



UNIVERSIDAD NACIONAL DE ROSARIO

FACULTAD DE CIENCIAS EXACTAS INGENIERÍA Y AGRIMENSURA

Departamento de Física

**SIMULACIÓN DE UN POTENCIAL ATOMÍSTICO MEDIANTE REDES
NEURONALES ARTIFICIALES**

Licenciatura en Física

MIRANDA INES LOPEZ

Rosario, Argentina.

Agosto, 2020

SIMULACIÓN DE UN POTENCIAL ATOMÍSTICO DE C-O/CU(110) MEDIANTE REDES
NEURONALES ARTIFICIALES

TRABAJO FINAL DE GRADO DE LA
FACULTAD DE CIENCIAS EXÁCTAS, INGENIERÍA Y AGRIMENSURA
PERTENECIENTE A LA UNIVERSIDAD NACIONAL DE ROSARIO
PARA OPTAR POR EL TÍTULO DE

Licenciada en Física

AGO 2020

autora
Miranda I. Lopez

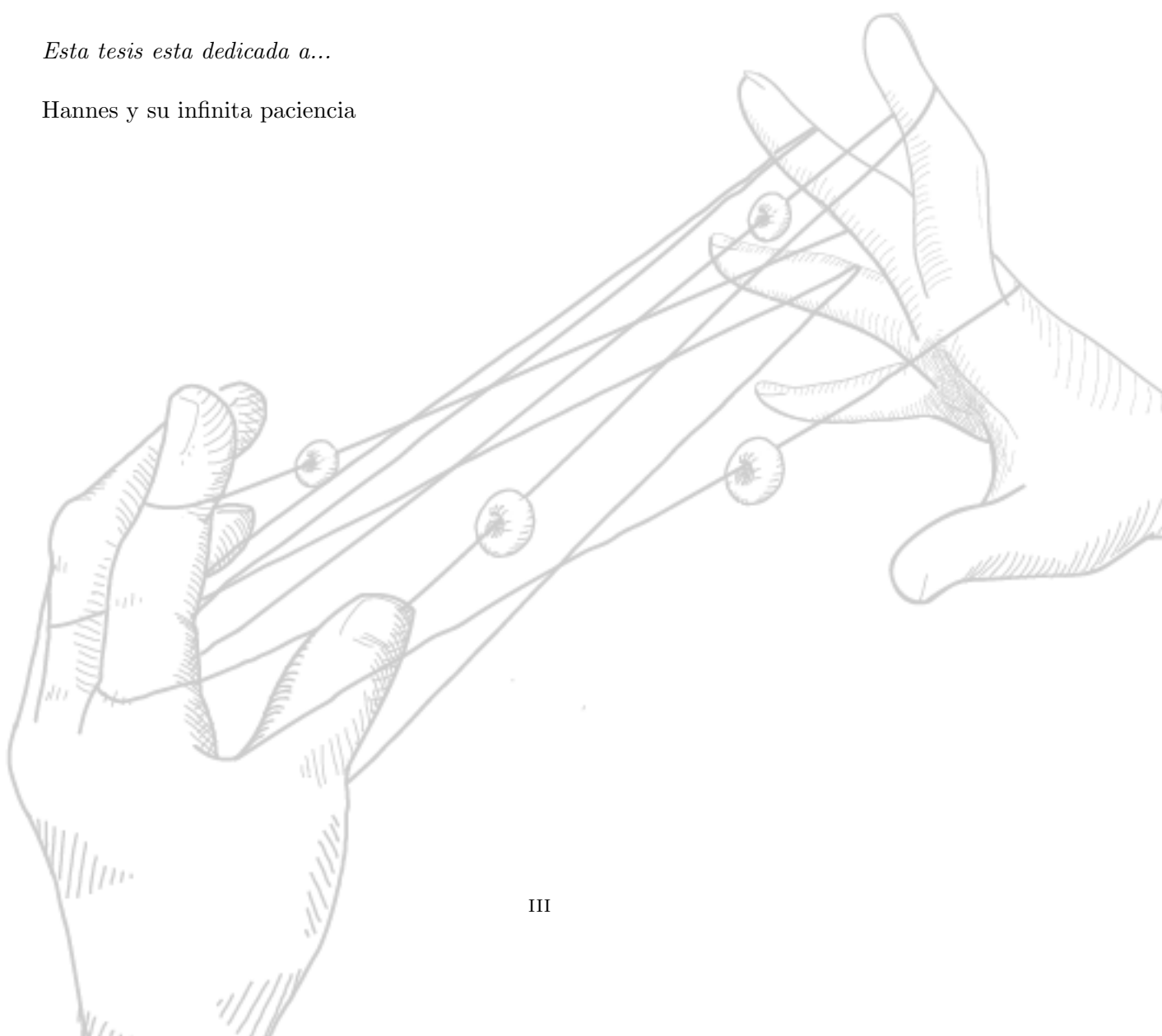
Director de tesina:
Dr. Maximiliano Ramos

Co-director de tesina:
Dr. H. Fabio Busnengo

Comité de evaluación:
Alejandra Martínez
Pablo Granitto
Maximiliano Ramos

Esta tesis esta dedicada a...

Hannes y su infinita paciencia



Agradecimientos

De todo corazón, les agradezco a mis directores, Dr. Maximiliano Ramos y Dr. Heriberto F Busnengo, en especial por toda la paciencia y conocimiento que dispusieron para el éxito de esta etapa y por la dedicación que mostraron desde el primer momento. Gracias por ayudarme a concretar mis proyectos. También quiero agradecer a los miembros de su grupo de investigación por la ayuda y cálidas discusiones, especialmente a Alejandra Martínez y Giulia Seminara por proveer los datos que conformaron la base de datos inicial de este trabajo y a Juan Manuel Lombardi por los códigos que dieron inicio al uso de redes neuronales en el grupo. Además quiero extender mi agradecimiento a Maite Alducin Ochoa y Alejandro Rivero Santamaría, del Centro de Física de Materiales (San Sebastián - España), por brindarnos las rutinas que hicieron posible acoplar el código de redes neuronales artificiales (AENET) con el código de dinámica molecular (GSRD) desarrollado localmente en el grupo del Dr. Busnengo. Finalmente agradecer a Nongnuch Artrith y Alexander Urban, de la Universidad de Columbia (New York - Estados Unidos), desarrolladores del código AENET que nos brindaron valiosas charlas y explicaciones acerca del uso de código y de desarrollo de potenciales basado en redes neuronales en general.

También quiero agradecer con mucho cariño a mi familia y amigos por que siempre estuvieron para escucharme y tolerarme en las (incontables) explosiones de nervios que me acompañaron todos los años de esta etapa.

Pero **muy especialmente**, a los amigos que me acompañaron en el camino

Nacho Fer Loren Gonza B Sol
Nati Gonza D Celi Lucho

Sin el apoyo, compañerismo y alegría que significaron, jamás hubiera podido lograrlo.

Por siempre, *gracias*.

Resumen

La investigación de los procesos físico-químicos en áreas como materia condensada, ciencias de materiales computacionales química o biología, por mencionar algunos ejemplos, requieren indefectiblemente un sólido entendimiento de las interacciones subyacentes a las complejas reacciones entre átomos y moléculas. En materia de representación de potenciales de interacción para simulaciones atomísticas, se han popularizado los métodos computacionales basados en tecnologías de inteligencia artificial, en particular: redes neuronales pre-alimentadas con entrenamiento supervisado. En este contexto, se aborda el estudio del sistema CO/Cu(110), mediante la implementación de herramientas propuestas para el desarrollo de potenciales atomísticos con redes neuronales pre-alimentadas.

Se implementó un método iterativo de desarrollo de varios potenciales atomísticos basados en redes neuronales (CO/Cu(110)), donde en cada iteración se amplió el tamaño de la base de datos (construida a partir de datos calculados con el método de teoría del funcional densidad). Con cada iteración, se identifican los puntos del espacio de configuraciones del sistema donde se estimó que el error del potencial atomístico neuronal era mayor, mediante analizar los resultados simulaciones de dinámica molecular. Las configuraciones representadas por estos puntos fueron incorporadas a la base de datos del sistema. En cada paso de entrenamiento se observó el comportamiento de la raíz del error cuadrático medio entre la energía del potencial neuronal y la de un subconjunto de los datos en la base de datos usado para entrenamiento (*train*) y entre la energía del potencial neuronal y un subconjunto de la base de datos usados para evaluación *test*. Se decidió optar por la representación artificial del potencial atomístico cuya raíz del error cuadrático medio de evaluación resultó ser menor.

Después de varias iteraciones, el potencial atómico obtenido fue utilizado en simulaciones de dinámica molecular para evaluar la probabilidad de adsorción del CO sobre la superficie de Cu(110) a $T_s = 0K$ y a $T_s = 90K$. Se presenta una disminución monótona de la probabilidad de adsorción de CO en Cu(110) en el rango de energías desde 0.01 eV hasta 2.0 eV, cuyo

comportamiento no cambia significativamente ante un cambio de temperatura superficial de $T_s = 0K$ a $T_s = 90K$. La probabilidad de adsorción para $90K$ están en excelente acuerdo con datos experimentales en las mismas condiciones.

La óptima descripción del sistema obtenida permite proyectar un futuro exitoso en la utilización de estos métodos en el campo de la dinámica de interacción de moléculas con superficies. Se espera un gran desarrollo y vastos avances en el campo de investigación de dinámicas moleculares gracias a los métodos de inteligencia artificial como el método aquí abordado.

Índice general

Agradecimientos	IV
Resumen	V
Índice de cuadros	IX
Índice de figuras	X
Indice de abreviaciones	XIII
I Introducción	1
II Metodología	7
Capítulo 1: Redes Neuronales Artificiales	8
1.1. Estructura base de una red neuronal	10
1.1.1. El Perceptrón	10
1.1.2. Función de activación	12
1.2. Arquitectura de las redes neuronales	13
1.2.1. Redes de propagación pre-alimentada	13
1.2.2. Forma funcional de las redes pre-alimentadas	15
1.3. Ajuste por minimización de la raíz del error cuadrático medio	17
Capítulo 2: Superficies de Energía Potencial con Redes Neuronales Artificiales	19
2.1. Características estructurales	19
2.2. Redes Neuronales Atomísticas	20
2.3. Funciones de Simetría	21
2.3.1. Forma funcional	21
2.3.2. Implementación: Los vectores descriptores o <i>huellas</i>	23
2.3.3. Elección de los parámetros de las FS	25

2.4. Superficies de Energía Potencial Neuronal	26
2.4.1. Arquitectura de las redes	26
2.4.2. Forma funcional de la raíz del error cuadrático medio	28
2.4.3. Sobreajuste	29
Capítulo 3: Construcción de la Superficie de Energía Potencial de C-O/Cu(110)	31
3.1. Instancia de entrenamiento	31
3.1.1. La base de datos	31
3.1.2. Proceso de Entrenamiento	34
3.1.3. Arquitectura de las RNAs	35
3.2. Instancia de análisis	36
3.2.1. Raíz del error cuadrático medio	36
3.2.2. Rasgos característicos del sistema	37
3.2.3. Trayectorias de dinámica molecular	38
III Resultados	42
Capítulo 4: Selección de las redes para la Superficie de Energía Potencial Neuronal	43
4.1. Muestreo de la base de datos inicial	43
4.1.1. Análisis de la base de datos inicial	44
4.2. Entrenamiento de redes con distintos conjuntos de parámetros de las FS	46
4.2.1. Selección de redes para optimización	48
4.3. Evaluación de los rasgos característicos del sistema	52
4.4. Efectos del incremento de configuraciones en la BD	54
4.5. Discrepancias entre redes con iguales vectores descriptores	59
4.6. Características estudiadas con la SEPN seleccionada	60
4.6.1. Rasgos característicos del potencial	61
4.6.2. Probabilidad de adsorción	62
Capítulo 5: Utilización de la Superficie de Energía Potencial Neuronal	64
5.1. Dinámica con temperatura	64
5.2. Probabilidad de Adsorción	66
5.3. Análisis del estado de las moléculas adsorbidas	68
5.3.1. Distribución de energías de interacción	68
5.3.2. Distribución espacial de las moléculas	70

5.3.3. Ángulo azimutal	71
5.3.4. Ángulo polar	72
5.3.5. Identidad de molécula	79
5.3.6. Zonas de adsorción	80
5.3.7. Energía cinética de la molécula	82
IV Conclusiones	85
Bibliografía.	89

Índice de cuadros

2.1. Pares y ternas de átomos diferentes posibles a una distancia menor al $R_c = 7\text{Å}$	23
2.2. Descripción de los casos de parámetros FS diferentes utilizados	23
4.1. Valores de la raíz cuadrática media de redes con distintas huellas.	48
4.2. Parámetros de las FS evaluados para el descriptor de dos cuerpos de las redes seleccionadas.	50
4.3. Parámetros de las FS evaluados para el descriptor de tres cuerpos de las redes seleccionadas.	51
4.4. Raíz del error cuadrático medio luego de la tercer iteración	56
4.5. Raíz del error cuadrático medio luego de la tercer iteración	59

Índice de figuras

1.	Resultados del estudio de rasgos característicos del sistema CO/Cu(110)	6
1.1.	Esquema de neuronas biológicas	9
1.2.	Esquema de un perceptrón	11
1.3.	Trazados ejemplo de funciones de activación	12
1.4.	Grafo ejemplo de red pre-alimentada individual	14
1.5.	Grafo ejemplo de red pre-alimentada multi-capa	15
2.1.	Radio de corte	20
2.2.	Arquitectura de una superficie de energía potencial	27
2.3.	Representación del ajuste de la SEPN sobre la SEP.	30
3.1.	Ciclo de construcción de la SEPN para CO/Cu(110)	32
3.2.	Arquitectura ejemplo con red RNA-Cu	36
3.3.	Orientaciones particulares de la molécula C-O	37
3.4.	Sitios característicos	38
4.1.	Figuras ejemplo de los histogramas analizados	45
4.2.	Efectos de las condiciones periódicas de contorno	46
4.3.	Zonas efectivas de las huellas descriptoras.	47
4.4.	Comparación de las desviaciones estándar.	49
4.5.	Estudio de rasgos característicos del sistema CO/Cu(110)	53
4.6.	Discrepancias de energías DM con distintas huellas	58
4.7.	Raíz del error cuadrático medio del entrenamiento de redes con igual huella . . .	60
4.8.	Rasgos característicos del sistema CO/Cu(110) con SEPN seleccionada	61
4.9.	Estudio de la probabilidad de adsorción para pocas trayectorias	63

5.1. Estudios de las configuraciones que describen la superficie a 90 K	66
5.2. Probabilidad de adsorción	68
5.3. Energía de interacción molécula-superficie final en configuraciones con adsorción.	70
5.4. Esquema del ángulo azimutal	71
5.5. Distribuciones inicial y final del ángulo azimutal	72
5.6. Esquema del ángulo polar	73
5.7. Distribuciones inicial y final del ángulo polar	75
5.8. Superficie de energía potencial según sitios	77
5.9. Posiciones de los átomos de C y O para energías de impacto bajas y altas a temperatura 0K	78
5.10. Radio internuclear C-O en configuraciones adsorbidas a t=0 ps y t=5 ps	80
5.11. Sitios de adsorción de la molécula	82
5.12. Distribución de energía cinética inicial y final de moléculas adsorbidas	84

Indice de abreviaciones

DM	Dinámica Molecular
AIMD	<i>Ab Initio Molecular Dynamics</i>
DFT	<i>Density Functional Theory</i>
SEP	Superficies de Energía Potencial
RNs	Redes Neuronales Artificiales
BD	Base de Datos
SEPN	Superficies de Energía Potencial Neuronales
SEPN AD	SEPN de Alta Dimensión
FS	Funciones de Simetría
RNA	Redes Neuronales Atómicas
CM	Centro de Masa

Parte I

Introducción

Introducción

Las simulaciones de Dinámica Molecular (DM) representan una herramienta de gran importancia para la investigación de la física de materia condensada, ciencias de materiales computacionales, química e incluso para las complejas reacciones biomoleculares [1] (ver por ejemplo referencias [2], [3], [4]). La calidad de dichas simulaciones, tradicionalmente, dependen de la eficacia de los campos de fuerzas o potenciales utilizados, generados a partir de resultados de cálculos de primeros principios y/o datos experimentales.

En el caso de implementar descripciones mecánico clásicas para el desarrollo de la dinámica, una de las dificultades fundamentales es el tamaño de los sistemas que es posible describir, ya que el costo computacional se ve muy afectado conforme la cantidad de átomos aumenta. Otra de las limitaciones que presentan las simulaciones de DM, son aquellas referidas a la cantidad de trayectorias posibles de integrar durante la simulación, por ejemplo, para estudiar propiedades estadísticas de los sistemas. Además estas trayectorias deben poder extenderse durante un periodo de tiempo considerable, dado que no todas las reacciones atómicas son procesos rápidos [5].

La eficacia de los estudios de DM depende de la eficiencia en la construcción de potenciales que representen lo más fielmente posible las complejas interacciones moleculares de sistemas de muchos cuerpos. Es decir, que la evaluación del potencial y su gradiente (las fuerzas) deben poder efectuarse con un costo computacional relativamente bajo, que permita la simulación de sistemas de muchas átomos durante tiempos relativamente largos (varios picosegundos).

Un importante y reciente avance de los estudios de DM, fue logrado con la mejora de los métodos *ab initio* (AIMD), los cuales computan las fuerzas actuantes sobre cada átomo a través de cálculos de estructura electrónica en cada paso de tiempo de la DM (obtenidos, generalmente, mediante el método del funcional densidad DFT) [1]. Si bien esta técnica ha sido utilizada con éxito para varios sistemas a lo largo de los últimos años [1], [6], la misma presenta limitaciones en cuanto a la duración de las trayectorias integradas durante la DM y al número de átomos en

el sistema, debido a su alto costo computacional. Los métodos tradicionales para sortear estas limitaciones consisten en desarrollar potenciales predefinidos con formulas analíticas sencillas. Se trata de funciones multidimensionales intrincadas (Superficies de Energía Potencial SEP o campos de fuerzas) que describen la interacción entre los átomos, las cuales dependen de sus posiciones en el espacio. Sin embargo, estos potenciales presentan ciertas deficiencias: El trabajo que se debe realizar para conseguir buenas representaciones es arduo y no poseen grandes cualidades de transferibilidad. Una vez obtenidos los campos para un dado sistema particular, este no resulta transferible cuando se le agrega complejidad al sistema, por ejemplo cuando se agrega un nuevo tipo de átomo. En contraste, los potenciales generados por redes neuronales aparecen como una alternativa muy buena a las herramientas tradicionales de construcción de las SEPs, facilitando los métodos de ajuste, lo cual confiere mayor eficiencia del costo computacional. Además aparentan ser prometedores en cuanto a las capacidades de transferibilidad del sistema [7].

Con los recientes avances tecnológicos en materia de aprendizaje automatizado (*Machine Learning*), una de las ideas más prometedoras en el campo de la dinámica molecular es la implementación de estos métodos para la representación de las SEPs. La utilización de redes neuronales para sistemas atomísticos data en menos de 20 años hasta ahora [8] pero sus aplicaciones estuvieron restringidas por cierta cantidad de años a sistemas de baja dimensionalidad. La principal dificultad era la de compatibilizar la forma genérica de las redes neuronales con la necesidad de que se satisfagan las propiedades de simetría que deben tener las SEPs (por ejemplo la permutación de átomos idénticos, periodicidad en caso de moléculas en presencia de superficies infinitas etc.). Para resolver estas complicaciones, en el año 2007, un importante aporte fue realizado por J. Behler y M. Parrinello [7], que permitió extender al campo de aplicación de los métodos de aprendizaje automatizado (en particular de redes neuronales) hacia los campos de superficies de energía potencial para sistemas de muchos átomos. A partir de ese momento, han surgido muchos trabajos utilizando el método propuesto por Behler y Parrinello y variantes del mismo, en el campo de interacción de moléculas con superficies (por ejemplo [9], [10]).

Motivados por estos últimos logros, en este trabajo se plantea como objetivo investigar la dinámica de la adsorción molecular del monóxido de carbono (CO) sobre la superficie Cu(110) a través de simulaciones de dinámica molecular utilizando potenciales de interacción atomísticos generados a través de métodos de redes neuronales artificiales.

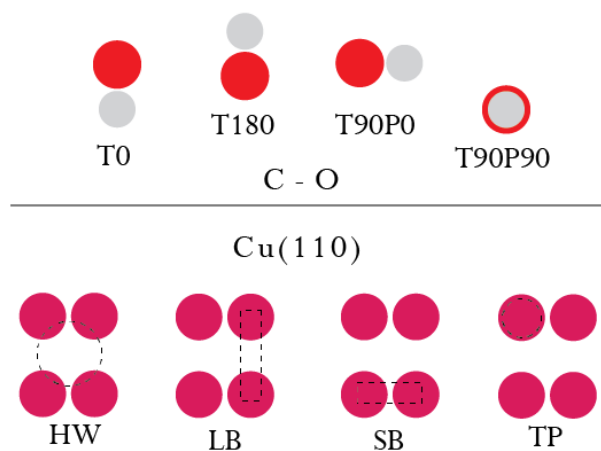
En concreto, las dificultades del sistema a abordar son las siguientes: El potencial de interacción de una molécula de CO con una superficie de Cu(110) es un potencial que depende de las 6 coordenadas que definen la posición de la molécula sobre la superficie (la posición del centro de masa y la orientación de la molécula y su distancia internuclear). Además, es necesario describir la dependencia de la energía de interacción con las posiciones de los átomos de Cu, dado que los mismos se desplazan en el tiempo dependiendo de la temperatura de la superficie del cristal. Por lo tanto la SEP depende de $6+3N$ coordenadas (donde N es el número de átomos de Cu móviles que se van a describir). En este trabajo N es igual a 18 átomos, lo que resulta en la función potencial a representar es función de 60 grados de libertad.

Como se muestra en la parte (a) de la Fig. 1, los átomos de la superficie de Cu(110) tienen una simetría rectangular. Allí se definen los sitios de alta simetría sobre la superficie: *Hollow*, *HW*, *Long Bridge*, *LB*, *Short Bridge*, *SB*, y *Top*, *TP*. La interacción depende fuertemente de la orientación de la molécula (*T0* denota una orientación perpendicular a la superficie con el átomo de C hacia la superficie, *T180* con el C hacia el vacío, *T90P0* denota una orientación paralela a la superficie en la dirección del *SB* y *T90P90* una orientación de la molécula paralela a la superficie pero en dirección del *LB*). En la parte (b) de la Fig. 1 se muestran los resultados de los cálculos DFT para esas cuatro orientaciones sobre los cuatro sitios mencionados más arriba [11].

Las principales características del potencial CO/Cu(110) descritas en la Fig. 1 son: tiene mínimos estables para la adsorción en configuraciones en donde la molécula se encuentra sobre sitios *TP* y *SB*, con sus átomos separados por su distancia de equilibrio en vacío (1.14\AA) y formando un ángulo de 0° con la vertical, es decir, la molécula se ubica verticalmente con el átomo de C más cerca de la superficie que el átomo de O. En esta misma configuración, al considerarse los sitios *LB* y *HW*, se observa una disminución de la energía potencial desde $Z_{cm} = 7\text{\AA}$ hasta $\sim 4.0\text{\AA}$. Cuando se estudia el comportamiento de la SEP para configuraciones en que la molécula se encuentra en posición horizontal, es decir con el vector internuclear paralelo a la superficie, en todos los sitios estudiados se observa una ligera disminución de la energía a medida que la molécula se aproxima desde $Z_{cm} = 7\text{\AA}$ hasta una distancia de $\sim 3.25\text{\AA}$ y para distancias más cortas el potencial aumenta y se vuelve repulsivo. En el caso en que el ángulo internuclear forma un ángulo de 180° con el eje z , es decir cuando el átomo de O se encuentra más cerca de la superficie que el átomo de C, también se observa una disminución de la energía potencial en los cuatro sitios considerados en el rango de $Z_{cm} = 7\text{\AA}$ hasta $Z_{cm} \sim 3,5\text{\AA}$. Es

decir que a largas distancias, $Z_{cm} \gtrsim 3,5\text{\AA}$ la interacción entre la molécula y la superficie es atractiva. Estas son las características principales que se intenta representar a partir del método de inteligencia artificial desarrollado para superficies de energía potencial atomística, el método de redes neuronales artificiales.

a- Sitios y orientaciones particulares del sistema



b- Presentación del potencial de interacción para los sitios y orientaciones de (a)

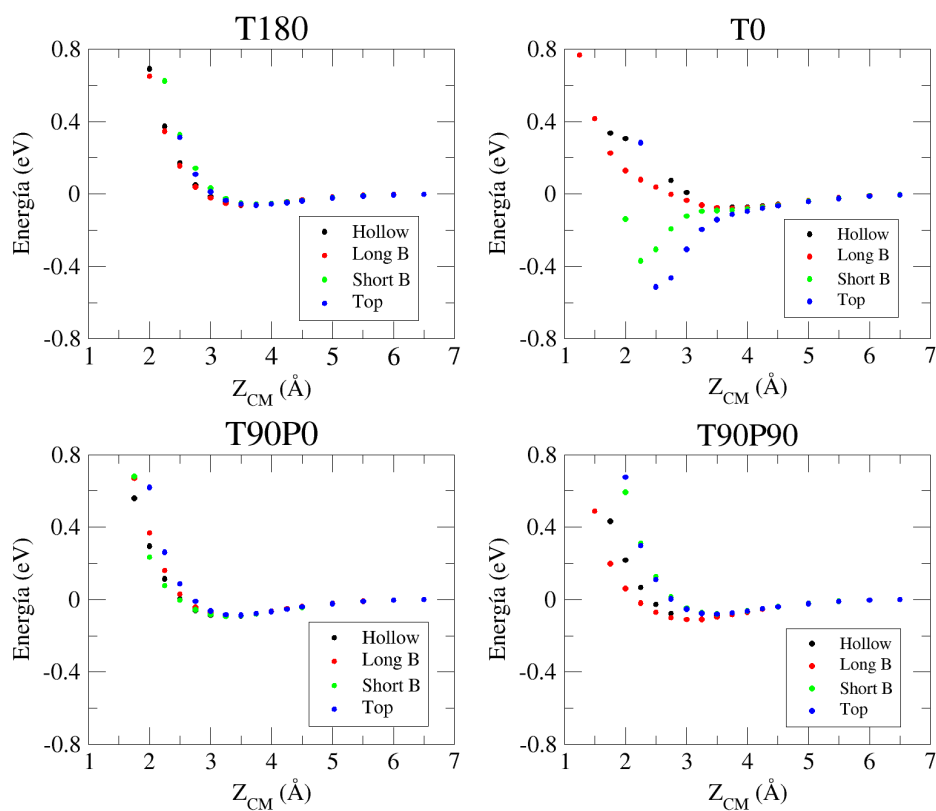


Figura 1 (a) Los círculos rojos (grises) representan a la molécula de oxígeno (carbono). La superficie de Cu se representa mediante los círculos rosa. (b) Los puntos de la figura representan los datos calculados con el método DFT que se conocen para el sistema CO/Cu(110).

Parte II

Metodología

Capítulo 1

Redes Neuronales Artificiales

Aprendizaje biológico

A principios del siglo XX, el estudio de los mecanismos de procesamiento de información del cerebro fue una fuente de inspiración para las ciencias de aquella época, y los resultados obtenidos han sido tan innovadores que su aplicación sigue en auge hasta ahora.

Sencillas tareas (en primera instancia) como el reconocimiento de un rostro familiar en un entorno desconocido, el control de la motricidad del cuerpo, clasificación de patrones, entre otras; son llevadas a cabo por el cerebro en rangos de tiempo de 100-200ms [12], mientras que (tareas incluso mas simples) requerirían tiempos de operación mas grandes aún para las más desarrolladas computadoras actuales. El reconocimiento de estas capacidades biológicas fueron fuente de inspiración para las primeras hipótesis sobre los hiper conectados y complejos, no obstante, altamente eficientes, mecanismos de procesamiento de información del cerebro.

El cerebro es un sistema de procesamiento de información altamente complejo, no lineal y de computo paralelizado [12]

El pionero en esclarecer el intrincado mecanismo del cerebro fue Santiago Ramón y Cajal, quien en el año 1911 introdujo la idea de neuronas como la unidad estructural del cerebro. Cada neurona del cerebro humano tiene propiedades particulares y especiales para recibir, procesar y transmitir señales electroquímicas a través de todas las interconexiones del sistema de comunicación del cerebro.

Del cuerpo de una neurona, como se esquematiza en la Fig. 1.1, se extienden las dendritas hacia otras neuronas de la red, donde reciben las señales de entrada. El punto de contacto o de conexión se llama sinapsis. Estas entradas son dirigidas al núcleo donde se les asigna un peso sináptico a sus estímulos y se los suma. Algunas combinaciones de estímulos tienden a

Esquematación de una neurona biológica

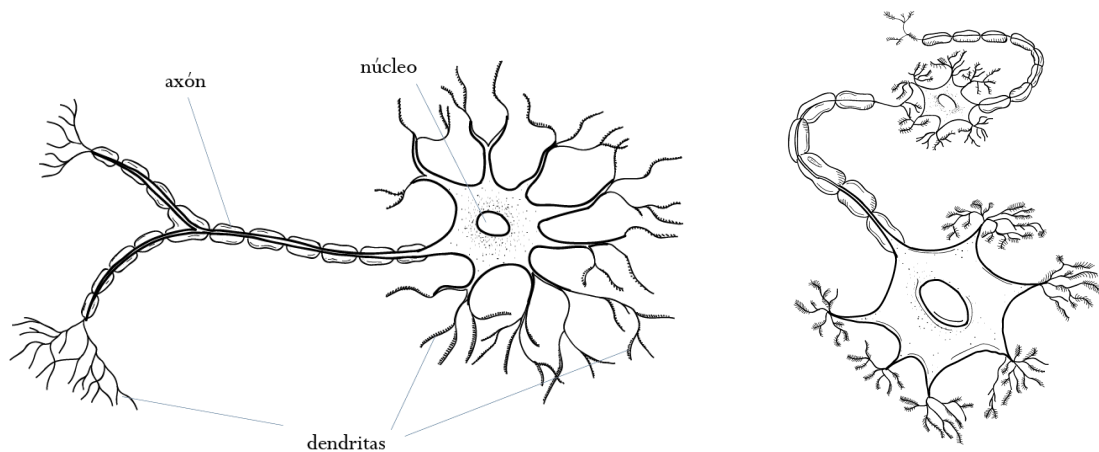


Figura 1.1 Las neuronas son la unidad elemental de los interconectados y altamente eficientes, mecanismos de procesamiento de la información del cerebro. Una neurona biológica general se esquematiza en el diagrama de la izquierda. Las dendritas interactúan con el entorno que las rodea y los estímulos que perciben por medio de la sinapsis neuronal, son comunicados al núcleo de la neurona. La respuesta del núcleo sobre este estímulo se comunica a las demás neuronas de la red a través del axón de la neurona.

excitar a la célula y otras, por el contrario, tienden a inhibirla. Esta respuesta del núcleo es transmitida hacia las demás neuronas a través del elemento comunicante del cuerpo neuronal: el axón. Cuando la excitación acumulada supera un valor umbral, la neurona envía una señal que estimula la activación de las demás neuronas de la red a través del axón. De esta forma, los estímulos de entrada viajan a través del sistema de neuronas interconectadas hasta convertirse en la respuesta de salida del cerebro sobre dicho estímulo.

