entrenar al modelo. La antepenúltima



**Figura 2.8.** Función ReLU

ma capa del modelo no utiliza función de activación, para esta capa se combinan linealmente a las neuronas de la capa anterior con los parámetros internos correspondientes, esto resultaba en un vector de dimensión 9000. Esto fue un detalle de gran importancia para el desempeño de SeismicNet, Para poder comparar esta añadidura a la arquitectura se decidió llamar a una de las arquitecturas de exploración TestNet3 para comparar los resultados y entender la razón de esta elección. En TestNet3 se utilizó a la función Relu en como función de activación en la penúltima capa de la red.

### 2.2.3. Capa de salida

Para las arquitecturas de exploración en las que se utilizó un modelo de regresión se utilizó una capa en la que se tenían como salida 2 números reales que servían para determinar la posiciones relativas de las fases sísmicas, esto se ha discutido previamente. Para SeismicNet y otras arquitecturas de exploración se utilizaron como salida matrices con la forma de 2.2, esta matriz representa las 3 dist. de probabilidad como se ha explicado previamente, para poder generar dist. de probabilidad se utilizaron 2 pasos. Primero se redimensiona al vector resultante de la penúltima capa a una matriz con las dimensiones correctas después se aplica la función softmax a las columnas esta matriz, la función softmax se utiliza para convertir un vector con entradas reales en una distribución de probabilidad para n clases donde n es la dimensión del vector al que se aplica softmax. La función softmax está definida como

$$x_i = \frac{e^{x_i}}{\sum_j e^{x_j}} \quad (2.5)$$

en este caso se aplica a cada columna del tipo 2.6 por eso el índice  $j$  llega hasta 3, de esta forma toda columna de la matriz resultante tiene una suma de sus elementos

igual a 1, consistente con la identificación de 3 clases: fase P, fase S y ruido, al final se escoge a la clase con la mayor probabilidad para clasificar a cada punto de la serie de tiempo.

#### 2.2.4. Función de error

La correcta elección de la función de error o función de costo es crucial en cualquier modelo de aprendizaje automático. La función de error determina el resultado del ajuste de los parámetros internos de la red, ya que es la función que minimiza con tal de encontrar parámetros óptimos, es una función de los parámetros internos de la red. Para las arquitecturas de regresión se utilizó la función de error cuadrático medio:

$$E(\vec{w}) = \frac{1}{2N} \sum_{n=1}^N \|\vec{y}(\vec{x}_n, \vec{w}) - \vec{t}_n\|^2. \quad (2.6)$$

donde  $\vec{y}$  es la predicción de la red,  $\vec{x}_n$  es el vector de entrada de la red,  $\vec{w}$  son los parámetros internos de la red y  $\vec{t}_n$  es la etiqueta de entrenamiento.

Para las arquitecturas de clasificación de clases múltiples e independientes se utilizaron funciones de error del tipo entropía cruzada, este tipo de error es el que usualmente se utiliza en arquitecturas de clasificación binaria y multiclas. Para la función de error de SeismicNet se utilizó una variación de la función de error entropía cruzada(*cross entropy*) específicamente se usa la siguiente función de error

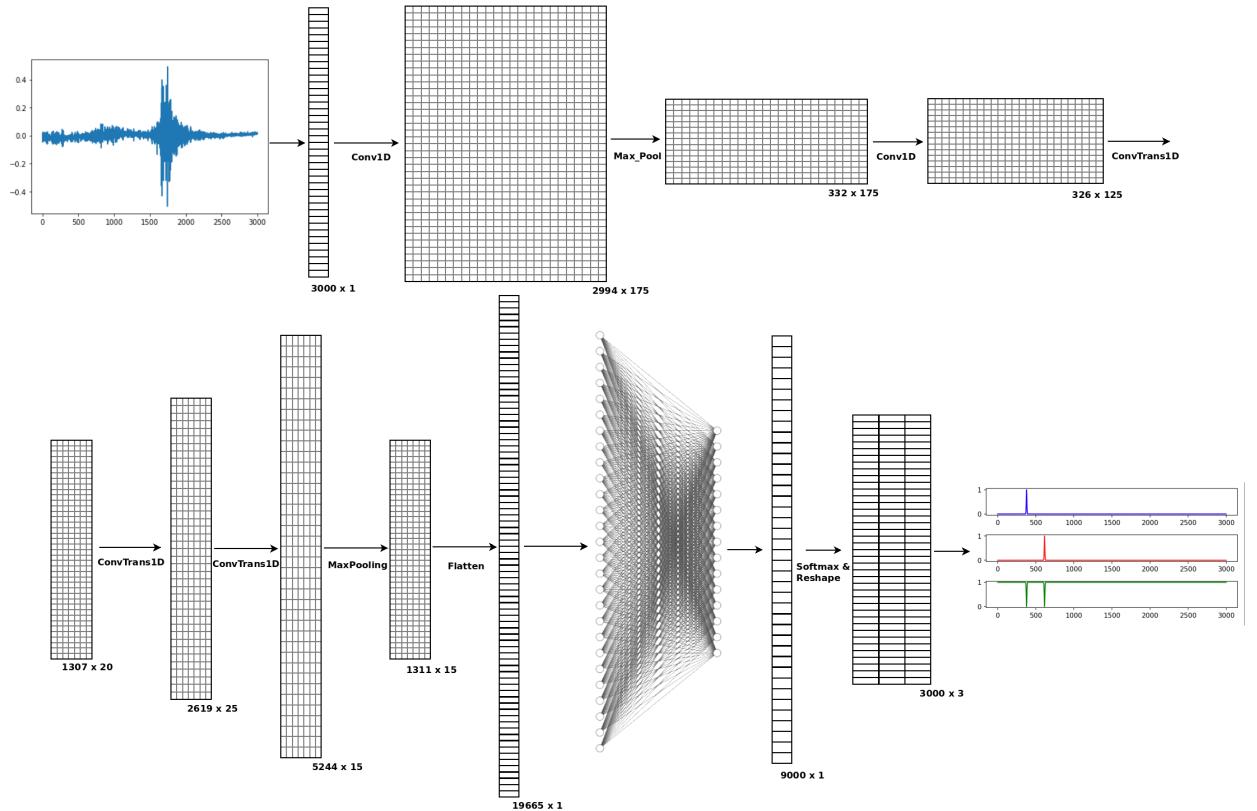
$$E(\hat{P}, \hat{Q}) = - \sum_i^N \sum_j^3 \sum_k^d P_{ijk} \ln Q_{ijk} \quad (2.7)$$

donde  $N$  es el numero total de muestras en el conjunto de entrenamiento,  $j$  se usa para etiquetar la distribución de probabilidad de la fase P, fase S y el ruido,  $d$  es la dimension del vector de entrada, para SeismicNet se utilizó  $d = 3000$ . El tensor  $P_{ijk}$  es el tensor que contiene las distribuciones de probabilidad del tipo 2.2 para todas las muestras del conjunto de entrenamiento,  $Q_{ijk}$  son las distribuciones de probabilidad que SeismicNet predice, en otras palabras la ultima capa de la red.

#### 2.2.5. Arquitectura de SeismicNet

En la figura 2.9 se aprecia la arquitectura de SeismicNet, se puede ver como cambian los datos en cada capa de la red, los tensores o matrices en la imagen son solo para fines ilustrativos y los tamaños de los diagramas no son proporcionales

a las dimensiones especificadas en la parte inferior de cada uno. Todo el código de SesimicNet ha sido subido a la red como un proyecto de acceso libre(open source) para que cualquier persona interesada en los detalles del desarrollo pueda tener acceso sin ningún inconveniente <https://github.com/Julio-Medina/SeismicNet>.



**Figura 2.9.** Arquitectura de SeismicNet, se pueden ver la forma que toman los datos en cada capa de procesamiento.

## 2.3. Implementación de SeismicNet

### 2.3.1. Herramientas de desarrollo

Entidades lineales y multilineales como los vectores, matrices y tensores<sup>1</sup> son ubicuas en el aprendizaje automático, de hecho estas entidades le dan su nombre a la librería de python utilizada en este proyecto: **TensorFlow** que se puede traducir como “flujo de tensores”. TensorFlow es un conjunto de librerías que funcionan

---

<sup>1</sup>Los tensores del aprendizaje automático no cumplen necesariamente las leyes de transformación que definen a un tensor en física o matemática, su uso es mas orientado a estructuras multidimensionales(e.g. arreglos cúbicos, tetra-dimensionales, o hiper-dimensionales de números).

con varios lenguajes de programación: python, javascript, C++ y otros. TensorFlow fue desarrollado por ingenieros de Google Brain y es una herramienta de acceso libre(*open source*) que es comúnmente utilizada en la actualidad por varios sectores de desarrollo tecnológico y por muchos investigadores. El desarrollo rápido de TensorFlow al igual que otras herramientas similares(PyTorch, MATLAB, Caffe) han impulsado el avance actual del aprendizaje automático.

Una de las grandes ventajas de utilizar TensorFlow en este proyecto es el soporte que tiene para paralelizar los procesos de una manera natural. El entrenamiento de una red neuronal en general es un proceso de computación intensiva: se utilizan ciclos enormes para las iteraciones de entrenamiento, se utilizan grandes cantidades de datos, se hacen millones de operaciones matriciales en cada paso, se utilizan métodos numéricos para optimizar el descenso del gradiente. Por estas razones es vital utilizar los recursos de computación de una manera eficiente. Otro aspecto interesante de TensorFlow es el soporte de hardware que ofrece, tiene versiones que corren específicamente en GPU(*graphic processor unit*) y TPU(*tensor processor unit*). Los GPU son tarjetas gráficas , estas tiene en general una gran cantidad de núcleos simples en contraste con un procesador (CPU) que tienen pocos núcleos complejos, los núcleos simples de una GPU estan diseñados para realizar operaciones matriciales a gran velocidad, ya que comúnmente se utilizan a los GPU en tareas de renderización de vídeo o la creación de ambientes virtuales tridimensionales(vídeojuegos), tareas que están colmadas de transformaciones lineales, productos punto y re-dimensionamiento de tensores. Para poner en contexto la diferencia entre los núcleos de un CPU y un GPU se tienen los siguientes ejemplos: unos de los mejores CPU en la actualidad es el AMD RYZEN ThreadRipper el cual tiene 64 núcleos con tecnología HyperThreading que hacen un total de 128 hilos o tareas, por otro lado se tiene la GPU Nvidia RTX3080 Ti de uso comercial con un número total de 10240 núcleos gráficos y 320 núcleos tensoriales.

