

Universidad de La Habana
Facultad de Matemática y Computación



Análisis Computacional de Coros de *Eleutherodactylus eileenae*: Inferencia de Redes desde Señales Bioacústicas

Autor: Daniel Machado Pérez

Tutores:

Dr. Roberto Mulet Genicio

Dr. Milton García Borroto

Trabajo de Diploma
presentado en opción al título de
Licenciado en Ciencia de la Computación



La Habana,
mayo 2025

github.com/DanielMPMatCom/Thesis.git

*"La utopía está en el horizonte
Camino dos pasos y se aleja dos pasos.
Camino diez y el horizonte se corre diez pasos más allá.
Entonces, ¿para qué sirve la utopía?
Para eso, sirve para caminar."*

Fernando Birri

Agradecimientos

El autor agradece al Dr. Roberto Alonso Bosch y la M.Sc. Mariam Curbelo Cruz por facilitar los datos de campo, compartir su amplia experiencia sobre el comportamiento de *Eleutherodactylus eileenae*, orientar la revisión de la literatura especializada y la revisión del manuscrito.

Resumen

Abstract

Índice general

Introducción	9
0.1. Motivación y justificación	10
0.2. Formulación del problema	11
0.3. Hipótesis	13
0.4. Objetivos	13
0.5. Propuesta de solución	14
0.6. Estructura del Documento	15
1. Antecedentes	16
1.1. Detección de individuos en grabaciones	17
1.1.1. Métodos clásicos de clasificación de señales acústicas	17
1.1.2. Localización y asignación de fuentes	18
1.2. Análisis de coros y comportamiento grupal	18
1.2.1. Patrones de formación de coros	19
1.2.2. Ecoacústica de coros	19
1.3. Modelos de interacción y causalidad	20
1.3.1. Modelos estadísticos de redes	20
1.3.2. Técnicas de inferencia de causalidad	21
1.4. Antecedentes específicos en <i>E. eileenae</i>	21
2. Métodos	23
2.1. Representación de audios con Mel-espectrogramas	25
2.2. Sincronización de los audios	26
2.3. Filtrado de ruido de fondo	28
2.4. Detección y asignación heurística de cantos de Colines a Mi- crófonos	29
2.4.1. Algoritmo basado en hipótesis de energías relativas .	30
2.4.2. Algoritmo basado en hipótesis de comportamiento .	31
2.5. Modelado de interacciones con el modelo de Ising	31

2.5.1. Fundamentos del modelo de Ising	32
2.5.2. Formulación del sistema acústico como red de espines	33
2.5.3. Inferencia por máxima verosimilitud y descenso por gradiente	33
3. Resultados e Interpretación	36
3.1. Evaluación de la sincronización	36
3.2. Exactitud de la detección y asignación micrófono-Colín . . .	36
3.3. Comparación entre algoritmos heurísticos	36
3.4. Análisis de las interacciones inferidas $\{J_{ij}\}$	36
3.5. Idoneidad del modelo de Ising	36
Conclusiones	37
Recomendaciones	38
Referencias Bibliográficas	39

Índice de figuras

1.	Ejemplar de <i>Eleutherodactylus eileenae</i>	10
2.	Distribución Geográfica de los Micrófonos.	12
2.1.	Mel-espectrograma de un fragmento del audio 20231021_190000h, donde se observan las notas “CO” (baja frecuencia) y “LIN” (alta frecuencia).	26
2.2.	Energía temporal del audio 20231021_190000h.	27
2.3.	Momentos de pico de energía de un fragmento de los audios del 20231021_190000 antes de sincronizar.	27
2.4.	Momentos de pico de energía de un fragmento de los audios del 20231021_190000 después de sincronizar.	28
2.5.	Mel-espectrograma del fragmento de audio en que aparece el ruido de un motor de automóvil.	29
2.6.	Esquema de ejecución del Algoritmo basado en energías relativas.	32

Introducción

Las técnicas de monitorización acústica automatizada se han convertido en una herramienta esencial para el estudio de comunidades de animales en su medio natural. En particular, el análisis del canto de los anuros ofrece información valiosa sobre sus estrategias de apareamiento, la dinámica de sus poblaciones y la salud de los ecosistemas en los que habitan [5]. La rana cubana *Eleutherodactylus eileenae* (Dunn, 1929), endémica de Cuba, constituye un modelo idóneo para investigar estos procesos, pues sus machos forman coros nocturnos cuyos patrones de interacción pueden reflejar aspectos tanto ecológicos como de comportamiento social [2].

E. eileenae, conocido comúnmente como “Colín”, es una especie perteneciente a la familia *Eleutherodactylidae*. Su distribución abarca desde la península de Guanahacabibes, en Pinar del Río, hasta la Sierra de Najasa, en Camagüey, ocupando diversos hábitats como bosques semideciduos, pinares y cafetales, hasta altitudes de aproximadamente 900 metros sobre el nivel del mar [2, 36, 17]. Durante el día, estos anuros se refugian en oquedades de rocas, troncos, hojarasca y bromelias; al anochecer, los machos ascienden a ramas y hojas, hasta alturas de 3 metros, para emitir sus vocalizaciones.

El canto característico de *E. eileenae* consta de dos notas diferenciadas: una primera “Co”, breve y de frecuencia menor, seguida de una segunda “Lin”, más prolongada y de frecuencia ligeramente superior. Esta estructura ha motivado el uso del término “Colín” para referirse tanto a cada canto individual como, coloquialmente, a la especie misma. Estas llamadas desempeñan un papel crucial en la comunicación intraespecífica, mediando aspectos relacionados con el estado fisiológico y la competencia territorial [2]. Además, se ha observado que los machos forman agrupaciones o “coros”, sincronizando sus cantos para aumentar la eficacia en la atracción de hembras y reducir el riesgo de depredación.

La presente investigación desarrolla un flujo computacional automatizado para procesar grabaciones de los cantos de los Colines y, a partir

de él, comenzar a analizar las interacciones acústicas asociadas a estos coros. El objetivo es comprender las estrategias de comunicación que emplea *E. eileenae* en su entorno natural. En la Figura 1 se muestra un ejemplar de esta especie, y en la Guía sonora de los anfibios de Cuba ¹ se puede encontrar una breve grabación de su canto y otros datos relevantes.



Figura 1: Ejemplar de *Eleutherodactylus eileenae*.

0.1. Motivación y justificación

La ecoacústica ha evolucionado desde la grabación pasiva hasta flujos computacionales que automatizan la detección de llamadas en ambientes ruidosos [1, 5]. Sin embargo, como resultado de la revisión bibliográfica realizada, no se conocen estudios en Cuba que hayan integrado estas técnicas con modelos de interacción y causalidad.

¹Grabación disponible en: https://www.fonozoo.com/fnz_detalle_registro_amphibia.php?id=97953&tipo_registro=1

El análisis manual de grabaciones de campo resulta lento, costoso y propenso a errores humanos, lo que limita la escala y reproducibilidad de los estudios bioacústicos. En trabajos previos sobre *E. eileenae*, los investigadores describieron manualmente las etapas de vocalización y realizaron muestreos puntuales de la actividad acústica y trófica de los machos en la Sierra del Rosario [2]. Sin embargo, no se abordó la sincronización automática de múltiples micrófonos ni la clasificación sistemática de cada llamado, pasos indispensables para escalar el análisis a largas series temporales y diferentes localidades.

Además, los coros de machos implican interacciones acústicas cuya dinámica no se comprende completamente: ¿cómo influyen la densidad de individuos, la configuración espacial de los micrófonos y el ruido ambiental en la estructura del coro? ¿Qué reglas simples de interacción subyacen en la sincronización de los “Colines”? En este contexto, un enfoque computacional reproducible y basado en algoritmos heurísticos permitiría procesar de manera eficiente decenas de horas de grabaciones, discriminando eventos de interés (cantos de Colines) de otros sonidos (otra fauna, viento, tráfico) y garantizando resultados comparables entre campañas de muestreo.

Los resultados de la presente tesis constituyen un primer paso hacia el desarrollo de un flujo completo que abarque, desde la adquisición y sincronización de señales, hasta la clasificación automática de cantos y el modelado de interacciones acústicas, a fin de superar las limitaciones de los métodos manuales y aportar herramientas robustas para la ecoacústica de anfibios en Cuba y regiones semejantes.

0.2. Formulación del problema

El presente trabajo se ocupa del procesamiento automático de un conjunto de grabaciones registradas en la Reserva de la Biosfera “Sierra del Rosario” con el objetivo de caracterizar los cantos de machos de *Eleutherodactylus eileenae* y reconstruir la dinámica de sus coros. Estas grabaciones, realizadas con nueve micrófonos omnidireccionales dispuestos en un área de aproximadamente 20 m de radio, abarcan ciclos de 58 minutos de registro seguidos de 2 minutos de descarga, durante tres noches consecutivas entre las 18:00 y las 06:00 horas. El entorno de captura se ve afectado por ruido de lluvia, viento, tráfico y la superposición de cantos de varios individuos, lo que plantea retos importantes tanto en la sincronización de señales como en la discriminación de los eventos acústicos relevantes frente a emi-

siones no deseadas. En la Figura 2 se muestra la distribución geográfica de los micrófonos en el área de estudio.