Sin embargo, la parte fundamental de este proceso es que la sinapsis entre neuronas no es estática. El cerebro tiene la habilidad, desde su formación muy temprana, para adaptarse al entorno que lo rodea a través de la experiencia que recibe de sus diferentes receptores, mediante ajustar la importancia con que pesa cada estímulo. Se denomina a esta capacidad como "plasticidad". En otras palabras, las respuestas que cada neurona transmite ante ciertos estímulos son variables a través de la experiencia: excitando, inhibiendo e incluso hasta creando nuevas sinapsis neuronales entre sus núcleos. A este proceso de adaptabilidad es a lo que comúnmente se le llama *aprendizaje*.

1.1. Estructura base de una red neuronal

1.1.1. El Perceptrón

Las Redes Neuronales Artificiales (RNs) fueron inspiradas por la organización del complejo sistema neuronal del cerebro humano. La mayoría de los modelos de las RNs presenta este funcionamiento básico de la neurona cerebral: una señal externa es recibida, procesada y una nueva señal es expulsada como respuesta. A diferencia de las redes neuronales biológicas, las RNs son funciones analíticas representadas por algoritmos computacionales.

En el caso de RNs con *entrenamiento supervisado*, la red aprende a predecir, e incluso a generalizar, una respuesta o salida sobre un determinado estímulo, a través de un proceso iterativo conocido como *entrenamiento*. En este proceso la red predice un resultado, sobre determinados valores de entrada, y lo optimiza iterativamente, hasta que la respuesta de salida coincide lo suficiente con un conjunto de datos ejemplo disponible para la red de antemano (llamado valores objetivo o *conjunto de datos ejemplo*). La red auto-ajusta sus parámetros para que la diferencia entre el valor objetivo buscado y el valor predicho por la red sea apropiada. Al finalizar el proceso (en el caso en que los valores objetivo-predicción estén adecuadamente cerca) se dice que la red neuronal ha aprendido la respuesta correcta a dicho estímulo.

Para las RNs, la unidad análoga a la neurona biológica es la unidad de procesamiento o nodo, denominada también *perceptrón*. Esta unidad básica fue la primer descripción algorítmica de una red neuronal. Rosenblatt, un psicólogo, fue quien inventó esta estructura e inspiró durante los años 1960s y 1970s a quienes estudiaban los campos de ingeniería, física y matemática, a dedicarse con devoción y esfuerzo a la investigación de los distintos aspectos de las redes neuronales [12].

El perceptrón, que se esquematiza en la Fig 1.2, recibe como estímulo de entrada valores numéricos representativos del sistema sobre el cual se desea aprender cierta propiedad o analizar sus características (un vector multidimensional $\mathbf{x} = \{x_j\}_{j=0}^N$).

La señal j -ésima se considera con cierta fuerza o, más precisamente, es sopesada por los parámetros del nodo i -ésimo $\mathbf{w}_i = \{w_{ij}\}_{j=0}^N$ y combinada con las demás entradas, normalmente, en una suma básica $\mathbf{w}_i \cdot \mathbf{x} = \sum_{j=0}^N w_{ij} \cdot x_j$. La suma de las entradas se modifica por una función de transferencia $\psi_i(\star)$ que deforma el hiper-plano generado por la combinación lineal de las entradas (la combinación $\mathbf{w}_i \cdot \mathbf{x}$). La salida de esta función se pasa directamente como salida del elemento procesador, resultando, para la respuesta del nodo i -ésimo:

El perceptrón

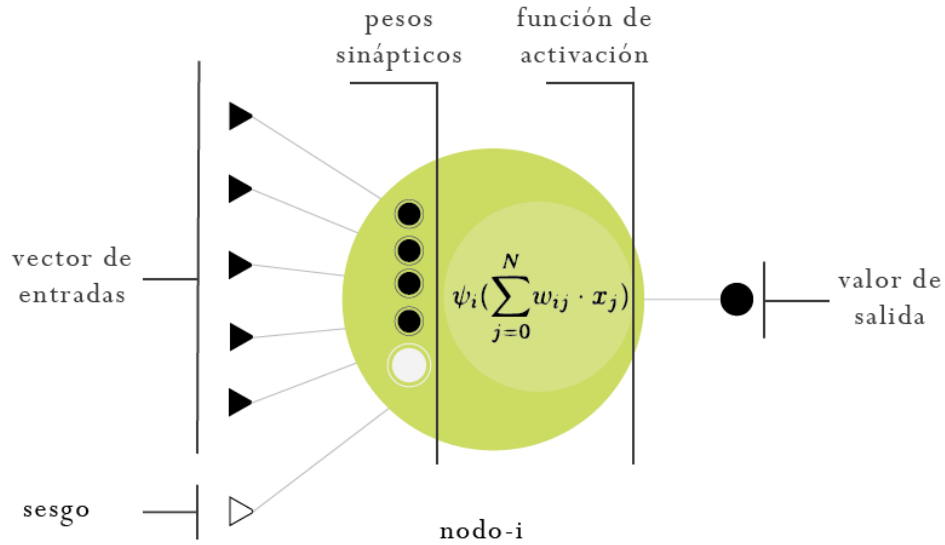


Figura 1.2 El perceptrón es la unidad básica de las redes neuronales. Los valores de entrada y el sesgo son procesados por los pesos sinápticos del perceptrón, o nodo, y por la función de activación. Comúnmente el procesamiento del nodo consiste en evaluar la función de activación en una suma básica resultado de la combinación lineal entre las entradas y los pesos.

$$y_i = \psi_i\left(\sum_{j=0}^N w_{ij} \cdot x_j\right) = \psi_i(\mathbf{w}_i \cdot \mathbf{x}). \quad (1.1)$$

La importancia de los parámetros de peso \mathbf{w}_i en la suma del campo de valores de entradas de la red, es fundamental. Son éstos parámetros los que la red auto-ajusta para determinar la relevancia con que cada una de las entradas debe considerarse al optimizar la respuesta de salida de la red.

En esta suma sopesada, uno de los términos particulares es conocido como factor de sesgo o *bias*. El parámetro de sesgo w_{0i} se encarga de aumentar o disminuir, de manera directa, la entrada en la función de activación, otorgándole mayor flexibilidad a la red para ajustar sus parámetros. Para cualquier red, se considera que este parámetro es el valor que pesa la entrada $x_0 = 1$, puesto que muchas veces la suma $\mathbf{w}_i \mathbf{x} = \sum_{j=0}^N w_{ij} x_j$ también se presenta como $\mathbf{w}_i \mathbf{x} = w_{i0} \cdot 1 + \sum_{j=1}^N w_{ij} x_j$. Otra forma de entender este factor de sesgo, es como la ordenada al origen del plano hiper-dimensional generado por el vector de entradas de la red.

1.1.2. Función de activación

La función de activación o transferencia $\psi_i(\star)$ denota el valor de salida del elemento procesador en respuesta al campo de valores generados por las entradas y los pesos. Las funciones de activación son funciones no-lineales, continuas, que se utilizan como imitación de los umbrales de activación de las neuronas biológicas.

Originariamente las funciones de activación eran representadas por la función escalón, en la búsqueda por imitar la señal binaria de las neuronas biológicas, pero varios tipos de formas funcionales pueden ser implementados. El uso de funciones continuas permite reemplazar la señal binaria de las RNs originales por valores continuos, necesarios para funciones tales como las concernientes a los sistemas atómicos.

En este contexto, dentro de los casos de implementación de funciones con variable continua, las funciones de activación más comúnmente usadas son la función Sigmoidea y la tangente hiperbólica, las cuales convergen a valores constantes para argumentos muy positivos o muy negativos, mientras que para valores intermedios, la función presenta una región no-lineal, que permite a la RNs representar funciones continuas de forma arbitraria. [13]

Trazados ejemplo de una función de activación

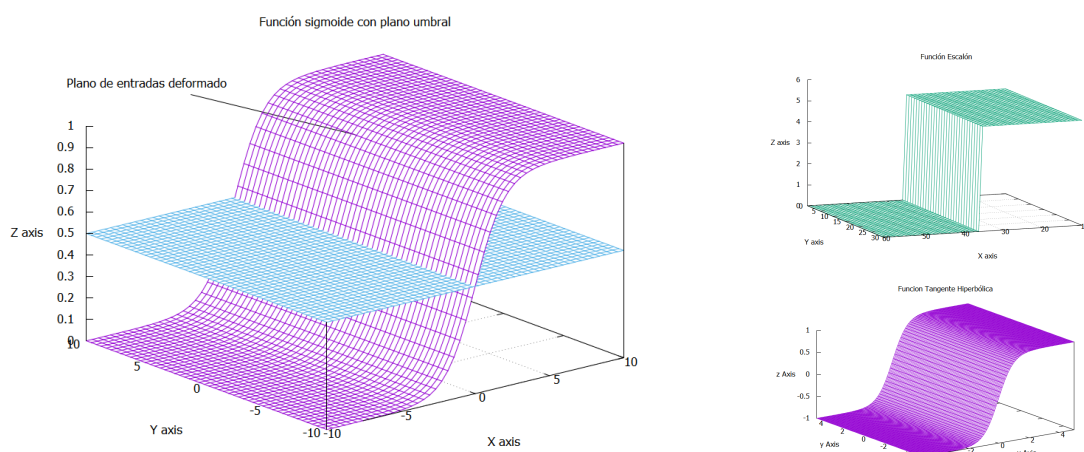


Figura 1.3 El caso ejemplo esquematizado en la figura es la interpretación geométrica de las funciones de activación utilizadas para redes neuronales con clasificación. Los resultados q se encuentren por encima del plano umbral (plano celeste), se los clasifica como un grupo y los que estén por debajo, se los clasifica en otro grupo. A la derecha, otros esquemas de funciones de activación implementadas.

El efecto de la función de activación es el de distorsionar el plano generado por las entradas de la neurona (*plano* $\mathbf{w}_i \cdot \mathbf{x}$). Esta distorsión, o más específicamente combinaciones lineales de esta distorsión, es la base para la eficaz representación de superficies de alta complejidad que ofrecen la redes neuronales.

1.2. Arquitectura de las redes neuronales

La infraestructura hiper conectada que caracteriza la habilidad y alta capacidad de aprendizaje de las redes neuronales, comienza cuando la unidad de procesamiento (el perceptrón) se combina con otros nodos; formando un conjunto de unidades de procesamiento que determinan una capa de la red. Dependiendo el tipo de red, las respuestas de esta capa de nodos ingresan como señal de entrada hacia otra capa de nodos, la cual procesa estas señales hacia otras capas adyacentes. La información de entrada atraviesa así la red de capas, hasta que la señal inicial se transforma en la respuesta de salida de la red.

1.2.1. Redes de propagación pre-alimentada

Dependiendo el tipo de conexión y la organización de las capas de nodos, existen distintos tipos de redes neuronales. El nombre pre-alimentada o *feed forward* hace referencia a que el flujo de información proveniente del vector de entradas o *input*, atraviesa la red únicamente en un solo sentido y los valores de salida de una capa son los valores de entrada de su capa inmediata posterior, pero nunca en el sentido contrario. Las capas de nodos que interactúan directamente con las señales de entrada y salida de la red, se las conoce como capas de entrada (*input layer*) y capas de salida (*output layer*), respectivamente. Cuando existen capas de nodos entre las capas de entrada y salida, que no interactúan directamente con las señales externas, se las llama con el nombre de capas ocultas (*hidden layer*).

El tipo de red neuronal pre-alimentada más simple conocido, consiste de una única capa individual de nodos (*single-layer*) cuyo grafo se esquematiza en la Fig. 1.4.

En el grafo de la capa-individual (Fig. 1.4) la red neuronal recibe 6 señales o valores estímulo externos, que actúan como entrada de los 3 nodos que conforman la capa individual. Los nodos, procesan la información ingresada y transmiten su respuesta como valores de salida de la red.

Grafo ejemplo de red pre-alimentada individual

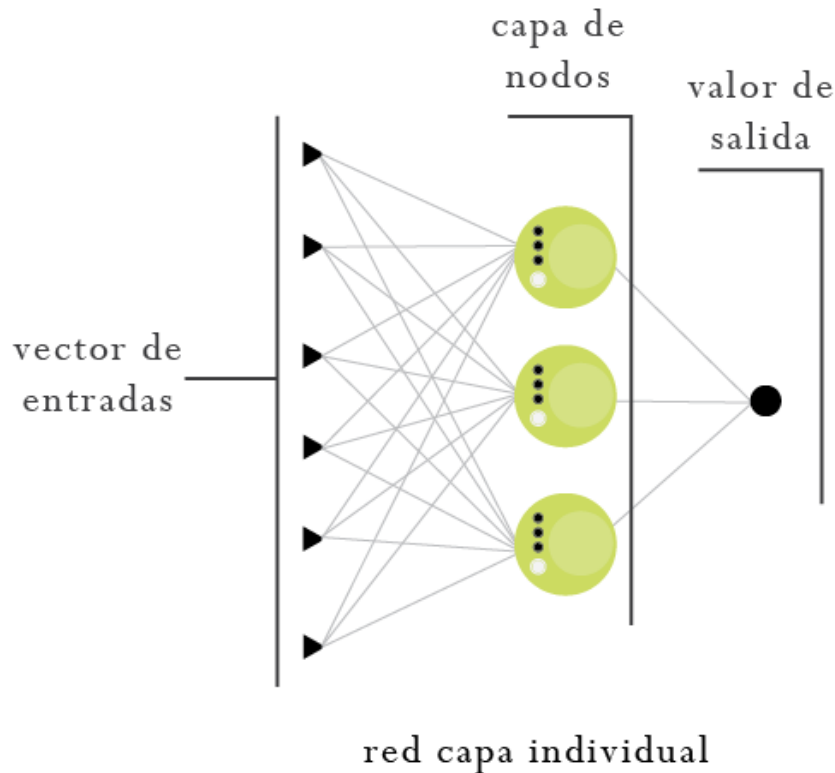


Figura 1.4 En la figura superior se esquematiza la red neuronal de capa individual. La información atraviesa las capas en un único sentido (redes prealimentadas) y el vector de entradas se transforma en un único valor de salida que se considera la respuesta de la red.

Al incorporar capas ocultas, se construye la red neuronal pre-alimentada multi-capas (*multilayer*). El grafo de una red pre-alimentada multi-capas, se presenta como ejemplo en la Fig. 1.5. Constituida por una serie A-B-C de nodos ordenados en tres capas.

En el grafo de la arquitectura multi-capas (ver Fig. 1.5) la red neuronal recibe 6 señales o valores estímulo externos, que actúan como entrada de los 3 nodos de la capa de entrada A. Los nodos procesan la información ingresada y transmiten su respuesta como valores de entrada hacia la primer (y única) capa oculta de la red (la capa B). Al ingresar a esta nueva capa, una vez más, la información es procesada en los 4 nodos correspondientes a la capa B, y las señales de respuesta son transmitidas hacia los últimos 2 nodos de la capa C, donde, finalmente, se efectúa la respuesta final de salida de la red.

Grafo ejemplo de red pre-alimentada multi-capa

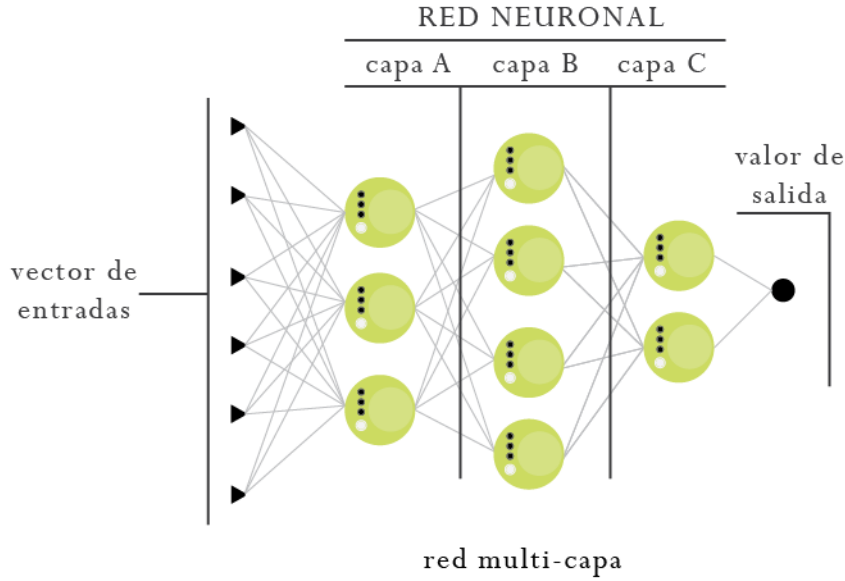


Figura 1.5 En la figura superior se esquematiza la red neuronal multi-capa A-B-C. De la misma forma que en la red de capa individual, la información atraviesa la red en un único sentido, salvo que en este caso la información se procesa en tres capas de nodos A-B-C. Luego de atravesar la red y ser procesada, lo *inputs* de la red se convierten en el valor de salida de esta (*outputs*).

1.2.2. Forma funcional de las redes pre-alimentadas

Comenzando por el perceptrón, la unidad elemental de las redes neuronales, la salida de un nodo j -ésimo con función de activación $\psi_j(\star)$, que recibe un vector \mathbf{x} de N estímulos, se calcula mediante

$$z_j = \psi_j \left(\sum_{i=0}^N w_{ji}^I \cdot x_i \right) \quad (1.2)$$

La expresión anterior (Eq. 1.2), se corresponde con un nodo j -ésimo de la capa de entradas (por ejemplo, de la capa A en Fig. 1.5) que sopesa el vector de entradas \mathbf{x} , mediante el vector de pesos inicial j -ésimo \mathbf{w}_j^I , donde el subíndice j indica que se trata del vector de pesos correspondiente al nodo j -ésimo y el supraíndice I que es el vector de pesos *inicial*. Luego, al agrupar este nodo con los restantes N_A nodos de la capa, se obtiene un arreglo de valores respuesta. A este vector de respuesta, interno a la red, se lo puede etiquetar como \mathbf{z}^A . Las componentes de este vector $\{z_j^A\}_{j=1}^{N_A}$ indican las salidas individuales de cada nodo j -ésimo contenido en la capa A que entran como estímulos de la primera capa oculta de una estructura multi-capa (capa B en Fig. 1.5).

$$\left\{ \psi_j \left(\sum_{i=0}^N w_{ji}^I x_i \right) \right\}_{j=1}^{N_A} = \{z_j^A\}_{j=1}^{N_A} = \mathbf{z}^A \quad (1.3)$$

A continuación, las respuestas \mathbf{z}^A se implementan como entradas de los N_B nodos de la capa oculta (capa B en Fig. 1.5). El vector de entradas -interno- \mathbf{z}^A se combina como una suma sopesada por los parámetros de la matriz $\mathbf{w}^{BA} = \{w_{kj}^{BA}\}_{k=1}^{N_B}$. Los supra índices BA indican que el vector sopesa las entradas procedentes de la capa A, en los nodos de la capa B. Los subíndices de las componentes $\{w_{kj}^{BA}\}_{k=1}^{N_B}$ hacen referencia a que se sopesa la entrada j-ésima (de la capa A) en el nodo k-ésimo (de la capa B).

$$\left\{ \psi_k \left(\sum_{j=0}^{N_A} w_{kj}^{BA} \cdot z_j^A \right) \right\}_{k=1}^{N_B} = \{z_k^B\}_{k=1}^{N_B} = \mathbf{z}^B \quad (1.4)$$

Nuevamente, al combinar el vector \mathbf{z}^B linealmente con la matriz de pesos \mathbf{w}^{CB} , se obtiene un vector de salidas de la capa C, de la forma

$$\left\{ \psi_l \left(\sum_{k=0}^{N_B} w_{lk}^{CB} \cdot z_k^B \right) \right\}_{l=1}^{N_C} = \{z_l^C\}_{l=1}^{N_C} = \mathbf{z}^C \quad (1.5)$$

Para finalizar, una vez más, las salidas de la capa C se combinan, mediante el vector de pesos final \mathbf{w}^F , pero en una suma sencilla durante el procesamiento del único nodo de la capa final, cuya función de activación, generalmente, se corresponde con la función identidad $\mathfrak{J}(\star)$, resultando

$$\mathfrak{J} \left(\sum_{l=0}^{N_C} w_l^F \cdot z_l^C \right) = y \quad (1.6)$$

La combinación de las Eqs. 1.3, 1.4, 1.5 y 1.6 se traduce en una función hiper dimensional sofisticada, algo extensa, pero relativamente sencilla y lineal, que arroja el valor de salida de una red multi-capas con estructura A-B-C. El valor determinado depende explícitamente del vector inicial de entradas \mathbf{x} y del conjunto de matrices peso $\{\mathbf{w}^I\}, \{\mathbf{w}^{BA}\}, \{\mathbf{w}^{CB}\}, \{\mathbf{w}^F\}$.

La forma funcional hiper dimensional y determinada que subyace al tipo de red pre-alimentada multi-capas ejemplificado en la Fig. 1.5 produce

$$y = \mathfrak{J} \left(\sum_{l=0}^{N_C} w_l^F \cdot \psi_l \left(\sum_{k=0}^{N_B} w_{lk}^{CB} \cdot \psi_k \left(\sum_{j=0}^{N_A} w_{kj}^{BA} \cdot \psi_j \left(\sum_{i=0}^N w_{ji}^I \cdot x_i \right) \right) \right) \right) \quad (1.7)$$

Este resultado es la base fundamental de la inmensa cantidad de aplicaciones de los algoritmos de redes neuronales como funciones de interpolación y ajuste de datos, ya que se pueden construir funciones hiper-dimensionales sofisticadas (tales como las SEP) mediante evaluación directa de la funcional que propone la red neuronal.

En este punto es importante señalar que el proceso de ajuste de parámetros de interpolación (conocido como *el entrenamiento de la red*) consiste en optimizar los valores del conjunto de parámetros peso $\{\mathbf{w}^I\}, \{\mathbf{w}^{BA}\}, \{\mathbf{w}^{CB}\}, \{\mathbf{w}^F\}$ de manera que se minimice la diferencia entre la respuesta de salida de la red y algún resultado ejemplo objetivo, del cual se sirve la red para ejercer su optimización.

Dado que el sistema cuenta con la forma funcional a partir de la cual se calculan los valores de salida de las redes, es posible obtener tanto las derivadas de los pesos sinápticos de la red, útiles para los procesos de optimización de funciones e interpolación de puntos, como las derivadas respecto de las coordenadas de entrada $\{x_i\}$, útiles para aplicaciones como la simulación de dinámicas moleculares.

No obstante, para la implementación de este tipo de redes neuronales en la representación de SEP, es menester realizar un pre-procesamiento de las coordenadas que definen la estructura atómica (de la cual depende la SEP) ya que de lo contrario las propiedades físico-químicas no logran ser debidamente satisfechas. Los detalles particulares del armado de SEP con redes neuronales artificiales se presentan en el siguiente capítulo.

1.3. Ajuste por minimización de la raíz del error cuadrático medio

De forma general, una red neuronal debe aprender las particularidades sobre el ambiente en el cual será implementada. Para esto, precisa conocer de antemano las respuestas deseadas sobre dicho ambiente. Para redes con aprendizaje supervisado, el sistema dispone de un conjunto de datos ejemplo sobre los cuales basará el ajuste de los parámetros de pesos y sesgos conocido como base de datos de entrenamiento o simplemente base de datos (BD).

La red se somete a un entrenamiento iterativo; el cual no es más que una interpolación de puntos en un espacio multidimensional definido por la BD, mediante un ajuste de los parámetros característicos de su forma funcional (el ajuste de los pesos sinápticos). Para llevar a cabo este

ajuste la red contrasta los resultados de sus predicciones con un conjunto de datos de entrada y salida definido por los datos de entrenamiento (la BD) en el entorno en análisis.

Al comienzo de la iteración, es decir, en el primer paso de entrenamiento, la red asigna valores aleatorios a los pesos sinápticos $\{\mathbf{w}\}_i$ y predice una respuesta para los valores de entrada $\{x_i\}_{i=0}^N$ (Ec. 1.7).

A continuación, los valores de los pesos sinápticos \mathbf{w} son optimizados mediante algún método de interpolación numérica que varía según el tipo de red y las particularidades del código computacional usado. Comúnmente se utiliza un método de minimización de la función de coste, es decir, del error que genera la red en su predicción. La señal de error e_j (del dato j -ésimo) es la diferencia entre la respuesta deseada y la respuesta actual de la red.

$$e_j = d_j - y_j \quad (1.8)$$

En este informe la función de coste que se minimiza para optimizar al conjunto de parámetros peso $\{\mathbf{w}\}_i$ es la raíz del error cuadrático medio σ entre los valores predichos y los valores de entrada:

$$\sigma = \sqrt{\sum_{j=1}^{N_{data}} \frac{e_j^2}{N_{data}}} \quad (1.9)$$

Capítulo 2

Superficies de Energía Potencial con Redes Neuronales Artificiales

Resumen

Las Superficies de Energía Potencial Neuronales (SEPN) tienen la particularidad de estar representadas por algoritmos de inteligencia artificial, más específicamente por algoritmos de aprendizaje automatizado, lo cual confiere a las computadoras la habilidad de descifrar la función potencial subyacente a los datos que, generalmente, se obtienen de cálculos de estructura electrónica. A diferencia de los potenciales atomísticos tradicionales, las formas funcionales de la SEPN se basan puramente en un procedimiento de ajuste matemático, por eso, al no estar basados en hipótesis físicas, es preciso estudiarlos en detalle.

2.1. Características estructurales

Las invariancias que presenta todo sistema químico del cual se desee representar su SEP son, por ejemplo:

- Invariancia por transformaciones rígidas de rotación y traslación del sistema atomístico
- Invariancia por intercambio de átomos de la misma especie

Configuraciones de átomos que difieran bajo estas transformaciones, deben arrojar los mismos valores numéricos de la energía potencial del sistema. Por esto es necesario que se efectúe un pre-procesamiento de los vectores de coordenadas atómicas, pues, en caso contrario, si la red recibe directamente las coordenadas atómicas como vector de entrada, ante un intercambio de

átomos de igual especie o ante una traslación rígida de la estructura, los valores numéricos de entrada en la red variarían y por ende, variaría el resultado de la energía final de la red.

Además determinar el valor de la SEPN por un procesamiento directo de las coordenadas atómicas no es conveniente para la creación de SEPN de Alta Dimensión (SEPN AD) ya que, por un lado, el número de valores de entrada de la red neuronal debe permanecer fijo, pero este no es necesariamente el caso al realizar simulaciones de DM como las que se implementan con SEPN AD. [14]

Entonces para satisfacer las simetrías pertinentes a los sistemas que describen las SEPN, es necesario efectuar un pre-procesamiento de las coordenadas atómicas antes de ser implementadas como valores de entradas en las redes.

Las **Funciones de Simetría (FS) o Descriptores**, son los entes matemáticos utilizados para el pre-procesamiento de las coordenadas atómicas sobre el sistema del cual se desea conocer la energía potencial atomística. Estos descriptores (del entorno químico del átomo) junto con la disgregación de la energía potencial en energías atómicas mediante el uso de **Redes Neuronales Atómicas (RNA)** son los elementos clave para el desarrollo de SEPNs AD, propuesto y desarrollado por J. Behler y colaboradores en Göttingen, Alemania. [13]

2.2. Redes Neuronales Atomísticas

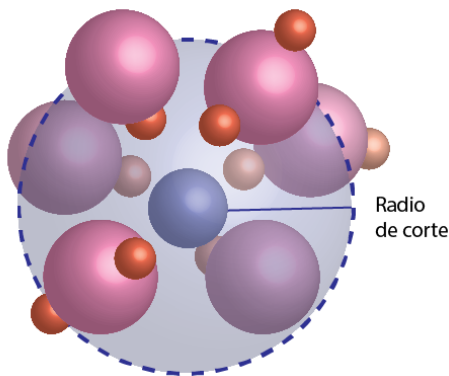


Figura 2.1 Esfera de corte.

En el esquema desarrollado por Behler-Parrinello [13], cada átomo interactúa con su entorno químico cercano y contribuye individualmente a la energía total del sistema de manera aditiva. Se implementa una red neuronal para calcular la energía de cada átomo del sistema. Lo que significa implementar un sistema de N redes neuronales atómicas (con $N =$ cantidad de átomos del sistema):

$$E_{conf} = \sum_{i=1}^N E_i(RNA_i) \quad (2.1)$$

Los átomos interactuantes que se consideran durante el cálculo de las energías atómicas individuales, son aquellos que quedan incluidos dentro de una *esfera de corte* de radio R_c centrada

en el átomo. El valor R_c se escoge de manera tal que todos los átomos que interactúan de manera significativa con el átomo central estén incluidos dentro de la esfera. [15]

Si el sistema se compone de elementos diferentes, la energía total se calcula como suma sobre todos los átomos de todos los elementos presentes. Aquellos elementos que sean al mismo tipo comparten la misma estructura de su RNA:

$$E_{conf} = \sum_{\nu=1}^N \sum_{i=1}^{N^\nu} E_i^\nu(RNA^\nu) \quad (2.2)$$

Donde N es el número de especies distintas en el sistema y N^ν el número de átomos de la especie ν .