Otras herramientas importantes en el desarrollo de este proyecto son librerías de python como:

- ObsPy: Librerías con utilerías de observación sísmica.
- NumPy: Procedimientos estándar de computación científica de alto desempeño.
- Matplotlib: Librerías para visualización de datos.
- Keras: Módulos de empaquetamiento de TensorFlow(*wrappers*)

- SciPy: Herramientas de computación científica.

El software de observación sísmica SEISAN también fue de gran ayuda, sobre todo para visualizar y entender los datos en su forma original, la interfaz gráfica para picar fases se uso para corroborar que las etiquetas obtenidas en el proceso de adquisición de datos fuesen correctas.

### 2.3.2. Costo computacional

Una de las razones por las que el uso generalizado de las redes neuronales y el aprendizaje automático no despegó desde que se crearon los conceptos teóricos fue la falta de hardware capaz de entrenar modelos exitosos. Hay varios factores que hacen del aprendizaje automático una disciplina con una demanda muy alta de poder de computo, requiriendo también una alta demanda de almacenamiento. Los factores principales para la necesidad de computación de alto desempeño en la implementación de modelos de aprendizaje automático se resumen a continuación:

- Se usan una gran cantidad de cálculos matriciales.
- Se usan métodos numéricos para optimizar funciones de miles y millones de variables.
- Los modelos exitosos se entrena con conjuntos enormes de datos.
- Se usan ciclos de iteración extensos para que la función de error del modelo sea lo mas pequeña posible.
- El procedimiento de retro-propagación de errores en una CNN, en el que se calculan las derivadas parciales de la función de error tiene que calcular derivadas con respecto a TODAS la neuronas del modelo.

Hay otros factores que también influyen en la complejidad tecnológica que plantean las redes neuronales. Unos de los mayores retos a la hora de entrenar un modelo es la posibilidad que en el proceso del descenso del gradiente se encuentre un mínimo local muy alejado del valor optimo que debe de tener la función de error para un desempeño aceptable de la CNN, esto se añade a la necesidades de procesamiento del entrenamiento de los modelos.

En este proyecto se tuvieron varios retos de este tipo, uno de los mayores retos fue el tiempo en los ciclos de entrenamiento, los modelos exitosos se entrenaban por varios días. Como se discutió anteriormente el uso de las herramientas de software y

hardware adecuadas mejoran las posibilidades de éxito de un proyecto de IA. En el desarrollo de SeismicNet se hizo uso de computadoras personales con procesadores de 8 núcleos y algunas configuraciones en las que se auxiliaba el entrenamiento con GPU, sin embargo se pudo notar el lento proceso de exploración de arquitecturas y el extenuante proceso de entrenamiento. Esto sugiere que el desarrollo de modelos mas robustos de aprendizaje automático se pueden llevar a cabo en infraestructuras de computo mas sofisticadas, como por ejemplo la computación en la nube, clusters de computación, servidores con varios GPU y TPU, etc.

### 2.3.3. Entrenamiento

Cuando se incrementan las neuronas, la capas y se aumenta la complejidad de una CNN es de esperarse que el costo computacional será mucho mas elevado. Por lo general se necesitan de varios ciclos de entrenamiento para que un modelo logre ajustar los parámetros internos y se logre minimizar la función de error. En algunas ocasiones los modelos entrenados no tienen los resultados esperados por lo que el entrenamiento es un proceso de constante depuración. Esto aunado a la naturaleza de los datos hicieron que el entrenamiento de SeismicNet no fuese un proceso con la celeridad deseada, sin embargo se utilizaron los recursos computacionales de tal manera de incrementar las posibilidades de éxito del proyecto. Unos de los obstáculos a superar fue que debido que utilizaron datos de 2 años, se tuvieron que hallar herramientas de programación que permitieran entrenar continuamente al modelo con la totalidad de los datos, en las primeras pruebas el equipo utilizado simplemente no pudo contener a todos los datos en memoria por lo que se tuvieron múltiples colapsos del sistema. El proceso de entrenamiento de SeismicNet duro alrededor de 9 semanas, en esta estimación no se toman en cuenta otras aspectos del proyecto. La variación del algoritmo del gradiente utilizada para entrenar SeismicNet es el algoritmo llamado Adam(*Adaptive moment estimation*). En ese algoritmo se usan tasas de aprendizaje variables para cada parámetro siendo optimizado, esto incrementa significativamente la velocidad de convergencia del modelo. El algoritmo Adam es una variante de el algoritmo Momentum. La función de error utilizada para entrenar al modelo fue 2.7 como se expuso en 2.2.4. Se agruparon los datos en grupos de 1500 elementos, estos grupos son denominados lotes en IA. Se utilizaron 75 ciclos de entrenamiento.

### 3. Resultados

El criterio que se ha elegido para determinar las predicciones exitosas de SeismicNet es una muy sencilla, todo arriba de fase en el que la diferencia temporal  $\Delta t$  entre la etiqueta del conjunto de entrenamiento y la predicción sea menor a 1 segundo se considera como un acierto en la predicción, una diferencia  $\Delta t < 1s$  en relación al análisis hecho por analistas experimentados es un muy buen resultado. Adicionalmente se tomaron en cuenta  $\Delta t < 2s$  y  $\Delta t < 3s$  para tener una perspectiva general del desempeño obtenido con SeismicNet. Para apreciar las predicciones del modelo se han escogido algunos resultados y se han agrupado en función a la diferencia  $\Delta t$  graficando la forma de onda del sismograma y las distribuciones de probabilidad resultantes de la última capa de SeismicNet. Adicionalmente a SeismicNet se presentan resultados de otras arquitecturas exploradas para comparar como mejora o empeora el desempeño de un modelo dependiendo de su arquitectura, a continuación se resumen las características de las diferentes arquitecturas de prueba(ver 2.2.1):

TestNet : Modelo investigado previamente en [29]. Es un modelo de regresión en el que se usan capas de convolución.

TestNet2 : Modelo en el que se explora el uso de capas de convolución bidimensionales.

TestNet3 : Modelo en el que se compara un cambio en la penúltima capa de SeismicNet(ver 2.2.2).

TestNet4 : Modelo en el que se explora el desempeño de SeismicNet y un modelo de menor resolución. Este modelo es análogo a SeismicNet con la diferencia que el vector de entrada tiene una dimensión de 2000.

SeismicNet : Modelo implementado en este trabajo de graduación.

### 3.1. Resultados tabulados

Se presentan los resultados tabulados para los distintos modelos, se han agrupado los datos del mes de enero 2021 y febrero 2021 por aparte. En la siguientes tablas se tienen los resultados para el promedio  $\bar{\Delta t}$  de las todas la predicciones del modelo,  $\bar{\Delta t}_P$  el promedio para la predicciones de fase P y  $\bar{\Delta t}_S$  para las predicciones de fase S. Se incluyen las desviaciones estándar correspondientes. El conjunto de muestras de prueba de mes de enero consiste de 944 trazas y para febrero se tienen 2052.

#### 3.1.1. TestNet

##### 3.1.1.1. Enero 2021

$\bar{\Delta t}$	12.39 s
$\sigma$	29.34 seg
$\bar{\Delta t}_P$	12.47 s
$\sigma_P$	26.49 s
$\bar{\Delta t}_S$	12.31 s
$\sigma_S$	31.95 s

**Tabla 3.1.** Medidas de tendencia central para modelo de prueba TestNet, datos de enero 2021.

	$\Delta t < 1\text{s}$		$\Delta t < 2\text{s}$		$\Delta t < 3\text{s}$	
Fases P identificadas	83	8.80 %	165	17.48 %	246	26.06 %
Fases S identificadas	166	17.58 %	335	35.49 %	453	47.99 %
Una fase identificada	228	24.15 %	433	45.87 %	558	59.11 %
Dos fases identificadas	21	2.22 %	67	7.10 %	141	14.94 %

**Tabla 3.2.** Contabilización de predicciones acertadas del modelo TestNet en los datos de enero 2021

### 3.1.1.2. Febrero 2021

$\bar{\Delta t}$	29.01 s
$\sigma$	78.40 seg
$\bar{\Delta t}_P$	28.83 s
$\sigma_P$	29.18 s
$\bar{\Delta t}_S$	29.18 s
$\sigma_S$	76.49 s

**Tabla 3.3.** Medidas de tendencia central para modelo de prueba TestNet, datos de febrero 2021.

	$\Delta t < 1s$	$\Delta t < 2s$	$\Delta t < 3s$			
Fases P identificadas	103	5.02 %	222	10.82 %	345	16.81 %
Fases S identificadas	317	15.45 %	546	26.61 %	731	35.62 %
Una fase identificada	399	24.15 %	684	33.33 %	905	44.10 %
Dos fases identificadas	21	1.02 %	84	4.09 %	171	8.33 %

**Tabla 3.4.** Contabilización de predicciones acertadas del modelo TestNet en los datos de febrero 2021

## 3.1.2. TestNet2

### 3.1.2.1. Enero 2021

$\bar{\Delta t}$	45.34 s
$\sigma$	44.09 seg
$\bar{\Delta t}_P$	42.63 s
$\sigma_P$	41.68 s
$\bar{\Delta t}_S$	48.05 s
$\sigma_S$	46.37 s

**Tabla 3.5.** Medidas de tendencia central para modelo de prueba TestNet2, datos de enero 2021.

	$\Delta t < 1\text{s}$		$\Delta t < 2\text{s}$		$\Delta t < 3\text{s}$	
Fases P identificadas	31	3.28 %	48	5.08 %	65	6.89 %
Fases S identificadas	9	0.95 %	24	2.54 %	37	3.92 %
Una fase identificada	38	4.02 %	63	6.67 %	84	8.90 %
Dos fases identificadas	2	0.21 %	9	0.95 %	18	1.91 %

**Tabla 3.6.** Contabilización de predicciones acertadas del modelo TestNet2 en los datos de enero 2021

### 3.1.2.2. Febrero 2021

$\overline{\Delta t}$	155.9 s
$\sigma$	326.76 seg
$\overline{\Delta t_P}$	158.72 s
$\sigma_P$	334.74
$\overline{\Delta t_S}$	153.09 s
$\sigma_S$	318.57 s

**Tabla 3.7.** Medidas de tendencia central para modelo de prueba TestNet2, datos de febrero 2021.

	$\Delta t < 1\text{s}$		$\Delta t < 2\text{s}$		$\Delta t < 3\text{s}$	
Fases P identificadas	36	1.75 %	74	3.61 %	102	4.97 %
Fases S identificadas	29	1.41 %	62	3.02 %	92	4.48 %
Una fase identificada	61	2.73 %	121	5.90 %	159	7.75 %
Dos fases identificadas	4	0.21 %	15	0.73 %	35	1.71 %