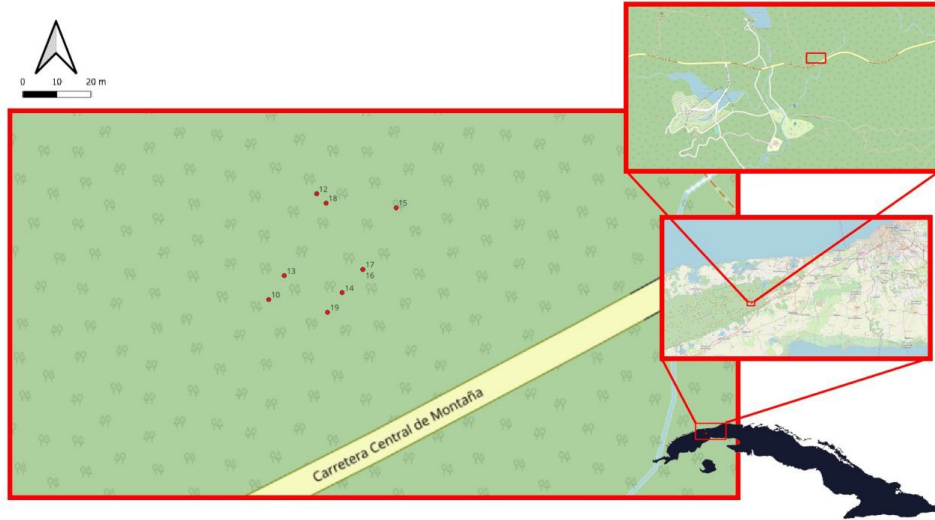


Figura 2: Distribución Geográfica de los Micrófonos.

En este ámbito, el problema central consiste en el diseño de un flujo heurístico capaz de sincronizar con precisión las nueve pistas de audio, detectar de manera fiable los instantes de emisión de cada canto (“Colín”) y asignar cada evento al micrófono (y por tanto al individuo) correspondiente; a partir de ello, se extraen parámetros temporales y espectrales de cada llamado para generar un registro estructurado que sirve de base a la inferencia de una red de interacción inspirada en el modelo de Ising [23]. Se espera que la sincronización automática alcance una resolución temporal tal que permita distinguir llamadas superpuestas de machos vecinos, y que la combinación de criterios de proximidad y similitud espectral garantice una alta exactitud en la etiqueta de cada canto, validada mediante muestreo manual. Finalmente, las secuencias temporales resultantes se emplean para estimar los parámetros de interacción J_{ij} , de modo que el modelo reproduzca con fidelidad las correlaciones de co-emisión observadas en el coro.

La complejidad del problema reside en la superposición de señales y en la variabilidad del nivel de ruido ambiental, lo cual exige emplear técnicas robustas de filtrado. Este trabajo opta por heurísticas de intensidad relativa

y correlación espectral para asignar cada “Colín” al micrófono y al individuo correspondiente, garantizando la reproducibilidad de los resultados y la comparabilidad con futuros estudios bioacústicos. Este estudio es el primero en Cuba en aplicar el modelo de Ising para describir interacciones acústicas en coros de anuros, siguiendo aproximaciones similares en redes neuronales y colonias de insectos [34, 27, 4, 32, 23].

0.3. Hipótesis

Se plantea como hipótesis principal que un flujo computacional basado en algoritmos heurísticos para la sincronización, detección y clasificación de cantos de Colines podrá reproducir con un alto grado de fidelidad las secuencias de emisión de cada individuo y, a partir de ellas, inferir las interacciones acústicas subyacentes a la formación de coros.

Esto fundamenta la viabilidad de un enfoque computacional replicable y escalable que, aplicado a los cantos de Colines, aporte nuevas perspectivas cuantitativas sobre la dinámica de los coros y sirva como base para futuros estudios comparativos en ecoacústica.

0.4. Objetivos

El objetivo general de esta tesis es desarrollar y validar un flujo heurístico que, mediante la detección y clasificación de los cantos de *E. eileenae*, permita reconstruir la dinámica de sus coros y cuantificar su sincronización y causalidad. Para ello se plantean los siguientes objetivos específicos:

1. Construir un *dataset* limpio y sincronizado:

- Sincronizar las nueve pistas de audio utilizando referencias comunes (picos de energía) y correlación cruzada. [14]
- Eliminar ruidos de fondo (lluvia, viento, tráfico) e interferencias no deseadas.

2. Detectar y asignar cada canto a su emisor:

- Desarrollar heurísticas de proximidad basadas en la intensidad relativa en cada micrófono.
- Validar la asignación mediante muestreo manual de segmentos y cálculo de exactitud.

- Evaluar la consistencia de los algoritmos de detección y asignación.
- Comparar el rendimiento de los algoritmos diseñados.

3. Inferir un modelo de interacciones acústicas:

- Formular la red de interacciones como un modelo de Ising con parámetros J_{ij} que cuantifiquen la co-emisión.
- Estimar los J_{ij} mediante el Principio de Máxima Verosimilitud [19] y algoritmos de descenso por gradiente.
- Evaluar la idoneidad del modelo de Ising mediante una comparación con el modelo Independiente en cuanto a su capacidad para predecir los patrones de cantos.

0.5. Propuesta de solución

Para dar respuesta al problema planteado, se propone un flujo de trabajo modular que integre cuatro etapas principales: sincronización, filtrado, asignación de emisores y modelado de interacciones. En primer lugar, la sincronización de las nueve pistas de audio se realiza mediante la aplicación de correlación cruzada sobre los histogramas de distribución de distancias temporales entre los picos de energía de los audios y un punto de referencia. A continuación, se aplica un filtro que consiste en una poda espectral basada en el percentil 99.9 de la amplitud por frecuencia. Con ello se busca atenuar ruido de baja (viento, tráfico) y alta frecuencia (insectos).

En la tercera fase se asigna cada llamado al micrófono más adecuado mediante heurísticas de intensidad relativa. Finalmente, la información temporal y espectral extraída de cada evento sirve para construir un grafo de interacciones, modelado con un enfoque de máxima verosimilitud en un sistema análogo al modelo de Ising. El ajuste de los parámetros J_{ij} se lleva a cabo mediante un algoritmo de descenso por gradiente.

Este flujo, además de automatizar completamente la extracción de datos, permite validar la hipótesis mediante métricas objetivas de precisión y reproducibilidad, y ofrece una herramienta extensible a otros sistemas de ecoacústica con configuraciones similares.

0.6. Estructura del Documento

Para explicar la solución propuesta y el flujo de trabajo, se diseñó la siguiente estructura del presente documento. En el Capítulo 1 se revisan los antecedentes y actualidad en cuanto a detección y clasificación de llamadas bioacústicas, con especial énfasis en métodos de sincronización y filtrado de señales. El Capítulo 2 detalla la metodología empleada: adquisición y preprocesamiento de las grabaciones, tanto el diseño de algoritmos heurísticos para la detección de los “Colines” como la asignación de cada canto a su emisor, y la formulación del modelo de interacciones basado en Ising con inferencia por máxima verosimilitud y descenso por gradiente. En el Capítulo 3 se presentan los resultados: evaluación de la precisión de sincronización, rendimiento de los detectores de canto, exactitud de la asignación micrófono-Colín y análisis de las interacciones inferidas $\{J_{ij}\}$. Además en las Conclusiones se discuten las aportaciones principales, las implicaciones ecológicas y computacionales y las limitaciones del estudio. Finalmente, en las Recomendaciones se presentarán sugerencias generales basadas en los resultados, destinadas a orientar mejoras metodológicas, proponer nuevas líneas de investigación y facilitar la aplicación práctica del enfoque desarrollado en estudios futuros.

Capítulo 1

Antecedentes

Las primeras grabaciones sistemáticas de sonidos de vida silvestre se realizaron en mayo de 1929, cuando Arthur A. Allen y Peter P. Kellogg registraron por primera vez el canto de especies neárticas en un parque de Ithaca, Nueva York [13]. Este hito marcó el nacimiento de la colección que hoy conocemos como Macaulay Library, el mayor archivo multimedia de fauna del mundo, con más de 2.6 millones de grabaciones de audio y cobertura de casi todas las especies de aves, así como registros de insectos, peces, anfibios y mamíferos [12].

Durante décadas, la bioacústica de campo dependió de hardware analógico y procesamiento manual de cintas, lo cual limitaba el volumen de datos y la reproducibilidad de los estudios [5]. Con la transición a sistemas digitales a principios de los años 2000, la monitorización acústica pasiva (PAM) emergió como herramienta clave en ecología y conservación, permitiendo grabaciones continuas y remotas que facilitan el muestreo a gran escala sin perturbar el comportamiento natural de los organismos [1].

Las organizaciones de conservación, entre ellas WWF, impulsaron la adopción de PAM en áreas protegidas, destacando la necesidad de protocolos robustos de preprocesamiento para atenuar ruido ambiental (como lluvia, viento y tráfico) y mejorar la relación señal-ruido [5]. Paralelamente, el desarrollo de índices acústicos automatizados, introducidos por Towsey et al. (2014), permitió estimar la riqueza de especies en paisajes sonoros complejos mediante métricas computables directamente de las grabaciones [43].