2.3. Funciones de Simetría

Las distribuciones espaciales de los átomos dentro de la esfera, limitadas por el radio de corte R_c , se describen a partir de funciones de simetría (FS). Estas, son funciones de muchos cuerpos encargadas de describir el entorno químico con el cual interactúa el átomo central de la esfera de corte, mediante un conjunto de valores numéricos ordenados en un vector de alta dimensión que caracterizan al ambiente atómico donde se encuentra sumergido ese átomo. [14] A este vector se lo conoce como *vector descriptor* o *vector de simetría* y a sus componentes se las llama *huellas* o *fingerprints*.

2.3.1. Forma funcional

Función de dos cuerpos: *La función radial*

La función radial describe las distancias radiales en el entorno cercano de cada átomo, [14] por ejemplo, en un átomo i -ésimo con los átomos vecinos de cada especie en el sistema, como ser, la especie ν :

$$D_i^2 = \sum_{j=1}^{N^\nu} e^{\eta(R_{ij}-R_s)^2} \cdot f_c(R_{ij}) \quad (2.3)$$

En la Ec. 2.3 R_{ij} representa la distancia entre el átomo central i -ésimo y su vecino j -ésimo de especie ν , η y R_s son parámetros que deben ser elegidos convenientemente (ver Sec. 2.3.2) y f_c es la función de corte:

$$f_c(R_{ij}) = \begin{cases} 0,5 \cdot \left[\cos\left(\frac{\pi R_{ij}}{R_c}\right) + 1 \right] & \text{for } R_{ij} \leq R_c \\ 0 & \text{for } R_{ij} > R_c \end{cases} \quad (2.4)$$

donde R_c es el radio de la esfera de corte.

Claramente, la función D_i^2 depende del átomo i -ésimo sobre el cual esta centrada, la especie ν que esta describiendo, la cantidad de átomos j -ésimos incluidos en la esfera de corte (N_ν), la distancia R_{ij} a la q se hallan los átomos vecinos, y por último, de los parámetros η y R_s que caracterizan la forma espacial del descriptor D^2 .

La función de corte decrece con el incremento de la distancia entre el átomo central i -ésimo y su vecino j -ésimo (R_{ij}), hasta anularse para valores mayores al radio de corte R_c . Mediante esta función se representa el hecho de que la descripción de los átomos considerados durante el calculo de las energías potenciales, se corresponde con átomos dentro del entorno químico cercano. Esta función es continua y suave durante todo su dominio, lo que permite que sea derivable (las derivadas son necesarias para calcular, luego, las fuerzas a partir de la SEPN). Además, se anula para valores iguales o mayores que R_c , tanto la función como su derivada.

Función de tres cuerpos: *La función angular*

En las funciones de tres cuerpos se describe el entorno atómico a partir del ángulo que forma el átomo central i -ésimo, en conjunto con dos átomos vecinos, j -ésimo de la especie ν y k -ésimo de la especie σ , de la siguiente manera:

$$D_i^3 = 2^{1-\zeta} \sum_{j \neq i}^{N_\nu} \sum_{k \neq i, j}^{N_\sigma} \left[(1 + \lambda \cdot \cos \theta_{ijk})^\zeta \cdot e^{-\eta(R_{ij}^2 + R_{ik}^2)} \cdot f_c(R_{ij}) \cdot f_c(R_{ik}) \right] \quad (2.5)$$

Con esta forma funcional, se asegura que los valores del descriptor tiendan suave y continuamente a cero, cuando alguno de los pares de distancias R_{ij} o R_{ik} supera el radio de corte R_c . El hecho de permitir ternas de átomos con distancias $R_{jk} \geq R_c$, permite a la función abarcar mayor cantidad de ángulos dentro de la esfera de corte, y por ende, describir en más detalle el entorno atómico (angular) en el cual se encuentra el átomo- i .

Al igual que en la función radial, la función de tres cuerpos depende de varios factores, tanto del átomo central i ; de las distancias R_{ij} y R_{ik} a los vecinos j y k ; de la cantidad de especies ν y σ incluidas en la esfera de corte (N_ν y N_σ); y del conjunto de parámetros que caracterizan la función ($\{\zeta, \lambda, \eta\}$)

2.3.2. Implementación: Los vectores descriptores o *huellas*

Debido a la complejidad que supone describir la construcción de los vectores descriptores de forma general, se ejemplifica la metodología implementada en la construcción de los vectores de simetría en función del sistema que compete al trabajo realizado en esta tesina: una molécula de C-O interactuando con una superficie metálica compuesta por 30 átomos de Cu(110).

El sistema CO/Cu(110) esta compuesto de 3 especies: C, O y Cu. De estas tres especies, como en la celda unidad hay un solo átomo de C y un solo átomo de O, los pares de átomos y ternas de átomos posibles son los siguientes:

Pares	Ternas
Cu-C	Cu-CuC
Cu-O	Cu-CuO
Cu-Cu	Cu-CuCu
C-O	Cu-CO
- -	C-CuO
- -	C-CuCu
- -	O-CuC
- -	O-CuCu

Cuadro 2.1 Pares y ternas de átomos diferentes posibles a una distancia menor al $R_c = 7\text{\AA}$.

Una vez identificados los pares y ternas de átomos que deben ser descriptos, se seleccionan los conjuntos de parámetros utilizados por cada *fingerprint* y para cada especie. Rememorando las Eqs. 2.3 y 2.5, los parámetros a seleccionar para cada *fingerprint* son los conjuntos $\{\eta, R_s\}$ y $\{\eta, \lambda, \xi\}$ respectivamente.

Para el caso del descriptor D^2 , se estudiaron dos casos: Se seleccionaron 4 valores de R_s y se mantuvo fijo el valor de η y, en el segundo caso, se seleccionaron 4 η distintos y se mantuvo fijo R_s . Para el caso del descriptor D^3 , se consideraron 2 valores de λ y 2 de ξ manteniendo fijo el parámetro η . Esta información se sintetiza en el siguiente cuadro:

	Descriptor	Variación	Cantidad	Fijo
Caso 1	D^2	R_s	(4)	η
	D^3	λ, ξ	(2,2)	η
Caso 2	D^2	η	(4)	R_s
	D^3	λ, ξ	(2,2)	η

Cuadro 2.2 Descripción de los dos casos cómo se varían los parámetros de las FS.

Las variaciones indicadas en el Cuad. 2.2 se mantuvieron para todos los pares y para todas las ternas de átomos presentadas en el Cuad. 2.1. No obstante, la dimensión del vector descriptor para un átomo de Cu no es la misma que la de los vectores descriptores de las especies C u O.

Por ejemplo, existen 3 pares diferentes que involucran a un átomo de Cu y para cada uno de esos pares tenemos 4 variaciones de R_s con η fijo (caso 1) o 4 variaciones de η con R_s fijo (caso 2). Entonces, para un átomo de cobre en el caso 1 o en el caso 2, existen 12 componentes del vector descriptor asociados a funciones de simetría de dos cuerpos (3 pares por 4 variaciones). Para el descriptor D^3 , existen 4 ternas de átomos que tienen como primer átomo a uno de Cu y para cada una ellas existen 4 combinaciones distintas de los parámetros $\{\eta, \lambda, \xi\}$ (2 variaciones de λ , 2 variaciones de ξ para un solo valor de η). Entonces para un átomo de cobre existen 16 componentes del vector descriptor asociados a funciones de simetría de tres cuerpos (4 ternas por 4 variaciones). Sumando las componentes de ambos descriptores (12 del D^2 y 16 del D^3) resulta que el vector descriptor \mathbf{D}_{Cu} tiene 28 componentes:

$$\mathbf{D}_{Cu} = \underbrace{(D_1, \dots, D_5, \dots, D_9, \dots)}_{CuC} \underbrace{(D_5, \dots, D_9, \dots)}_{CuO} \underbrace{(D_9, \dots)}_{CuCu} \underbrace{(D_{13}, \dots, D_{17}, \dots, D_{21}, \dots, D_{25}, \dots)}_{Cu-CuC} \underbrace{(D_{17}, \dots, D_{21}, \dots, D_{25}, \dots)}_{Cu-CO} \underbrace{(D_{21}, \dots, D_{25}, \dots)}_{Cu-CuCu} \underbrace{(D_{25}, \dots)}_{Cu-CO} \quad (2.6)$$

Para el caso de C y O, se construyen los vectores \mathbf{D}_C y \mathbf{D}_O siguiendo la misma lógica, pero teniendo en cuenta que los tipos de distancias diferentes entre pares de átomos para ambas especies son solo 2 para cada especie (lo que significa que las componentes del descriptor D^2 serán 8) y las ternas de ángulos diferentes que se pueden formar son, también para ambas especies, sólo 2 (lo que origina que las componentes del descriptor D^3 sean solo 8). De esta forma, tanto el vector descriptor de C como el vector descriptor de O, tienen 16 componentes, de la forma:

$$\mathbf{D}_C = \underbrace{(D_1, \dots, D_5, \dots)}_{CCu} \underbrace{(D_5, \dots, D_9, \dots)}_{CO} \underbrace{(D_9, \dots, D_{13}, \dots)}_{C-CuO} \underbrace{(D_{13}, \dots)}_{C-CuCu} \quad (2.7)$$

$$\mathbf{D}_O = \underbrace{(D_1, \dots, D_5, \dots)}_{OCu} \underbrace{(D_5, \dots, D_9, \dots)}_{OC} \underbrace{(D_9, \dots, D_{13}, \dots)}_{O-CuC} \underbrace{(D_{13}, \dots)}_{O-CuCu} \quad (2.8)$$

En las ecuaciones de 2.6 a 2.8, se ha omitido la componente D_0 , que para los átomos de todas las especies vale $D_0 = 1$ pues corresponde al sesgo de las redes neuronales, de acuerdo a lo discutido al finalizar la Sec. 1.1.1.

A partir de estos ejemplos, se observa como las cantidades de conjuntos de parámetros $\{\eta, R^s\}$ y $\{\eta, \lambda, \xi\}$, junto con la cantidad de especies ν distintas del sistema y sus posibles combinaciones de pares y ternas de átomos determinan la dimensión de cada vector descriptor.

Construidos de esta forma, los vectores descriptores contienen la información numérica del entorno químico que experimenta cada átomo en una configuración dada del sistema. Los valores numéricos de las componentes de estos vectores dependen de la cantidad de átomos dentro de la esfera de corte, pero la dimensión de cada vector, no. Una vez determinados los vectores descriptores, su dimensión no cambia.

Es importante señalar el hecho de que cada uno de los átomos de la configuración presenta su propio vector de simetría \mathbf{D}'_i , donde la letra griega ν hace alusión a la especie atómica e i al número de átomo de dicha especie.

La metodología aplicada consiste en seleccionar el conjunto de parámetros $(\{\eta, R_s\}, \{\eta, \lambda, \xi\})$ que otorgue a las FS (es decir, a los vectores descriptores) la mayor sensibilidad en cuanto a la distribución de distancias y ángulos presentes en el conjunto ejemplo de entrenamiento de las redes. Es por esta razón que un estudio previo de los datos de entrenamiento resulta imperante, con el objetivo de orientar la elección del conjunto de parámetros espaciales adecuados que se deben implementar para las distribuciones de pares y ternas de átomos presentes.

Para la posterior evaluación de la energía potencial mediante la SEPN, es necesario contar con los parámetros de las funciones de simetrías correspondiente a los diferentes vectores descriptores tanto como los pesos sinápticos optimizados durante el entrenamiento de las redes neuronales.

2.3.3. Elección de los parámetros de las FS

Para el estudio preliminar de la base de datos y la adecuada elección de los parámetros de las FS, es posible seguir algunas recomendaciones solicitadas en la bibliografía [14]; puesto que, la selección de descriptores aún se implementa mayormente de un modo empírico. Las recomendaciones son

- Las funciones de simetría deben cubrir el espacio de configuraciones de un modo imparcial. La extensión espacial de la función con menor zona efectiva debe ser seleccionada en base a la distancia interatómica más corta del conjunto de datos de referencia.
- Deben realizarse estudios sobre el rango de valores que sondean las funciones de simetría tanto como el rango determinado por la normalización de estos valores (es decir todos los

valores divididos la desviación estándar de los resultados de la función de simetría). Si las diferencias entre el máximo y el mínimo valor que adopta un descriptor es muy pequeña, la función de simetría no estaría diferenciando hábilmente las diferentes estructuras, además, la adecuación de la red se tornaría numéricamente inestable ya que trataría de asignar energías estructurales diferentes a valores de la función de simetría similares. Durante el estudio de los valores normalizados, se asegura que no exista un valor dentro del rango de valores de la función de simetría que este siendo excesivamente descripto por esta.

- Entornos atómicos diferentes deben arrojar vectores de funciones de simetría diferentes. Bajo este contexto es necesario analizar sistemáticamente el conjunto de datos de entrada en contraste con las descripciones de las FS, para identificar si existen pares de estructuras diferentes cuyos vectores de simetría sean similares. En ese caso, se debe incorporar nuevas funciones de simetría que diferencien estas estructuras.
- El conjunto de funciones de simetría debe persistir siendo lo mas pequeño posible, en orden de incrementar la eficiencia de los cálculos computacionales de la red.

2.4. Superficies de Energía Potencial Neuronal

2.4.1. Arquitectura de las redes

Una vez establecidos los descriptores, se procede con la construcción del sistema de redes neuronales para la confección de la SEPN. Para mejor entendimiento de este esquema (desarrollado por el grupo de investigación en Göttingen [13], [14], [16]) se ilustra un diagrama de la arquitectura de redes neuronales en la Fig. 2.2, que representa la obtención de la energía potencial para una configuración genérica k -ésima.

En la Fig. 2.2 se sintetizan mediante filas la información concerniente a un átomo de la configuración (el total de filas se corresponde con el total de átomos presentes en la configuración) y, mediante columnas, las instancias que en conjunto determinan la energía potencial para dicha configuración.

Todas las entradas de la primer columna (triángulos negros) representan a los vectores posición de cada átomo en una configuración dada. Estos vectores se implementan conjuntamente como entradas de la segunda columna, en donde se calculan las componentes de los vectores

Arquitectura de redes para la construcción de una SEPN

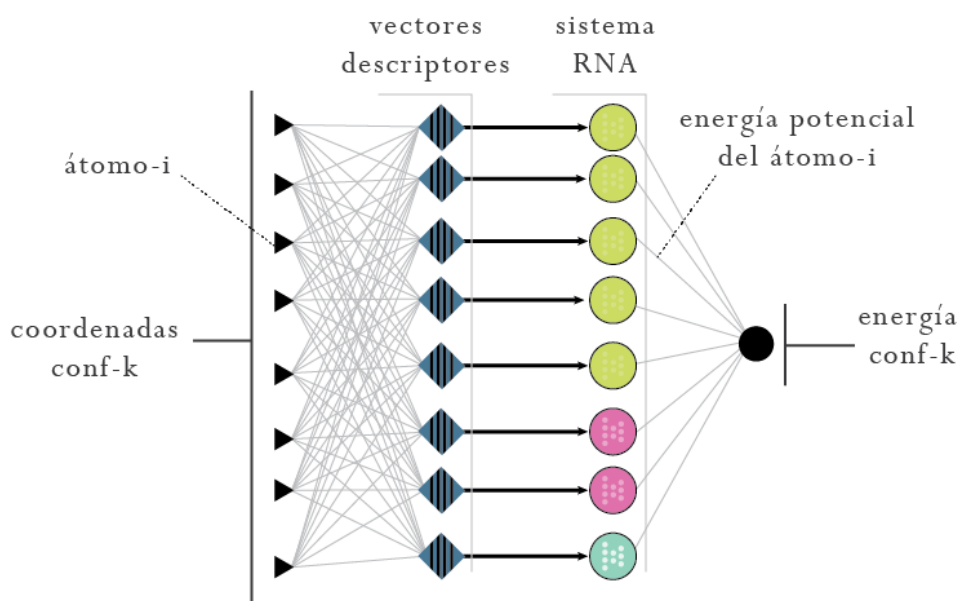


Figura 2.2 La primera columna de triángulos negros de la izquierda representa los vectores coordenados que indican la posición de cada átomo en una configuración dada. Las componentes de los vectores posición se utilizan como entradas de la columna de rombos azules, que representa a los vectores que sirven de descriptores del entorno químico de cada átomo. Las componentes numéricas de cada vector, se utilizan en las correspondientes RNA (columna de círculos verdes) que procesan la información y calculan la energía individual que cada átomo aporta al sistema. El color de los círculos representa las RNA de cada especie diferente. Finalmente la energía de la configuración se calcula sumando todas las energías individuales de los átomos, es decir, sumando todas las salidas de cada RNA (esquema propuesto por Behler-Perrinello).

descriptores con el método detallado en la sección anterior para cada átomo (Eqs. 2.6, 2.7 y 2.8).

Una vez computados los vectores descriptores del entorno químico de cada átomo, estos valores se implementan como entradas hacia la RNA de la especie atómica correspondiente; es decir, el vector descriptor dado para una fila se utiliza como vector de entradas hacia la RNA de esa misma fila. En este punto es donde se identifica que la SEPN se obtiene mediante un “sistema de redes neuronales atómicas”, en vez de ser calculada mediante *una única red*; ya que cada átomo de la configuración (o sea, cada fila del esquema) utiliza una RNA, cuya arquitectura cambia conforme se trate de una especie diferente, constituyendo un sistema de RNA como constructoras de la SEPN.

Las RNAs son las encargadas de determinar la contribución de cada átomo a la energía potencial del sistema mediante la Ec. 1.7. Por ejemplo, la energía individual de un átomo

cualquiera del sistema, cuya RNA presenta la arquitectura del ejemplo citado en la Ec. 1.7, y donde sus entradas corresponden a las componentes del vector descriptor \mathbf{D} (Eq. 2.6 o 2.7 o 2.8), queda determinada por

$$E = \mathfrak{J} \left(\sum_{l=0}^{N_C} w_l^F \cdot \psi_l \left(\sum_{k=0}^{N_B} w_{lk}^{CB} \cdot \psi_k \left(\sum_{j=0}^{N_A} w_{kj}^{BA} \cdot \psi_j \left(\sum_{i=0}^N w_{ji}^I \cdot D_i \right) \right) \right) \right) \quad (2.9)$$

donde los subíndices que identifican al átomo en cuestión se han omitido para facilitar la lectura.

Cabe enfatizar nuevamente que en el sistema implementado para las SEPN se optimizan los conjuntos $\{\{\mathbf{w}^I\}, \{\mathbf{w}^{BA}\}, \{\mathbf{w}^{CB}\}, \{\mathbf{w}^F\}\}_\nu$ con $\nu = 1, \dots, N$ (donde N es la cantidad de especies diferentes) de parámetros pesos.

Finalmente, la energía total de la configuración k -ésima resulta de sumar las distintas energías atómicas individuales (Eq. 2.9) de lo que se obtiene, para la última columna del diagrama de la Fig. 2.2.

$$E_{conf-k} = \sum_{\nu=1}^N \sum_{i=1}^{N^\nu} E_i^\nu(\mathbf{D}_i^\nu) \quad (2.10)$$

En la simulación de SEP por redes neuronales, el hecho de que la entradas de las RNAs sean las componentes de cada descriptor es fundamental. De esta forma, la incorporación (o eliminación) de átomos en el entorno químico delimitado por la esfera de corte, repercute únicamente en el valor numérico que adoptan las componentes D_i de los vectores descriptores de cada átomo, pero no modifica la dimensión determinada de cada vector. Por lo que la RNA recibe siempre la misma cantidad de entradas, indistintamente de los átomos que fueron considerados durante la simulación. Esto último, además, se traduce en el hecho de que el método es expansible a sistemas atómicos de mayor o menor tamaño que aquel para el cual la SEPN fue entrenada.

Si bien queda expuesto que la SEPN se construye mediante un *sistema de redes neuronales*, a lo largo de este informe los términos *sistema de redes* y *la red* se utilizan de forma indistinta para referirse a la SEPN.

2.4.2. Forma funcional de la raíz del error cuadrático medio

Al comienzo del proceso, es decir en el *paso-0* del entrenamiento iterativo, se asignan valores aleatorios a todos los conjuntos de parámetros peso de las RNAs que componen una SEPN (es decir, a los $\{\mathbf{w}\}$). Para todas las configuraciones se calcula la energía potencial E_k^{SEPN} y a fin

de evaluar la calidad del ajuste hecho por la red, se compara con el valor de la energía E_k^{DFT} correspondiente a la misma configuración, mediante:

$$\sigma_{paso-i} = \sqrt{\sum_{k=1}^{N_{conf}} \frac{(E_k^{SEPN} - E_k^{DFT})^2}{N_{conf}}} \quad (2.11)$$

Específicamente, cada paso en las iteraciones sucesivas que componen al entrenamiento de una SEPN, arroja un nuevo valor para σ .

Se implementa el método de optimización por propagación hacia atrás de los parámetros $\{\mathbf{w}\}$, a partir de minimizar σ_{paso-i} . De esta forma, se proponen nuevos conjuntos de parámetros pesos de las RNAs y se calcula $\sigma_{paso-(i+1)}$. Este proceso se realiza sucesivamente durante una cierta cantidad de pasos hasta obtener el conjunto de parámetros pesos optimizado y, por ende, la SEPN entrenada.

2.4.3. Sobreajuste

Las distintas configuraciones de la BD que se utilizan para el entrenamiento de las redes neuronales, repercuten directamente en la precisión que puede alcanzar la SEPN. Aquellos puntos pertenecientes a la BD, son las áreas del espacio multi-dimensional de configuraciones del sistema cuya energía potencial la red aprende con mejor precisión, dado que los conjuntos de parámetros peso se ajustan respecto a estos mismos puntos. (Ver Fig. 2.3). Sin embargo, las SEPN deben generalizar adecuadamente sus predicciones hacia zonas inexploradas por la BD ya que durante la aplicación de las SEPN en DM no existen restricciones sobre los puntos del espacio de configuración visitados durante la simulación.

Si todos los puntos que representan las configuraciones existentes en la BD, son implementados para interpolar la SEPN se corre el riesgo de caer en errores por sobre ajuste de los datos; un error en donde los datos de referencia son ajustados correctamente pero la predicción sobre resultados desconocidos (la habilidad de generalización de la red), no es razonable.

Para que el método de convergencia sea mas eficiente, las metodologías consultadas en la bibliografía [13] sugieren reservar un subconjunto de datos de la BD (escogido de forma aleatoria) el cual se denomina subconjunto para evaluación o *test* (por su nombre en inglés). Este conjunto de configuraciones de *test* no es tenido en cuenta durante el proceso de optimización de los parámetros de la red. Entonces, la red optimiza sus parámetros con un conjunto reducido de las configuraciones de la BD (denominado subconjunto para entrenamiento o *train*) y

Esquema de interpolación de la SEPN sobre la BD

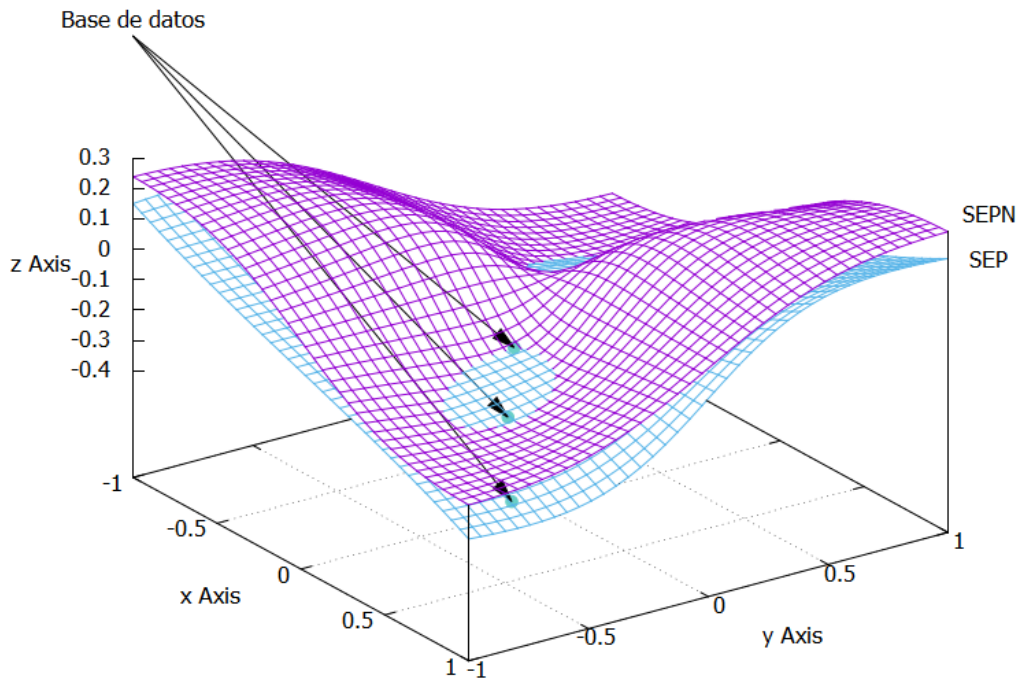


Figura 2.3 En el gráfico se observa una superficie interpoladora que representa a la SEPN (superficie violeta) ajustando la superficie SEP (superficie celeste) subyacente a la base de datos. A priori la SEPN predice valores más próximos a los datos de entrada para configuraciones cercanas a las que están en la BD usada en el entrenamiento. Las superficies de la figura dependen de 2 variables y se las ilustra solamente a modo ilustrativo. Las SEP y las SEPN son superficies multi-dimensionales que dependen de $3N$ variables con N siendo el número de átomos considerado en la celda de simulación.

contrasta, luego, el alcance de su habilidad generalizadora prediciendo el resultado de las configuraciones reservadas en el subconjunto *test* (cuya energía potencial es desconocida por la red hasta entonces). Ambos procesos, el de entrenamiento y el de evaluación, son efectuados automáticamente durante la instancia de entrenamiento de la SEPN y, al final del entrenamiento, se reportan tanto la raíz del error cuadrático medio (σ) como el error absoluto medio (MAE) de ambos subconjuntos, *test* y *train*. De esta forma, es posible detectar los casos en que la flexibilidad de la red comienza a ser exagerada y se vuelve contraproducente. Como en cualquier proceso de ajuste, la clave es que la red sea suficientemente flexible para ser capaz de describir adecuadamente los datos de entrada, pero no exageradamente flexible como para predecir con error significativo datos intermedios producto, por ejemplo, de oscilaciones espurias.

Capítulo 3

Construcción de la Superficie de Energía Potencial de C-O/Cu(110)

Resumen

La metodología que se implementa para la construcción de la SEPN concerniente al sistema C-O/Cu(110), consiste en realizar iteraciones sistemáticas sobre un conjunto de redes neuronales, seleccionadas a partir de sus σ_{test} de entre un conjunto de redes con diferentes parámetros de FS. Estas iteraciones están compuestas por una instancias de entrenamiento y distintas instancias de análisis de algunos observables físicos, característicos del sistema. El objetivo es optimizar las SEPN hasta converger a una representación coherente de la SEP buscada.

A partir de los estudios y de los resultados obtenidos en las instancias de análisis, se determina cuál ha sido la representación más óptima del potencial y se procede a implementar el potencial y estudiar las propiedades físicas del sistema atómico descripto.

3.1. Instancia de entrenamiento

3.1.1. La base de datos

La base de datos de referencia (BD) es el paso inicial de los entrenamientos de redes neuronales con entrenamiento supervisado. Para la construcción de potenciales de energía atomísticos, la BD es un conjunto ejemplo, muy amplio, de configuraciones que indican la energía potencial del sistema en función de sus posiciones atómicas. Las configuraciones representan las diferentes situaciones en que se puede encontrar el sistema atómico y, de esta manera, cómo varía la energía potencial del mismo.

En el caso de estudio de este informe, se trata de un sistema molécula superficie donde se muestrean situaciones en las que la molécula se encuentra lejos y cerca de la superficie, con su vector internuclear a su distancia de equilibrio, a distancias mayores y menores, y con diferentes orientaciones respecto de la superficie. Así como también se incluyen configuraciones correspondientes a la molécula oscilando en cada uno de sus modos de vibración y ubicada con su CM lejos de la superficie. La BD provee la información a través de la cual se optimizan los valores de los pesos sinápticos de la red.

Ampliación sistemática de la BD

Ciclo de construcción de la SEPN para CO/Cu(110)



Figura 3.1 El ciclo implementado para la construcción de la SEPN consiste en dos grandes instancias, una de entrenamiento y otra de análisis de resultados. Durante la instancia de análisis se determina la calidad de las SEPNs que están siendo entrenadas mediante analizar principalmente el observable físico de convergencia: la PBA. Este ciclo se repite hasta que las propiedades físicas de la SEP para CO/Cu(110) son simuladas eficientemente.

Debido a que las configuraciones descritas en la BD representan un pilar importante en la precisión que pueden alcanzar la SEPN, se implementa como metodología partir desde un conjunto de configuraciones inicial (4290 configuraciones), seleccionado *a priori* de manera intuitiva, pero proyectando la incorporación sistemática de configuraciones particulares conforme se realizan entrenamientos sucesivos de un determinado grupo de redes neuronales. El objetivo de esta estrategia, es adicionar información al conjunto de datos ejemplo particularmente en

zonas que presenten baja población de muestreo y en donde, además, las energías potenciales predichas por las redes presenten mayor discrepancia entre sí (Fig. 3.1).