**Tabla 3.8.** Contabilización de predicciones acertadas del modelo TestNet2 en los datos de febrero 2021

### 3.1.3. TestNet3

#### 3.1.3.1. Enero 2021

$\overline{\Delta t}$	12.74 s
$\sigma$	32.84 seg
$\overline{\Delta t_P}$	12.05 s
$\sigma_P$	27.42 s
$\overline{\Delta t_S}$	13.42 s
$\sigma_S$	37.50 s

**Tabla 3.9.** Medidas de tendencia central para modelo de prueba TestNet3, datos de enero 2021.

	$\Delta t < 1s$		$\Delta t < 2s$		$\Delta t < 3s$	
Fases P identificadas	218	23.09 %	352	37.29 %	438	46.40 %
Fases S identificadas	359	38.03 %	559	59.22 %	633	67.06 %
Una fase identificada	472	50.00 %	647	68.54 %	706	74.79 %
Dos fases identificadas	104	11.02 %	264	27.97 %	366	38.77 %

**Tabla 3.10.** Contabilización de predicciones acertadas del modelo TestNet3 en los datos de enero 2021

#### 3.1.3.2. Febrero 2021

$\overline{\Delta t}$	190.39 s
$\sigma$	522.65 s
$\overline{\Delta t_P}$	203.43 s
$\sigma_P$	553.94
$\overline{\Delta t_S}$	177.35 s
$\sigma_S$	489.37 s

**Tabla 3.11.** Medidas de tendencia central para modelo de prueba TestNet3, datos de febrero 2021.

	$\Delta t < 1\text{s}$		$\Delta t < 2\text{s}$		$\Delta t < 3\text{s}$	
Fases P identificadas	262	12.77 %	457	22.27 %	578	28.17 %
Fases S identificadas	555	27.05 %	861	41.96 %	978	47.66 %
Una fase identificada	660	32.16 %	961	46.83 %	1060	51.66 %
Dos fases identificadas	156	7.60 %	357	17.40 %	497	24.22 %

**Tabla 3.12.** Contabilización de predicciones acertadas del modelo TestNet3 en los datos de febrero 2021

### 3.1.4. TestNet4

#### 3.1.4.1. Enero 2021

$\bar{\Delta t}$	13.51 s
$\sigma$	30.96 s
$\bar{\Delta t}_P$	12.51 s
$\sigma_P$	24.89 s
$\bar{\Delta t}_S$	14.51 s
$\sigma_S$	36.02 s

**Tabla 3.13.** Medidas de tendencia central para modelo de prueba TestNet3, datos de enero 2021.

	$\Delta t < 1\text{s}$		$\Delta t < 2\text{s}$		$\Delta t < 3\text{s}$	
Fases P identificadas	275	29.13 %	404	42.80 %	477	50.53 %
Fases S identificadas	350	37.08 %	543	57.52 %	629	66.63 %
Una fase identificada	504	53.39 %	676	71.61 %	735	77.86 %
Dos fases identificadas	120	12.71 %	271	28.71 %	371	39.30 %

**Tabla 3.14.** Contabilización de predicciones acertadas del modelo TestNet4 en los datos de enero 2021

### 3.1.4.2. Febrero 2021

$\bar{\Delta t}$	217.09 s
$\sigma$	619.53 s
$\bar{\Delta t}_P$	235.01 s
$\sigma_P$	665.32
$\bar{\Delta t}_S$	199.16 s
$\sigma_S$	570.07 s

**Tabla 3.15.** Medidas de tendencia central para modelo de prueba TestNet4, datos de febrero 2021.

	$\Delta t < 1\text{s}$		$\Delta t < 2\text{s}$		$\Delta t < 3\text{s}$	
Fases P identificadas	350	17.06 %	501	24.42 %	611	29.78 %
Fases S identificadas	617	30.07 %	873	42.54 %	998	48.64 %
Una fase identificada	789	38.45 %	1009	49.17 %	1109	54.04 %
Dos fases identificadas	179	8.72 %	364	17.74 %	500	24.37 %

**Tabla 3.16.** Contabilización de predicciones acertadas del modelo TestNet4 en los datos de febrero 2021

### 3.1.5. SeismicNet

#### 3.1.5.1. Enero 2021

$\bar{\Delta t}$	8.34 s
$\sigma$	24.01 s
$\bar{\Delta t}_P$	6.80 s
$\sigma_P$	18.14 s
$\bar{\Delta t}_S$	9.88 s
$\sigma_S$	28.71 s

**Tabla 3.17.** Medidas de tendencia central para modelo SeismicNet, datos de enero 2021.

	$\Delta t < 1\text{s}$		$\Delta t < 2\text{s}$		$\Delta t < 3\text{s}$	
Fases P identificadas	543	57.52 %	655	69.39 %	700	74.15 %
Fases S identificadas	483	51.16 %	686	72.67 %	737	78.07 %
Una fase identificada	728	77.12 %	813	86.12 %	833	88.24 %
Dos fases identificadas	298	31.57 %	528	55.93 %	604	63.98 %

**Tabla 3.18.** Contabilización de predicciones acertadas del modelo SeismicNet en los datos de enero 2021

### 3.1.5.2. Febrero 2021

$\bar{\Delta t}$	24.50 seg
$\sigma$	67.56 seg
$\bar{\Delta t}_P$	24.34 seg
$\sigma_P$	71.33
$\bar{\Delta t}_S$	24.66 seg
$\sigma_P$	63.58 seg

**Tabla 3.19.** Medidas de tendencia central para modelo de prueba SeismicNet, datos de febrero 2021.

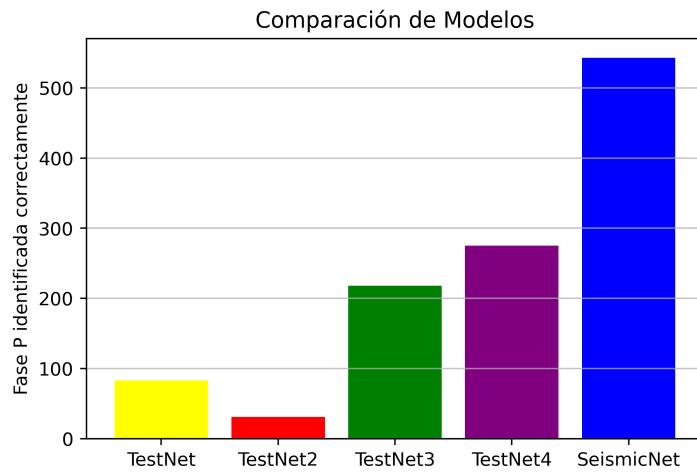
	$\Delta t < 1\text{s}$		$\Delta t < 2\text{s}$		$\Delta t < 3\text{s}$	
Fases P identificadas	763	37.18 %	949	46.25 %	1039	50.63 %
Fases S identificadas	876	42.69 %	1161	56.58 %	1277	62.23 %
Una fase identificada	1172	57.12 %	1376	67.06 %	1445	70.42 %
Dos fases identificadas	467	22.76 %	734	35.77 %	871	42.45 %

**Tabla 3.20.** Contabilización de predicciones acertadas del modelo SeismicNet en los datos de febrero 2021

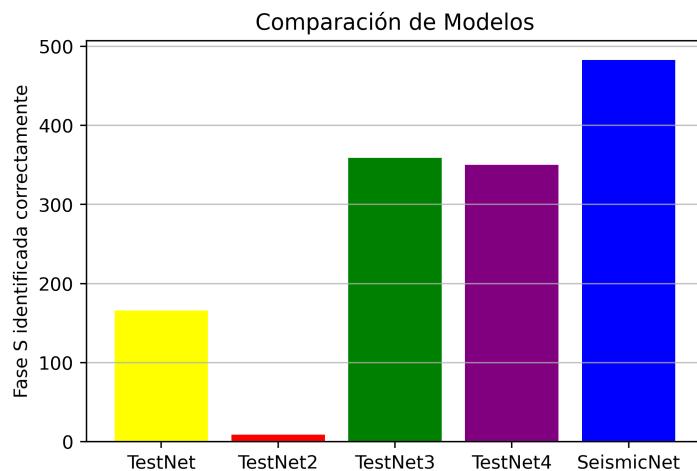
## 3.2. Comparaciones

Para tener una comparación visual del desempeño de los distintos modelos se han hecho gráficas de barras de resultados de las predicciones acertadas, es decir  $\Delta t < 1\text{s}$ .

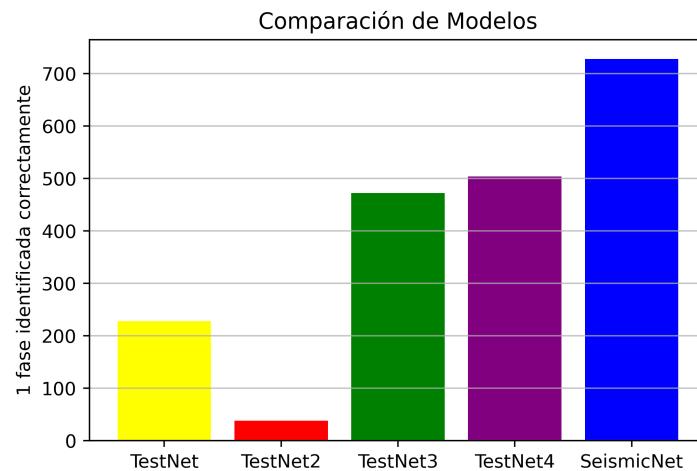
### 3.2.1. Enero 2021



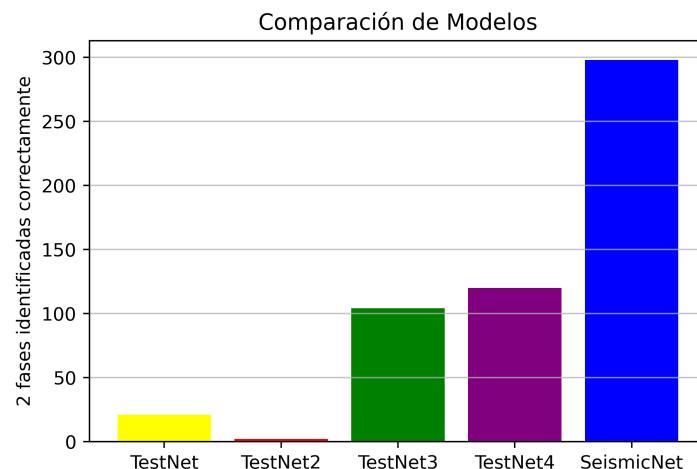
**Figura 3.1.** Gráfico de barras en el que se comparan los resultados de los distintos modelos evaluados. En este caso se han contabilizado el numero de predicciones de fase P con  $\Delta t < 1s$  entre la predicción y la etiqueta de entrenamiento.