La creación de grandes conjuntos de datos de referencia, como AnuraSet para llamadas de anuros neotropicales, estandarizó la evaluación de algoritmos de identificación en entornos ruidosos, facilitando la compara-

ción de metodologías y la reproducibilidad de los resultados [9]. Revisiones como la de Stowell (2022) han descrito las oportunidades y desafíos de incorporar redes neuronales profundas en bioacústica, estableciendo una hoja de ruta para futuros avances computacionales [39].

En el ámbito del modelado de interacciones, Ōta et al. (2020) cuantificaron los mecanismos de sincronización en coros de ranas a partir de características dinámicas de las llamadas, demostrando la aplicabilidad de enfoques de red para estudiar la co-emisión y la causalidad acústica [28]. Asimismo, Kalra (2024) examinó la percepción de señales en ambientes complejos, proporcionando métodos de localización pasiva que mejoran la atribución de fuentes en PAM [25].

Estos avances tecnológicos y metodológicos forman el fundamento de la presente investigación, que adapta flujos computacionales y modelos estadísticos de interacción al estudio de los coros de *Eleutherodactylus eileenae* en su entorno natural cubano.

1.1. Detección de individuos en grabaciones

La detección precisa de emisores individuales en grabaciones de campo constituye un paso esencial para cualquier estudio bioacústico que aspire a caracterizar patrones de comportamiento a nivel de individuos. Este proceso implica la identificación de eventos acústicos relevantes en señales, la discriminación de llamadas frente a ruido ambiental y, cuando es posible, la atribución de cada emisión a su emisor de origen. A continuación, se revisan primero los métodos clásicos de clasificación de señales basados en análisis espectral y extracción de características, y luego se exponen enfoques de localización pasiva y heurísticas de asignación que permiten estimar la procedencia de las llamadas.

1.1.1. Métodos clásicos de clasificación de señales acústicas

La clasificación de vocalizaciones animales en grabaciones se fundamenta en la extracción de características espectrales a partir de la Transformada Rápida de Fourier (FFT) [11]. Sobre esta base, la conversión a coeficientes cepstrales en frecuencia mel (MFCC) ha demostrado ser especialmente eficaz para representar texturas acústicas complejas, tal como lo estableció Mermelstein (1976) en el ámbito de voz humana y luego adaptado a bioacústica [26]. Zhang et al. (2021) emplearon la descomposición de espectrogramas mel y la fusión de modelos para clasificar escenas acústi-

cas, alcanzando mejoras significativas en la precisión de clasificación [49]. Xie et al. (2016) extendieron este enfoque integrando detección de eventos acústicos con aprendizaje multi-etiqueta sobre espectrogramas, lo que permitió discriminar actividades de canto de ruido de lluvia con resultados prometedores [46]. Estos métodos clásicos forman la columna vertebral de los flujos heurísticos de detección, siendo especialmente adecuados cuando solo se dispone de una única pista de audio.

1.1.2. Localización y asignación de fuentes

Aunque la verdadera localización tridimensional de emisores requiere múltiples canales sincronizados, diversos autores han explorado técnicas de triangulación y diferencias de tiempo de llegada (TDOA) para estimar direcciones de origen. Spiesberger y Fristrup (1990) detallaron un método de TDOA que calcula desplazamientos de fase entre pares de micrófonos para inferir la dirección de la fuente sonora [37]. Cuando solo se cuenta con grabaciones monofónicas, se han diseñado heurísticas basadas en la intensidad relativa de la señal y la similitud espectral. Kalra (2024) demostró que, al comparar la energía de una llamada en momentos consecutivos y su correlación espectral con plantillas predefinidas, es posible asignar eventos a individuos cercanos con una precisión alta [25]. Estas heurísticas aprovechan la atenuación del sonido en el espacio y la coherencia espectral para discriminar entre vocalizaciones de distintos machos dentro de una misma grabación, constituyendo una alternativa pragmática cuando no se dispone de redes de micrófonos múltiples.

1.2. Análisis de coros y comportamiento grupal

El estudio de los coros de anuros combina observaciones de campo con análisis cuantitativos para desentrañar cómo las interacciones individuales dan lugar a patrones colectivos de canto. Este enfoque implica medir la distribución espacial y temporal de las vocalizaciones, identificar picos de actividad sincronizada y describir las reglas de proximidad y fase que los machos siguen al unirse al coro. A continuación se revisan los principales hallazgos sobre la formación de coros y las herramientas ecoacústicas desarrolladas para caracterizar su estructura y dinámica.

1.2.1. Patrones de formación de coros

En *Eleutherodactylus coqui*, Woolbright (1985) registró que los individuos ascienden al anochecer a perchas de la vegetación, donde emiten series regulares de “co-qui” alternadas con períodos de silencio, y observó que la densidad de cantores influye en la cercanía entre perchas, reduciéndose la distancia media en noches de alta actividad [45]. Stewart y Pough (1983) encontraron en la misma especie que la localización de los machos en el espacio de canto está relacionada con la disponibilidad de refugios y la competencia territorial, ajustando su posición en respuesta a la presencia de vecinos [38]. Drewry y Rand (1983) describieron patrones similares en *E. coqui* en Puerto Rico, donde la sincronización de picos vocales ocurre en respuesta a variaciones de luz y temperatura, con crestas al anochecer y al amanecer [15].

En el caso de *Eleutherodactylus eileenae* en Cuba, Alonso et al. (2001) analizaron la actividad acústica de machos cantores y revelaron dos picos de emisión al anochecer y al amanecer, así como desplazamientos verticales en la vegetación que enfatizan la dimensión espacial de la formación de coros [2]. Estos hallazgos concuerdan con los patrones descritos en otras especies de *Eleutherodactylus*, donde la competencia acústica y las condiciones microambientales determinan la estructura del coro [42]. Más recientemente, Òta et al. (2020) cuantificaron dinámicamente las interacciones entre pares de individuos mediante análisis de series temporales de llamadas, demostrando que algunos machos actúan como “iniciadores” de pulsos colectivos, un comportamiento que puede describirse con modelos de red de interacción [28].

1.2.2. Ecoacústica de coros

La ecoacústica de coros se apoya en el uso de índices acústicos y herramientas computacionales para detectar y caracterizar la actividad colectiva sin necesidad de seguimiento visual. Boelman et al. (2007) introdujeron el Bioacoustic Index (BI), calculado como la energía acumulada en la banda de frecuencias de interés [6], mientras que Pieretti et al. (2011) complementaron este enfoque con el Acoustic Complexity Index (ACI), que mide variaciones espectrales dentro de ventanas temporales para estimar la riqueza y la intensidad coral, mostrando alta correlación con el número de machos activos en grabaciones de bosques tropicales, y demostraron que ambos índices permiten resaltar los períodos de mayor sincronización en coros de anfibios [30]. Además, Farina (2018) estableció un marco cuantitativo para

la ecoacústica que enfatiza la integración de índices acústicos en estudios de dinámica comunitaria, subrayando su utilidad para evaluar patrones de sincronización en coros de anuros y otros sistemas biológicos [18].

En 2021, Gan et al. propusieron un método de reconocimiento de coros basado en espectrogramas falsos en color y aprendizaje automático, logrando identificar patrones de emisión colectiva con una precisión del 77% en grabaciones de campo de especies tropicales [20]. Brodie et al. (2020) evaluaron la eficacia de múltiples índices acústicos y confirmaron que la combinación de ACI, BI y Net Spectrum ofrece una representación robusta de la estructura coral a lo largo de ciclos nocturnos [7].

1.3. Modelos de interacción y causalidad

En ciencias biológicas y ecológicas, comprender cómo las interacciones locales entre individuos dan lugar a fenómenos colectivos constituye un desafío central. Los modelos estadísticos de redes, en particular las formulaciones de máxima entropía equivalentes al modelo de Ising, ofrecen un marco potente para describir la sincronización y la causalidad en sistemas de canto animal. En esta sección se revisan primero las bases y aplicaciones de los modelos de Ising en redes biológicas, y luego se presentan las principales técnicas de inferencia de causalidad adaptadas a series temporales acústicas.

1.3.1. Modelos estadísticos de redes

El modelo de Ising, originado en física estadística para describir ferromagnetismo, se reformuló como un problema de máxima entropía para ajustar distribuciones observadas de estados binarios con correlaciones dadas. Schneidman et al. (2006) demostraron que un modelo de Ising con interacciones par a par reproduce con alta fidelidad las correlaciones de actividad en poblaciones de neuronas retinianas, sin necesidad de términos de orden superior [34]. Mora y Bialek (2011) ampliaron esta perspectiva, argumentando que sistemas biológicos (desde redes neuronales hasta colonias de insectos) operan cerca de puntos críticos de fase, un rasgo que maximiza la capacidad de información y la sensibilidad a perturbaciones externas [27].