Los extensos conjuntos de configuraciones que conforman la BD de este trabajo, son calculados mediante el método de estructura electrónica conocido como Teoría del Funcional Densidad [17], [18] (DFT, por sus siglas en inglés). Estos cálculos fueron obtenidos gracias al trabajo de Giulia Seminara, como parte de su trabajo de tesis doctoral bajo la dirección de la Dra. Alejandra Martínez. Las características técnicas de estos cálculos se detallan a continuación:

Características Técnicas

Las configuraciones que constituyen la BD, se componen por las coordenadas de cada átomo de una superficie de cobre Cu(110) (con 5 capas de 6 átomos cada una) y, de los átomos de una molécula libre de monóxido de carbono C-O localizada por encima de la superficie. Inicialmente se cuenta con 4290 configuraciones en la BD. La información allí presente se detalla, para cada configuración, en el siguiente orden

- la energía potencial del sistema (eV)
- la posición (x, y, z) de los 30 átomos de Cu
- la posición (x, y, z) del átomo de C
- la posición (x, y, z) del átomo de O

La estructura atómica se delimita por una celda prismática de base rectangular de dimensiones $[\mathbf{a}_1; \mathbf{a}_2; \mathbf{a}_3]$ con

$$\begin{aligned}\mathbf{a}_1 &= (7,960; 0; 0) \text{ \AA} \\ \mathbf{a}_2 &= (0; 7,508; 0) \text{ \AA} \\ \mathbf{a}_3 &= (0; 0; 18,770) \text{ \AA}\end{aligned}$$

Para el cálculo de la DM, la estructura cristalina dentro de la celda, dada por la configuración sobre la cual se lleva a cabo el proceso dinámico, se reproduce a lo largo de las direcciones x e y ; simulando una superficie extendida infinitamente. Estas condiciones se conocen como condiciones periódicas de contorno (*PBC* por sus siglas en inglés) y se implementan durante toda la simulación.

Las condiciones de equilibrio del sistema, son aquellas donde la energía de interacción molécula-superficie es mínima. Estos estados pueden ser, cuando la molécula esta “muy lejos” de

la superficie, y por lo tanto no existe interacción significativa entre estas o cuando la molécula se encuentra cerca de la superficie pero atrapada por el pozo del potencial de interacción. Según la orientación relativa de la molécula de CO sobre los sitios particulares de la estructura cristalina de Cu(110), este pozo es más o menos profundo, por lo que la molécula tiende a quedar atrapada adherida a la superficie en mayor o menor medida con posibilidades mas altas o mas bajas de superar la barrera de potencial y alejarse de la superficie.

Teoría del funcional densidad: DFT

Los cálculos por DFT, arrojan un amplio número de configuraciones para el sistema determinado por la molécula de CO y la superficie de Cu(110). En donde la molécula se posiciona con diferentes orientaciones y estados vibracionales, y a distintas distancias respecto de la superficie de cobre.

Como el eje central es el estudio de la adsorción del monóxido de carbono por la superficie de cobre, las configuraciones que constituyen la BD contienen, en su amplia mayoría, ejemplos donde el monóxido de carbono se posiciona en regiones de interacción con la superficie con altas probabilidades de adsorción. Los procesos de adsorción de la molécula son de baja energía y su probabilidad es sensible a la forma del potencial de interacción molécula-superficie a largas distancias (entre 4 y 6 Å). En esa zona dominan las interacciones de Van der Waals que se describen en la base de datos, con la funcional usada en los cálculos DFT, razonablemente.

No obstante es necesario incorporar a la BD algunos casos excepcionales, en donde la molécula se ha alejado de la superficie o, incluso, casos en los que se ha disociado; a modo tal de informar a la red cuál debería ser la predicción deseada para tal evento.

Únicamente al comienzo del proceso, y para determinar el conjunto de redes neuronales que serán entrenadas iterativamente, se estudia la información vigente en la base de datos y se determinan los conjuntos de parámetros caracterizadores de las FS, que servirán como funciones descriptoras de los entornos químicos atómicos a las RNAs en las SEPNS.

3.1.2. Proceso de Entrenamiento

El proceso de entrenamiento consiste en optimizar, de forma iterativa, los conjuntos de pesos sinápticos $\{\mathbf{w}\}$ de cada RNA, mediante un método implementado por el algoritmo computacional de entrenamiento *aenet*. Básicamente, el método consiste en minimizar el error entre las energías predichas por la red y los valores ejemplo de la BD; a través de minimizar la raíz del error

cuadrático medio (Ec. 2.11). Al converger adecuadamente este error, se dice que la red ha sido entrenada y queda determinada según el conjunto optimizado de parámetros peso $\{\mathbf{w}\}$; que se almacenan como un archivo de salida del algoritmo computacional *aenet*. El objetivo fundamental del entrenamiento es que la red aprenda lo suficiente, para desarrollar la habilidad de generalización en sus predicciones.

En esta instancia, se realizaron entrenamientos sobre conjuntos de redes que diferían por la elección de sus parámetros en las FS, para estudiar sus repercusiones sobre la precisión de las SEPNS. También, se indaga el efecto que se genera sobre la convergencia de los errores (σ_{train} y σ_{test} particularmente) al incorporar configuraciones en la BD de entrenamiento. Por último, se estudiaron SEPNS generadas por redes con los mismos conjuntos de parámetros de las FS, para investigar el efecto que la determinación aleatoria de los conjuntos de evaluación (*test*) y de entrenamiento (*train*), y la inicialización aleatoria de los parámetros de las redes ($\{\mathbf{w}\}$), al comienzo de cada entrenamiento, podía generar en la precisión de las SEPNS.

3.1.3. Arquitectura de las RNAs

Las RNAs que se utilizan comparten la misma arquitectura para todas las especies. Estas redes se componen de 2 capas ocultas con 10 nodos cada una, y la función de activación $\psi(\star)$ que se implementa es $\tanh(\star)$. Un esquema de esta arquitectura se observa en la Fig. 3.2, donde se ejemplifica la RNA implementada para el caso un átomo de Cu.

Todas las configuraciones de la BD se conforman siempre por la misma cantidad de átomos. Es decir, como se mencionó anteriormente, las configuraciones están compuestas por 30 átomos de Cu, 1 átomo de C y 1 átomo de O. Por lo tanto, la energía de una configuración genérica k , mediante implementar la Eq. 2.10, resulta

$$E_{conf-k} = \sum_{i=1}^{30} E_i^{Cu}(\mathbf{D}_i^{Cu}) + E_1^C(\mathbf{D}_1^C) + E_1^O(\mathbf{D}_1^O) \quad (3.1)$$

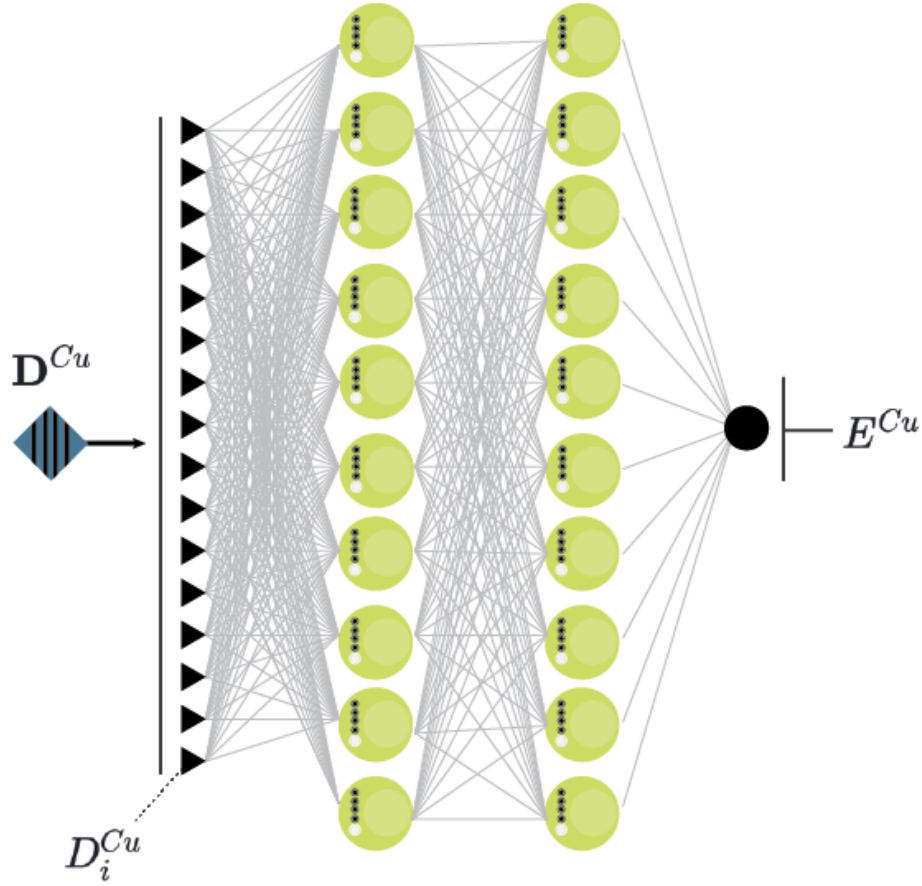


Figura 3.2 En la imagen se ejemplifica la arquitectura de la red RNA utilizada para la especie Cu. Se puede observar que las entradas de esta red son las componentes D_i^{Cu} del vector \mathbf{D}^{Cu} . Todas las RNAs implementadas para cada elemento comparten la misma estructura: 2 capas ocultas con 10 nodos cada una con función de activación $\tanh(\star)$

3.2. Instancia de análisis

3.2.1. Raíz del error cuadrático medio

Uno de los análisis utilizados para identificar la SEPN que mejor esta describiendo la BD de entrenamiento, es el estudio de la curva de la raíz del error cuadrático medio (σ_{paso-i} de la Ec. 2.11):

$$\sigma_{paso-i} = \sqrt{\sum_{k=1}^{N_{conf}} \frac{(E_k^{SEPN} - E_k^{DFT})^2}{N_{conf}}} \quad i = 1, N_{pasos} \quad (3.2)$$

Básicamente esta curva reporta valores promedio de la calidad con la que las energías de todas las configuraciones han sido predichas por determinado conjunto de parámetros pesos $\{\mathbf{w}\}_{\nu=1}^3$, de las ternas RNA_{Cu} , RNA_C y RNA_O , en cada paso de entrenamiento.

A partir de esta curva, es posible hallar aquel conjunto de $\{RNA\}$ cuya σ haya convergido al menor valor. En consecuencia, esto determina a las ‘mejores’ representaciones de la SEPN, es decir, las representaciones que han ajustado con mayor precisión al conjunto de datos de referencia inicial, y se escogen como potenciales candidatas para la representación final de la SEPN.

3.2.2. Rasgos característicos del sistema

Se conoce que la DM de adsorción de la molécula de C-O con la superficie de Cu, depende ampliamente de cómo es la interacción entre estos elementos para diferentes posiciones de la molécula relativas a la superficie. Por lo tanto, una manera de analizar la precisión de la SEPN entrenada, es manteniendo la distancia internuclear fija y la orientación de la molécula fija y, en función de la coordenada Z del CM de la molécula, analizar los valores predichos por la red respecto a los valores obtenidos utilizando DFT; cuando la molécula se sitúa en cuatro sitios particulares.

Orientaciones particulares de la molécula de C-O estudiadas

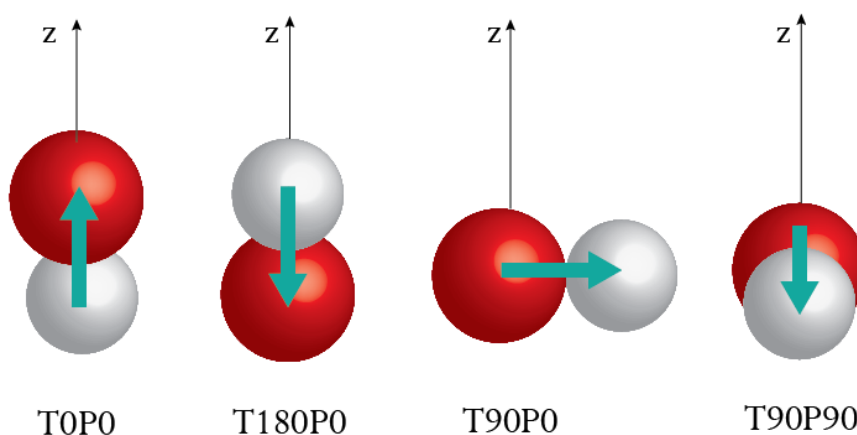


Figura 3.3 Se estudian cuatro orientaciones particulares de la molécula. De las cuales se conoce el potencial que debe describir la molécula cuando el C se encuentra en la posición vertical. En ese caso, la molécula se adsorbe por el pozo de potencial que encuentra al interactuar con la superficie de Cu. Cada esfera color rojo (blanco) representa un átomo de oxígeno (carbono).

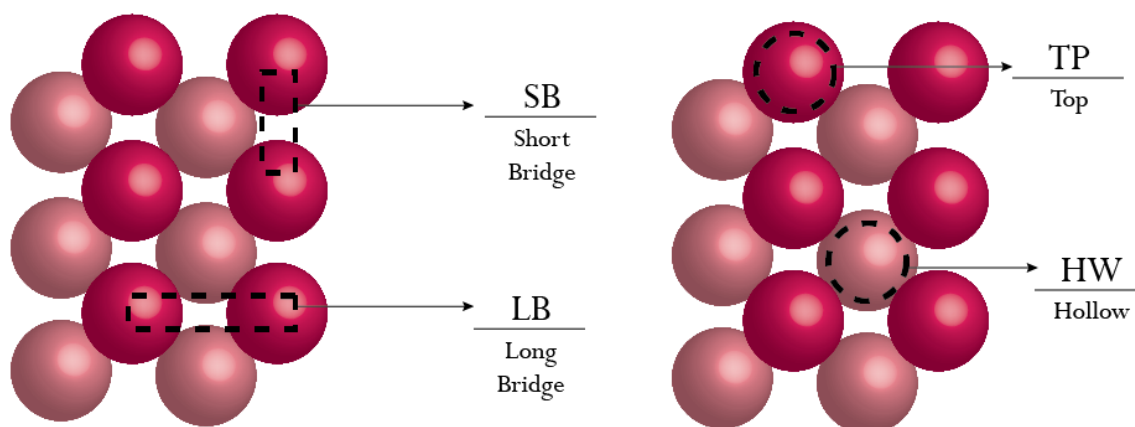


Figura 3.4 La figura representa una vista superior sobre la superficie de Cu. Los distintos sitios se determinan según las posiciones adyacentes de los átomos de Cu en la red cristalina. Como la estructura atómica que adopta la superficie es una red cristalina FCC se generan los sitios TP, HW, SB y LB. Las esferas rosa oscuro (claro) representan un átomo de Cu ubicado en la primera (segunda) capa atómica.

Las orientaciones que adopta la molécula se describen en la Fig. 3.3 y las etiquetas de cada orientación se asignan por los valores que adquiere el ángulo polar θ y el azimutal ϕ que forma el eje internuclear de la molécula (con dirección C-O) con un eje vertical Z perpendicular a la superficie y con el eje X, fijado en la dirección de primeros vecinos (dirección del SB) en la superficie, respectivamente. Los sitios, se detallan en la Fig 3.4. Sus nombres se asignan por sus siglas en inglés *hollow* (HW), *top* (TP), *short bridge* (SB) y *long bridge* (LB).

Con este análisis se verifica como las redes neuronales reproducen la dependencia de la energía con la distancia molécula-superficie, para diferentes orientaciones de la molécula. En particular, cuando la molécula esta ubicada en diferentes sitios (TP HW LB SB) y con orientación paralela y perpendicular a la superficie, con el O para abajo y con el O para arriba. Estos son rasgos característicos de la SEP de interacción de la molécula con la superficie, sobre los cuales es preciso que las redes cuenten con una buena descripción para luego hacer dinámicas moleculares del sistema.

3.2.3. Trayectorias de dinámica molecular

De las SEPN entrenadas más óptimas, se procede a generar distintas simulaciones de DM clásica utilizando un código desarrollado en el grupo de investigación: *GSRD-dyn*. Esta dinámica

simula la trayectoria que recorre la molécula de CO cuando es arrojada hacia la superficie metálica, con orientaciones determinadas al azar, para determinadas energías cinéticas iniciales.

Las trayectorias de DM se obtienen mediante integrar en el tiempo las ecuaciones de movimiento de las coordenadas cartesianas de cada uno de los átomos del sistema. Las fuerzas necesarias para integrar las ecuaciones de movimiento, se obtienen por derivación directa de la SEPN (debido a que se cuenta con su expresión funcional explícita: Eq. 3.1). De lo que resulta la fuerza implementada sobre el átomo i de la especie ν :

$$F_{i,\nu} = \nabla_{i,\nu}(E^{SEPN}) \quad (3.3)$$

Análisis de observable físico: Probabilidad de Adsorción

Otra de las propiedades analizadas durante la construcción y optimización de la SEPN, es evaluar la probabilidad estadística de adsorción de la molécula de CO cuando esta impacta sobre la superficie de Cu(110), en los rangos de energías entre 0.0 eV a ~ 2.0 eV.

La probabilidad de adsorción de la molécula, se calcula a partir de los resultados que se obtienen al realizar la integración de las trayectorias de DM durante un cierto periodo de tiempo (5ps). Ésta se obtiene a partir del cociente entre la cantidad de moléculas consideradas “adsorbidas” sobre la cantidad total de trayectorias integradas.

Durante la integración de la trayectoria, el algoritmo computacional reconoce los distintos procesos que pueden ocurrir según la relación que hay entre la posición del CM de la molécula, la dirección y sentido del vector momento lineal con la que esta se mueve (respecto de la superficie) y el vector internuclear de la molécula. La definición de estos procesos se detalla a continuación:

- Reflexión: Las moléculas se etiquetan como reflejadas cuando el CM de la molécula retorna a la altura inicial desde la cual fue lanzada (luego de impactar con la superficie) y además el momento lineal de su CM apunta hacia arriba, de forma que la proyección en el eje perpendicular a la superficie sea positiva (es decir el vector apunta hacia el vacío, alejándose de la superficie). Este estado de la molécula es particularmente preciso y de fácil caracterización: Z_{cm} igual al valor inicial y componente en z del momento lineal positiva (sobre un eje perpendicular que se aleja de la superficie).
- Adsorción disociativa: Otra de las posibilidades es que la molécula sea etiquetada como adsorción disociativa. En este evento la molécula se encuentra con su CM cercano a

la superficie y el radio internuclear se ha estirado una cierta proporción especificada al comienzo de la DM (en el caso de este informe se considera un valor cercano al doble del estado de equilibrio). Éste es también un estado preciso de la molécula,

Si en algún determinado momento, antes de los 5ps, el programa registra alguno de los procesos definidos en los puntos anteriores, la DM automáticamente se frena y etiqueta a la molécula bajo el estado de reflexión o adsorción disociativa. En caso de que se completen los 5ps, y no se hayan conseguido ninguno de los dos estados, entonces se define que la molécula ha sido “adsorbida molecularmente”. Es decir, la componente Z del CM de la molécula aún no ha alcanzado nuevamente su valor inicial y la distancia internuclear no es tal como para que la molécula haya perdido su identidad como molécula.

Sin embargo un proceso singular posible ocurre cuando la molécula se acerca a la superficie, cede su energía cinética al chocar con las moléculas de Cu, y permanece atrapada durante un tiempo considerable en las cercanías de la superficie (es decir queda absorbida molecularmente). Pero, pasado un tiempo $< 5ps$, consigue la energía cinética suficiente como para desorberse y comienza su proceso alejándose de la superficie. Si el tiempo de integración final de la DM se alcanza antes de que la molécula haya adoptado las condiciones necesarias para ser considerada “Reflejada”, su CM se halla a en un estado intermedio entre ser considerada “Adsorbida” y “Reflejada” pero no adopta ninguno de los dos, por lo que el programa la etiqueta como “Adsorbida”. Este evento es más difícil de precisar, pero como los átomos de la superficie son móviles, es totalmente posible que estos choquen con los átomos de la molécula y la molécula adquiera la suficiente energía cinética como para desorberse del pozo de potencial.

Efectos del incremento de configuraciones en la BD

En esta instancia, se analiza un proceso iterativo en donde se observa el efecto que presenta el hecho de agregar configuraciones a la BD de entrenamiento, sobre las predicciones que las redes neuronales hacen del potencial.

Para esto, las configuraciones que se incorporan sistemáticamente a la BD se seleccionan mediante simular una DM con una de las SEPN seleccionadas y se estudia cuanto difieren las restantes SEPN para las configuraciones visitadas durante dicha DM. Aquellas configuraciones donde se hayan reportado mayores discrepancias de energía entre las SEPNs seleccionadas, se agregan a la BD (en otros términos, se las considera como zonas del espacio cuya descripción es compleja y cuya energía debe ser calculada por DFT e incorporada a la BD de entrenamiento

de las SEPN). El hecho de que las redes no coincidan en su predicción de la energía potencial, implica que esa zona del espacio de configuraciones del sistema no posee la cantidad de muestreos óptimos en la BD que permitan su correcta interpretación, por lo que más muestras deben ser incorporados para que las redes cuenten con mayor información sobre esta región.

Con esta metodología, se intenta discernir la elección más apropiada para el conjunto de parámetros de las FS entre un conjunto de SEPN cuyos errores en las predicciones de energía durante el entrenamiento convergen a valores similares,

Parte III

Resultados

Capítulo 4

Selección de las redes para la Superficie de Energía Potencial Neuronal

En el margen de la construcción del potencial SEPN-CO/Cu(110), este capítulo explora las características particulares del desarrollo del potencial neuronal y las propiedades físicas relevantes del sistema en estudio descritas por el mismo, con el objetivo de que la representación de la energía de interacción del sistema CO/Cu(110) mediante la SEPN sea lo más precisa posible.

Para eso, en primer instancia, se realiza un estudio sobre los efectos de variar los conjuntos de parámetros que definen las FS de los vectores descriptores (detallados en la Sec. 2.3.2), con el fin de determinar el conjunto de parámetros más óptimo; es decir, aquel conjunto de parámetros que minimice los errores de entrenamiento y de evaluación: σ_{train} y σ_{test} (ver Sec. 3.2.1, en particular la Eq. 3.2).

4.1. Muestreo de la base de datos inicial

Previo al proceso de entrenamiento de las redes neuronales, se realiza un muestreo de las distancias y de los ángulos entre pares y ternas de átomos que había en las 4290 configuraciones de la base de datos inicialmente. Estudiar las distribuciones presentes en la BD, ayuda a discernir los posibles conjuntos de parámetros característicos de las FS, que minimizarán los errores en las predicciones de las SEPN sobre dicha BD, pero no determina exactamente cuál es el criterio a adoptar para su adecuada elección.

Con el fin de discernir los posibles conjuntos óptimos de parámetros de las FS, se confeccionó una celda expandida (conformada por cuatro subceldas equivalentes) replicando las posiciones de los átomos de la celda original en cada subcelda, teniendo en cuenta condiciones periódicas de contorno (PBC por sus siglas en inglés). Sobre la celda expandida se realizaron los estudios de

las distribuciones de distancias interatómicas y ternas de ángulos distintos que se encontraban dentro de la esfera delimitada por el $R_c = 7 \text{ \AA}$.

En la Fig. 4.1, se exhiben los resultados de algunos histogramas obtenidos, los cuales muestran la distribución de distancias del C-O y del Cu-C y de ángulos para el Cu-CO, presentes entre las 4290 configuraciones de la BD. En el caso del par C-O, la mayor muestra de distancias oscilan alrededor de $\sim 1.5 \text{ \AA}$ y $\sim 6 \text{ \AA}$ pues la amplitud de aparición es más grande alrededor de estos valores. En la figura Cu-C la distribución de distancias es más uniforme. Ambas distancias no presentan ningún caso en los que la distancia sea menor a $\sim 0.5 \text{ \AA}$ debido a la repulsión atómica característica de las interacciones atómicas. En la figura Cu-CO, los histogramas presentan mayoría de casos en los ángulos $\sim 0^\circ$ y también entre 60° y 80° pues en estos rangos de ángulos los histogramas presentan frecuencias de aparición más elevadas.

4.1.1. Análisis de la base de datos inicial

Los resultados más llamativos de estas distribuciones, fueron conferidos a los histogramas de las distancias interatómicas del par C-O.

En el histograma puede observarse una gran cantidad de casos donde las distancias internucleares de la molécula C-O oscilan alrededor de su distancia de equilibrio en el vacío ($\sim 1.14 \text{ \AA}$). Este resultado es razonable, ya que las configuraciones de la BD-0 se componen de gran cantidad de configuraciones ejemplo donde el vector internuclear de la molécula oscila levemente en las proximidades de su distancia de equilibrio.

Por otro lado, entre 6 y 7 \AA , se observan gran cantidad de casos para los pares de distancias C-O. Si bien la BD no fue intesamente muestreada en estas zonas del espacio de configuraciones, es decir, con distancias C-O de $\sim 6 \text{ \AA}$, se concluye que estos resultados son consecuencia de la implementación de las *PBC* durante la confección de los histogramas (ver Fig. 4.2). Es decir, como la celda utilizada en la confección de la BD presenta lados con dimensiones $\sim 7,5 \text{ \AA}$ y $\sim 7,9 \text{ \AA}$, para ciertas orientaciones de la molécula es razonable pensar que la distancia entre el C de una molécula y el O de su imagen periódica puede llegar a encontrarse por debajo del radio de corte y, a efectos de esto, ampliar la cantidad de casos donde la distancia C-O se hallase entre 6 y 7 \AA . Este resultado también puede observarse en la figura referida a los ángulos Cu-CO (es decir, vértice en Cu y extremos en C y O) donde también se observan una amplia cantidad de casos con ángulos entre 60° y 120° . Estos histogramas sustentan la hipótesis antes mencionada, ya que la amplia cantidad de ángulos que están siendo tenidos en cuenta pueden corresponderse

Histogramas de distancias y ángulos

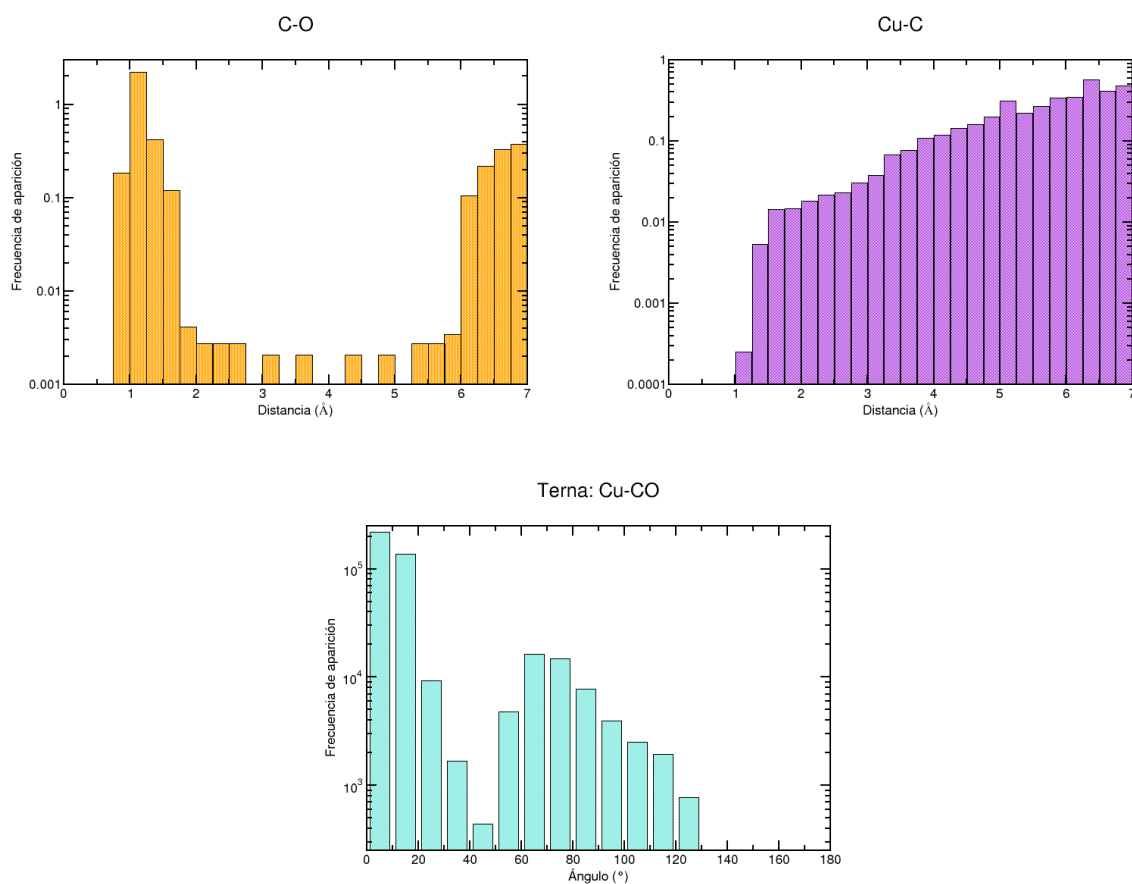


Figura 4.1 En las figuras superiores, se observan dos distribuciones de distancias (C-O y Cu-C) y en la figura inferior se observa una distribución de ángulos (de la terna Cu en el vértice con C-O en los extremos). Los histogramas indican la muestra de distancias C-O y Cu-C, y el ángulo Cu-CO presentes entre las 4290 configuraciones de la BD-0.

también con los ángulos construidos por la imagen periódica de algunos átomos (para mayor ilustración de esta situación referirse a la Fig. 4.2).

En conclusión, el resultado de aplicar condiciones periódicas de contorno genera que se contabilicen más cantidad de casos dentro de las limitaciones del radio de corte, en donde las distancias y/o ángulos superan a las cantidades verdaderamente presentes en la BD.

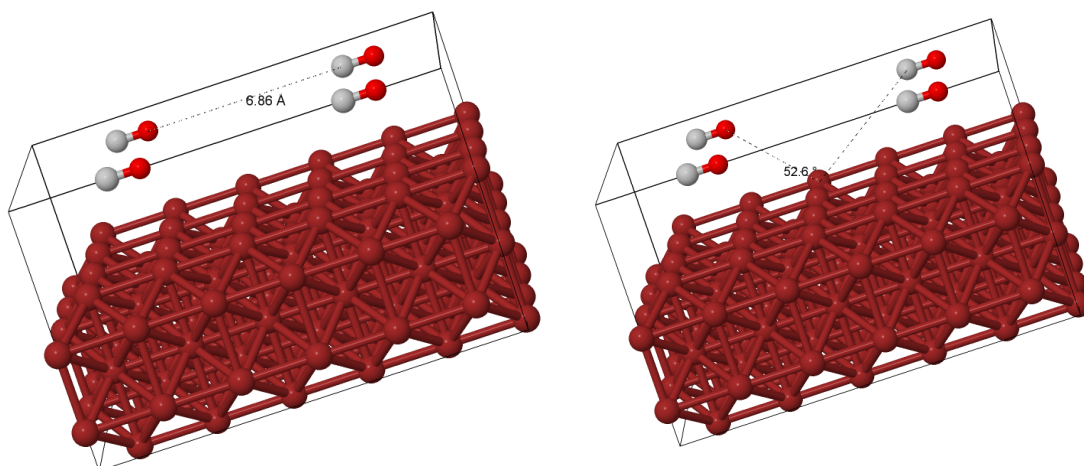


Figura 4.2 En la figura se observan dos configuraciones de la BD-0 representadas por sus posiciones atómicas, mediante una herramienta de visualización de estructuras atómicas. En estos ejemplos, el efecto de aplicar PBC provoca que se contabilicen casos ficticios generados por la imagen periódica de la molécula de C-O en las subceldas de la celda expandida.