**Figura 3.2.** Gráfico de barras en el que se comparan los resultados de los distintos modelos evaluados. En este caso se han contabilizado el número de predicciones de fase S con  $\Delta t < 1s$  entre la predicción y la etiqueta de entrenamiento.

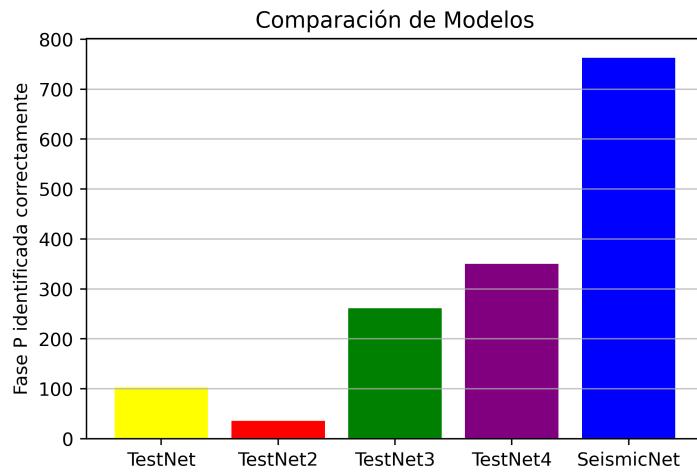


**Figura 3.3.** En este gráfico de barras se aprecian los resultados en los que se han contabilizado las predicciones en cualquiera de las 2 fases sísmicas S o P.

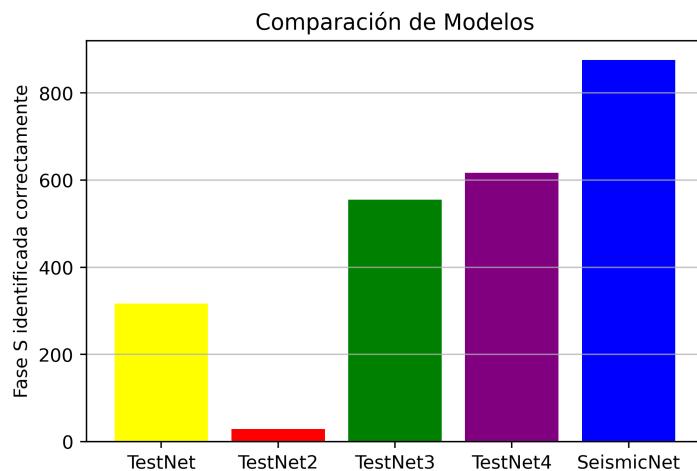


**Figura 3.4.** El criterio mas riguroso en el que se han contabilizado predicciones correctas del modelo en las 2 fases sísmicas, se tuvo una localización de fases sísmicas completamente acertada en las muestras.

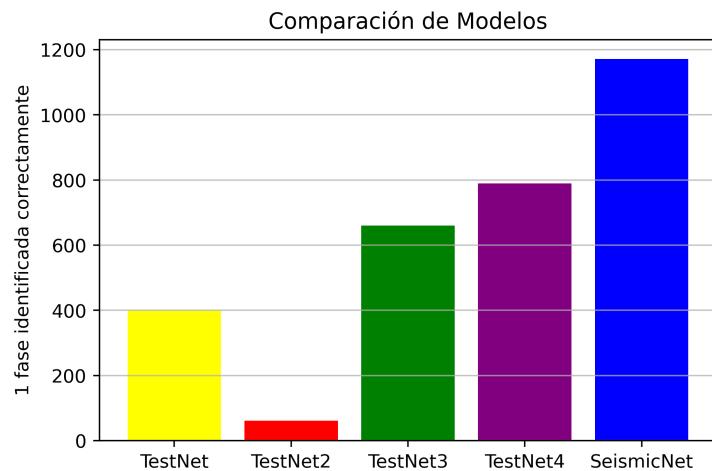
### 3.2.2. Febrero 2021



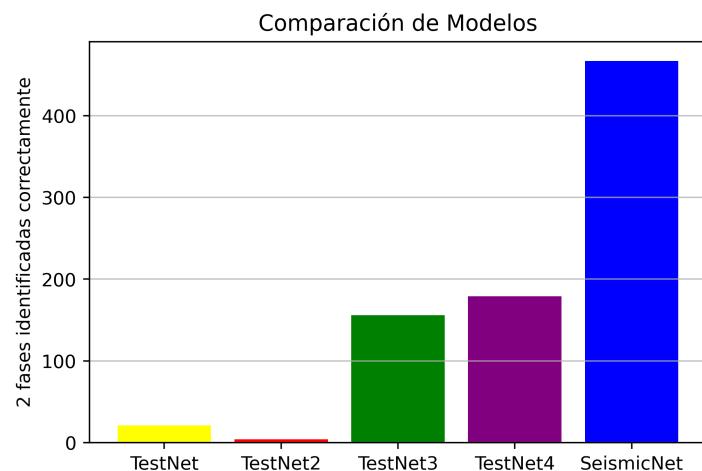
**Figura 3.5.** Gráfico de barras en el que se comparan los resultados de los distintos modelos evaluados. En este caso se han contabilizado el numero de predicciones de fase P con  $\Delta t < 1s$  entre la predicción y la etiqueta de entrenamiento.



**Figura 3.6.** Gráfico de barras en el que se comparan los resultados de los distintos modelos evaluados. En este caso se han contabilizado el número de predicciones de fase S con  $\Delta t < 1s$  entre la predicción y la etiqueta de entrenamiento.



**Figura 3.7.** En este gráfico de barras se aprecian los resultados en los que se han contabilizado las predicciones en cualquiera de las 2 fases sísmicas S o P.

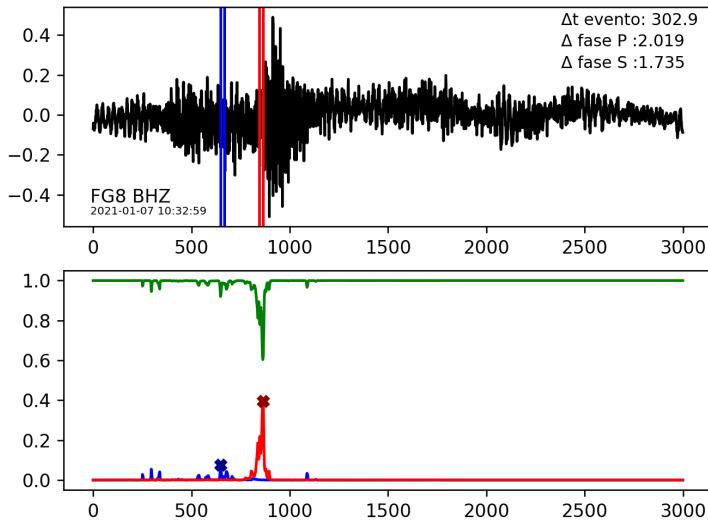


**Figura 3.8.** El criterio mas riguroso en el que se han contabilizado predicciones correctas del modelo en las 2 fases sísmicas, se tuvo unas caracterización completamente acertada en las muestras.

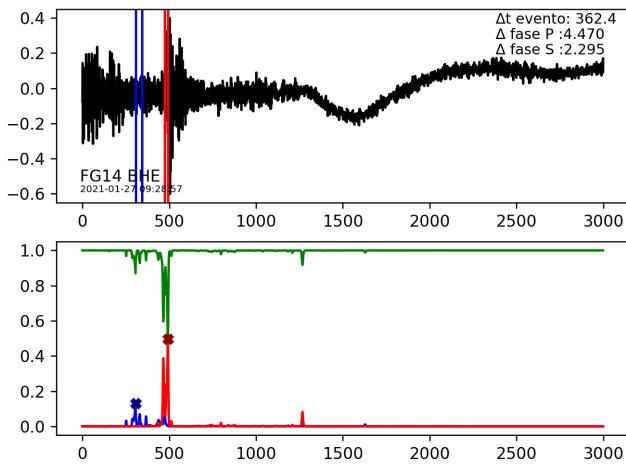
### 3.3. Predicciones de SeismicNet

Para apreciar los resultados de las predicciones de SeismicNet se han escogido varios resultados al azar y se han agrupado dependiendo si se han tenido aciertos entre las etiquetas y la predicción: en ninguna picada de fase, 1 picada de fase y 2 picadas de fase. En las gráficas se presenta el sismograma(muestra de entrada), las picadas de fase recolectadas(etiquetas de entrenamiento), las picadas de fase y las distribuciones de probabilidad(P, S y ruido) predichas por SeismicNet. En la parte superior derecha de todas las muestras presentadas se indica en segundos el intervalo temporal del evento sísmico  $\Delta t$  evento, también se dan las diferencias temporales entre las predicciones de la fase P y fase S y las etiquetas recolectadas.

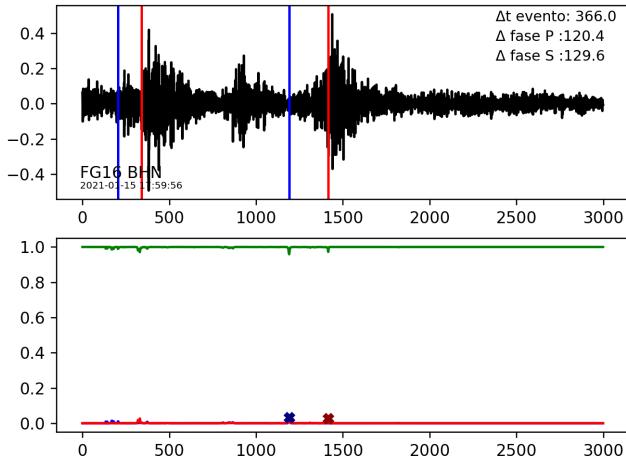
#### 3.3.1. Sin Aciertos



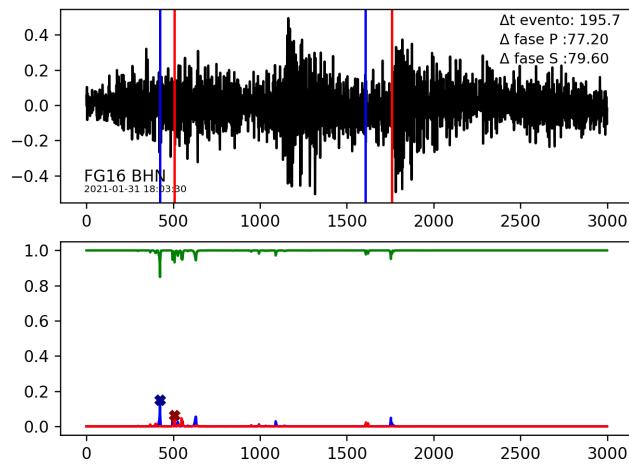
**Figura 3.9.** Muestra en la que no se tuvo coincidencia para ninguna de las fases sísmicas. En la parte superior aparece el sismograma de la muestra, en azul se aprecian las picadas de fase P(etiqueta y predicción) y en rojo las picadas de fase S . En la parte de abajo se ven las distribuciones de probabilidad resultantes de SeismicNet, el punto máximo que se aprecia en la gráfica marcado con una X de las distribuciones de probabilidad sirve para identificar el arribo de fase correspondiente, es decir generar una predicción. En la parte inferior izquierda del sismograma se identifica el código de la estación sísmica y el canal de la muestra, abajo aparece la hora del evento sísmico.