Para inferir los parámetros de interacción J_{ij} en sistemas fuera de equilibrio, Zeng et al. (2013) desarrollaron un método de máxima verosimilitud para modelos de Ising asíncronos, derivando reglas de aprendizaje que

dependen únicamente de las correlaciones temporales de los datos [48]. Chau Nguyen et al. (2017) revisaron un amplio espectro de métodos de inferencia inversa (incluyendo pseudoverosimilitud y aproximaciones de mensaje) que escalan eficientemente a grandes redes biológicas, destacando aplicaciones en neurociencia y ecología [10]. Estos enfoques han permitido reconstruir estructuras de interacción tanto en poblaciones de neuronas (Tkačik et al., 2006) como en sistemas de canto colectivo, donde las “conexiones” cuantifican la probabilidad de co-emisión entre pares de individuos [41].

1.3.2. Técnicas de inferencia de causalidad

Más allá de la correlación, establecer relaciones causales en series temporales acústicas requiere herramientas específicas. El enfoque clásico de Granger (1969) define la causalidad en términos predictivos: una serie X *cause*^a otra Y si el pasado de X mejora la predicción de Y más allá de lo que hace el propio pasado de Y [21]. Porta y Faes (2015) ampliaron este marco en el contexto de sistemas fisiológicos y redes neuronales, describiendo cómo la causalidad de Wiener-Granger puede aplicarse a señales biológicas con metodologías que consideran tanto ruido como dinámicas no lineales [31].

Para capturar relaciones más complejas, se han empleado medidas de información direccional como la entropía de transferencia (Schreiber, 2000), que cuantifica flujos de información sin asumir linealidad [35]. Más recientemente, el método de Convergent Cross Mapping (Sugihara et al., 2012) ha mostrado robustez al inferir causalidad en sistemas dinámicos de alta dimensionalidad, al reconstruir espacios de fase conjuntos y medir la habilidad de uno de los vectores de recuperación del otro [40]. Estas técnicas, que serán utilizadas en trabajos futuros, proporcionan un conjunto de herramientas complementarias para analizar la causalidad acústica en coros de *E. eileenae*, permitiendo distinguir interacciones directas de sincronizaciones inducidas por factores ambientales comunes.

1.4. Antecedentes específicos en *E. eileenae*

La investigación acerca de la biología y ecología de *Eleutherodactylus eileenae* arrancó fundamentalmente con estudios taxonómicos y de distribución. Schwartz (1958) describió por primera vez la especie en la región occidental de Cuba, señalando su afinidad con el subgénero *Eleutherodacty-*

lus auriculatus y detallando caracteres morfológicos diagnósticos [36]. Posteriormente, Estrada (1984, 1994) amplió el conocimiento de su rango geográfico, documentando nuevas localidades en Sancti Spíritus y otros municipios centrales, así como preferencias de hábitat que incluyen bosques semideciduos, pinares y cafetales hasta los 900 m s. n. m. [16, 17]. Estas obras sentaron las bases para comprender la distribución espacial de *E. eileenae* y su adaptación a distintos sustratos, como bromelias y hojarasca.

En el plano de la bioacústica y la ecología trófica, Alonso et al. (2001) llevaron a cabo muestreos en la Reserva de la Biosfera “Sierra del Rosario”, donde cuantificaron patrones diel y estacionales de vocalización de los machos y relacionaron la actividad acústica con la alimentación mediante análisis de contenido estomacal [2]. Encontraron dos picos de canto, al anochecer y al amanecer, coincidentes con desplazamientos verticales en la vegetación, y demostraron que durante los períodos de depresión vocal los machos se alimentan, consumiendo principalmente himenópteros y arácnidos. Este estudio fue el primero en integrar datos acústicos y tróficos de *E. eileenae*, estableciendo el marco experimental y metodológico que la presente tesis extiende mediante flujos computacionales automatizados y modelado de interacciones acústicas.

Capítulo 2

Métodos

El canto de *E. eileenae* ocurre durante la noche, cuando los machos se agrupan para vocalizar, con picos de actividad en los meses cálidos y lluviosos del año, y en las horas de la tarde-noche y la madrugada. Los datos empleados en esta investigación fueron obtenidos por un grupo de investigación liderado por el Dr. Roberto Alonso, de la Facultad de Biología de la Universidad de La Habana, mediante un conjunto de grabaciones de campo realizadas en la Reserva de la Biosfera “Sierra del Rosario”. Luego de identificar ciertos individuos de Colín y los sitios en los que habitualmente se ubicaban para emitir sus cantos, se colocaron nueve micrófonos (cada uno a aproximadamente un metro de su individuo más cercano) cuya activación remota permitió registrar la información acústica del coro. Todos los dispositivos, del mismo modelo y omnidireccionales, fueron activados simultáneamente para grabar durante 58 minutos consecutivos, seguidos por un intervalo de 2 minutos destinado a la descarga y almacenamiento de los datos adquiridos. Este ciclo se repitió cada hora, desde las 18:00 hasta las 06:00 del día siguiente, durante tres noches consecutivas, entre el 20 y el 23 de octubre de 2023. Como resultado, se conformó un *dataset* estructurado en el que, por cada hora de grabación, se dispone de los nueve archivos correspondientes a los nueve micrófonos utilizados. Tal como se indicó previamente, los dispositivos fueron ubicados dentro de un área de aproximadamente 20 metros de radio.

Debido a que los micrófonos eran omnidireccionales y estaban situados relativamente cercanos entre sí, en cada uno se registraron no solamente los cantos del individuo de Colín más próximo, sino también cantos de especímenes cercanos, sonidos del ambiente natural de la zona, e incluso ruidos de agentes artificiales como autos pasando por una carretera aledaña.

Esto planteó una serie de dificultades para el análisis de los datos, como la de eliminar la información ruidosa, y la de distinguir y extraer en cada audio los cantos del Colín que corresponden al micrófono en cuestión.

De esta forma, se impuso emplear un método automático para “limpiar” los audios. Además, para poder obtener las secuencias de cantos y estudiar los coros, surgió la necesidad de diseñar procesos para discriminar correctamente estos datos, y obtener una asignación Micrófono-Colín. También se debía verificar la correcta sincronización de las grabaciones, pues a pesar de la activación remota y simultánea, el posible error de *hardware* podría comprometer la precisión de los resultados. En este capítulo se describen los métodos utilizados para resolver tanto estos problemas como el de modelar las interacciones del sistema.

Con el objetivo de simplificar las explicaciones, se decidió asignar una letra entre la “a” y la “i” a cada uno de los micrófonos, de modo que la correspondencia entre la notación original y la nueva qued de la siguiente forma:

- Micrófono 10: a
- Micrófono 12: b¹
- Micrófono 13: c
- Micrófono 14: d
- Micrófono 15: e
- Micrófono 16: f
- Micrófono 17: g
- Micrófono 18: h
- Micrófono 19: i

Por lo tanto, una grabación en específico, por ejemplo la del 21 de octubre de 2023 a las 19:00 horas, registrada por el micrófono “h”, será referida como “20231021_190000h”.

Para la grabación se emplearon dispositivos AudioMoth, desarrollados como placas de circuito impreso de 58x48x18mm (80g con baterías), basados en un microcontrolador ARM Cortex-M4F con unidad de punto flotante y un micrófono MEMS bottom-ported. Cada unidad muestreó audio

¹En los datos originales no existe micrófono 11.

sin compresión WAV a velocidades de hasta 384kHz, almacenando los datos en tarjetas microSD de hasta 32GB. Gracias a su excepcional eficiencia energética (con consumos de 10-25mW durante la adquisición y 80μW en reposo), resultaron idóneos para despliegues prolongados alimentados por tres pilas AA, manteniendo la sincronización y calidad requeridas para el estudio de los coros de *E.eileenae* [22].

2.1. Representación de audios con Mel-espectrogramas

Para el análisis de las grabaciones se transformó cada archivo de audio en su correspondiente Mel-espectrograma, una representación tiempo-frecuencia que aproxima la percepción auditiva humana mediante la escala mel [49]. Un Mel-espectrograma se obtiene aplicando una ventana deslizante al audio, calculando la FFT en cada segmento, mapeando las bandas de frecuencia a la escala mel y, finalmente, tomando el logaritmo de la energía resultante [49]. Esta representación enfatiza las variaciones espectrales relevantes para el canto de *E.eileenae* y facilita la detección de los pulsos “CO” y “LIN” en las grabaciones.

Para la generación de los Mel-espectrogramas se emplearon los siguientes parámetros, optimizados para cubrir el rango de frecuencias de los cantos de Colín (aprox. 1.8-3.1 kHz) [3] y garantizar resolución temporal suficiente para distinguir emisiones consecutivas:

- `hop_length` = 512 (desplazamiento de ventana)
- `n_fft` = 6096 (número de puntos para calcular la FFT)
- `n_mels` = 128 (número de bandas mel a utilizar)
- `f_min` = 1600 Hz (frecuencia mínima mapeada)
- `f_max` = 4096 Hz (frecuencia máxima mapeada)
- `sample_rate` = 96000 Hz (tasa de muestreo de las grabaciones)

En la Figura 2.1 se muestra un Mel-espectrograma de un fragmento del audio 20231021_190000h. En él pueden identificarse claramente los instantes de emisión de la nota baja “CO” (fragmento en torno a la 20 banda de mel) y de la nota alta “LIN” (bandas intensas por encima de la 60).