4.2. Entrenamiento de redes con distintos conjuntos de parámetros de las FS

En la búsqueda por captar la mayor cantidad de detalles del entorno químico de los átomos en las diferentes configuraciones que constituyen la BD, las funciones que describen los entornos atómicos (las FS) deben ser hábilmente sensibles a las distribuciones de distancias entre pares de átomos y de ángulos entre ternas de estos. Para ello, en este capítulo, se investiga el efecto de los parámetros que determinan la forma funcional de las FS (detallados en la Sec. 2.3.2) con el objetivo de seleccionar el conjunto de parámetros de las FS que minimice el error de entrenamiento y de evaluación (σ_{train} y σ_{test}). Se recuerda al lector que la selección del conjunto de parámetros de las FS determina vectores descriptores diferentes y esto deriva en SEPN diferentes.

Puntualmente, se contrastan las zonas efectivas de las FS de los vectores descriptores (\mathbf{D}_2 y \mathbf{D}_3) con los histogramas de distancias y ángulos expuestos en la sección anterior (de la forma ejemplificada en la Fig. 4.3). De esta manera, es posible seleccionar conjuntos de parámetros $\{R_s, \eta\}$ y $\{\eta, \lambda, \xi\}$, en principio, coherentes. Sin embargo, por simple inspección de los resultados de la Fig. 4.3, no es posible determinar unívocamente cuál es el conjunto de parámetros de las FS adecuado para conseguir el menor valor en σ_{train} y σ_{test} durante el entrenamiento de las

Zonas efectivas de las FS de distintas redes

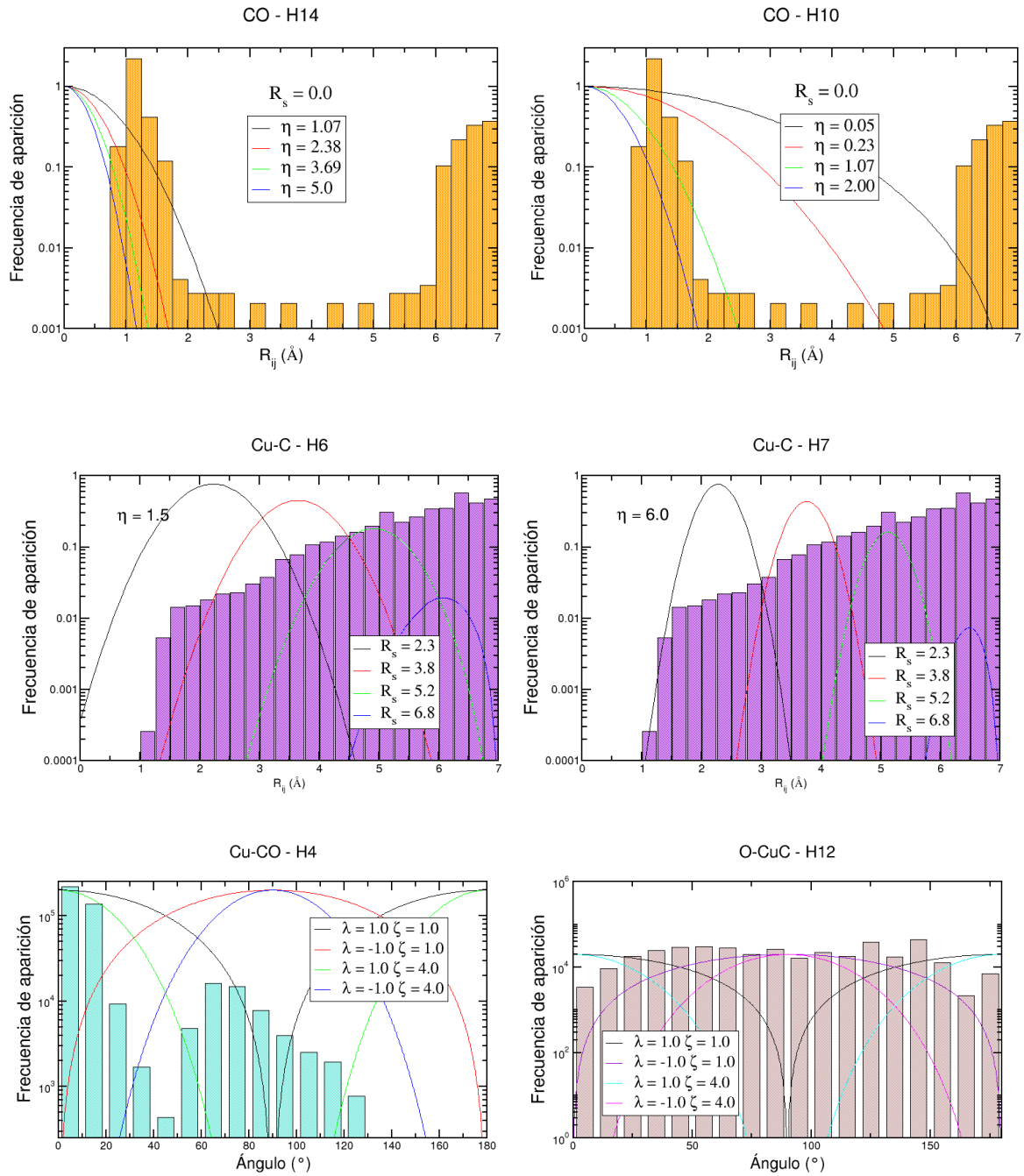


Figura 4.3 Los paneles superiores representan ejemplos en donde se comparan las zonas efectivas de algunas de las huellas estudiadas con las distribuciones de distancias y ángulos que conforman la BD-0, al momento inicial de los entrenamientos de las SEP.N.

SEP.Ns. Por esto, se conviene evaluar 15 conjuntos de parámetros levemente diferentes (variando los parámetros del modo indicado en la Tabla 2.2) cuyos valores son seleccionados a partir inspecciones como las ejemplificadas en la Fig. 4.3. Como no se esperan grandes variaciones de

la raíz del error cuadrático medio debido a que las variaciones de los parámetros de las FS no son abruptamente diferentes, con estos 15 conjuntos, se entrenan las SEPNs en un entrenamiento corto de 1000 pasos, pues se observó que luego de 1000 pasos el carácter de las raíces de los errores cuadráticos medios entre las redes persistía inalterado (es decir, la red cuyo σ era el menor en el paso 1000, persistía siendo el menor en el paso 30 000) y se determina, a partir de inspeccionar los errores σ_{train} y σ_{test} , cuales son los conjuntos de parámetros que determinan las mejores SEPN.

4.2.1. Selección de redes para optimización

Redes que difieran en sus parámetros de FS (o en sus huellas), generarán RNAs diferentes, y por ende, SEPN diferentes; pues sus parámetros peso optimizados convergerán a valores distintos. No obstante, es importante remarcar que las condiciones iniciales aleatorias con las que se seleccionan los subconjuntos para *train* y *test* y la inicialización de los parámetros de las RNA también provoca SEPNs diferentes; este punto será estudiado en la sección posterior.

En base al análisis de parámetros presentado en la sección precedente, se entrenaron las redes con los 15 distintos grupos de parámetros seleccionados por simple inspección de la Fig. 4.3. Se escogieron las redes que reportaban el menor σ_{test} en el paso 1000 como las redes candidatas para los subsecuentes procesos de optimización y análisis de la SEPN. En la Tabla 4.1 y en la Fig. 4.3 se presentan los menores valores de σ_{train} y σ_{test} del conjunto de las 15 SEPNs.

Nro. de configuraciones en la BD		4290
Pasos de entrenamiento		1000
Red	σ_{train}	σ_{test}
H13	9.9E-04	1.5E-03
H10	1.0E-03	1.6E-03
H12	1.2E-03	1.8E-03
H9	1.2E-03	1.9E-03
H11	1.2E-03	2.2E-03
H8	1.0E-03	2.4E-03
H7	3.0E-03	4.1E-03
H6	3.3E-03	5.8E-03

Cuadro 4.1 En la tabla se muestran las raíces del error cuadrático medio del conjunto de redes dentro de las 15 estudiadas cuyos valores resultaron ser menores. Se observa que los menores valores reportados corresponden a las redes H10, H12, y H13. Por lo tanto se continua optimizando la SEPN descrita por este conjunto de redes.

A partir de los resultados expuestos, se prosiguió con el entrenamiento de las redes cuyos conjuntos de parámetros de FS fueron etiquetados como H10, H12 y H13, para ser implementadas

Selección de redes para re-entrenamiento

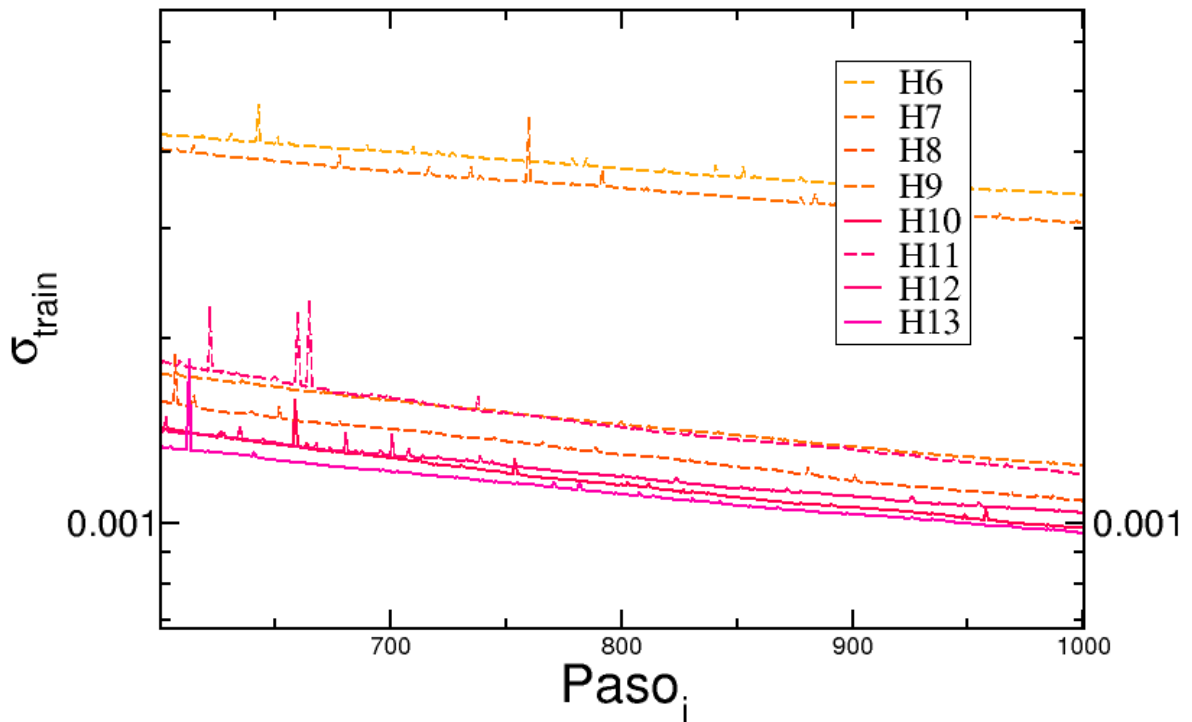


Figura 4.4 La comparación de los σ_{train} posiciona a las redes cuyos parámetros descriptores fueron etiquetados como H10, H12 y H13 como las SEPN candidatas a ser optimizadas y analizadas para describir la SEP.

como potenciales candidatas para simular la SEP del sistema CO/Cu(110). A partir de este momento, se descartaron las restantes redes y se prosiguió con el proceso de optimización y análisis iterativo de este conjunto de redes seleccionadas, cuyos valores de los parámetros de las FS se detallan en los Cuad. 4.2 y 4.3.

Valores implementados en los parámetros espaciales de los descriptores de 2 cuerpos

Distancia	D_2	H10				H12				H13				
Cu	Cu	η	0.05	0.10	0.23	0.40	0.05	0.10	0.23	0.50	0.05	0.10	0.23	0.50
		R_s	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	C	η	0.05	0.23	1.07	2.00	0.05	0.23	1.07	2.00	0.05	0.23	1.07	2.00
		R_s	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	O	η	0.05	0.23	1.07	2.00	0.05	0.23	1.07	2.00	0.05	0.23	1.07	2.00
		R_s	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Distancia	D_2	H10				H12				H13				
C	Cu	η	0.05	0.23	1.07	2.00	0.05	0.23	1.07	2.00	0.05	0.23	1.07	2.00
		R_s	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	O	η	0.05	0.23	1.07	2.00	0.05	0.23	1.07	2.00	0.05	0.23	1.07	3.00
		R_s	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Distancia	D_2	H10				H12				H13				
O	Cu	η	0.05	0.23	1.07	2.00	0.05	0.23	1.07	2.00	0.05	0.23	1.07	2.00
		R_s	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	C	η	0.05	0.23	1.07	2.00	0.05	0.23	1.07	2.00	0.05	0.23	1.07	3.00
		R_s	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Cuadro 4.2 Parámetros de las FS evaluados para el descriptor de dos cuerpos de las redes seleccionadas.

Cuadro 4.3 Parámetros de las FS evaluados para el descriptor de tres cuerpos de las redes seleccionadas.

Valores implementados en los parámetros espaciales de los descriptores de 3 cuerpos

Vértice	Extremos	D_3	H10			H12			H13					
			η	λ	ξ	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Cu	Cu	η	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		λ	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0
		ξ	1.0	1.0	4.0	4.0	1.0	1.0	4.0	4.0	1.0	1.0	4.0	4.0
	C	η	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		λ	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0
		ξ	1.0	1.0	4.0	4.0	1.0	1.0	4.0	4.0	1.0	1.0	4.0	4.0
C	η	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	
	λ	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	
	ξ	1.0	1.0	4.0	4.0	1.0	1.0	4.0	4.0	1.0	1.0	4.0	4.0	

Vértice	Extremos	D_3	H10			H12			H13				
			η	λ	ξ	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
C	Cu	η	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		λ	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0
		ξ	1.0	1.0	4.0	4.0	1.0	1.0	4.0	4.0	1.0	1.0	4.0
	O	η	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		λ	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0
		ξ	1.0	1.0	4.0	4.0	1.0	1.0	4.0	4.0	1.0	1.0	4.0

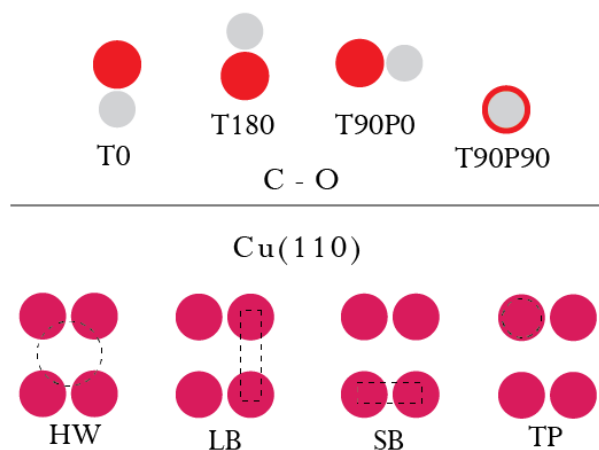
Vértice	Extremos	D_3	H10			H12			H13				
			η	λ	ξ	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
O	Cu	η	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		λ	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0
		ξ	1.0	1.0	4.0	4.0	1.0	1.0	4.0	4.0	1.0	1.0	4.0
	C	η	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
		λ	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0
		ξ	1.0	1.0	4.0	4.0	1.0	1.0	4.0	4.0	1.0	1.0	4.0

4.3. Evaluación de los rasgos característicos del sistema

En la Fig. 4.5 se muestran gráficas de la energía potencial CO-Cu(110) en función de la coordenada Z_{cm} , para la molécula a su distancia internuclear de equilibrio en el vacío y para las cuatro orientaciones mostradas en el esquema superior, y cuando el centro de masas de la molécula se encuentra sobre los sitios TP, HW, SB y LB, obtenidos mediante cálculos DFT y evaluando la SEPN. Los círculos representan los valores obtenidos mediante cálculos DFT y la distancia entre puntos adyacentes es de 0.25 \AA , las líneas continuas representan los valores obtenidos evaluando la SEPN y las separación entre puntos adyacentes es de 0.05 \AA . Puede observarse que las SEPN seleccionadas reproducen cualitativamente el comportamiento del sistema en todos los casos, ya que la SEPN presenta mínimos estables para la adsorción en las configuraciones correctas y la profundidad de los pozos de energía es adecuada. Se observa también que la SEPN es atractiva a largas distancias molécula superficie y repulsiva para $Z_{cm} \lesssim 2,1 \text{ \AA}$. Además no se observan oscilaciones espúreas en la región comprendida entre los datos DFT, típicas de potenciales sobre-ajustados. El acuerdo cuantitativo también es muy bueno ya que las diferencias entre los valores obtenido por DFT y la SEPN en ningún caso son mayores a 0.04 eV .

A partir de las figuras que describen las energías predichas por la red conforme la molécula de C-O se acerca a la superficie de Cu, es posible confiar en el ajuste que las redes seleccionadas realizan sobre las configuraciones características del estudio de la adsorción de CO sobre Cu(110).

a- Sitios y orientaciones particulares del sistema analizadas.



b- Predicciones de energía de las SEPN seleccionadas para los sitios y orientaciones de (a) en contraste con los datos DFT.

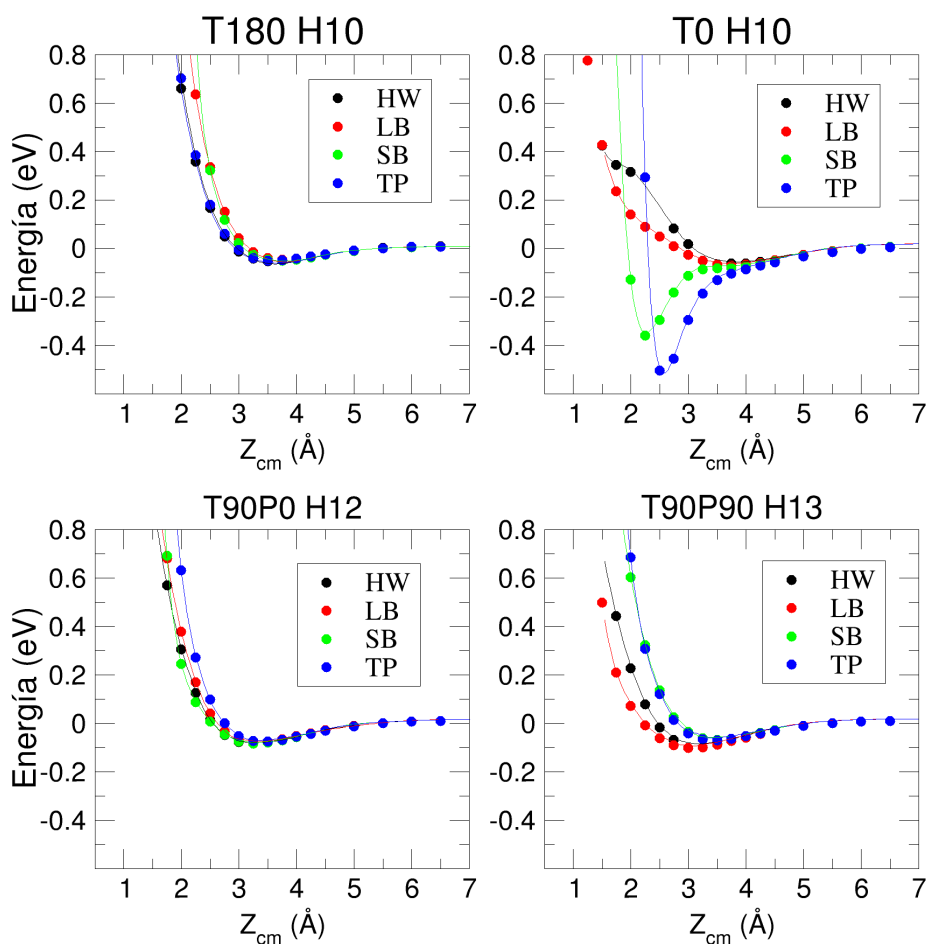


Figura 4.5 Energía potencial predicha por las SEPN en función de la coordenada Z_{CM} del sistema en contraste con la energía DFT. Los círculos representan los valores DFT y la distancia entre puntos adyacentes es 0.25 Å. Las líneas continuas representan los valores obtenidos evaluando las SEPN y la separación entre puntos adyacentes es de 0.05 Å. Cada color indica que el CM de la molécula se encuentra sobre un sitio de la superficie de Cu, detallado en el esquema superior. En este esquema se detallan también las 4 orientaciones de la molécula que se implementaron.

4.4. Efectos del incremento de configuraciones en la BD

Una vez seleccionados los conjuntos de parámetros de las FS de las SEPN con menor σ_{train} y σ_{test} , se procedió a estudiar el efecto de incorporar configuraciones en la BD en zonas donde las SEPNs discrepan más en la predicción de la energía del sistema; con el objetivo de hallar cuál de los conjuntos H10, H12 o H13 puede ser considerado el conjunto de parámetros más óptimo para constituir la SEPN representante del potencial CO/Cu(110).

Particularmente, como el proceso principal del estudio es la adsorción molecular de CO en la superficie Cu(110), se determinan las nuevas configuraciones a incorporar mediante un proceso que consiste en integrar cierto número de DM en donde la molécula es lanzada hacia la superficie con distintas energías de incidencia, de la siguiente manera:

- Con una sola de las redes (H12) se integraron 50 trayectorias para cada una de las energías: 0.01 eV, 0.2 eV, 0.4 eV, 1.0 eV, 1.5 eV y 2.0 eV, es decir, en total se integraron 300 trayectorias de DM y para cada una de estas trayectorias, se registran las configuraciones visitadas cada 50 pasos de tiempo.
- Con las restantes SEPN, se predicen las energías potenciales sobre las configuraciones visitadas durante las DM del punto anterior.
- Por cada trayectoria, se elige la configuración para la cual la discrepancia entre las energías potenciales predichas por las redes es máxima, lo que significa que se elige la configuración en donde es máxima la diferencia entre el mínimo valor de energía predicho por una de las SEPN y el máximo valor, predicho por otra. Como el total de trayectorias es 300, se eligen 300 configuraciones.
- Se evalúa la energía de estas configuraciones con discrepancia máxima a través de cálculos DFT y se las incorporan a la BD de entrenamiento.
- Se entrenan todas las redes con esa nueva BD ampliada y se repite el mismo proceso nuevamente.

Este proceso arroja los resultados con los que se confeccionan las figuras de la Fig. 4.6.

En la Fig. 4.6 (paneles (a)), las curvas indican los valores de la energía potencial predichos por las SEPN en las configuraciones de mayor discrepancia. Las primeras 50 configuraciones corresponden a trayectorias con energía de incidencia 0.01 eV, mientras que las últimas 50

configuraciones (de la 250-300) corresponden a la energía 2.0 eV, y las energías intermedias se ordenan de forma creciente (de la 50 a la 250). Se observa que conforme las configuraciones varían de la 1 a la 300, (es decir que la energía de incidencia aumenta de 0.01 eV a 2.0 eV) la energía de interacción donde las redes más discrepan aumenta (de ~ 0.25 eV a ~ 1.25 eV). Indicando que las configuraciones con mayores discrepancias son las que corresponden a mayores energías de incidencia. Lógicamente, debido a que la molécula no posee suficiente energía cinética inicial, es esperable no encontrar energías potenciales muy altas entre las configuraciones de baja energía. Por otro lado, para energías de 2.0 eV, la molécula posee mayor energía cinética inicial, por lo que las zonas del espacio de configuraciones que puede explorar tienen la posibilidad de ser zonas de más alta energía. No necesariamente la configuración donde las redes más discrepen debe ser una configuración de alta energía. Pero las figuras presentan una tendencia a que, conforme la molécula dispone de la posibilidad de acceder a configuraciones más energéticas, también tiende a discrepar más en esas regiones.

Sin embargo, las energías de las configuraciones con mayor discrepancia aparentemente no varían significativamente al incorporar las nuevas configuraciones a la BD, pues las curvas que describen los paneles (a) comparando las 3 iteraciones son similares.

En los paneles (b) y (c), se observa efectivamente la diferencia entre el menor y el mayor valor de la energía ($\Delta E_{max} = E_{SEPN,max} - E_{SEPN,min}$, simbolizado con círculos negros) para cada una de las 300 configuraciones. Las líneas rojas son los valores medios de las ΔE_{max} por cada energía de incidencia. Los paneles (b) se corresponden con las diferencias de energía calculadas a partir de los datos de los paneles (a), respectivamente para cada iteración. Estas trayectorias fueron inicializadas con las mismas condiciones iniciales (en cuanto a la semilla de los algoritmos aleatorios que determinan las condiciones iniciales de la DM).

En los paneles (c) también se detallan las ΔE_{max} de otras 300 configuraciones seleccionadas como configuraciones de discrepancia máxima, pero en los paneles (c) las condiciones iniciales de la DM han sido seleccionadas con distintos valores semilla que las DM de los paneles (a). Las energías de las redes implementadas para calcular las discrepancias de los paneles (c) no se exhiben en la figura.

En los paneles (b) y (c), se observa que el valor medio de las diferencias de energía aumenta conforme aumentan las energías de incidencia. Esto significa que la SEPN tiende a ser más precisa a más bajas energías (0.01 eV \sim 0.6 eV), mientras que a más altas energías el valor medio de las diferencias tiende a ser mayor, por lo que es recomendable muestrear en mayor medida

esta zona del espacio de configuraciones, mediante incorporar los resultados de los cálculos DFT en configuraciones de alta energía.

Se puede observar que la disminución del valor medio de las diferencias, conforme se avanza de la iteración 1 a la 3, en los paneles (c) no es significativa, e incluso para las trayectorias de (b) aumenta. Por lo que de generarse una disminución en las discrepancias máximas de energía al incorporar muestras en las zonas de alta discrepancia, esta disminución no parece ser muy eficiente.

Para escoger la red cuyos parámetros de las FS resultan ser más óptimos, se estudian los resultados de las raíces del error cuadrático medio (σ_{train} y σ_{test}) por cada vez que más configuraciones eran agregadas a la BD. Los resultados obtenidos se confeccionan en los siguientes cuadros:

Iteración no. 1				
Configuraciones en la BD				4170
Pasos de entrenamiento				30000
Red	σ_{train}	MAE $_{train}$	σ_{test}	MAE $_{test}$
H12	2.1E-04	1.2E-04	4.5E-04	1.9E-04
H10	1.9E-04	1.1E-04	7.6E-04	2.3E-04
H13	2.4E-04	1.5E-04	9.8E-04	3.2E-04

Iteración no. 2				
Configuraciones en la BD				4370
Pasos de entrenamiento				30000
Red	σ_{train}	MAE $_{train}$	σ_{test}	MAE $_{test}$
H12	2.9E-04	1.9E-04	6.6E-04	3.5E-04
H13	3.0E-04	1.9E-04	7.4E-04	3.3E-04
H10	3.1E-04	2.0E-04	1.1E-03	4.4E-04

Iteración no. 3				
Configuraciones en la BD				4670
Pasos de entrenamiento				30000
Red	σ_{train}	MAE $_{train}$	σ_{test}	MAE $_{test}$
H12	3.5E-04	2.4E-04	9.5E-04	4.2E-04
H13	4.9E-04	3.4E-04	1.0E-03	5.0E-04
H10	5.2E-04	3.6E-04	1.1E-03	5.0E-04

Cuadro 4.4 Los resultados de las raíces del error cuadrático medio bajo el efecto de incorporar más cantidad de configuraciones a la BD de entrenamiento en las zonas donde los potenciales presentaban mayores discrepancias en cuanto a la predicción de energía potencial.

Dicho esto, se selecciona la red que alcanzó el mayor grado de convergencia, de manera robusta se puede decir que fue la red H12 por los Cuad. 4.4, para investigar los efectos de la aleatoriedad en la inicialización de los parámetros peso y la conformación de los conjuntos de

train y *test* sobre el grado de convergencia y precisión de la SEPN resultante. Para responder a este cuestionamiento, en la sección siguiente, se observa qué sucedería si en vez de entrenar redes que difieran en los parámetros de sus FS, se entrena siempre la misma red (en cuanto a iguales parámetros de su FS).