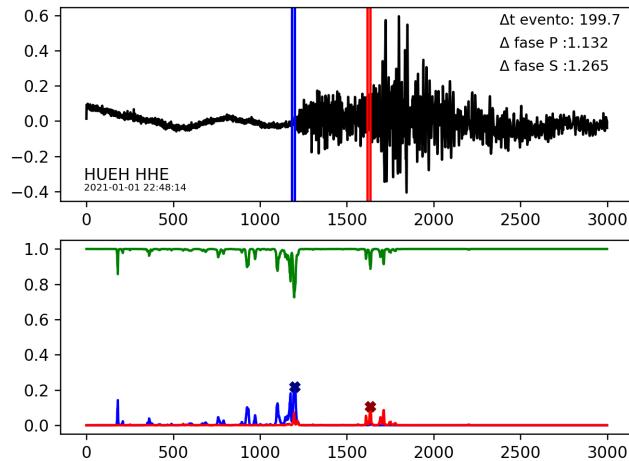
**Figura 3.10.** Muestra sin coincidencias. Se ven varios picos en las distribuciones resultantes cercanos al arribo de fase recolectado, sin embargo las posiciones de los máximos de las distribuciones y los arribos difieren por mas de 1 segundo.



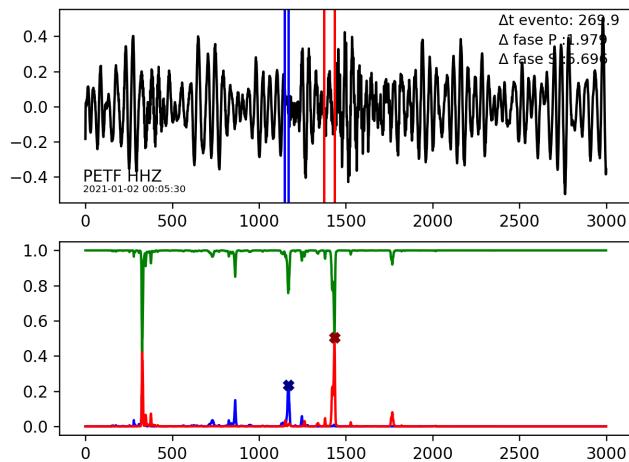
**Figura 3.11.** Muestra sin coincidencias. La predicciones de SeismicNet a veces están muy alejadas de las etiquetas recolectadas, en este caso se puede apreciar ruido considerable con naturaleza de actividad volcánica, esto tiende a confundir al modelo ya que no se tomó en cuenta esta posibilidad. El modelo esta diseñado para reconocer un evento por traza y la actividad volcánica tiende a confundir al modelo, ya que lo toma como otro evento en la misma traza. Como se pueden ver los picos de probabilidad son muy pequeños, el modelo tuvo problemas al caracterizar una muestra con varios eventos posibles, pero se puede ver las predicciones están cercanas al tercer evento de la traza.



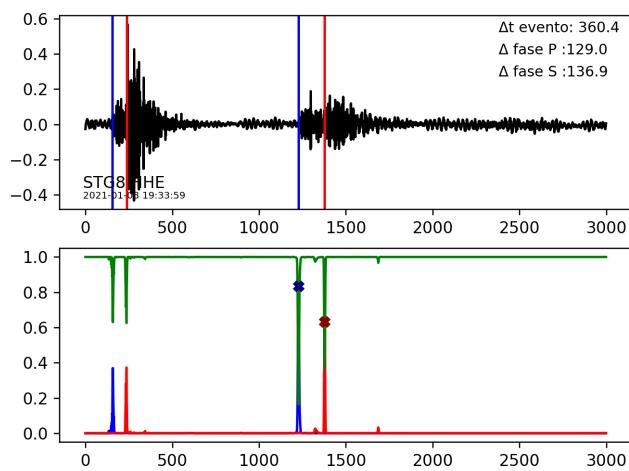
**Figura 3.12.** Muestra sin coincidencias. En este caso también se tienen varios eventos sísmicos en la misma traza, los picos en las dist. de prob. son pequeños. Hay algunas detecciones del modelo cerca de las etiquetas sin embargo las detecciones previas tienen una probabilidad mayor que hizo que las predicciones finales del modelo sean desacertadas.



**Figura 3.13.** Muestra sin coincidencias. Para estas predicciones se tienen varios picos en las dist. de prob. se puede ver claramente como los picos mas elevados están bastante cerca de las etiquetas sin embargo las diferencias para las 2 fases son mayores a 1 segundo.



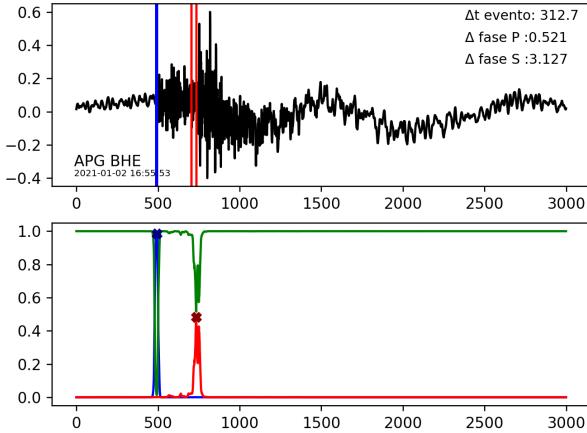
**Figura 3.14.** Muestra sin coincidencias. En este caso el evento sísmico no es muy notorio en la traza, hay varios picos en la dist. de prob., los máximos están cercanos a las etiquetas pero sin la suficiente precisión.



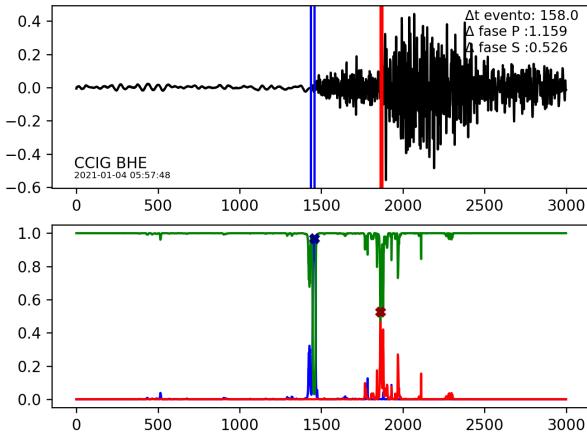
**Figura 3.15.** Muestra sin coincidencias. Se tienen picos de las dist. de prob para los dos eventos en la traza, se tienen máximos en el segundo evento, para esta muestra se ha etiquetado el primer evento resultando en una predicción errónea.

### 3.3.2. Un acierto

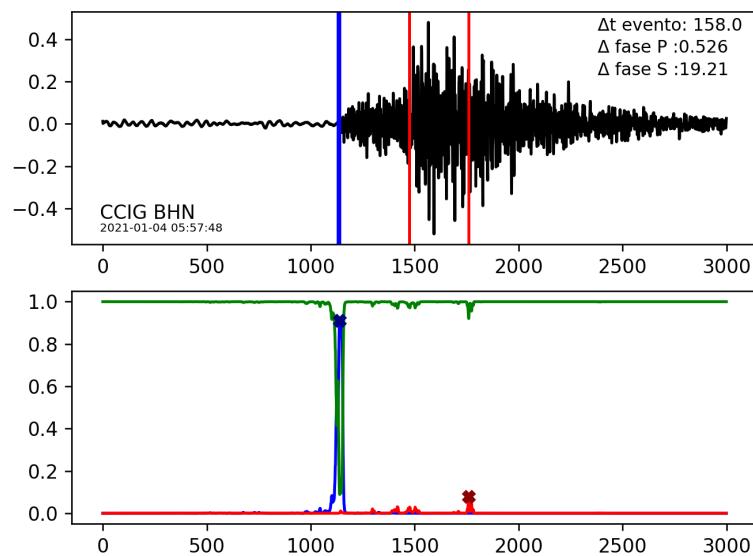
En esta sección se presentan resultados en los que se tuvo coincidencia en una de las fases.



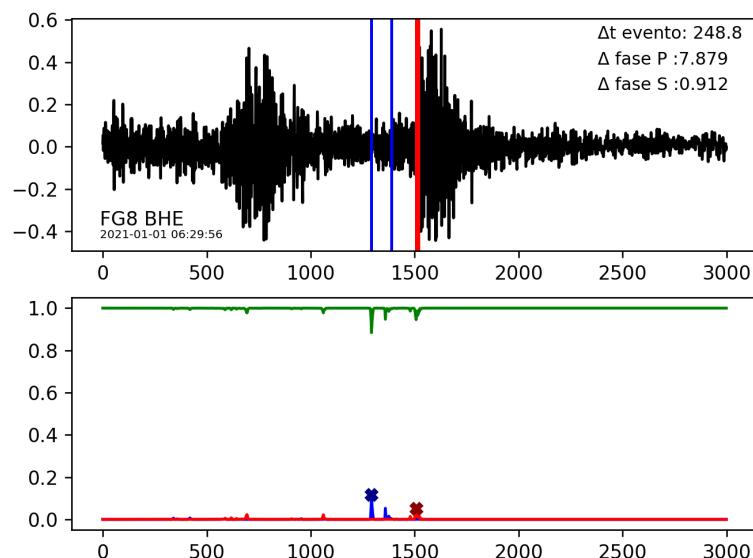
**Figura 3.16.** Muestra en la que se tiene coincidencia para la picada de fase P, la predicción de la fase S esta cercana a la etiqueta pero hay una diferencia de mas de 3 segundos, se puede notar que el pico de la fase P es bastante elevado, por lo regular se tiene este comportamiento para buenas predicciones.