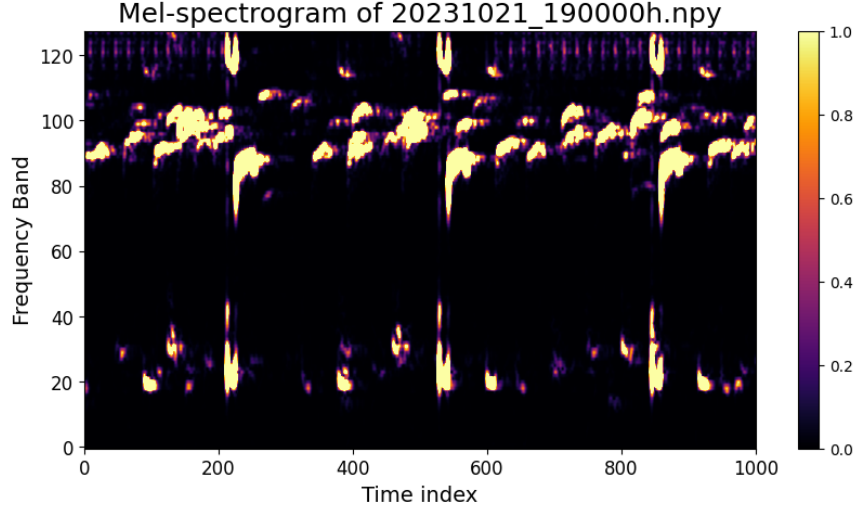


Figura 2.1: Mel-espectrograma de un fragmento del audio 20231021_190000h, donde se observan las notas “CO” (baja frecuencia) y “LIN” (alta frecuencia).

2.2. Sincronización de los audios

La sincronización de las grabaciones de audio constituyó una etapa crítica para asegurar la coherencia temporal necesaria en el análisis posterior de los cantos de *E. eileenae*. Aunque los nueve dispositivos de grabación fueron activados simultáneamente, cada uno presentaba desfases mínimos, posiblemente debido a pequeñas variaciones en el hardware y el error de la activación remota, los cuales resultaban significativos para el procesamiento posterior. Para corregir estos desfases, se implementó un procedimiento automático basado en señales de energía derivadas de los Mel-espectrogramas de cada canal.

En primer lugar, se calculó la energía temporal de cada Mel-espectrograma, obteniendo en cada instante j , la suma de los cuadrados de las intensidades en las n bandas de frecuencia:

$$E_j = \sum_{i=0}^n a_{ij}^2 \quad \forall j, \quad (2.1)$$

donde a_{ij} es la intensidad en la banda i y el tiempo j , y E_j representa la energía total en dicho instante [24]. En la Figura 2.2 se aprecia la evolución de E_j para el canal “20231021_190000h”.

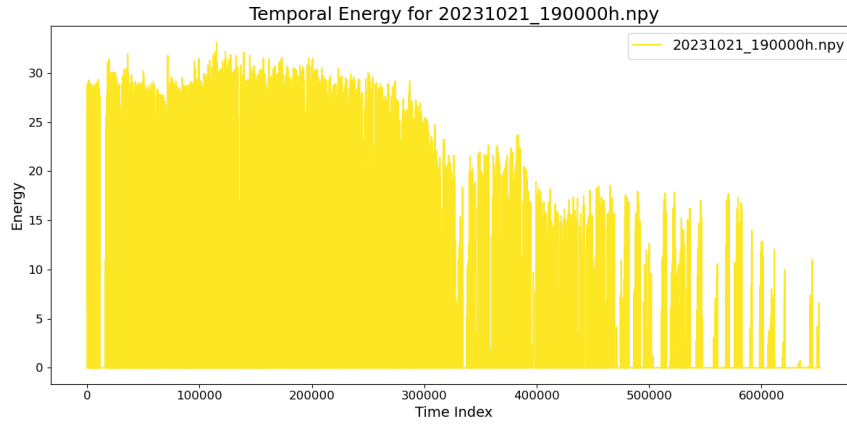


Figura 2.2: Energía temporal del audio 20231021_190000h.

A continuación, se detectaron los picos locales de E_j mediante una función que localiza máximos relativos en una secuencia discreta, ajustando parámetros de prominencia y distancia mínima para evitar falsas detecciones debidas a ruido residual [44]. En la Figura 2.3 se muestran los momentos de pico de energía de un fragmento del audio 20231021_190000h.

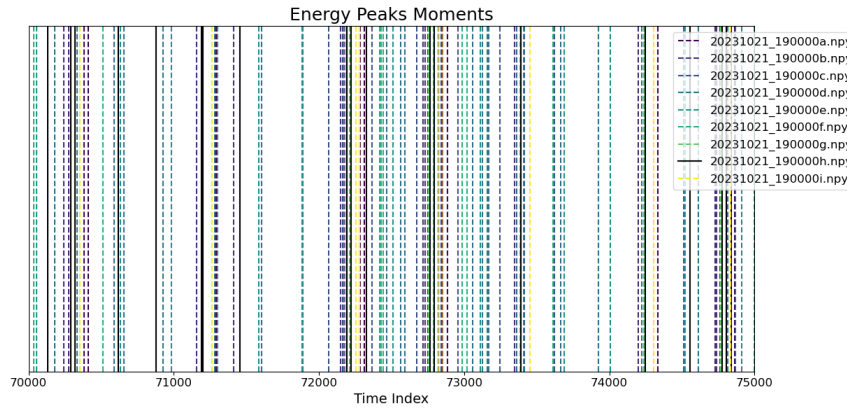


Figura 2.3: Momentos de pico de energía de un fragmento de los audios del 20231021_190000 antes de sincronizar.

Para estimar los desfases entre canales, se construyó un histograma de los tiempos de pico de cada audio, relativo al inicio de la grabación. Luego, se aplicó correlación cruzada [14] entre los histogramas de pares de canales,

desplazando uno respecto al otro y calculando el coeficiente de correlación máxima para cada posible *lag*. Este enfoque permitió identificar el desfase óptimo que maximiza el solapamiento de las estructuras temporales de canto.

Una vez determinados los desfases relativos, se alinearon todos los canales utilizando un canal de referencia global (en este caso, el canal “20231021_190000h”, que presentaba la señal más limpia) aplicando un desplazamiento de muestras correspondiente al *lag* calculado. La Figura 2.4 ilustra los picos de energía tras la alineación, mostrando una coincidencia temporal precisa. Finalmente, estos desfases se incorporaron a los archivos originales, de modo que todos los audios quedaran sincronizados con un error mínimo.

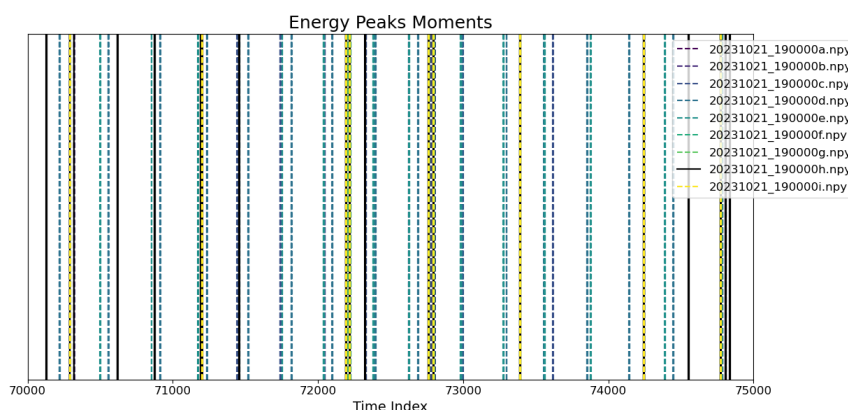


Figura 2.4: Momentos de pico de energía de un fragmento de los audios del 20231021_190000 después de sincronizar.

2.3. Filtrado de ruido de fondo

Para garantizar que las características acústicas extraídas representaran de manera fidedigna la emisión vocal de los Colines, se implementó una etapa previa de eliminación de ruido dirigida a atenuar tanto interferencias puntuales (por ejemplo, vocalizaciones de otras especies o clics mecánicos) como ruidos persistentes de banda ancha (motores de vehículos, canto de grillos, viento). En la Figura 2.5 se muestra el Mel-espectrograma de un fragmento en el que se registró el ruido generado por el paso de un automóvil cercano.