Discrepancias máximas de energía para las DM de redes con distinta huella

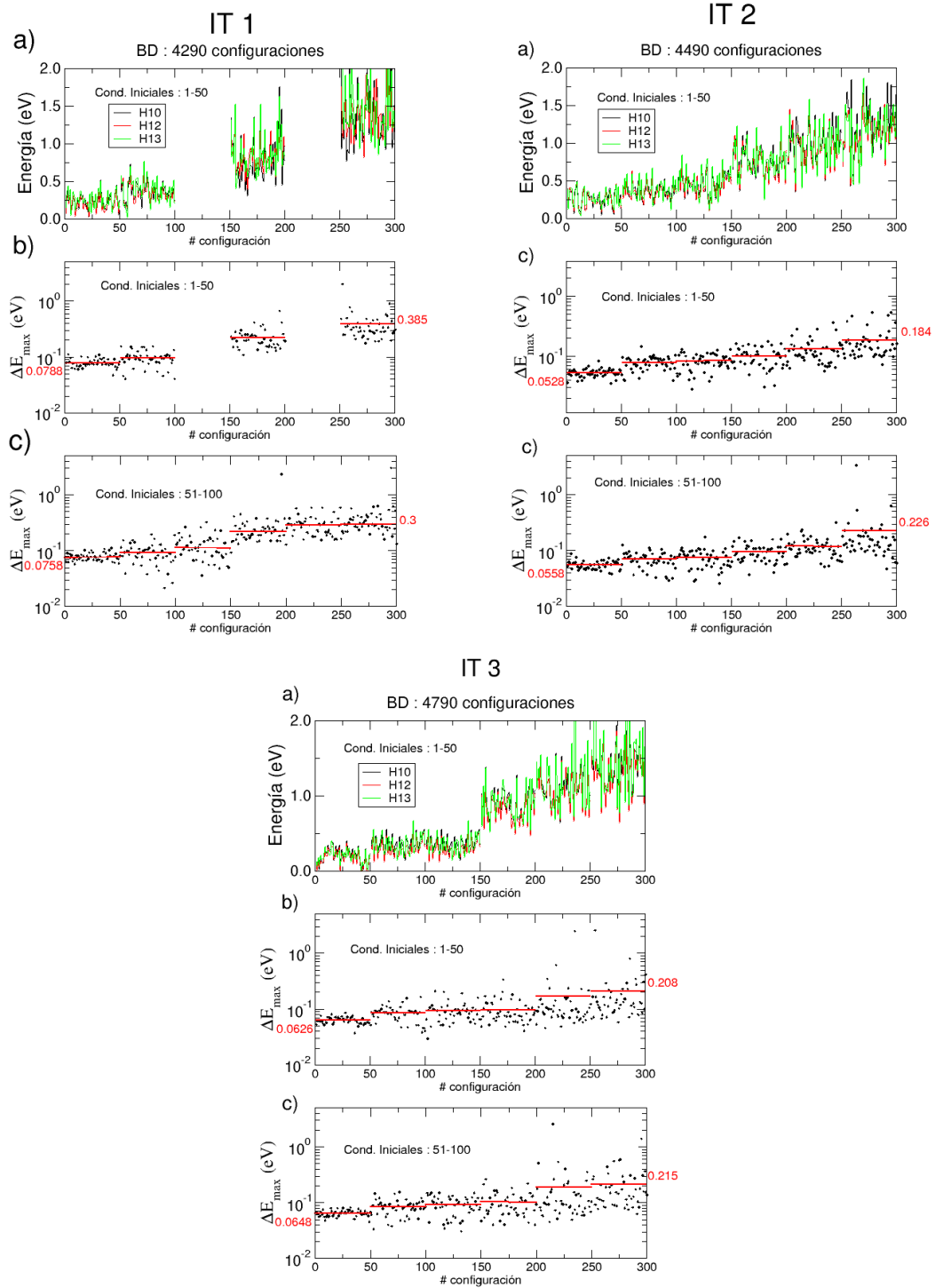


Figura 4.6 Las energías de incidencia se ordenan de menor a mayor sobre el eje x desde la configuración 1 a la 300. En los paneles (a), las curvas negra, roja y verde representan para las redes H10, H12 y H13, las energías predichas en las configuraciones donde se identificó la mayor discrepancia. Los paneles (b), se construyen mediante calcular la diferencia que existe entre el menor valor de energía y el mayor valor de energía registrado en los paneles (a) y (b) para cada configuración, y las líneas rojas representan el valor medio de estas diferencias para cada energía de incidencia. Los paneles (c) se construyen de igual manera que (b), pero a partir de los resultados de DM cuyas condiciones iniciales fueron diferentes de las de los paneles (a) (los resultados implementados para el cálculo no se exponen en esta figura).

4.5. Discrepancias entre redes con iguales vectores descriptores

En este análisis, se utiliza el conjunto de parámetros huella seleccionado (el de la H12) y se entrenan 3 nuevas redes neuronales con estos parámetros. A estas redes se las etiqueta como CI_{m-n} donde m indica el número de iteración y n es simplemente un índice que diferencia a una red de la otra.

Nuevamente, se investiga si el proceso de aumentar sucesivamente el número de configuraciones en la BD, es efectivo para mejorar la convergencia de la SEPN. Para ello, se realizaron tres iteraciones más con las redes CI_{m-n} mediante el mismo proceso de incremento de la BD que se detalló en las secciones precedentes (continuando el incremento con la BD de 4670 configuraciones) y se observa el grado de convergencia y precisión de la SEPN resultante:

Iteración no. 4				
Configuraciones en la BD				4670
Pasos de entrenamiento				30000
Red	σ_{train}	MAE_{train}	σ_{test}	MAE_{test}
CI1-3	5.0E-04	3.4E-04	9.1E-04	4.9E-04
CI1-1	3.5E-04	2.4E-04	9.5E-04	4.2E-04
CI1-2	3.9E-04	2.6E-04	4.5E-03	6.5E-04

Iteración no. 5				
Configuraciones en la BD				4970
Pasos de entrenamiento				30000
Red	σ_{train}	MAE_{train}	σ_{test}	MAE_{test}
CI2-5	3.4E-04	2.3E-04	6.4E-04	3.4E-04
CI2-4	3.9E-04	2.6E-04	8.6E-04	4.0E-04
CI2-6	3.3E-04	2.3E-04	3.7E-03	5.4E-04

Iteración no. 6				
Configuraciones en la BD				5270
Pasos de entrenamiento				30000
Red	σ_{train}	MAE_{train}	σ_{test}	MAE_{test}
CI3-9	3.7E-04	2.5E-04	5.7E-04	3.4E-04
CI3-7	4.8E-04	3.2E-04	1.2E-03	4.9E-04
CI3-8	4.6E-04	3.2E-04	1.6E-03	5.5E-04

Cuadro 4.5 En el cuadro se muestran los valores de la raíz del error cuadrático medio y del Error Absoluto Medio (MAE) que se obtuvieron luego de realizar tres iteraciones del entrenamiento de las SEPN, es decir, luego de incorporar más configuraciones a la BD de entrenamiento en zonas del espacio de configuraciones donde las redes presentaban mayores discrepancias.)

En vista de los resultados obtenidos, por observación del valor σ_{test} , se deduce que no es posible seguir aumentando la convergencia de las redes, debido a que al incorporar configuraciones de la iteración 5 a la 6, el σ_{test} a incrementado. Esto indica que es posible que la

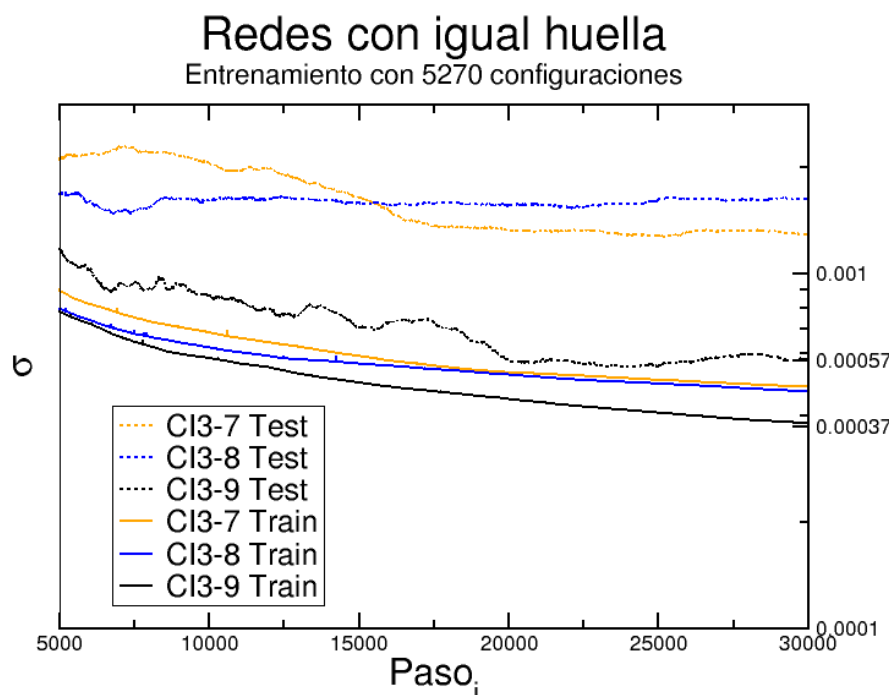


Figura 4.7 Cada curva de la figura indica para las tres redes entrenadas, el comportamiento de los valores σ_{train} y σ_{test} calculados durante cada paso del entrenamiento. De la raíz del error cuadrático medio analizado durante la última instancia de entrenamiento de las redes con igual parámetros de sus FS (conjunto H12), se deduce que la red que mejor esta describiendo la SEP es la red CI3-9

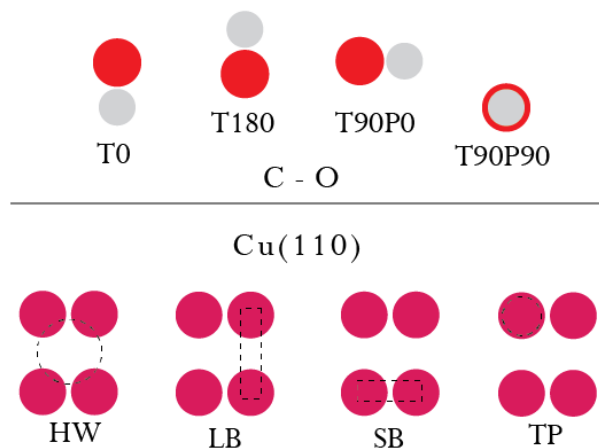
convergencia en la precisión de la SEP N se encuentre en los límites de la zona de *overfitting* por lo que se decide implementar la red del último entrenamiento (CI3- n) para la cual el valor de σ_{test} ha resultado menor. Se selecciona el sistema de RNAs de la red CI3-9 como representantes de la SEP N del sistema CO/Cu(110) y se la implementa para el estudio de las propiedades físicas, en particular, el fenómeno de adsorción.

4.6. Características estudiadas con la SEP N seleccionada

A partir de la elección de la SEP N dispuesta en la sección anterior, se procede a corroborar que las propiedades fundamentales del sistema en estudio se encuentren correctamente satisfechas. Para eso se disponen de dos de las propiedades esenciales del sistema, por un lado se observa la curva que describe el potencial neuronal conforme la molécula se acerca a la superficie (del mismo modo que se estudió en la Sec. 4.3) y por otro lado se estudia la probabilidad de adsorción molecular del monóxido de carbono sobre la superficie de cobre (110).

4.6.1. Rasgos característicos del potencial

a- Sitios y orientaciones particulares del sistema



b- Predicciones de energía de la SEP-NH12 para los sitios y orientaciones de (a) durante la IT3

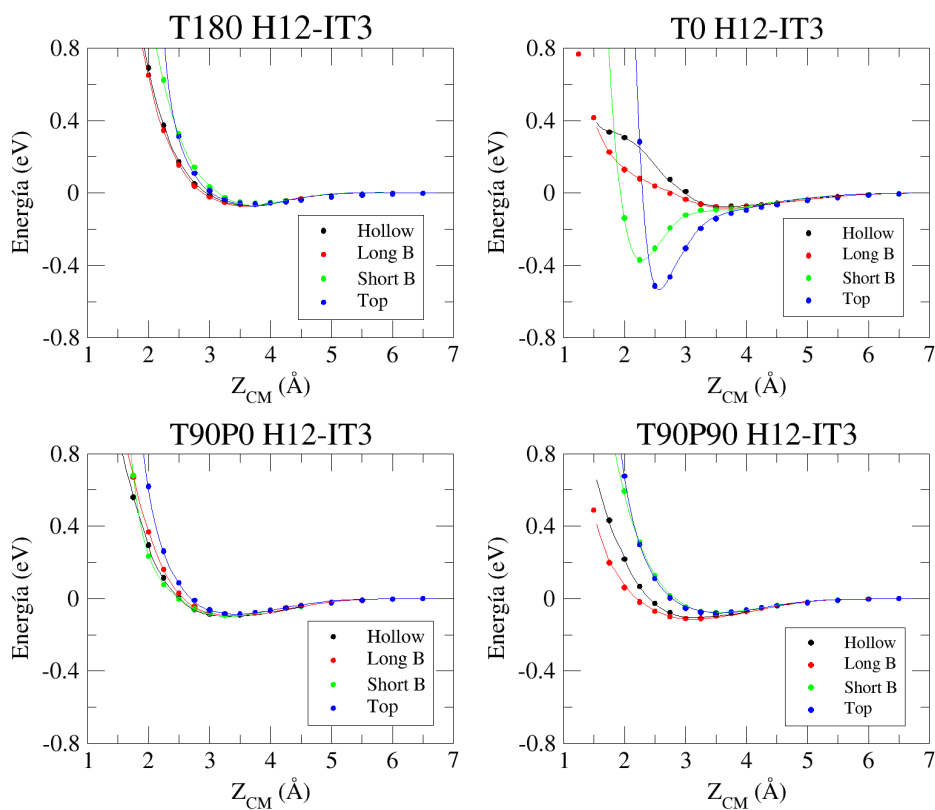


Figura 4.8 Resultados de las predicciones de las SEP-N entrenadas durante la iteración 3. Los puntos representan los cálculos hechos mediante DFT y las líneas los valores predichos por la SEP-N H12 sobre esas configuraciones. La distancia entre puntos adyacentes DFT es de 0.25 \AA y la de SEP-N es de 0.05 \AA .

Del mismo modo que el análisis hecho durante la Sec. 4.3, en donde se estudiaron las predicciones de las SEPN entrenadas sobre los rasgos característicos ya conocidos del sistema, es decir, se estudió que las SEPN describieran correctamente las orientaciones preferenciales para la adsorción molecular y que esta misma ocurra en los sitios correspondientes. Este estudio se realizó reiteradas veces sobre la red H12 que resultó ser la red seleccionada para la representación de la SEP de CO/Cu(110).

Como puede verse en la Fig. 4.8 donde nuevamente se muestran gráficas de la energía potencial CO/Cu(110) en función de la coordenada Z_{cm} , para la molécula a su distancia internuclear de equilibrio en el vacío y para las cuatro orientaciones mostradas en el esquema superior, y cuando el centro de masas de la molécula se encuentra sobre los sitios TP, HW, SB y LB, obtenidos mediante cálculos DFT y evaluando la SEPN; las predicciones de la SEPN aparentan ser correctas en contraste con los datos DFT (representados por líneas las primeras y por puntos los segundos), los mínimos estables que presenta la SEPN se encuentran en las configuraciones correctas y la profundidad de los pozos de energía es la adecuada. Una vez más la diferencia entre los valores DFT y la SEPN en ningún caso es mayor a 0.04 eV. Por lo que se deduce que la SEPN H12 describe respectivamente bien el potencial estudiado.

4.6.2. Probabilidad de adsorción

Otra de las características del sistema que fueron analizadas para determinar la precisión de la SEPN H12, fue el estudio de la probabilidad de adsorción inicial (PBA).

La PBA se define, para una dada energía, como la cantidad de casos en que las trayectorias de DM han finalizado con la molécula adsorbida, sobre la cantidad de trayectorias totales integradas para esa energía (tal y como se lo definió en la Sec. 3.2), según los resultados obtenidos con el código *GSRD-dyn*.

La figura que se obtiene durante las sucesivas ampliaciones de la BD para la probabilidad de adsorción, se detalla en la Fig. 4.9. Esta figura muestra la probabilidad inicial de adsorción estimada con la red H12 por cada vez que nuevas configuraciones eran agregadas a la BD.

Se observa que la probabilidad de adsorción, conforme se agregan configuraciones a la BD, converge a la curva delimitada por las curvas H12-IT5 y H12-IT6. Como en la Fig. 4.9 se observa que en las últimas ampliaciones de la BD (IT4 a IT6) la forma de las curvas PBA es similar, se puede estimar que la probabilidad ha convergido y en base a este resultado, se concluye que el potencial definido con los parámetros de FS de la red H12 ha alcanzado un estado óptimo

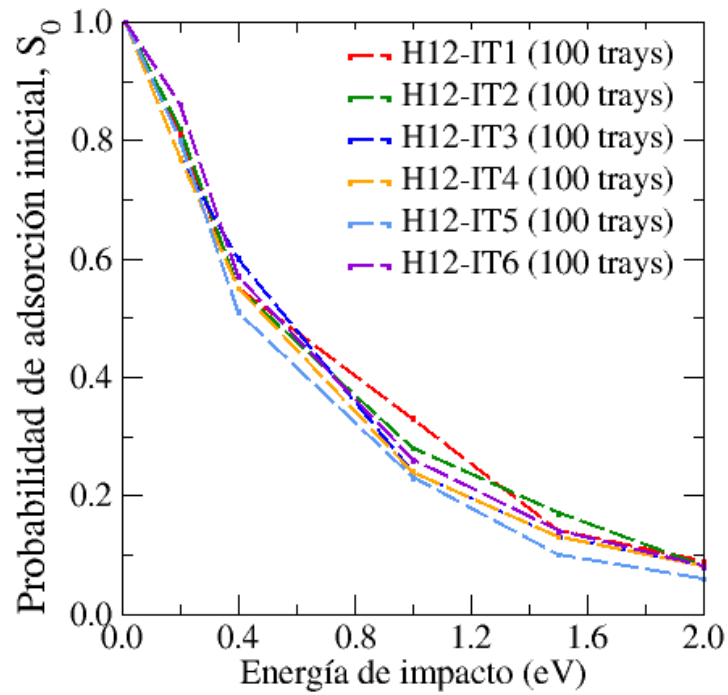


Figura 4.9 Cada línea de la figura representa la curva de PBA en función de la energía de incidencia de la SEP N H12, durante las iteraciones de la IT1 a la IT6. Las energías de incidencia evaluadas son: 0.01eV, 0.20eV, 0.40eV, 1.00eV, 1.50eV y 2.00eV. La cantidad de trayectorias integradas para la estadística es 100. La PBA se calcula mediante la relación entre la cantidad de trayectorias adsorbidas sobre la cantidad de trayectorias integradas.

para poder ser implementado como SEP N del sistema CO/Cu(110) para simular DM con más cantidad de energías y durante más cantidad de trayectorias.

Capítulo 5

Utilización de la Superficie de Energía Potencial Neuronal

5.1. Dinámica con temperatura

El modelo implementado para representar las vibraciones de los átomos en la red cristalina, propone que los átomos de Cu son representados como un objeto clásico, por lo que sus modos de oscilación representan un continuo de energías de oscilación. La temperatura asignada a la superficie permite calcular el valor medio de la energía interna de los átomos mediante la ecuación

$$E_{int} = N \cdot \frac{3}{2} \cdot k_B \cdot T_s \quad (5.1)$$

En el caso de la superficie representada, se termalizan solamente las primeras tres capas de Cu, por lo que el valor $N = 18$, a una temperatura $T=90$ K, resulta $E_{int} = 0,023$ eV.

Para la realización de DM en superficies con temperatura es preciso generar configuraciones que simulen las vibraciones de los átomos de la red cristalina según la temperatura establecida y de forma físicamente coherente. Por esta razón, se analizan las curvas generadas a partir de las configuraciones que simulan la vibración de los átomos a la temperatura $T_s = 90K$, en las Figs. 5.1.

En el panel (a) se reporta la energía cinética y la energía potencial de vibración de los átomos de la red cristalina de todas las capas móviles (las tres capas termalizadas). Con el histograma, se puede ver que los átomos adoptaron energías potenciales y cinéticas en el rango de energías de 0 a 0.4 eV. El pico de ambas distribuciones coincide con el valor delimitado por la línea de trazos vertical. Este valor indica la E_{int} calculada en la Ec. 5.1. Al coincidir los picos del histograma con la línea de trazos, se concluye que la mayoría de los átomos de Cu presentan

energías iguales al valor calculado para la energía interna a temperatura $T=90\text{K}$. Además, dado que ambas curvas coinciden, se verifica que la simulación cumple adecuadamente el Principio de equipartición de la energía [19].

El panel (b) representa la distribución de temperaturas de las configuraciones generadas. Éstas abarcan un rango de 0K a 180K , y se calculan a partir de todos los átomos de las tres capas móviles. En esta figura se observa que la distribución presenta su pico alrededor de 90K , lo que significa que la temperatura calculada para la mayoría de las configuraciones es cercana al valor deseado de 90K .

En el panel (c) se ilustra, para cada capa atómica, la distribución de las desviaciones respecto a la posición de equilibrio de cada átomo, en las direcciones del eje x , y y z entre todas las configuraciones generadas. El pico de las distribuciones sugiere que la última capa (línea verde) es la que menos desplazamiento presenta en todas las direcciones, puesto que el pico del histograma es siempre el más alto y centrado en cero. Mientras que la primera capa (línea negra) disminuye su desplazamiento alrededor de la zona de equilibrio conforme cambia la dirección del eje x al z (ya que el punto máximo del histograma alrededor de cero, aumenta del eje x al z). De estos dos puntos se infiere que, tal como es de esperarse, los átomos de la capa superficial se encuentran menos ligados que los átomos de las capas atómicas internas. Además se puede concluir que el potencial es isotrópico, es decir, que los átomos pueden moverse en las tres direcciones con igual probabilidad.

A partir de estos resultados es posible corroborar que la temperatura de la superficie a 90K se representa correctamente en la simulación. Se concluye que el potencial implementado es coherente con esa situación, y reproduce situaciones que tienen deformaciones físicas acordes a la temperatura simulada.

Análisis de las propiedades térmicas generadas para la superficie de Cu(110) a 90 K

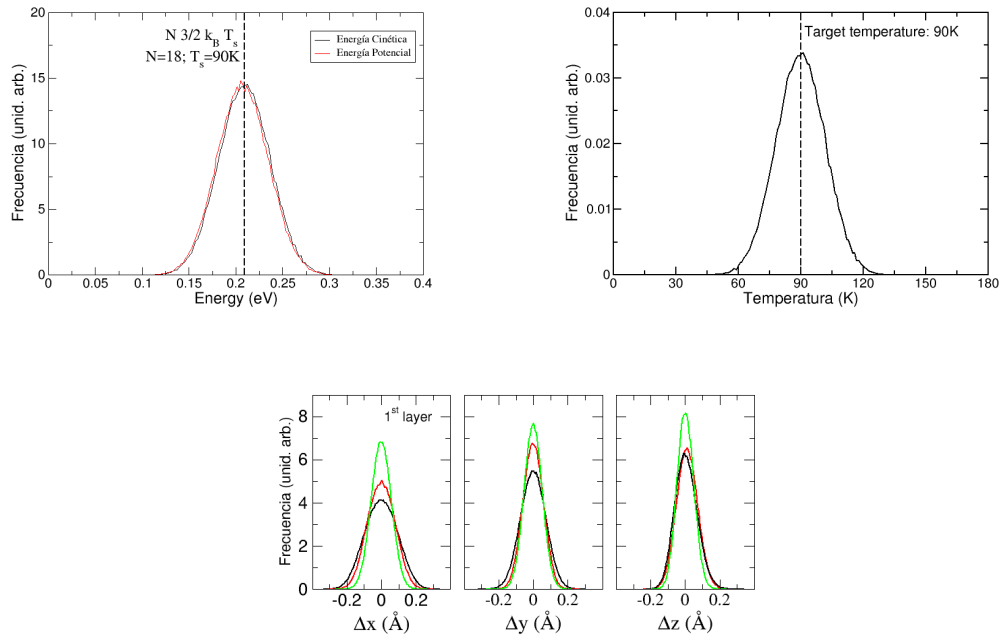


Figura 5.1 En (a), la curva que se observa representa la energía cinética y la energía potencial de las moléculas que componen la superficie. Se puede ver como el principio de equipartición de energía se satisface correctamente para todas las vibraciones de los átomos. En (b), se observa la distribución de temperaturas de los átomos que componen la superficie, el pico se encuentra sobre la temperatura deseada, indicando que el valor medio de la temperatura de la superficie es el correcto. Para el panel (c), se observa como los átomos de Cu oscilan alrededor de su posición de equilibrio tanto en la dirección x (primer panel), y (segundo panel) como z (tercer panel). La primer, segunda y tercer capas se representan por los colores negro, rojo y verde respectivamente

5.2. Probabilidad de Adsorción

Una vez definida la SEPN más óptima, ésta se implementa en diferentes simulaciones de DM con mayor número de trayectorias y para mayor número de valores de energía entre 0 eV y 2.0 eV.

Puntualmente, se simularon trayectorias de 0.01 eV, 0.1 eV, 0.2 eV, 0.4 eV, 0.6 eV, 0.8 eV, 1.0 eV, 1.25 eV, 1.5 eV, 1.75 eV y 2.0 eV. Para las energías entre 2.0 eV y 0.6 eV se simularon 1000 trayectorias y para las restantes energías, 200 trayectorias debido a las limitaciones del costo computacional de cada simulación. Con los resultados de las simulaciones, se calculó la probabilidad de adsorción estadística (PBA) a partir del cociente entre la cantidad de moléculas finales reportadas como adsorbidas y la cantidad de trayectorias integradas, por cada energía (según la definición dada en la Sec. 3.2). El resultado de este análisis se observa en la Fig. 5.2.

A partir de los resultados obtenidos, se observa una marcada disminución de la probabilidad de adsorción al aumentar la energía de incidencia. La PBA es aproximadamente 1 para energías $\sim 0,0$ eV y disminuye hasta un valor $\sim 0,1$ para energías cercanas a 2.0 eV. Este comportamiento de la curva resulta ser cuantitativamente bueno comparado con la forma sugerida por los valores en las zonas donde se dispone de datos experimentales. Además, al aumentar la temperatura de la superficie de 0K a 90K, no se observa una variación significativa de la PBA.

Mediante experimentos de haces moleculares supersónicos se ha establecido que la adsorción molecular de CO sobre Cu(110) no presenta variaciones significativas en el rango de 0K a 120K [20]. Por lo que se considera que la SEPN entrenada está describiendo coherentemente los resultados experimentales. La curva de PBA conserva su comportamiento al cambiar la temperatura y el hecho de calcular la PBA para más energías permite obtener una curva suave.

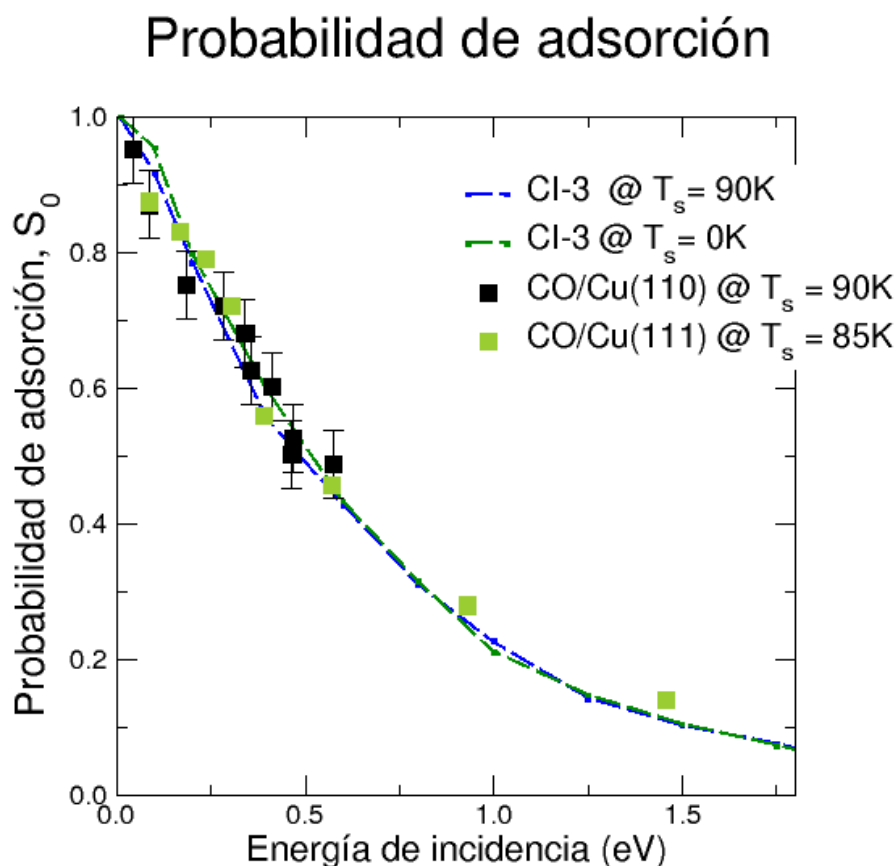


Figura 5.2 Se contrastan los resultados de valores experimentales sobre la probabilidad de adsorción de la molécula. Estos datos se contrastan con la curva que se genera mediante realizar DM implementando la SEPN para varias energías. Se observa que en todos los casos los valores obtenidos en la simulación se encuentra dentro del rango de error de los valores experimentales.

5.3. Análisis del estado de las moléculas adsorbidas

5.3.1. Distribución de energías de interacción

Para una configuración dada, la energía de interacción se calcula mediante comparar la energía del sistema con la energía que se obtiene al trasladar la molécula, rígidamente, hasta una zona donde la interacción con la superficie sea despreciable. De esta comparación (que implica restar las energías de la molécula cercana a la superficie y la molécula trasladada rígidamente hasta una zona de interacción despreciable con la superficie) es posible estimar cuantitativamente la energía provocada por la interacción molécula-superficie sin tener en cuenta las energías propias del estado de vibración de la molécula, ni de los átomos de la superficie.

Los análisis hechos sobre esta energía se reportan en la Fig. 5.3. El panel (a) (panel (b)) representa la distribución de energías de interacción de las configuraciones a $t=5$ ps, en donde la DM reportó un evento de adsorción, con la temperatura de la superficie a $T_s = 0K$ ($T_s = 90K$). Cada histograma está normalizado por la PBA de cada energía (es decir que el área de los histogramas representa la PBA calculada para esa energía correspondiente). Debido a que las PBA son diferentes (y disminuyen conforme la energía de incidencia aumenta) se observa como las áreas de los histogramas disminuyen correspondientemente a este resultado, siendo mayor el área correspondiente a la energía de incidencia igual a 0.01 eV y menor, a la de 2.0 eV. Las energías de interacción reportadas, se encuentran entre el rango de $\sim -0,75$ eV hasta $\sim -0,35$ eV. Aproximadamente, la mayoría de los casos reportados en todo el rango de energías se ubica entre las energías $\sim -0,6$ eV y $\sim -0,5$ eV. Si bien este resultado no se corresponde exactamente con la profundidad del pozo de potencial estudiado en la Sec. 2.1 (particularmente en la Fig. 4.5), en donde el sistema presenta un estado estable de energía del orden de $\sim -0,45$ eV, se estima que la diferencia de energías entre los estados estables del panel (a) de la Fig. 5.3 y los del sitio TP y SB de la Fig. 4.5 se deben a que existen diferencias respecto al estado en que se considera la energía interna de la molécula al interactuar con el potencial. Esto último, se refiere a que en las curvas del potencial analizadas en la Fig. 4.5 se considera a la molécula en su estado de mínima energía (es decir con los átomos C y O a la distancia de equilibrio) mientras que, en el resultado de la Fig. 5.3, en donde la mayoría de energías de interacción son (en valor absoluto) más elevadas que las de la Fig. 4.5, no hay restricción sobre el estado de la molécula (salvo por la energía total disponible), provocando que en ciertos casos, la energía estable del pozo sea (en valor absoluto) mayor al pozo estudiado en la Sec. 2.1.