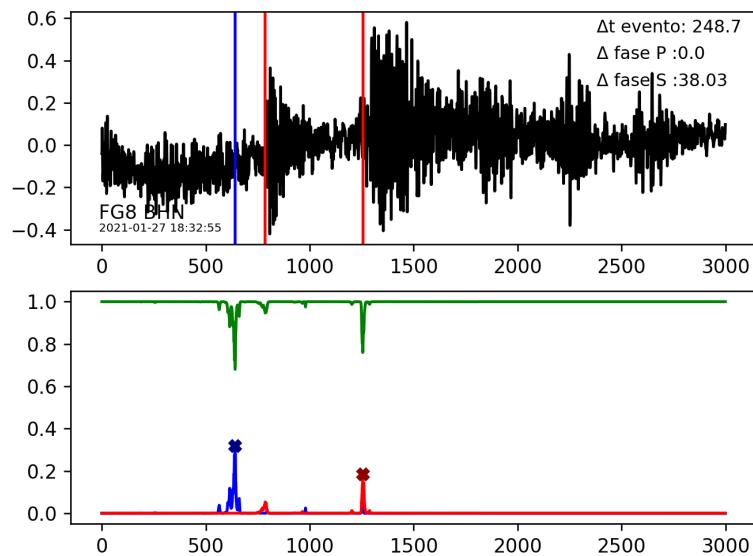
**Figura 3.17.** En las predicciones de esta muestra se puede ver que el modelo acertó en la fase S y se equivocó en la fase P, se puede ver varios picos en la dist. de prob. de la fase S, es normal que la fase S tenga varios picos de este tipo, en general las fases S plantean un reto ya que vienen superpuestas con las ondas P por lo que su correcta caracterización es mas complicada.



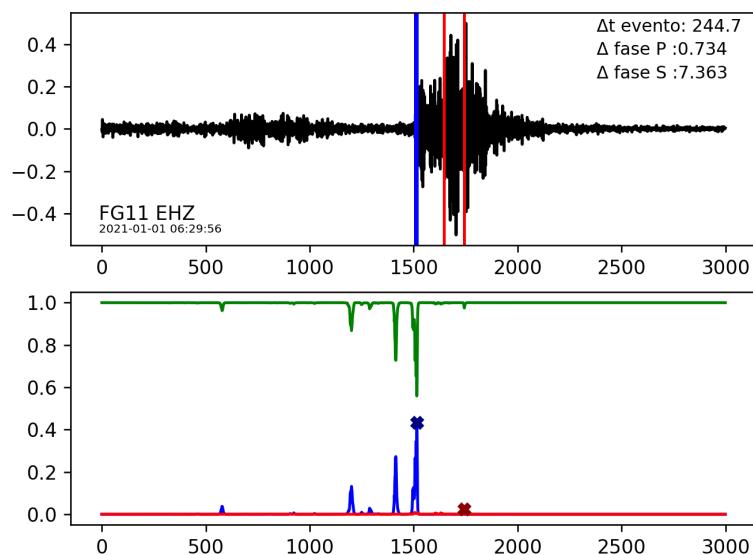
**Figura 3.18.** En este caso se tiene una predicción correcta para la fase P y para la fase S se tiene una predicción totalmente desacertada, como se aprecia la predicción esta muy lejos de la etiqueta, la diferencia es de mas de 19 segundos, resultados como este son los que se pueden mejorar con la utilización de mas datos en el entrenamiento del modelo.



**Figura 3.19.** Se tiene una picada de fase S predicha exitosamente a pesar de que tienen 2 eventos en la traza. Los picos de prob. son bajos para las picadas predichas.



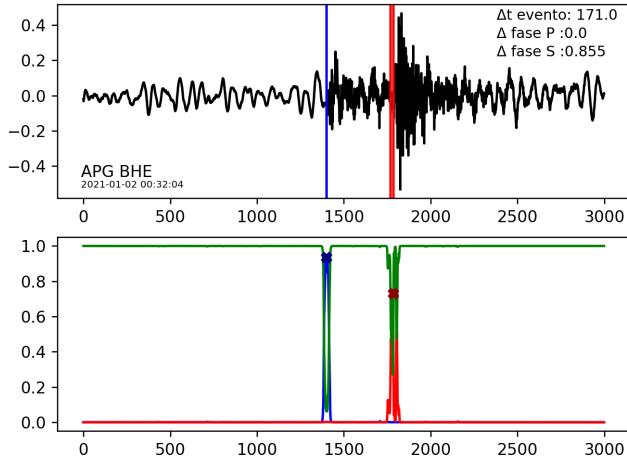
**Figura 3.20.** Aunque se tiene una predicción perfecta para la fase P, el modelo identificó en la traza un arribo posterior de fase S como la predicción para la traza, de nuevo se tiene el caso con varias señales sísmicas en una sola traza.



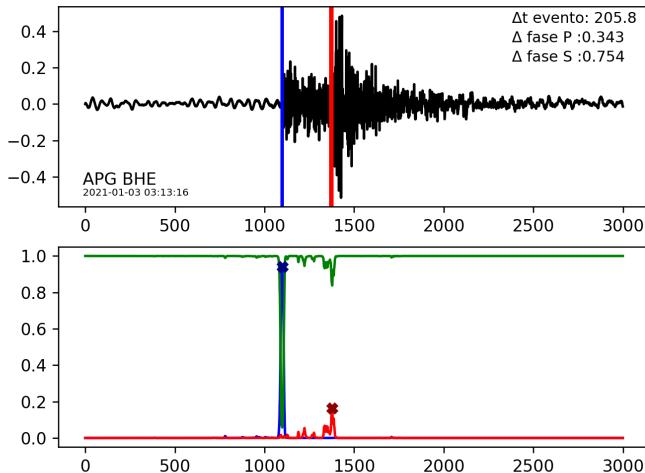
**Figura 3.21.** Para esta muestra no se pudo identificar a la fase S correctamente, el único pico de probabilidad en la dist. de fase S está en la posición equivocada. Hay varios picos en para la fase P, el máximo corresponde a una identificación correcta.

### 3.3.3. Dos aciertos

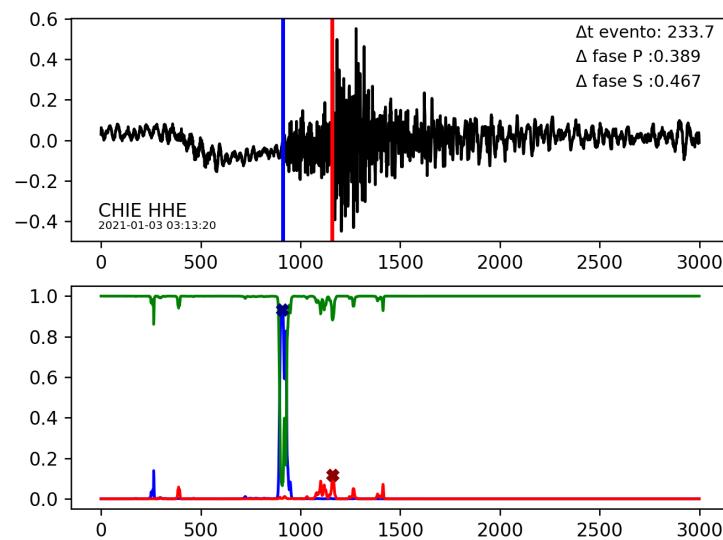
Se presentan muestras en las que se tuvo predicciones acertadas para las 2 fases sísmicas.



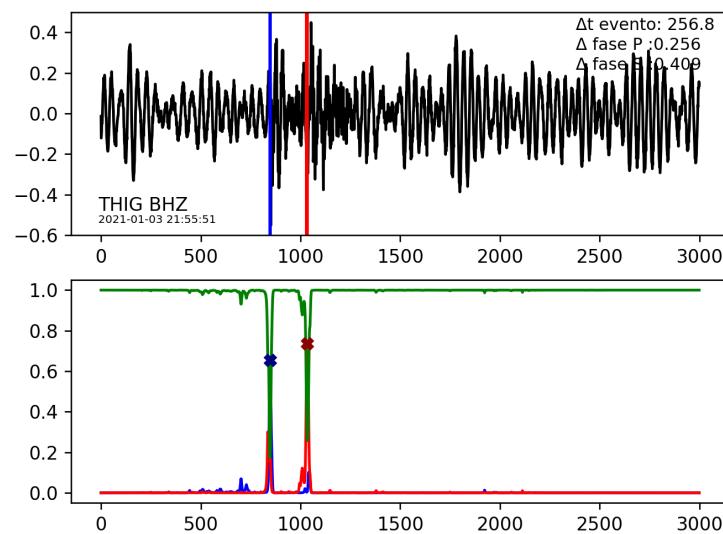
**Figura 3.22.** Muestra en la que se identificaron correctamente a las 2 fases sísmicas, para las 2 fases se tiene  $\Delta t < 0$ , en el caso de la onda P no hay diferencia entre la etiqueta y la predicción.



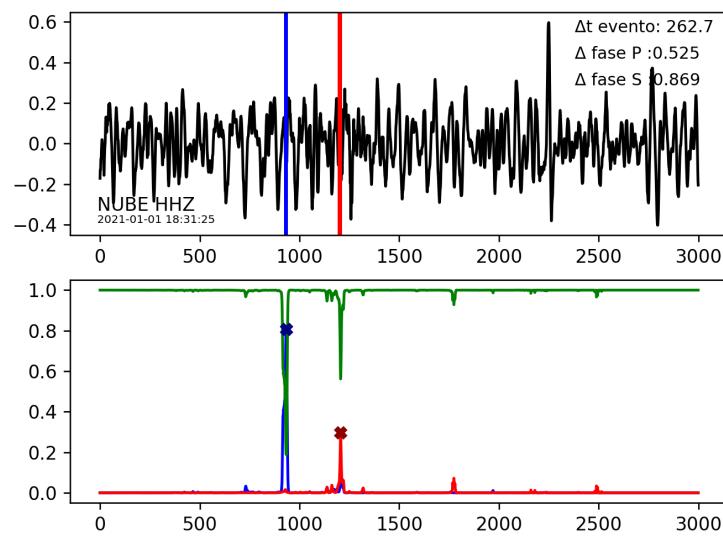
**Figura 3.23.** Se tienen dos aciertos, se puede notar que para la distribución de fase S hay varios picos de probabilidad, el máximo corresponde bastante bien con la posición de la etiqueta recolectada.