La técnica empleada consistió en una poda espectral basada en el per-

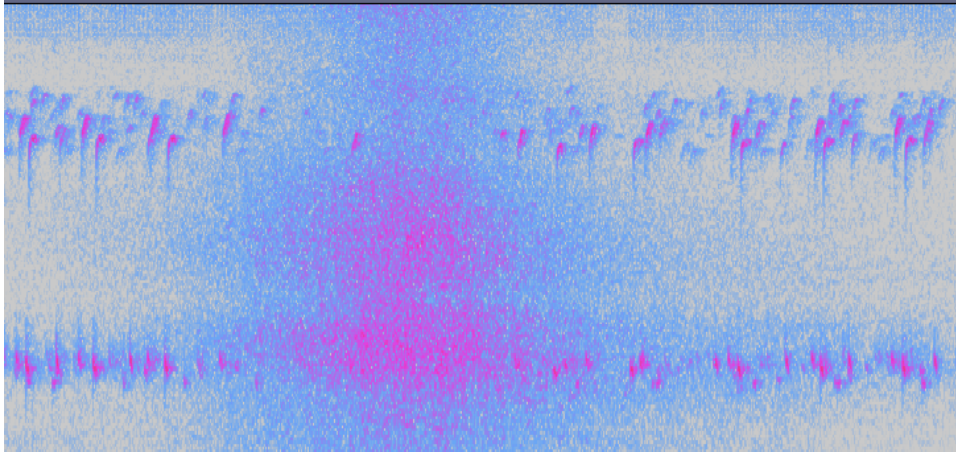


Figura 2.5: Mel-espectrograma del fragmento de audio en que aparece el ruido de un motor de automóvil.

centil 99.9 de la amplitud por frecuencia. Para cada grabación se calculó el espectrograma de potencia y, en cada banda de frecuencia, se determinó el valor correspondiente al percentil 99.9 de la amplitud a lo largo de todo el período analizado. A continuación, se aplicó un umbral dinámico que descartó todas las componentes espectrales por debajo de ese nivel, reteniendo solo los eventos acústicos cuyas intensidades superaron el fondo de ruido común.

Este procedimiento permitió suprimir eficazmente las componentes de baja energía asociadas al ruido ambiental no informativo y conservar las estructuras espectrales dominantes de los cantos de *E. eileenae*. De esta forma, el flujo de detección heurística se alimentó de señales con un cociente señal-ruido significativamente mejorado, requisito indispensable para la posterior identificación precisa de los instantes de emisión de los “Colines” y la asignación a cada micrófono.

2.4. Detección y asignación heurística de cantos de Colines a Micrófonos

Para obtener un *dataset* con el que fuera posible llevar a cabo los estudios de comportamiento de los coros, se hizo imprescindible diseñar procesos automáticos capaces de extraer, a partir de las grabaciones sincroniza-

das y filtradas, las series temporales de emisión de cada Colín, asignando correctamente cada canto a su respectivo micrófono. Solo disponiendo de estas secuencias individuales se puede avanzar hacia un análisis riguroso de las interacciones acústicas y las dinámicas de grupo, ya que cualquier estudio de causalidad o modelado de red requiere un sistema de eventos bien definido y libre de ambigüedades en la identificación de los emisores. En esta sección se presentan dos algoritmos heurísticos que, partiendo de Mel-espectrogramas, automatizaron la detección de instantes de canto y su asignación espacial, constituyendo la base metodológica para el posterior estudio de las redes de interacción en el coro de *E. eileenae*.

2.4.1. Algoritmo basado en hipótesis de energías relativas

El método propuesto se sustenta en la suposición de que para un micrófono, los cantos del individuo más cercano deben ser registrados con las mayores energías relativas a dicho micrófono. Por lo tanto, si en cada uno se identifican los momentos de mayor energía y se elimina la información de esos cantos de las demás grabaciones, se obtendrán por separado los datos que se buscan. Para explotar esta idea se dividió cada grabación en dos bandas de interés (baja frecuencia ("CO") y alta frecuencia ("LIN")) y se procesaron en paralelo. En cada banda se computaron nuevamente las energías temporales con la fórmula (2.1). A partir de estos vectores de energía, el algoritmo aplicó la siguiente estrategia:

1. Se creó un *array* de ceros para cada micrófono, que almacenaría la energía del canto asignado en cada instante.
2. Mientras existieran valores no nulos en los vectores de energía originales:
 - Se seleccionó aleatoriamente uno de los nueve canales. (Esta selección aleatoria permite que si dos cantos pertenecientes a individuos distintos se multiplexan en tiempo, se guarde la información de ambos en el *dataset* final.)
 - Se halló la posición j^* de la energía máxima en ese canal.
 - Se guardó el valor de dicha energía en la posición j^* del *array* de resultados para ese micrófono.
 - Se estableció a cero el valor de la energía en la posición j^* de todos los vectores, eliminando así la contribución de ese canto para las demás grabaciones.

2.5. MODELADO DE INTERACCIONES CON EL MODELO DE ISING31

3. El proceso se repitió hasta que todos los vectores de energía quedaron vacíos.
4. Finalmente, se almacenaron los *arrays* resultantes como secuencias de cantos por micrófono.

Este procedimiento, ilustrado en el pseudocódigo del Algoritmo 1 y en la Figura 2.6, garantizó que cada instante de canto fuera asignado al micrófono cuya energía fuera dominante. La complejidad temporal resultó $O(n^2)$, con n igual al número de muestras en el eje temporal del Mel-espectrograma.

Algorithm 1 Colin Sequence by Energy

Require: Temporal energies $\{E^{(k)}\}_{k=1}^9$, length N

Ensure: Calls sequences $\{S^{(k)}\}_{k=1}^9$

```

1: Initialize  $S^{(k)} \leftarrow [0, \dots, 0]$  for  $k = 1, \dots, 9$ 
2: while  $\exists k, j \mid E_j^{(k)} > 0$  do
3:   Select randomly  $k \in \{1, \dots, 9\}$ 
4:    $j^* \leftarrow \arg \max_j E_j^{(k)}$ 
5:    $S_{j^*}^{(k)} \leftarrow E_{j^*}^{(k)}$  ▷ Save maximum value to file
6:   for each  $k' = 1, \dots, 9$  do ▷ Erase value from all arrays
7:      $E_{j^*}^{(k')} \leftarrow 0$ 
8:   end for
9: end while
10: return  $\{S^{(k)}\}$ 

```

2.4.2. Algoritmo basado en hipótesis de comportamiento

2.5. Modelado de interacciones con el modelo de Ising

El análisis de las dinámicas colectivas de los Colines requiere un marco formal capaz de capturar las dependencias entre los individuos a partir de observaciones de su comportamiento acústico. En esta sección se describe el uso del modelo de Ising como herramienta para representar y cuantificar las interacciones entre machos vocalizantes, siguiendo un enfoque basado en inferencia estadística y física computacional.

2.5. MODELADO DE INTERACCIONES CON EL MODELO DE ISING32

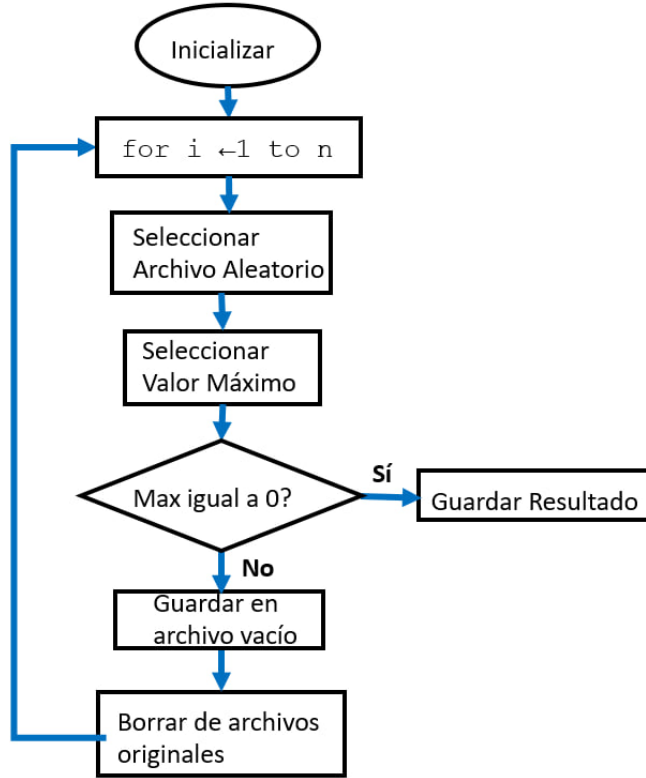


Figura 2.6: Esquema de ejecución del Algoritmo basado en energías relativas.

2.5.1. Fundamentos del modelo de Ising

El modelo de Ising fue concebido por Wilhelm Lenz e introducido por Ernst Ising en 1925 como una simplificación del comportamiento colectivo de espines en materiales ferromagnéticos [23]. En su formulación clásica, cada componente del sistema está representado por una variable binaria $\sigma_i \in \{-1, +1\}$, que puede interpretarse como un estado de activación (por ejemplo, presencia o ausencia de vocalización). Las interacciones entre pares de componentes se codifican mediante parámetros J_{ij} , que determinan si las unidades tienden a sincronizarse o a inhibirse mutuamente, mientras que un parámetro h_i refleja la predisposición individual (o sesgo) del componente i a estar activo.