En el panel (b) (de la Fig. 5.3) se observa la energía de interacción para las configuraciones finales de DM con temperaturas superficiales de $T_s = 90K$. La construcción de los histogramas es similar que en la figura adyacente: cada distribución está normalizada según la PBA de la energía que representa. Del mismo modo, y en concordancia con el hecho de que las PBA no varían significativamente para rangos de temperatura entre 0 K y 120 K, el área de estas distribuciones disminuye conforme aumenta la energía de incidencia (el área de la energía 0.01 eV es mayor al área de la energía 2.0 eV); representando el hecho de que la PBA de la energía 0.01 eV es mayor a la de 2.0 eV. A su vez, en la mayoría de los casos las energías de interacción finales coinciden con el rango de $\sim -0,6$ eV hasta $\sim -0,5$ eV. Similarmente al panel adyacente, la frecuencia de aparición de los casos donde las energías de interacción rondan los $-0,5$ eV \sim

$-0,6$ eV explicita el hecho de que cuando el estado de la molécula no es restringido a ser el de mínima energía, el pozo de adsorción, que en el caso de mínima de energía de la molécula es de $\sim -0,45$ eV, se torna más profundo y adopta valores de entre $-0,5$ eV $\sim -0,6$ eV.

Energía de interacción molécula-superficie

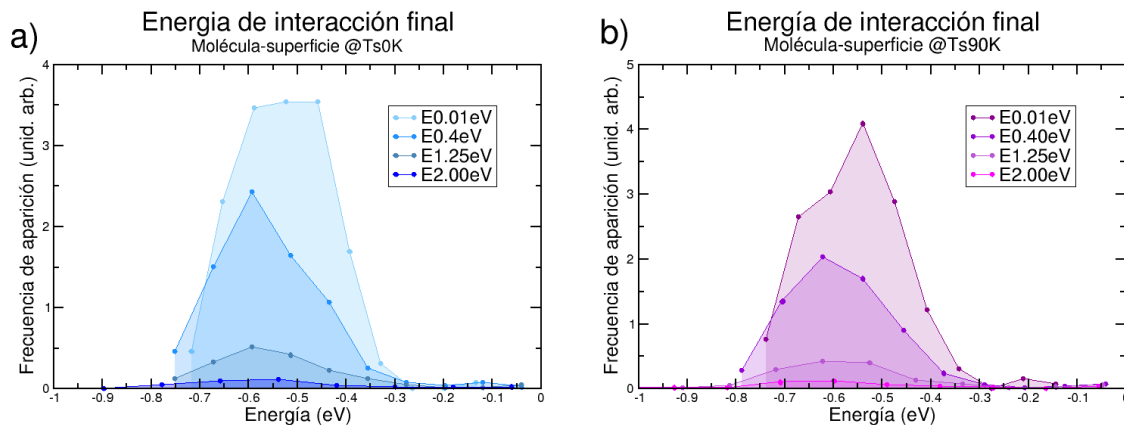


Figura 5.3 Se observa distribución de energías de interacción molécula-superficie, calculada para las configuraciones finales de los eventos de adsorción con superficies a temperatura 0 K (90 K) en el panel izquierdo (derecho). El área de cada superficie representa la probabilidad de adsorción estadística obtenida con cada energía. La distribución presenta un pico en la zona de -0.6 eV ~ -0.5 eV debido al pozo de potencial característico del sistema.

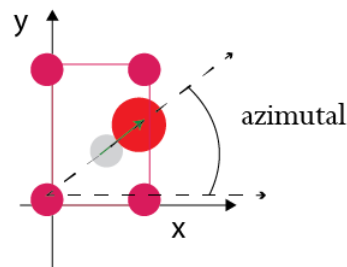
Además, de la figura, también se infiere que al crecer la energía de incidencia, la campana de distribuciones en los histogramas tiende a ensancharse; indicando que, al poseer más energía de incidencia, más cantidades de moléculas tienen la posibilidad de adsorberse con mayor fuerza en el pozo de potencial. Esto se deduce pues la amplitud máxima de todos los histogramas se desplaza levemente hacia valores más negativos conforme la energía de incidencia aumenta.

5.3.2. Distribución espacial de las moléculas

En esta sección se analizan las distribuciones espaciales de las configuraciones cuya DM finalizó en estado de adsorción molecular. Se caracterizan los ángulos polares y azimutales que forman el eje internuclear de la molécula con los ejes de referencia solidarios a la superficie, las posiciones atómicas de los átomos y del CM de la molécula sobre la superficie y la distancia interatómica entre los átomos de C-O al comienzo y al final del proceso.

5.3.3. Ángulo azimutal

El ángulo azimutal es el ángulo que forma el eje internuclear de la molécula con la dirección positiva del eje x , solidaria a la dirección que coincide con el primer vecino de la red cristalina de la superficie de Cu (es decir en el sistema de referencias solidario a la superficie, el eje x es paralelo a la dirección donde se encuentra el SB de la red cristalina del Cu).



En la Fig. 5.5 se reportan las distribuciones del ángulo azimutal, calculado a partir de las posiciones atómicas de las configuraciones donde se reportó la adsorción molecular. Estos histogramas han sido normalizados, de manera que el área total es igual a 1 para todos los casos.

Los paneles (a) y (c), representan las DM a temperatura superficial $T_s = 0K$, y los (b) y (d), las DM a temperaturas de $T_s = 90K$. En los paneles (a) y (b) se ilustran las distribuciones de ángulos azimutales reportadas en las configuraciones iniciales (es decir al comienzo de la DM en $t = 0ps$), para varias energías de incidencia diferentes. Los paneles (c) y (d), aquellas distribuciones calculadas a partir de las configuraciones finales (en $t = 5ps$), también para diferentes energías de incidencia.

Todos los histogramas aparentan una distribución uniforme de ángulos azimutales, pues los histogramas no presentan mayoría de casos en ningún ángulo particular. Las oscilaciones que se observan se adjudican al hecho de que es necesario integrar una amplia cantidad de DM para que el muestreo presente una curva uniforme; lo cual resulta ser ineficiente y costoso computacionalmente. En ese aspecto, al no ser eficiente realizar la cantidad de trayectorias de DM necesarias para que el histograma resulte uniforme, éste queda condicionado por la baja estadística que es posible conseguir sin sacrificar la eficiencia de la simulación computacional implementada.

De la distribución uniforme de ángulos azimutales tanto para temperaturas superficiales a $T_s = 0K$ como a $T_s = 90K$, y para las configuraciones en $t = 0ps$ como en $t = 5ps$, se desprende que no existe orientación azimutal preferencial de las moléculas adsorbidas. Es decir que el potencial no realiza ninguna fuerza sobre la molécula para modificar la orientación azimutal del

Figura 5.4 Esquema del ángulo azimutal. La esfera roja (gris) representa al átomo de Oxígeno (Carbono). Las esferas rosas representan a los átomos de Cu.

eje de internuclear del C-O al interactuar con la superficie. Por ende, la proyección en el plano paralelo a la superficie del vector internuclear de la molécula puede apuntar tanto en la dirección del eje x como en la dirección del eje y y en todas las direcciones intermedias entre los ejes.

Distribución del ángulo azimutal

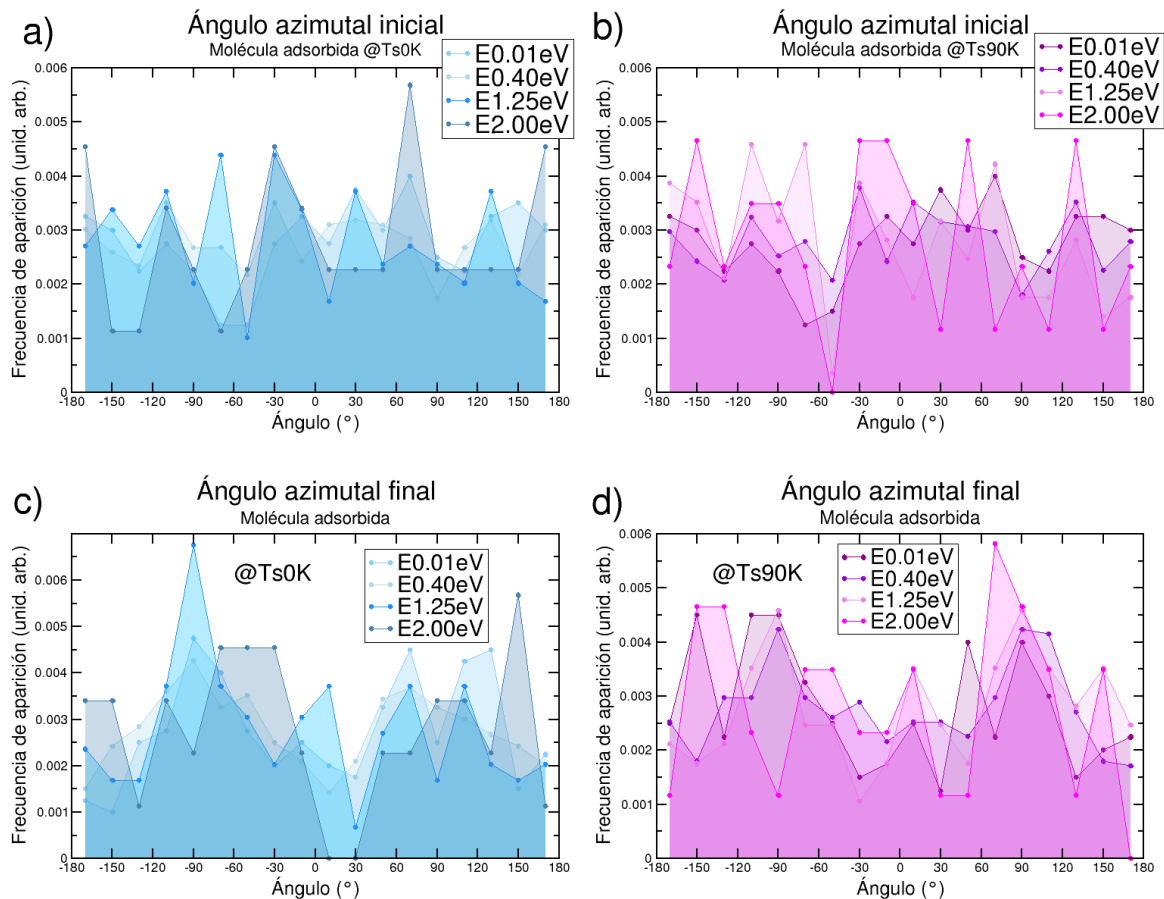


Figura 5.5 Los paneles representan las distribuciones del ángulo azimutal que forma la proyección del eje internuclear de la molécula sobre el plano paralelo a la superficie. Se representan las distribuciones para los eventos de adsorción en $t=0$ ps -(a) y (b)- y en $t=5$ ps -(c) y (d)-, y con la superficie a temperatura 0 K (90 K) para los histogramas con color azul (rosa).

5.3.4. Ángulo polar

El ángulo polar es el ángulo que forma el vector internuclear de la molécula (cuyo sentido apunta en la dirección C-O) con el eje perpendicular a la superficie (positivo hacia el vacío).

En la Fig. 5.7, se reportan las distribuciones del ángulo polar, calculadas a partir de las posiciones atómicas de cada configuración para DM con eventos de adsorción. En el panel (a), se observa la distribución de ángulos polares para las configuraciones iniciales (tiempo $t = 0ps$)

a $T_s = 0K$, y en el panel (b), aquellas para $t = 0ps$ y $T_s = 90K$. En los paneles inferiores se detallan las distribuciones de ángulos polares para las configuraciones finales (en el tiempo $t = 5ps$), para $T_s = 0K$ en el panel (c) y para $T_s = 90K$ en el panel (d); y en todos los casos para las energías de incidencia indicadas en la figura. Nuevamente los histogramas son normalizados acorde a la PBA de cada energía. Por lo que el área de cada histograma, es proporcional a la PBA de dicha energía (resultando mayor el área de la energía 0.01 eV y menor el de la energía 2.0 eV, por el carácter de sus PBA).

En primera instancia, se observa que no hay diferencias significativas entre los paneles en cuanto a las temperaturas superficiales de $T_s = 0K$ y de $T_s = 90K$. Este resultado es coherente con las conclusiones previamente mencionadas, en donde el comportamiento de los procesos de adsorción de la molécula de C-O sobre Cu(110), no varía para rangos de temperaturas entre 0 K y 120 K. Sin embargo, existen diferencias significativas entre los histogramas correspondientes a las configuraciones en $t=0$ ps (paneles superiores) con los de $t=5$ ps (paneles inferiores).

En $t=0$ ps, las distribuciones presentan formas que reproducen la función $sen(\star)$. La forma sinusoidal de la distribución, tiene su origen en las condiciones necesarias para realizar un muestreo uniforme de las coordenadas esféricas (que son las coordenadas implementadas para describir la orientación de la molécula C-O). Debido a que todas las orientaciones de la molécula en el espacio deben ser igualmente probables y, a que la coordenada que se barre uniformemente es la que corresponde al ángulo ϕ (ángulo azimutal), es preciso realizar un muestreo no uniforme del ángulo θ (ángulo polar) evitando favorecer alguna dirección particular involuntariamente (como ocurriría si ambos ángulos fueran muestreados uniformemente). Es decir, la forma $sen(\star)$ de los muestreos del ángulo polar, corresponden con las particularidades del sistema de coordenadas (esférico) utilizado y con el hecho de que se muestrea uniformemente el ángulo azimutal (Fig. 5.5).

Para energías de incidencia entre 0,01 eV \sim 0,40 eV, la amplitud máxima de la curva corresponde con el ángulo polar $\sim 90^\circ$, mientras que para energías ~ 1.25 eV el ángulo

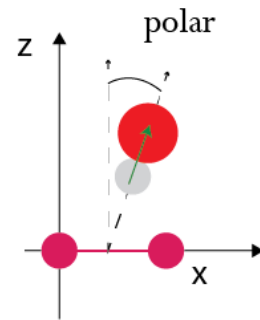


Figura 5.6 Esquema del ángulo polar.

La esfera roja (gris) representa al átomo de Oxígeno (Carbono). Las esferas rosas representan a los átomos de Cu.

correspondiente a la amplitud máxima disminuye ($a \sim 60^\circ$), y, casi se vuelve imperceptible para la energía de incidencia de 2.00eV. Este comportamiento se puede ver por que la amplitud máxima de las distribuciones para energías de incidencia 0.01 eV se encuentra alrededor de 90° mientras que aquella para energías mayores 1.25 eV se va desplazando hacia valores alrededor de 60° .

El hecho de que la amplitud de las distribuciones se desplace hacia la izquierda conforme la energía de incidencia aumenta, efectivamente se le adjudica a una propiedad física del sistema: Cuando la energía de incidencia aumenta, las fuerzas que tienden a re-orientar la molécula para ser adsorbida, comienzan a ser menos eficientes y se tiende a precisar que la trayectoria comience con el C hacia abajo (o a la misma altura que el O), para conseguir un evento de adsorción molecular. De lo contrario, la molécula se refleja debido a una interacción repulsiva con la superficie por tener el O hacia abajo. Lo que significa que al aumentar la energía de incidencia, las moléculas que mayormente tienden a ser adsorbidas son las que comienzan la DM con el C más cercano a la superficie. Esto genera que el ángulo polar adopte valores iniciales entre $0^\circ \leq \theta \leq 90^\circ$) conforme aumenta la energía de incidencia, como se puede observar por el desplazamiento hacia la izquierda de la amplitud máxima en los histogramas de los paneles superiores de la Fig. 5.7.

En los paneles inferiores ((c) y (d)) de la Fig. 5.7, correspondientes a $t=5$ ps (configuración final), al adsorberse la molécula, el ángulo polar adopta un valor preferencial, con mucha población de ángulos $\sim 30^\circ$. Este resultado se observa pues la amplitud máxima de todas las distribuciones, se halla ubicada en los alrededores del ángulo polar $\sim 30^\circ$, tanto para temperaturas superficiales $T_s = 0K$ como $T_s = 90K$ y en todas las energías de incidencia. Teniendo en mente la orientación de la molécula definida por la coordenada polar (Fig. 5.6), ángulos polares próximos a 30° indican que el átomo de C finaliza la trayectoria más cerca de la superficie que el O (para las trayectorias donde ha ocurrido adsorción molecular). Este resultado sugiere que la interacción molécula-superficie provoca una fuerza que tiende a re-orientar la molécula de manera que el C se encuentre más cerca de la superficie que el O, para poder ser adsorbida.

Para ilustrar esta característica del potencial de interacción, se confecciona la Fig. 5.8. En esta figura se utiliza la SEPN para predecir la energía potencial cuando el Z_{cm} de la molécula se acerca hacia la superficie, manteniendo la orientación de la molécula y la distancia internuclear

Distribuciones del ángulo polar

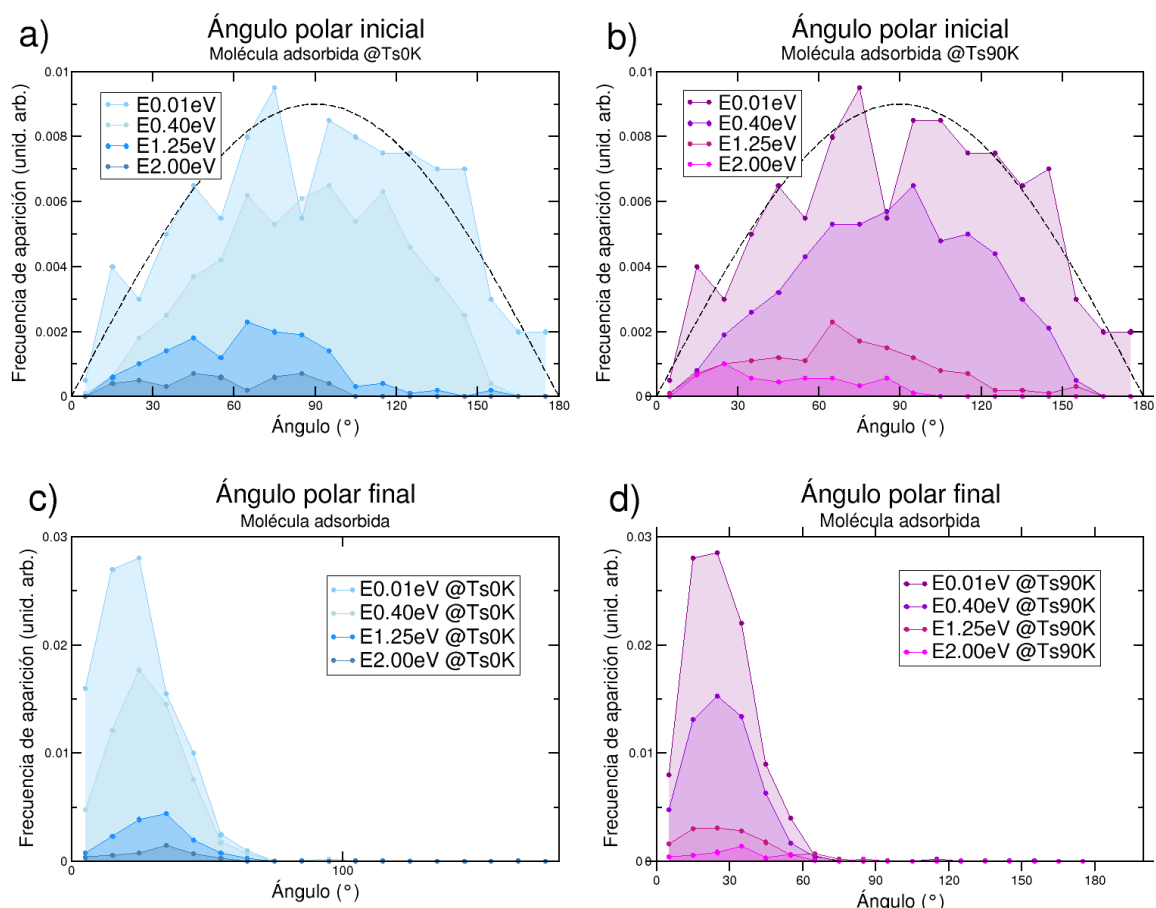


Figura 5.7 En los histogramas de la figura se observan las distribuciones del ángulo que forma el eje internuclear de la molécula con el eje perpendicular al plano de la superficie. Los distintos colores del histograma indican diferentes energías de incidencia. Las áreas de las curvas son proporcionales a las PBA de la energía y los colores azul (rosa) representan temperaturas superficiales 0 K (90 K)

fijas (en T0, T180 y T90P0 la primera y a la distancia de equilibrio la segunda) sobre los cuatro sitios característicos de la superficie (HW, TP, SB, LB).

Al observar esta figura, en el panel (a), se puede ver que cuando la molécula se aproxima a la superficie sobre un sitio *hollow*, disminuyendo la coordenada Z_{cm} hasta llegar a 4 Å, cualquier orientación de la molécula es equivalente para el potencial, pero al llegar a distancias menores de 4 Å, la energía es más baja cuando la molécula se orienta paralela a la superficie (T90P0). Esto provoca que cuando la molécula se acerca a la superficie con alguna inclinación, al acercarse a la distancia de 4 Å, se produzca un torque que intentará alinearla paralelamente a la superficie. Si se observa ahora el panel (b), que representa el sitio *top*, allí sucede algo similar: Hasta descender la molécula a 4 Å, la orientación de la molécula no es diferenciada, pero a distancias menores,

la curva que representa la orientación vertical con el C hacia abajo (T0) presenta un pozo de potencial significativamente más profundo que las restantes orientaciones. Esta característica del potencial, genera una re-orientación del ángulo polar de la molécula con orientaciones distintas a T0 (con el C hacia abajo) cuando la molécula desciende sobre el sitio TP. Una situación similar ocurre en el caso del sitio SB, como se observa en el panel (c) de la Fig. 5.7. Un pozo de potencial, aunque levemente menor, provoca un torque significativo sobre la orientación vertical de la molécula, que tiende a alinear el eje internuclear en la dirección vertical y con el C hacia abajo. El efecto de esta interacción con el potencial se traduce en valores del ángulo polar cercanos a $0^\circ \sim 30^\circ$, como puede observarse en los paneles inferiores de la Fig. 5.7.

Recapitulando, en la Fig. 5.8, se evidencia cómo, para distancias menores a 4 Å, la SEPN tiene una tendencia a generar un momento de fuerza sobre la molécula; que tiende a rotarla debido al pozo de potencial que existe para una orientación particular de ésta, cuando se acerca a la superficie, y se halla ubicada sobre un sitio *top* o *short bridge*. Este torque orienta a la molécula con el C por debajo del O (es decir el ángulo polar adopta valores entre $0^\circ \sim 90^\circ$) puesto que, como es bien sabido, cuando la energía potencial se aparta de la dirección de mínima energía, se genera una fuerza potencial que tiende a impulsar que el sistema adopte naturalmente la configuración de mínima energía.

A propósito de los resultados anteriores, se diagrama las posiciones finales de los átomos de la molécula C-O en las DM donde se ha reportado eventos de adsorción molecular, para una energía igual a 0.4 eV y para otra igual a 1.25 eV. Estos resultados se observan en la Fig. 5.9. En el eje vertical se muestra la coordenada z de cada uno de los átomos de la molécula medida a partir de la capa superficial de Cu, mientras que en el eje de abscisas se muestra el número de configuración correspondiente en la simulación.

En la Fig. 5.9 se observa una marcada población de moléculas adsorbidas para energías más bajas, y una disminución de la cantidad de moléculas adsorbidas para energías más altas. Debido a la amplia población de puntos en los paneles (a) y (c) (energía de incidencia 0.4 eV) en contraste con la menor población de puntos en los paneles (b) y (d) (energía de incidencia 1.25 eV). Este resultado concuerda con el hecho de que la PBA disminuye conforme aumenta la energía de incidencia.

En los paneles superiores se observa una marcada diferencia entre el comportamiento a baja energía (panel izquierdo) y a alta energía (panel derecho). Por un lado vemos que los átomos de C tienden a iniciar la DM preferentemente por debajo del átomo de O, para los eventos

Molécula sobre distintos sitios - Red CI3-3

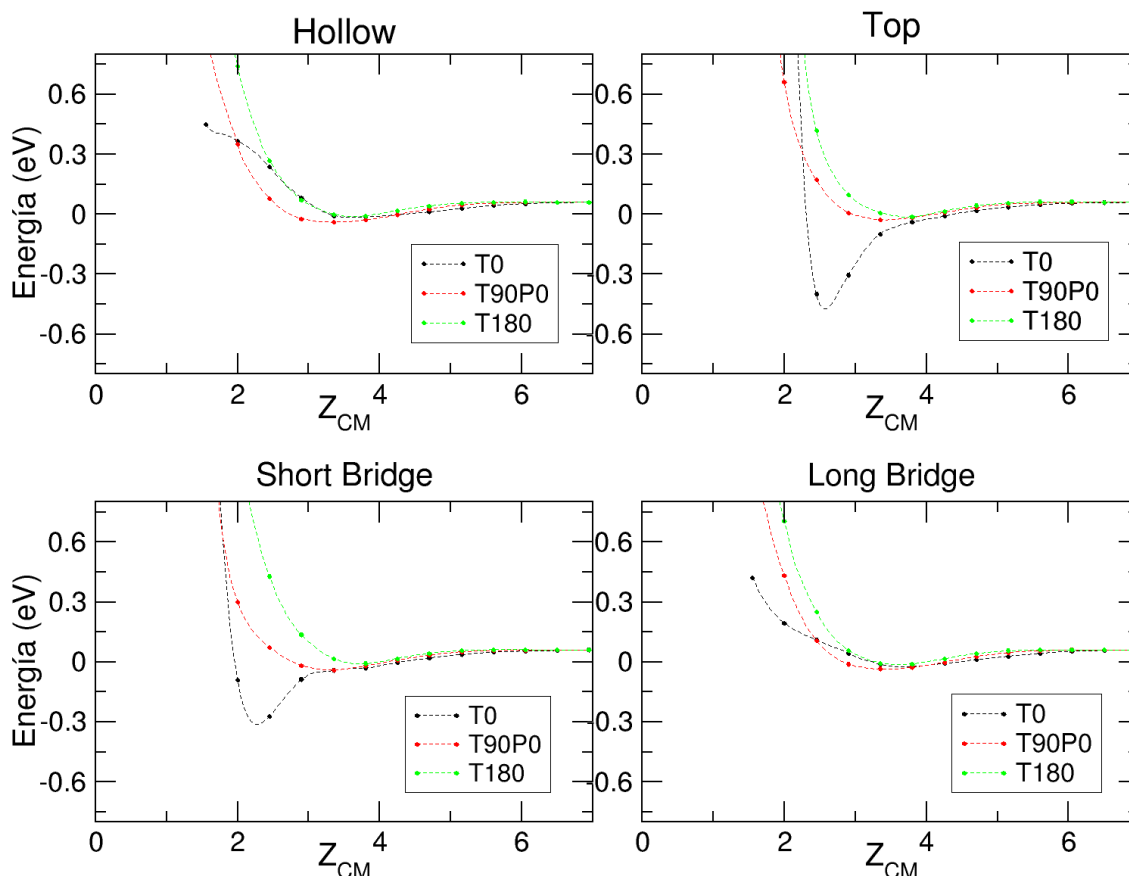


Figura 5.8 En cada cuadro de la figura se analiza la energía de interacción de la molécula cuando ésta interactúa con la superficie ubicada sobre los sitios HW (a), TP (b), SB (c) y LB (d). Cada curva representa una orientación particular de la molécula T0 (negro), T90P90 (rojo) y T180 (verde), cuyo eje internuclear se encuentra a la distancia de equilibrio. Los puntos representan valores calculados con DFT y las líneas de puntos son las energías calculadas con la red.

de adsorciones de alta energía (panel (b)); debido a que la componente Z de los átomos del O adopta preferentemente valores ~ 8.5 Å o superiores, mientras que la coordenada Z de los átomos de C se encuentra alrededor de ~ 8.0 Å. En contraste, las configuraciones iniciales de baja energía (panel (a)) presenta una gran cantidad de configuraciones en las que el átomo de C está por debajo del átomo de O así como también configuraciones en las que el átomo de C se encuentra por arriba del átomo de O, pues la coordenada Z de ambos átomos oscila uniformemente entre 8.0 Å y 8.5 Å. Este resultado concuerda con lo que se indicó previamente en los párrafos precedentes: las fuerzas potenciales que re-orientan a la molécula con el átomo de C hacia abajo pierden eficiencia conforme la molécula inicializa la DM con mayor energía cinética,

Posicionamiento de los átomos de C y O sobre la superficie de Cu(110)

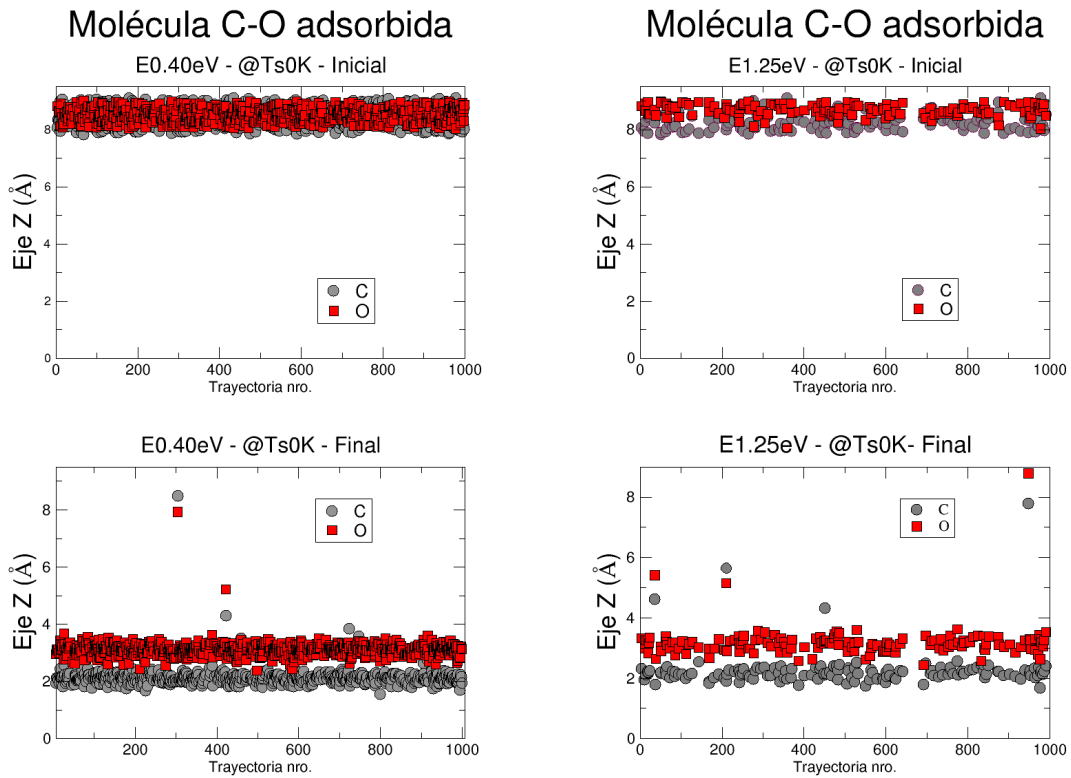


Figura 5.9 En las figuras se observan la DM a $t=0$ ps (paneles superiores) y a $t=5$ ps (paneles inferiores) de los átomos de carbono (oxígeno) representados con un punto gris (cuadrado rojo). La superficie de Cu se posiciona paralela al eje de las abscisas y en la coordenada $z = 0$ (pero no se representa explícitamente en la figura).

por lo que las configuraciones iniciales de adsorción con energías ~ 2.0 eV precisan comenzar la DM con el átomo de C preferentemente debajo del átomo de O.