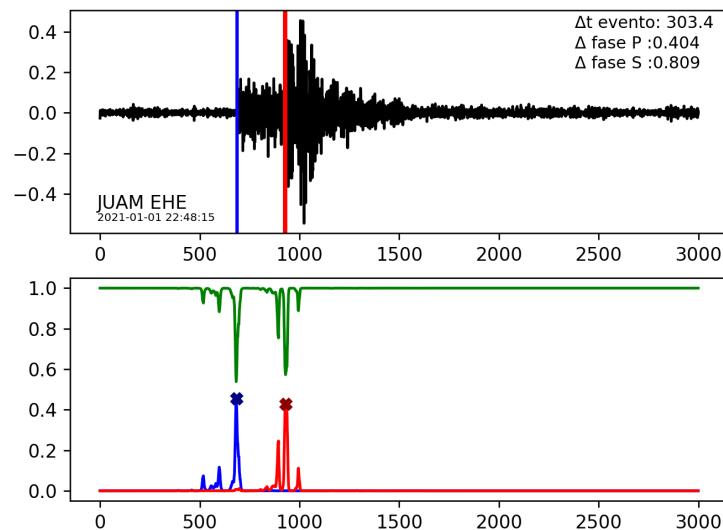
**Figura 3.24.** En este caso se pueden ver varios picos en las 2 distribuciones de probabilidad los picos mas altos coinciden con las etiquetas.



**Figura 3.25.** En esta muestra el evento sísmico no se ve a simple vista, para hacer un análisis de picadas de fase seria necesario aplicar acercamientos y filtros, SeismicNet reconoce exitosamente a las 2 fases.



**Figura 3.26.** En esta predicciones correctas de SeimicNet la muestra es visiblemente ruidosa, i.e. las amplitudes de la señal sísmicas son comparables o menores que el ruido, esto significa que el modelo aprende a reconocer otros parámetros como los cambios en el contenido de frecuencias de las ondas.



**Figura 3.27.** Se tiene varios picos cercanos a los máximos de las 2 distribuciones, las predicciones son acertadas.

### 3.4. Discusión de resultados

Uno de los resultados mas evidentes es el hecho que se superó el desempeño obtenido previamente, el aumento en la complejidad de la arquitectura del modelo de SeismicNet fue uno de los aspectos fundamentales para las mejoras, también se tiene el hecho que se utilizó un conjunto de datos mucho mas grande. En general utilizar mas datos en un modelo de CNN disminuye el fenómeno de sobre-ajuste discutido anteriormente. La forma de la etiquetas generadas(ver 2.1.2.4) fue una de las partes mas importantes del diseño de SeismicNet, esto le dio más expresividad al modelo, como se ve en los resultados presentados la red entrenada aprendió características claves de las señales sísmicas, se obtuvieron buenos resultados en muestras ruidosas y se ven picos de probabilidad en la posición correcta para trazas en las que se tienen varios eventos sísmicos(enjambres sísmicos), sin embargo el modelo tiende a confundir el ruido en estaciones volcánicas con eventos sísmicos, una mejora evidente al entrenar modelos posteriores seria depurar el uso de estaciones volcánicas con el fin de tener señales sísmicas menos ruidosas.

SeismicNet fue capaz de encontrar 57.52 % fases P y 51.16 % fases S con una diferencia  $\Delta t < 1s$  entre la etiqueta recolectada y la predicción del modelo para el conjunto de datos de enero 2021, para  $\Delta t < 2$  se obtuvieron 69.69 % fases P y 72.67 % fases S identificadas correctamente. Para el mes de febrero 2021 se tienen desempeños inferiores, esto se debe a que en el mes de febrero se tuvo un fenómeno conocido como enjambre sísmico, en dicho enjambre se dan una gran numero de eventos sísmicos en secuencias, esto hace difícil que en el procesamiento de datos utilizado actualmente se puedan segmentar eventos individuales. Esto se ve reflejado en algunos de lo resultados presentados. Como se ve en las muestras en las que se tienen 2 coincidencias, el modelo entrenado no tiene problema al identificar eventos en los que la señal sísmica es bastante clara, es decir que se tienen buenas predicciones con eventos de magnitud considerable, como se mencionó anteriormente las estaciones volcánicas presentan un reto debido a que en estas estaciones la actividad sísmica es mucho mas variada ya que no solo se tienen eventos de naturaleza tectónica sino que también se registran eventos de naturaleza propia de la actividad volcánica.

De las predicciones se pudo notar que el modelo aprendió a diferenciar entre el ruido y las señales sísmicas pero aún puede mejorar la precisión de la identificación de fases, este es un factor a tomar en cuenta antes de desplegar el modelo en cualquier flujo de trabajo real. Como se puede ver en los resultados de las distintas arquitecturas utilizadas el modelo se fue afinando con cada cambio en los hiper-

parámetros, en trabajos posteriores se pueden investigar modelos mas robustos y el uso de infraestructuras de computo mas avanzadas(TPU,GPU, Cloud computing, etc..) Los resultados de las predicciones de SeismicNet sugieren que el desarrollo de sistemas basados en técnicas recientes de aprendizaje automático tienen un gran potencial para auxiliar en las tareas de procesamiento de datos en geofísica en donde se utilizan grandes conjuntos de datos. Las mejoras en estos sistemas se dan en un continuo proceso en el que simplemente se pueden alimentar modelos establecidos con mas datos o se pueden crear rutinas de exploración de arquitecturas con el fin de tener sistemas cada vez mas “inteligentes”. Hay gran diversidad de problemas en geofísica en los que es aplicable el uso de aprendizaje automático y redes neuronales. Una de las características de la geofísica es el uso intensivo de datos, esto da el espacio perfecto para el uso de la IA en varios problemas, como se expuso en este trabajo se tiene la aplicación a la identificación de fases sísmicas. Otras aplicaciones interesantes en geofísica por mencionar algunas son: sistemas de alerta temprana, redes neuronales en las que se clasifican eventos para identificar explosiones y señales producidas por humanos, también se tienen aplicaciones variadas en geodesia y tomografía de suelos, predicción del movimiento del suelo, deformación no inercial y muchas otras([22]).

## CONCLUSIONES

1. Se desarrolló un modelo de red neuronal(SeismicNet) que mejoró considerablemente los resultados en las predicciones de fases sísmicas obtenidas con anterioridad.
2. El uso de distintas arquitecturas de exploración de red neuronal mejora el desempeño de los modelos, es decir es fundamental para optimizar hiperparámetros.
3. La calidad de los datos de entrada para el entrenamiento de una red neuronal es fundamental para optimizar el desempeño de los modelos resultantes.
4. SeismicNet tuvo un mejor desempeño en el conjunto de datos menos ruidoso, es decir los datos del mes de enero 2021<sup>1</sup>.
5. Se tuvieron resultados similares para las predicciones de SeismicNet para la fase P y la fase S en los conjuntos de datos evaluados.

---

<sup>1</sup>Durante el mes de febrero 2021 se presentó un enjambre sísmico, esto implica que el conjunto de datos de febrero es mas ruidoso, en varias formas de onda se tienen varios eventos en una sola traza.



## RECOMENDACIONES

1. Usar mas datos. Todo modelo de aprendizaje automático mejora al utilizar conjuntos de datos mas grandes, esto ya que al entrenar a un modelo cada vez con mas casos se van afinando mejor los parámetros que sirven para reconocer patrones generales del problema a resolver.
2. Explorar mas modelos(ModelSearch, Quantum Machine Learning, etc.). El desarrollo de nuevas técnicas y algoritmos innovadores para encontrar arquitecturas idóneas para un problema en específico, es una campo que actualmente esta en auge y cada día hay mas artículos donde se estudian nuevas maneras de abordar problemas de gran variedad.
3. Explorar otro tipo de modelos: redes recurrentes, redes residuales, transformadores, aprendizaje de reforzamiento etc. En este trabajo se utilizaron varias arquitecturas basadas en los filtros de convolución, sin embargo hay otras arquitecturas que han demostrado tener buenos resultados con señales de origen natural como lo son los sismogramas.
4. Como se ha mencionado con anterioridad la calidad de los datos dicta finalmente el desempeño de un modelo, esto sugiere que al tener una red sísmica mas robusta se pueden ir mejorando los modelos entrenados, por ejemplo al tener suficientes estaciones sísmicas es posible prescindir de estaciones en sitios volcánicos con el fin de disminuir el ruido presente en las formas de onda, esto ayuda para tener una mejor caracterización de eventos sísmicos.



## BIBLIOGRAFÍA

- [1] Hamed Habibi Aghdam, Elnaz Hahani Heravi. *Guide to Convolutional Neural Networks, A Practical Application to Traffic-Sign Detection and Classification.* Springer, 2017.
- [2] Shun-Ichi Amari, Kejiro Maginu. *Statistical Neurodynamics of Associative Memory.* [https://doi.org/10.1016/0893-6080\(88\)90022-6](https://doi.org/10.1016/0893-6080(88)90022-6)
- [3] Yasaman Bahri, Jonathan Kadmon, Jeffrey Pennington, Sam S. Schoenholz, Jascha Sohl-Dickstein, Surya Ganguli. *Statistical Mechanics of Deep Learning.* <https://www.annualreviews.org/doi/pdf/10.1146/annurev-conmatphys-031119-050745>
- [4] Christopher M. Bishop. *Pattern Recognition and Machine Learning.* Springer, 2006.
- [5] C.M. Bishop, I.T. Nabney. *Pattern Recognition and Machine Learning: A Matlab companion.* Springer. 2008.
- [6] Dimitri Bourilkov. *Machine and Deep Learning Applications in Particle Physics.* <https://arxiv.org/pdf/1912.08245.pdf>
- [7] Chris H. Chapman. *FUNDAMENTAL OF SEISMIC WAVE PROPAGATION.* Cambridge University Press, 2004.
- [8] Cirano De Dominicis, Irene Giardina. *Random fields and Spin Glasses. A Field Theory Approach.* Cambridge University Press. 2006.
- [9] Viktor Dotsenko. *An Introduction to the Theory of Spin Glasses and Neural Networks.* World Scientific 1994.
- [10] E. Fol, J. Coello de Portugal, G. Franchetti, R. Tomas. *OPTICS CORRECTIONS USING MACHINE LEARNING IN THE LHC.* Disponible en <https://accelconf.web.cern.ch/ipac2019/papers/thprb077.pdf>.