El conjunto de estados del sistema está dado por el vector binario $\sigma = (\sigma_1, \dots, \sigma_N)$, y la probabilidad de observar una configuración determinada

2.5. MODELADO DE INTERACCIONES CON EL MODELO DE ISING ³³

se modela según la distribución de Boltzmann:

$$P(\sigma) = \frac{1}{Z} \exp \left(\sum_{i < j} J_{ij} \sigma_i \sigma_j + \sum_i h_i \sigma_i \right),$$

donde Z es la función de partición que normaliza la distribución:

$$Z = \sum_{\sigma} \exp \left(\sum_{i < j} J_{ij} \sigma_i \sigma_j + \sum_i h_i \sigma_i \right).$$

Esta formulación permite representar redes de interacciones con estructuras arbitrarias, haciendo del modelo de Ising una herramienta versátil que ha sido empleada exitosamente en áreas como neurociencia, genética, economía y análisis de redes sociales [10].

2.5.2. Formulación del sistema acústico como red de espines

Para modelar los coros de *E. eileenae*, se asumió que cada uno de los nueve individuos registrados corresponde a un nodo del sistema, y su comportamiento acústico en un instante dado queda representado por un espín $\sigma_i \in \{-1, +1\}$, donde $+1$ indica que el individuo está emitiendo un canto y -1 que permanece en silencio.

Las configuraciones σ fueron extraídas a partir de las secuencias temporales de detección previamente construidas, dividiendo el tiempo en intervalos discretos (frames) de longitud fija. Cada configuración refleja el estado coral del sistema en un instante, y el conjunto de configuraciones observado se utilizó como base para inferir los parámetros $\{J_{ij}\}$ y $\{h_i\}$ que describen la estructura de interacción y la propensión individual a cantar.

Aunque el sistema de cantos no necesariamente satisface los criterios de equilibrio termodinámico, se adoptó esta suposición como primera aproximación, lo cual permitió el uso de herramientas formales del marco de física estadística [48]. Esta simplificación habilita una descripción cuantitativa de las correlaciones observadas, útil para identificar influencias directas entre individuos del coro.

2.5.3. Inferencia por máxima verosimilitud y descenso por gradiente

La tarea de inferir los parámetros del modelo a partir de observaciones consiste en encontrar los valores de J_{ij} y h_i que maximizan la verosimilitud del conjunto de configuraciones observadas:

2.5. MODELADO DE INTERACCIONES CON EL MODELO DE ISING 34

$$\mathcal{L}(J, h) = \sum_{m=1}^M \log P(\sigma^{(m)}),$$

donde $\{\sigma^{(m)}\}_{m=1}^M$ son las configuraciones registradas.

El gradiente de esta función de verosimilitud con respecto a los parámetros tiene la forma:

$$\frac{\partial \mathcal{L}}{\partial J_{ij}} = \langle \sigma_i \sigma_j \rangle_{\text{datos}} - \langle \sigma_i \sigma_j \rangle_{\text{modelo}},$$

$$\frac{\partial \mathcal{L}}{\partial h_i} = \langle \sigma_i \rangle_{\text{datos}} - \langle \sigma_i \rangle_{\text{modelo}},$$

donde los términos $\langle \cdot \rangle_{\text{datos}}$ se obtienen como promedios sobre las configuraciones empíricas, y los términos $\langle \cdot \rangle_{\text{modelo}}$ se calculan a partir de la distribución del modelo.

En este trabajo, gracias al tamaño reducido del sistema (9 nodos, $2^9 = 512$ configuraciones), fue posible calcular de forma exacta la función de partición y las expectativas modeladas, evitando el uso de muestreo estocástico como Monte Carlo [29]. Esto permitió implementar un algoritmo de descenso por gradiente exacto, complementado con regularización L_2 sobre los pesos J_{ij} para evitar sobreajuste.

La energía asociada a una configuración se definió como:

$$E(\sigma) = - \sum_{i < j} J_{ij} \sigma_i \sigma_j - \sum_i h_i \sigma_i,$$

lo cual permitió expresar su probabilidad como:

$$P(\sigma) = \frac{1}{Z} \exp(-E(\sigma)).$$

El algoritmo iterativo actualizó los parámetros siguiendo las reglas:

$$h_i \leftarrow h_i + \eta (\langle \sigma_i \rangle_{\text{datos}} - \langle \sigma_i \rangle_{\text{modelo}}), \quad (2.2)$$

$$J_{ij} \leftarrow J_{ij} + \eta (\langle \sigma_i \sigma_j \rangle_{\text{datos}} - \langle \sigma_i \sigma_j \rangle_{\text{modelo}}) - \eta \lambda J_{ij}, \quad (2.3)$$

donde η es la tasa de aprendizaje y λ el coeficiente de regularización.

El proceso iterativo se detalla a continuación:

Algorithm 2 Gradient Descent for Ising Model Parameter Inference

Require: Observed configurations $\{\sigma^{(1)}, \dots, \sigma^{(M)}\}$, learning rate η , number of iterations T , regularization parameter λ

- 1: Initialize $h_i \leftarrow 0, J_{ij} \leftarrow 0$ for all i, j
- 2: **for** $t \leftarrow 1$ to T **do**
- 3: Compute empirical averages:

$$\langle \sigma_i \rangle_{\text{data}}, \quad \langle \sigma_i \sigma_j \rangle_{\text{data}}$$

- 4: Enumerate all configurations $\sigma \in \{-1, +1\}^9$
- 5: **for** each configuration σ **do**
- 6: Compute energy:

$$E(\sigma) = - \sum_{i < j} J_{ij} \sigma_i \sigma_j - \sum_i h_i \sigma_i$$

- 7: Compute unnormalized probability:

$$\tilde{P}(\sigma) = \exp(-E(\sigma))$$

- 8: **end for**
- 9: Normalize to obtain $P(\sigma)$
- 10: Compute model expectations:

$$\langle \sigma_i \rangle_{\text{model}}, \quad \langle \sigma_i \sigma_j \rangle_{\text{model}}$$

- 11: Update parameters:

$$h_i \leftarrow h_i + \eta (\langle \sigma_i \rangle_{\text{data}} - \langle \sigma_i \rangle_{\text{model}})$$

$$J_{ij} \leftarrow J_{ij} + \eta (\langle \sigma_i \sigma_j \rangle_{\text{data}} - \langle \sigma_i \sigma_j \rangle_{\text{model}}) - \eta \lambda J_{ij}$$

- 12: **end for**
-

Este procedimiento garantizó la convergencia a un conjunto de parámetros que maximizan la verosimilitud del modelo, y proporciona una primera reconstrucción de la red de interacciones acústicas dentro del coro.

Capítulo 3

Resultados e Interpretación

- 3.1. Evaluación de la sincronización
- 3.2. Exactitud de la detección y asignación micrófono-Colín
- 3.3. Comparación entre algoritmos heurísticos
- 3.4. Análisis de las interacciones inferidas $\{J_{ij}\}$
- 3.5. Idoneidad del modelo de Ising

Conclusiones

Recomendaciones

Referencias Bibliográficas

- [1] Miguel A Acevedo, Carlos J Corrada-Bravo, Héctor Corrada-Bravo, Luis J Villanueva-Rivera, and T Mitchell Aide. Automated classification of bird and amphibian calls using machine learning: A comparison of methods. *Ecological Informatics*, 4(4):206–214, 2009. (Citado en las páginas 10 y 16).
- [2] Roberto Alonso, Ariel Rodríguez-Gómez, and Alberto R Estrada. Patrones de actividad acústica y trófica de machos cantores de *eleutherodactylus eileenae* (anura: Leptodactylidae). *Revista española de herpetología*, 15(2001):45–52, 2001. (Citado en las páginas 9, 11, 19 y 22).
- [3] Roberto Alonso Bosch and Ariel Rodríguez Gómez. Llamadas del repertorio vocal de *eleutherodactylus eileenae* (anura: Leptodactylidae) del occidente de cuba. *Cuadernos de Herpetología*, 14, 2001. (Citado en la página 25).
- [4] William Bialek. *Biophysics: searching for principles*. Princeton University Press, 2012. (Citado en la página 13).
- [5] Daniel T Blumstein, Daniel J Mennill, Patrick Clemens, Lewis Girod, Kung Yao, Gail Patricelli, Jill L Deppe, Alan H Krakauer, Christopher Clark, Kathryn A Cortopassi, et al. Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. *Journal of Applied Ecology*, 48(3):758–767, 2011. (Citado en las páginas 9, 10 y 16).
- [6] Natalie T Boelman, Gregory P Asner, Patrick J Hart, and Roberta E Martin. Multi-trophic invasion resistance in hawaii: bioacoustics, field surveys, and airborne remote sensing. *Ecological Applications*, 17(8):2137–2144, 2007. (Citado en la página 19).