En los paneles inferiores, se ve claramente como las moléculas adsorbidas tienden a reorientarse adoptando un estado final en donde el átomo de C se encuentre más próximo a la superficie de Cu (independientemente de la energía de incidencia), tal y como puede verse en los paneles inferiores de la Fig. 5.9. Al finalizar la trayectoria, cuando se registran eventos de adsorción, el átomo de C se encuentra siempre más cerca de la superficie, que el átomo de O.

Otro de los eventos que se aprecian entre estas figuras, son los eventos de desorción. Mediante la Fig. 5.9 puede apreciarse como algunas de las moléculas consideradas como adsorbidas en verdad se encuentran en un estado que fácilmente puede pensarse como “camino ” a ser consideradas reflejadas. A estas moléculas se las considera adsorbidas por las limitaciones del tiempo de integración finito que se implementa para integrar las trayectorias de DM. Esto sucede

pues el estado de *adsorción* no es un estado bien definido como pueden ser los estados de *reflexión* y el de *disociación*. En este caso particular las moléculas no se hallan en las proximidades de la superficie (la componente Z de los átomos de C y O se encuentra por encima del rango 2 Å a 4 Å) pero como el intervalo de integración límite, de 5ps, es finito, se interrumpe el proceso “camino a la desorción” de la molécula, generando que el código la considere como adsorbida sin ésta hallarse debidamente en las proximidades de la superficie, pero tampoco bajo las condiciones necesarias para ser considerada una molécula reflejada ni disociada. Sin embargo, el número de trayectorias en los que se da dicha situación es, en todos los casos, mucho más bajo que el número de eventos de adsorción molecular, y la variación de la PBA debida a conteo erróneo de estas trayectorias es despreciable.

5.3.5. Identidad de molécula

En esta sección, se muestra que la molécula aún conserva su identidad como tal al transcurrir los 5 ps de la DM donde se reportó un evento de adsorción. Para esto, se realiza un histograma con las distribuciones de distancias internucleares para todas las energías y se exponen dos de estas figuras como ejemplo en la Fig. 5.10.

Como puede observarse en la Fig. 5.10 los histogramas son similares tanto para diferentes energías como para los estados inicial y final de la molécula durante la DM con adsorción molecular. Esto se observa ya que los histogramas muestrean uniformemente las distancias internucleares (eje de las abscisas). Además, en todos los paneles, el rango de distancias internucleares es de 1.1 Å a 1.19 Å en el caso de las configuraciones iniciales, y de 1.1 Å a 1.3 Å en el caso de las configuraciones finales. En principio, este rango coincide con la distancia de equilibrio de la molécula de C-O (1.1412 Å en vacío), con lo cual los histogramas de esta figura indican que la molécula se encuentra siempre vibrando alrededor de su distancia de equilibrio. La diferencia entre los límites de 1.19 Å y 1.3 Å que se observa para el caso de configuraciones iniciales con configuraciones finales, se adjudica al hecho de que al estudiarse las moléculas que han sido adsorbidas, estas tienden a relajarse debido a la interacción con la superficie, permitiendo que la molécula adquiera distancias internucleares levemente mayores que la distancia de equilibrio en vacío. Con este análisis es posible observar que durante el momento $t=0$ ps y al momento $t=5$ ps, la molécula no ha perdido su identidad como tal, dado que para considerarse disociada los átomos de la molécula deberían alcanzar una separación de al menos 2 veces la distancia de equilibrio en vacío, es decir, $\sim 2,3$ Å.

Molécula adsorbida

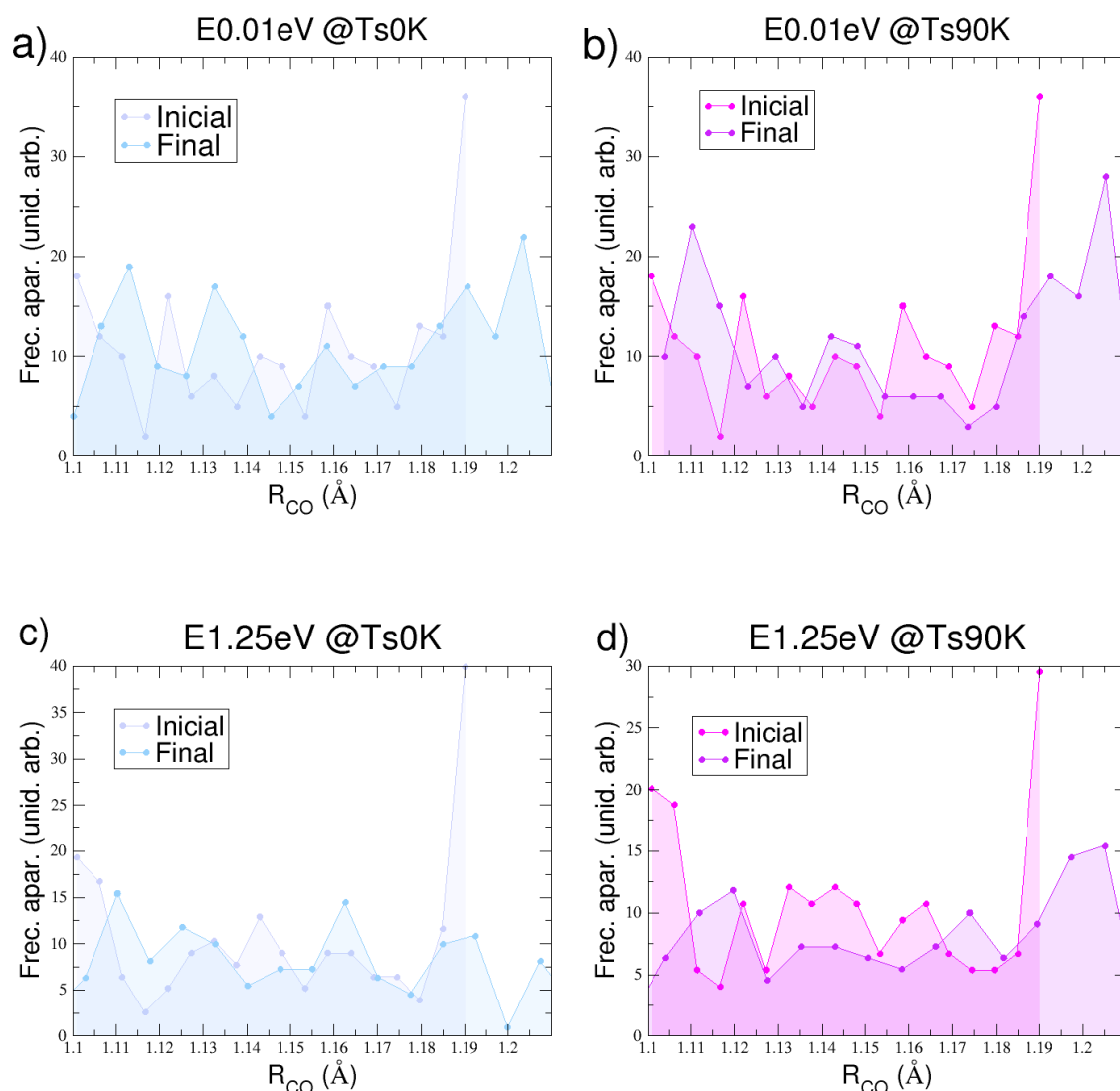


Figura 5.10 En la figura se observa la distribución de distancias del radio internuclear entre C-O para dos energías diferentes, representativas de las DM con baja energía de incidencia (paneles (a) y (b)) y con altas energías de incidencia (paneles (c) y (d)). Los colores de los histogramas hacen referencia a la temperatura a la cual se simuló la superficie de Cu, las figuras con tonalidad de azules (rosas) representa temperaturas superficiales de 0K (90 K).

5.3.6. Zonas de adsorción

Otro de los aspectos interesantes que se estudiaron fueron las zonas de adsorción de las moléculas de C-O sobre la superficie de Cu(110). En la Fig. 5.11 se observan dos ejemplos con energías de incidencia 0.4 eV y 1.25 eV de las posiciones del CM de la molécula en las DM reportadas con adsorción molecular, para $t=0$ ps y $t=5$ ps.

En los paneles (a) y (b), es decir para el comienzo de la DM a tiempo $t=0$ ps, se observa una distribución uniforme de los CM de la molécula sobre toda la superficie. Esto es así, pues los puntos rojos que describen las posiciones del CM se hayan distribuidos uniformemente sobre la región de puntos morados (átomos de Cu) independientemente de la energía de incidencia. Al finalizar la DM, para tiempos $t=5$ ps (paneles (c) y (d)), se puede observar como los CM de la molécula tienden a posicionarse en las zonas cercanas a los Cu de la superficie. Esta tendencia es más marcada para el caso de bajas energías ~ 0.4 eV de incidencia (panel (a)) que para energías más altas ~ 1.25 eV (panel (b)). La baja estadística con la que se cuenta para altas energías de incidencia (puesto que la PBA es menor) provoca que no exista una amplia cantidad de muestras para que las poblaciones de moléculas adsorbidos alrededor de los Cu sea marcadamente diferente a las demás regiones, sin embargo hay una leve tendencia a que los átomos que se adsorben se desplacen de los sitios HW, como puede observarse ya que en el panel (d) casi no hay puntos negros en dichos sitios. No obstante, en los paneles de baja energía de incidencia (paneles (a) y (c)), se observa marcadamente como las poblaciones más densas de moléculas adsorbidas se distribuyen en torno a los átomos de Cu, con algunas moléculas posicionadas en la zona de SB y casi ninguna en la zona de LB y HW. Lo cual concuerda con los sitios en que la energía es más estable (ver Fig. 5.8).

De estos resultados, se infiere que la molécula al adsorberse presenta tendencia a posicionarse cerca de los átomos de Cu, cuando la energía de incidencia es baja. Para energías de incidencia más altas, no es posible concluir si la zona de adsorción es más densa cerca de los Cu o si la molécula tiene menos restricciones en cuanto al sitio donde queda adsorbida, salvo por una leve tendencia a evadir los sitios HW.

A partir de este resultado, se deduce que, en el caso de bajas energías de incidencia, la molécula no posee la suficiente energía interna como para trasladarse a lo largo de la superficie, por lo que se ubica en los sitios donde existe mayor pozo de potencial; en cambio para las moléculas más energéticas adquiere la capacidad de permanecer próxima a la superficie pero con cierta movilidad en el plano paralelo a la superficie, por lo que las zonas de adsorción para altas energías de incidencia tienden a ser menos centralizadas que las obtenidas para bajas energías de incidencia.

CM de la molécula de C-O adsorbida

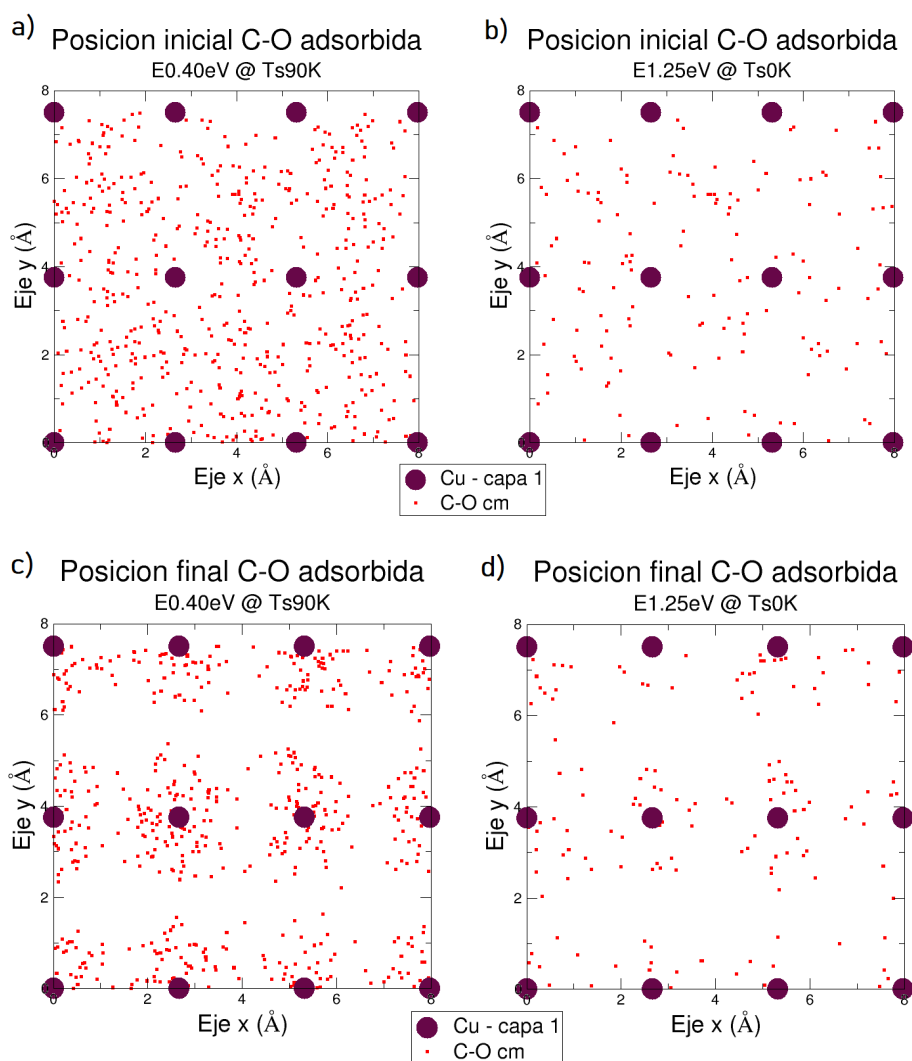


Figura 5.11 En la figura se observa la vista superior de la superficie de Cu(110) esquematizada por los círculos color púrpura, cuyo ordenamiento en el plano xy representa únicamente la primera capa de átomos. Los puntos rojos representan el CM de la molécula de C-O en $t=0$ ps (paneles (a) y (b)) y en $t=5$ ps (paneles (c) y (d)). Se analizan las zonas de adsorción en dos casos de energías de incidencia diferentes, 0.4 eV (izquierda) y 1.25 eV (derecha), para representar los comportamientos a bajas energías y a altas energías de incidencia, respectivamente.

5.3.7. Energía cinética de la molécula

En la Fig. 5.12 se ilustran las energías cinéticas que se calculan en las configuraciones iniciales ($t=0$ ps) y finales ($t=5$ ps) de los eventos de adsorción para algunas de las DM estudiadas.

En ellas se ve como la distribución de energía cinética comienza en el valor correspondiente a la energía de impacto asignada y el ancho que se observa se corresponde con la energía cinética de vibración, la cual (dependiendo la trayectoria) puede variar desde 0 hasta el total de la energía

vibracional usada (la *zero point energy* o *ZPE* en la figura). El alcance de la ZPE depende cuanta energía vibracional es cinética y cuanta es energía potencial (pero la suma de ambas siempre corresponde con el valor de ZPE).

Al final de la DM, es decir luego de los 5 ps, la distribución de energía cinética recorre el rango de 0 a ~ 0.5 eV, con un pico en las zonas cercanas al valor ~ 0.25 eV. Este resultado muestra claramente que las moléculas no han sido termalizadas pues, de lo contrario, el valor medio de la energía cinética debería ser:

$$2 \cdot \frac{3}{2} \cdot k_B \cdot 90K = 0,023eV \quad (5.2)$$

Como la amplitud máxima del histograma oscila alrededor de ~ 0.25 eV, no existe forma de que el valor medio de esta energía sea del orden de 0.023 eV, como indica que debiera ser el caso si las moléculas han logrado termalizarse por interactuar con la superficie. Esto significa que las moléculas que están adsorbidas molecularmente transcurridos 5 ps, aún no se han termalizado con los átomos de la superficie a través del intercambio de energía con las vibraciones de la superficie, que es lo que se tiene en cuenta en la simulación. Referido a este aspecto, se menciona el hecho de que, en este trabajo, no se han modelizado los intercambios de energía entre los átomos de la molécula y los electrones del metal, que en el caso de la adsorción molecular, significan un canal importante de intercambio de energía para la termalización. [21]

Distribución de energías cinéticas inicial y final de moléculas adsorbidas

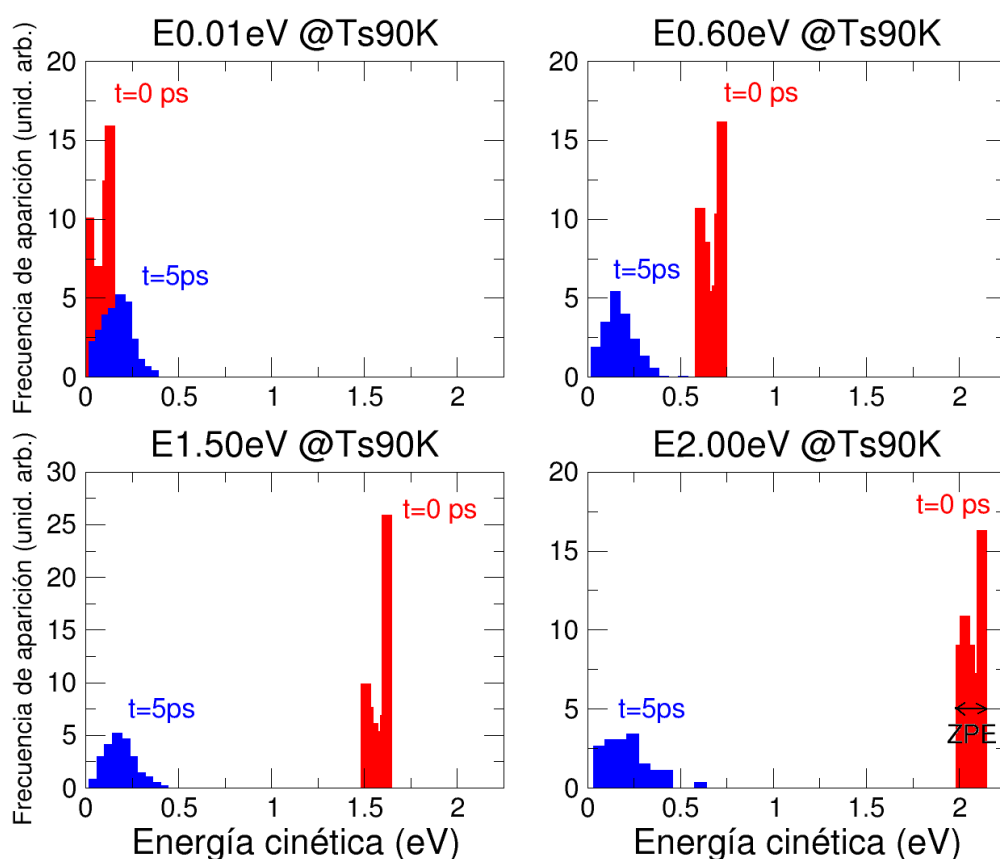


Figura 5.12 Los paneles superiores representan la distribución de energías cinéticas de la molécula C-O al comienzo de la DM ($t=0$ ps, histograma color rojo) y al finalizar la DM ($t=5$ ps, histograma color azul), para varias energías de incidencia diferentes: 0.01 eV (panel (a)), 0.6 eV (panel (b)), 1.5 eV (panel (c)) y 2.0 eV (panel (d)). Como la energía cinética de la molécula queda determinada por el valor ZPE y la energía de traslación de la molécula, el ancho de las distribuciones de energías cinéticas iniciales presenta un ancho igual al valor ZPE, como se indica en la figura.

Parte IV

Conclusiones

Conclusiones

La representación del potencial atómico de CO/Cu(110) mediante redes neuronales artificiales pre-alimentadas, ha retribuido una vasta cantidad de resultados satisfactorios. Las conclusiones de este trabajo se dividen en dos tópicos: una en cuanto a la metodología implementada y la otra en cuanto al análisis de las propiedades del sistema por medio de utilizar la metodología propuesta para redes neuronales como superficies de energía potencial atómica.

Redes neuronales artificiales

- Los parámetros de las FS que presentaron una mejor convergencia de la desviación estándar, tanto en el conjunto de entrenamiento como en el de evaluación, son los correspondientes a los Cuad 4.2 y 4.3.
- Sin embargo, no se encuentran marcadas diferencias entre los resultados predichos por redes cuya distinción se base en diferencias poco sustanciales de los parámetros espaciales de las FS (como ocurre entre los parámetros espaciales de los Cuad. 4.2 y 4.3). Por otro lado, se observó que las discrepancias de energía en donde las redes diferían más entre sí, no disminuyen significativamente conforme se incorporan más muestras (en dichas zonas) a la base de datos.
- La SEPN ajusta con mayor precisión los datos correspondientes a configuraciones de más bajas energías de incidencia ($0.01 \text{ eV} \sim 0.60 \text{ eV}$), mientras que a más altas energías el valor medio de las diferencias tiende a ser mayor. Por lo que es recomendable muestrear con más ejemplos en la base de datos de entrenamiento las zona del espacio de configuraciones donde se reporten mayores energías de incidencia.
- La red, en cuanto a los parámetros espaciales, que resultó ser la más óptima para ser implementada como representante de la SEP de interacción, fue la que corresponde con los parámetros H12 detallados en el Cuad. 4.2 y 4.3.

- De las propiedades físicas conocidas para el sistema CO/Cu(110) se concluye que el método de inteligencia artificial implementado en la construcción de la SEPN da resultados excelentes. La SEPN confeccionada reproduce con precisión los experimentos, la geometría del sistema y los procesos de intercambio de energía a un costo computacional menor que los métodos de AIMD. De esto es posible concluir que el método implementado es altamente eficiente para desarrollar potenciales de interacción atómicos similares al CO/Cu(110) para energías de incidencia en el rango 0.01 eV \sim 2.0 eV y para temperaturas de la superficie entre 0K y 90K.

Superficie de energía potencial CO/Cu(110)

- El comportamiento de la PBA obtenido a partir de las simulaciones realizadas utilizando las SEPN para CO/Cu(110), concuerda cualitativa y cuantitativamente con los resultados experimentales. Se presenta una disminución monótona de la PBA en el rango de energías desde 0.01 eV hasta 2.0 eV.
- Por otro lado la PBA no presenta variaciones apreciables para las temperaturas de la superficie consideradas, es decir 0K y 90K. Lo cual también concuerda con los resultados experimentales.
- La SEPN obtenida describe adecuadamente el comportamiento de la superficie a temperatura finita, dado que en las simulaciones a $T_s=90\text{K}$ se cumple con el principio de equipartición de la energía y los átomos se desplazan de forma isotrópica en las tres direcciones.
- Se observa que la adsorción de la molécula de CO sobre la superficie Cu(110) es no disociativa.
- Las moléculas de CO se adsorben sobre la superficie preferentemente sobre los sitios top (TP) y en menor medida sobre sitios *short-bridge* (SB).
- Las moléculas adsorbidas presentan una orientación preferencial en cuanto al ángulo polar, $\theta \lesssim 30^\circ$, es decir, que las moléculas se adsorben con el átomo de C más cercano a la superficie que el átomo de O.
- Las moléculas adsorbidas no presentan una orientación preferencial en cuanto al ángulo azimutal.

- La interacción molécula superficie presenta un mecanismo de *steering* que reorienta las moléculas para favorecer la adsorción, el cual es más eficiente a bajas energías de incidencia.
- El tiempo de integración total de 5 ps ha demostrado no ser suficiente para la termalización del sistema. Por otro lado la cantidad de eventos en los que la molécula se ha desorbido sin alcanzar las condiciones de reflexión es despreciable, por lo cual no se esperan cambios significativos en la PBA al aumentar el tiempo de simulación. Aún así, se sugiere que en estudios posteriores se integre durante un tiempo suficiente para alcanzar la termalización, lo cual permitirá estimar coeficientes de difusión de CO sobre Cu(110) así como la tasa de intercambio de energía entre la molécula y la superficie.

Finalmente, los alcances de la implementación de algoritmos de inteligencia artificial para representar potenciales atomísticos aparenta un futuro prometedor debido a la eficiencia y precisión con la que fueron desarrollados los cálculos y a las propiedades del sistema obtenidas.

Bibliografía

- [1] D. Marx y J. Hutter. “Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods”. En: *Cambridge University Press, New York, USA* (2009).
- [2] Stefano Piana; Alpan Raval et al. “Assessment of the utility of contact-based restraints in accelerating the prediction of protein structure using molecular dynamics simulations”. En: *Protein Science* 25 (2016), págs. 19-29.
- [3] N. Lu y M. Khorshidi. “Mechanisms for Soil-Water Retention and Hysteresis at High Suction Range”. En: *J. Geotech. Geoenviron. Eng* 8 141 (2015).
- [4] Dr. Kai Yan; Dr. Tuhina Adit Maark; Alireza Khorshidi et al. “The Influence of Elastic Strain on Catalytic Activity in the Hydrogen Evolution Reaction”. En: *Angewandte Chemie* 55 6175-6181 (2016).
- [5] J. C. Greer; R. Ahlrichs; e I. V. Hertel. “Proton - transfer in ammonia cluster cations – Molecular - dynamics in a self-consistent field”. En: *Zeitschrift fur Physik D* 18:413–426 (1991).
- [6] R. Car y M. Parrinello. “Unified Approach for Molecular Dynamics and Density-Functional Theory”. En: *Phys. Rev. Lett.* 55, 2471 (1985).
- [7] J. Behler y M. Parrinello. “Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces”. En: *Physical Review Letters* 98 (2007).
- [8] S. Lorenz; A. Groß y M. Schefflera. “Representing high-dimensional potential-energy surfaces for reactions at surfaces by neural networks”. En: *Chemical Physics Letters* 395 (2004), págs. 210-215.
- [9] Marcos del Cueto; Xueyao Zhou; Linsen Zhou et al. “New Perspectives on CO₂-Pt(111) Interaction with a HighDimensional Neural Network Potential Energy Surface”. En: *J. Phys. Chem. C* 10.1021/acs.jpcc.9b10883 (2020).

- [10] Jun Li Bin Jiang y Hua Guo. “High-Fidelity Potential Energy Surfaces for Gas Phase and Gas-Surface Scattering Processes from Machine Learning”. En: *J. Phys. Chem.* 10.1021/acs.jpcclett.0c00989 (2020).
- [11] G Seminara; A Martinez y H F Busnengo. “(Trabajo en preparación)”. En: () .
- [12] Simon Haykin. *Neural Networks and Learning Machines*. Vol. 3er Ed. 2009.
- [13] J. Behler. “First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems”. En: *Angewandte Chemie* 56 (2017), págs. 12828-12840.
- [14] J. Behler. “Constructing High-Dimensional Neural Network Potentials:A Tutorial Review”. En: *International Journal of Quantum Chemistry* 115 (2015), págs. 1032-1050.
- [15] J. Behler. “Atom-centered symmetry functions for constructing high-dimensional neural network potentials”. En: *J. Chem. Phys* 134 (2011).
- [16] J. Behler. “Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations”. En: *Phys. Chem. Chem. Phys.* 13 (2011), págs. 17930-17955.
- [17] P. Hohenberg y W. Kohn. “Inhomogeneous Electron Gas”. En: *Phys. Rev.* 136 (1964), B864.
- [18] W. Kohn y L. J. Sham. “Self-Consistent Equations Including Exchange and Correlation Effects”. En: *Phys. Rev.* 140 (1965).
- [19] Brian Cowan. “Topics in Statistical Mechanics”. En: *Imperial College Press* (2005), p. 37.
- [20] M. Kunat; Th Becker.; Ch. Boas et al. “Adsorption dynamics of CO on Cu(110): a molecular beam study”. En: *Surface Science* 474 (2001), págs. 114-128.
- [21] J.I. Juaristi; R. Díez Muiño; M. Blanco-Rey et al. “Electronic Friction Dominates Hydrogen Hot-Atom Relaxation on Pd(100)”. En: *Phys. Rev. Lett.* 112 (2014).