- [11] Alexander I. Galushkin. *Neural Networks Theory*. Springer, 2007.
- [12] Aurélien Géron. *Hands-On Machine Learning with Scikit-Learn and TensorFlow, Concepts, Tools and Techniques to Build Intelligent Systems*. O'REILLY.
- [13] Stephen Grossberg. *Nonlinear Neural Networks: Principles, Mechanisms, and Architectures*. [https://doi.org/10.1016/0893-6080\(88\)90021-4](https://doi.org/10.1016/0893-6080(88)90021-4)
- [14] Jiuxiang Gu, Zhenhua Wang, Jason Kuen Liyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, Xingxing Wang, Li Wang, Gang Wang, Jianfei Cai, Tsuhan Chen. *Recent Advances in Convolutional Neural Networks*. Disponible en <https://arxiv.org/pdf/1512.07108.pdf>
- [15] Trevor Hastie, Robert Tibshirani, Jerome Friedman. *The Elements of Statistical Learning*. Springer. 2nd Edition.
- [16] Jens Havskov, Lars Ottemöller. *Routine Data Processing in Earthquake Seismology*. Springer, 2010.
- [17] John Hertz, Anders Krogh, Richard G. Palmer. *INTRODUCTION TO THE THEORY OF NEURAL COMPUTATION*. Lecture Notes Volume I. Santa Fe Institute. Studies in the Sciences of Complexity. CRC Press 2018.
- [18] J.J. Hopfield. *Neural Networks and physical systems with emergent collective computational abilities*. <https://doi.org/10.1073/pnas.79.8.2554>
- [19] E. Ising. *Contributions to the theory of ferromagnetism*. Ph.D. thesis, and Z. Phys. **31**,253 (1925).
- [20] Raban Iten, Tony Metger, Henrik Wilming, Lídia del Rio, Renato Renner. *Discovering Physical Concepts with Neural Networks*. PHYSICAL REVIEW LETTERS. [10.1103/PhysRevLett.124.010508](https://doi.org/10.1103/PhysRevLett.124.010508)
- [21] Yuehaw Khoo, Jianfeng Lu, Lexing Ying. *Solving PDE problems with uncertainty using neural-networks*. <https://arxiv.org/pdf/1707.03351.pdf>
- [22] Qingkai Kong, Daniel T. Trugman, Zachary E. Ross, Michael J. Bianco, Brendan J. Meade, Peter Gersoft. *Machine Learning in Seismology: Turning Data into Insights*. Seismological Research Letters. <https://doi.org/10.1785/0220180259>.

- [23] Mario Kenn. *Physics Insights from Neural Networks*. <https://physics.aps.org/articles/pdf/10.1103/Physics.13.2>
- [24] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. *ImageNet Classification with Deep Convolutional Neural Networks*. <http://www.cs.toronto.edu/~fritz/absps/imagenet.pdf>
- [25] W. Lenz. *Contributions to the understanding of magnetic properties in solid bodies*. Physikalische Zeitschrift **21**, 613 (1920).
- [26] W.A. Little. *The Existence of Persistent States in the Brain*. [https://doi.org/10.1016/0025-5564\(74\)90031-5](https://doi.org/10.1016/0025-5564(74)90031-5)
- [27] Le Lu, Xiaosong Wang, Gustavo Carneiro, Lin Yang. *Deep Learning and Convolutional Neural Networks for Medical Imaging and Clinical Informatics*. Springer, 2019.
- [28] Warren S. McCulloch, Walter H. Pitts. *A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY*. <http://www.cse.chalmers.se/~coquand/AUTOMATA/mcp.pdf>
- [29] Julio A. Medina. *PRÁCTICA FINAL SUPERVISADA: Análisis de Red Neuronal para detección de fases sísmicas*. disponible en [https://github.com/Julio-Medina/Practica\\_final/blob/main/FinalReport.pdf](https://github.com/Julio-Medina/Practica_final/blob/main/FinalReport.pdf)
- [30] Umberto Michelucci. *Advanced Applied Deep Learning, Convolutional Neural Networks and Object Detection*. Apress, 2019.
- [31] Kevin P. Murphy. *Machine Learning A Probabilistic Perspective*. The MIT Press, Cambridge, Massachusetts. London, England. 2012.
- [32] Hidetoshi Nishimori. *Statistical Physics of Spin Glasses and Information Processing An Introduction*. International Series of Monographs on Physics No.111, Oxford University Press. 2001.
- [33] ObsPy. A python Framework for Seismology. <https://docs.obspy.org/>
- [34] Lars Ottemöller, Peter Voss, Jens Havskov. *SEISAN EARTHQUAKE ANALYSIS SOFTWARE*. [En linea] [seis.geus.net/software/seisan/seisan.pdf](http://seis.geus.net/software/seisan/seisan.pdf)

- [35] Lars Ottemöller, Peter Voss, Jens Havskov. *SEISAN EARTHQUAKE ANALYSIS SOFTWARE FOR WINDOWS, SOLARIS, LINUX and MACOSX*, 2017. <http://seis.geus.net/software/seisan/seisan.pdf>
- [36] Marc Mezard, Giorgio Parisi, Miguel Angel Virasoro. *SPIN GLASS THEORY AND BEYOND*. World Scientific Lecture Notes in Physics Vol. 9. 1987.
- [37] David Pfau, James S. Spencer, Alexander G. D. G. Matthews & W. M. C. Foulkes. *Ab initio solution of the many-electron Schrödinger equation with deep neural networks*. <https://journals.aps.org/prresearch/pdf/10.1103/PhysRevResearch.2.033429>
- [38] D.J. Rohde, M.J. Drinkwater, M.R. Gallagher, T. Downs, M.T. Doyle. *Applying machine learning to catalogue matching in astrophysics*. doi:10.1111/j.1365-2966.2005.08930.x
- [39] Frank Rosenblatt. *PRINCIPLES OF NEURODYNAMICS, PERCEPTRONS AND THE THEORY OF BRAIN MECHANISM*<https://apps.dtic.mil/dtic/tr/fulltext/u2/256582.pdf>
- [40] Jonathan Schmidt, Mário R. G. Marques, Silvana Botti, Miguel A. L. Marques. *Recent advances and applications of machine learning in solid-state materials science*. <https://www.nature.com/articles/s41524-019-0221-0.pdf>
- [41] Patrice Y. Simard, Dave Steinkraus, John C. Platt. *Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis*. IEEE Computer Society. Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003). <http://cognitivemedium.com/assets/rmnist/Simard.pdf>.
- [42] Seth Stein, Michael Wysession. *An Introduction to Seismology, Earthquakes and Earth Structure*. Blackwell Publishing, 2003.
- [43] Jonathan Tompson, Kristofer Schlachter, Pablo Sprechmann, Ken Perlin. *Accelerating Eulerian Fluid Simulation With Convolutional Networks*. <https://arxiv.org/pdf/1607.03597.pdf>
- [44] Ivan Vasilev, Daniel Slater, Gianmario Spacagna, Peter Roelants, Valentino Zocca. *Python Deep Learning*. Packt. Second Edition. 2019.

- [45] Bernard Widrow. *GENERALIZATION AND INFORMATION STORAGE IN NETWORKS OF ADALINE "NEURONS"*. <https://isl.stanford.edu/~widrow/papers/c1961generalizationand.pdf>
- [46] Jack Woollam, Andreas Rietbrock, Angel Bueno, Silvio De Angelis. *Convolutional Neural Network for Seismic Phase Classification, Performance Demonstration over a Local Seismic Network*. Seismological Research Letters. <https://doi.org/10.1785/0220180312>.
- [47] Rikiya Yamashita, Mizuho Nishio, Richar Kinh Gian Do, Kaori Togashi. *Convolutional neural networks: an overview and application in radiology*. Insights into Imaging. <https://doi.org/10.1007/s13244-018-0639-9>
- [48] Weiqiang Zhu, Gregory Beroza. *PhaseNet: A Deep-Neural-Network-Based Seismic Arrival Time Picking Method*. Geophysical Journal International. <https://arxiv.org/abs/1803.03211>



## A. Apéndice

### A.1. Método de Rélicas

Un “truco” para lidiar con el logaritmo que aparece en el promedio apagado 1.43 consiste en el conocido como método rélicas, todo se basa en la siguiente relación elemental:

$$\ln Z = \lim_{n \rightarrow 0} \frac{Z^n - 1}{n} \quad (\text{A.1})$$

que equivalentemente se puede reescribir como

$$Z^n \approx n \ln Z + 1, \quad \text{para } n \approx 0 \quad (\text{A.2})$$

gracias a la observación anterior el promedio del logaritmo es reducido al promedio de  $Z^n$ , como sigue. Primero se escribe explícitamente la expresión  $F_n$  introducida en 1.53

$$F_n = -\frac{1}{\beta n} \ln Z_n = -\frac{1}{\beta n} \ln \left( \int dJP(J) Z_J^n \right) \quad (\text{A.3})$$

donde  $Z_n = \overline{Z_J^n}$  (1.54), usando A.2 teniendo en mente que después se tomara el límite en el que es valida la sustitución se tiene que:

$$F_n = -\frac{1}{\beta n} \ln \left( \int dJP(J) \{1 + n \ln Z_J\} \right) \quad (\text{A.4})$$

de la normalización se sabe que  $\int dJP(J) = 1$ , entonces se tiene que

$$F_n = -\frac{1}{\beta n} \ln \left( 1 + n \int dJP(J) \ln Z_J \right) \quad (\text{A.5})$$

utilizando  $\langle \ln Z_J \rangle = \overline{\ln Z_J} = \int dJP(J) \ln Z_J$  se puede reescribir la expresión anterior como:

$$F_n = -\frac{1}{\beta n} \ln \left( 1 + n \overline{\ln Z_J} \right) \quad (\text{A.6})$$

tomando la serie de Taylor para el logaritmo exterior se obtiene

$$F_n = -\frac{1}{\beta n} (n \overline{\ln Z_J} + \mathcal{O}(n^2) + \dots) \quad (\text{A.7})$$

tomando el límite se obtiene que

$$\lim_{n \rightarrow \infty} F_n = -\frac{1}{\beta n} \lim_{n \rightarrow \infty} Z_n = -\frac{1}{\beta} \overline{\ln Z_J} = -\frac{1}{\beta} \langle \langle \ln Z_J \rangle \rangle \quad (\text{A.8})$$

consistente con el resultado en 1.54