- [7] Sheryn Brodie, Slade Allen-Ankins, Michael Towsey, Paul Roe, and Lin Schwarzkopf. Automated species identification of frog choruses in environmental recordings using acoustic indices. *Ecological Indicators*, 119:106852, 2020. (Citado en la página 20).
- [8] Ryan Calsbeek, Francisco Javier Zamora-Camacho, and Laurel B Symes. Individual contributions to group chorus dynamics influence access to mating opportunities in wood frogs. *Ecology Letters*, 25(6):1401–1409, 2022.
- [9] Juan Sebastián Cañas, María Paula Toro-Gómez, Larissa Sayuri Moreira Sugai, Hernán Darío Benítez Restrepo, Jorge Rudas, Breyner Posso Bautista, Luís Felipe Toledo, Simone Dena, Adão Henrique Rosa Domingos, Franco Leandro de Souza, et al. A dataset for benchmarking neotropical anuran calls identification in passive acoustic monitoring. *Scientific Data*, 10(1):771, 2023. (Citado en la página 17).
- [10] H Chau Nguyen, Riccardo Zecchina, and Johannes Berg. Inverse statistical problems: from the inverse ising problem to data science. *Advances in Physics*, 66(3):197–261, 2017. (Citado en las páginas 21 y 33).
- [11] James W Cooley and John W Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of computation*, 19(90):297–301, 1965. (Citado en la página 17).
- [12] Cornell Lab of Ornithology. The birth of natural sound recording. All About Birds website, n.d. “Allen and Kellogg recorded a Song Sparrow, a House Wren and a Rose-breasted Grosbeak on May 18, 1929.”. (Citado en la página 16).
- [13] Cornell Lab of Ornithology. Early milestones (1920-1950). Macaulay Library website, n.d. “1929: Allen and Kellogg made the first recordings of wild birds in Ithaca.”. (Citado en la página 16).
- [14] Luciano da F Costa. Comparing cross correlation-based similarities. *arXiv preprint arXiv:2111.08513*, 2021. (Citado en las páginas 13 y 27).
- [15] George E Drewry and A Stanley Rand. Characteristics of an acoustic community: Puerto rican frogs of the genus *eleutherodactylus*. *Copeia*, pages 941–953, 1983. (Citado en la página 19).
- [16] AR Estrada. Nuevas localidades para *eleutherodactylus eileenae dunn* (amphibia: Anura: Leptodactylidae) de cuba. *Misc. Zool., Academia de Ciencias de Cuba*, 22(4), 1984. (Citado en la página 22).

- [17] AR Estrada. Herpetofauna de la cuenca banao-higuanojo, sancti spíritus, cuba. *Revista de la Academia Colombiana de Ciencias*, 19:353–360, 1994. (Citado en las páginas 9 y 22).
- [18] Almo Farina. Ecoacoustics: A quantitative approach to investigate the ecological role of environmental sounds. *Mathematics*, 7(1):21, 2018. (Citado en la página 20).
- [19] RA Fisher. The maximum–likelihood–method. *Messenger in Mathematics*, 41:155–160, 1912. (Citado en la página 14).
- [20] Hongxiao Gan, Jinglan Zhang, Michael Towsey, Anthony Truskinger, Debra Stark, Berndt J Van Rensburg, Yuefeng Li, and Paul Roe. A novel frog chorusing recognition method with acoustic indices and machine learning. *Future Generation Computer Systems*, 125:485–495, 2021. (Citado en la página 20).
- [21] Clive WJ Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*, pages 424–438, 1969. (Citado en la página 21).
- [22] Andrew P Hill, Peter Prince, Evelyn Piña Covarrubias, C Patrick Doncaster, Jake L Snaddon, and Alex Rogers. Audiomoth: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment. *Methods in Ecology and Evolution*, 9(5):1199–1211, 2018. (Citado en la página 25).
- [23] Ernst Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik*, 31(1):253–258, 1925. (Citado en las páginas 12, 13 y 32).
- [24] Daniel Jurafsky and James H Martin. Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition. (Citado en la página 26).
- [25] Lata Kalra. *Signal Perception in Complex Acoustic Environments: A Study on Treefrog Communication*. PhD thesis, University of Minnesota, 2024. (Citado en las páginas 17 y 18).
- [26] Paul Mermelstein. Distance measures for speech recognition, psychological and instrumental. *Pattern recognition and artificial intelligence*, 116:374–388, 1976. (Citado en la página 17).

- [27] Thierry Mora and William Bialek. Are biological systems poised at criticality? *Journal of Statistical Physics*, 144:268–302, 2011. (Citado en las páginas 13 y 20).
- [28] Kaiichiro Ota, Ikkyu Aihara, and Toshio Aoyagi. Interaction mechanisms quantified from dynamical features of frog choruses. *Royal Society open science*, 7(3):191693, 2020. (Citado en las páginas 17 y 19).
- [29] Daniel Peña Sánchez de Rivera. Deducción de distribuciones: el método de monte carlo. *Fundamentos de Estadística. Alianza Editorial, Madrid. pp*, pages 220–300, 2001. (Citado en la página 34).
- [30] Nadia Pieretti, Almo Farina, and Davide Morri. A new methodology to infer the singing activity of an avian community: The acoustic complexity index (aci). *Ecological indicators*, 11(3):868–873, 2011. (Citado en la página 19).
- [31] Alberto Porta and Luca Faes. Wiener–granger causality in network physiology with applications to cardiovascular control and neuroscience. *Proceedings of the IEEE*, 104(2):282–309, 2015. (Citado en la página 21).
- [32] A Reyes, M Curbelo, F Tejera, A Rivera, G Simon, O Ramos, MS Turner, and E Altshuler. Transmission of danger information past physical barriers by ants. *arXiv preprint arXiv:1904.03236*, 2019. (Citado en la página 13).
- [33] Ignacio Sánchez-Gendríz and Linilson Rodrigues Padovese. A methodology for analyzing biological choruses from long-term passive acoustic monitoring in natural areas. *Ecological Informatics*, 41:1–10, 2017.
- [34] Elad Schneidman, Michael J Berry, Ronen Segev, and William Bialek. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*, 440(7087):1007–1012, 2006. (Citado en las páginas 13 y 20).
- [35] Thomas Schreiber. Measuring information transfer. *Physical review letters*, 85(2):461, 2000. (Citado en la página 21).
- [36] Albert Schwartz. A new frog of the auriculatus group of the genus eleutherodactylus from western cuba. *Herpetologica*, 14(2):69–77, 1958. (Citado en las páginas 9 y 22).

- [37] John L Spiesberger and Kurt M Fristrup. Passive localization of calling animals and sensing of their acoustic environment using acoustic tomography. *The american naturalist*, 135(1):107–153, 1990. (Citado en la página 18).
- [38] Margaret M Stewart and F Harvey Pough. Population density of tropical forest frogs: relation to retreat sites. *Science*, 221(4610):570–572, 1983. (Citado en la página 19).
- [39] Dan Stowell. Computational bioacoustics with deep learning: a review and roadmap. *PeerJ*, 10:e13152, 2022. (Citado en la página 17).
- [40] George Sugihara, Robert May, Hao Ye, Chih-hao Hsieh, Ethan Deyle, Michael Fogarty, and Stephan Munch. Detecting causality in complex ecosystems. *science*, 338(6106):496–500, 2012. (Citado en la página 21).
- [41] Gasper Tkacik, Elad Schneidman, Michael J Berry II, and William Bialek. Ising models for networks of real neurons. *arXiv preprint q-bio/0611072*, 2006. (Citado en la página 21).
- [42] Daniel S Townsend and Margaret M Stewart. Reproductive ecology of the puerto rican frog *eleutherodactylus coqui*. *Journal of Herpetology*, pages 34–40, 1994. (Citado en la página 19).
- [43] Michael Towsey, Jason Wimmer, Ian Williamson, and Paul Roe. The use of acoustic indices to determine avian species richness in audio-recordings of the environment. *Ecological Informatics*, 21:110–119, 2014. (Citado en la página 16).
- [44] Pauli Virtanen, Ralf Gommers, Travis Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. Nelson, Eric Jones, Robert Kern, Eric Larson, CJ Carey, İbrahim Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Damien Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, José Quintero, Charles R. Harris, Anne M. Archibald, António H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. Scipy signal processing (scipy.signal). <https://docs.scipy.org/doc/scipy/reference/signal.html>, 2020. (Citado en la página 27).

- [45] Lawrence L Woolbright. Patterns of nocturnal movement and calling by the tropical frog *eleutherodactylus coqui*. *Herpetologica*, pages 1–9, 1985. (Citado en la página 19).
- [46] Jie Xie, Towsey Michael, Jinglan Zhang, and Paul Roe. Detecting frog calling activity based on acoustic event detection and multi-label learning. *Procedia Computer Science*, 80:627–638, 2016. (Citado en la página 18).
- [47] Jie Xie, Mingying Zhu, Kai Hu, Jinglan Zhang, Harry Hines, and Ya Guo. Frog calling activity detection using lightweight cnn with multi-view spectrogram: A case study on kroombit tinker frog. *Machine Learning with Applications*, 7:100202, 2022.
- [48] Hong-Li Zeng, Mikko Alava, Erik Aurell, John Hertz, and Yasser Rou-di. Maximum likelihood reconstruction for ising models with asynchronous updates. *Physical review letters*, 110(21):210601, 2013. (Citado en las páginas 21 y 33).
- [49] Tao Zhang, Guoqing Feng, Jinhua Liang, and Tong An. Acoustic scene classification based on mel spectrogram decomposition and model merging. *Applied Acoustics*, 182:108258, 2021. (Citado en las páginas 18 y 25).