



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA  
INGENIERÍA ELÉCTRICA-PROCESAMIENTO DIGITAL DE SEÑALES

MAPEO ACÚSTICO DE UN AMBIENTE UTILIZANDO UNA RED DISTRIBUIDA DE  
SENSORES AD-HOC

TESIS  
QUE PARA OPTAR POR EL GRADO DE:  
DOCTOR EN INGENIERÍA

PRESENTA:  
EMILIANO EHECATL GARCÍA UNZUETA

TUTORES PRINCIPALES:  
CALEB ANTONIO RASCÓN ESTEBANÉ, IIMAS  
PAUL ERICK MÉNDEZ MONROY, IIMAS  
MIEMBROS DEL COMITÉ TUTOR:  
JIMENA OLVERES MONTIEL, FACULTAD DE INGENIERÍA  
EDGAR GARDUÑO ÁNGELES, IIMAS

CIUDAD UNIVERSITARIA, CD.MX, DICIEMBRE-2023



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



**JURADO ASIGNADO:**

Presidente: Dr. Garduño Ángeles Edgar

Secretario: Dr. Velasco Segura Roberto

1er. Vocal: Dr. Rascón Estebané Caleb A.

2do. Vocal: Dr. Méndez Monroy Paul Erick

3er. Vocal: Dra. Olveres Montiel Jimena

Lugar o lugares donde se realizó la tesis: IIMAS

**TUTORES DE TESIS:**

Dr. Caleb Antonio Rascón Estebané

Dr. Paul Erick Méndez Monroy



## **Agradecimientos**

A la Universidad Nacional Autónoma de México (UNAM) por permitirme formar parte de su comunidad, así como al Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (IIMAS) por permitirme cursar el Posgrado de Ingeniería y desarrollar el proyecto de investigación en sus instalaciones.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada para la realización del Doctorado en Ingeniería Eléctrica.

A todos los miembros del comité tutorial: la Dra. Jimena Olveres Montiel, Dr. Edgar Garduño Ángeles, y en particular a mis tutores, Dr. Caleb Antonio Rascón Estebané y Dr. Paul Erick Méndez Monroy.

A mi familia, la inspiración para superarme siempre: Katherin Ruiz, mi esposa y compañera en este viaje y a Beto, mi hijo y cómplice.

Agradecimiento especial para todos los guías que han sido parte de este camino y que no nos acompañan más: mis padres, Ana y Francisco, cuya guía ayudó a formar mi carácter y de quienes aprendí los valores que ahora rigen mi vida, y también al Físico Andrés Valentín Porta Contreras, mi primer tutor en este viaje académico y de quien aprendí las herramientas que me han llevado a concluir mi formación académica.



# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Abstract . . . . .	1
1.2. Resumen del proyecto . . . . .	1
1.3. Objetivo . . . . .	3
1.4. Motivación . . . . .	3
1.5. Contribución . . . . .	5
1.6. Estructura de la tesis . . . . .	5
1.7. Abreviaciones . . . . .	6
<b>2. Marco teórico</b>	<b>7</b>
2.1. Análisis del ambiente acústico . . . . .	7
2.2. Formador de haz . . . . .	8
2.3. Red Inalámbrica de Sensores Acústicos (WASN) . . . . .	9
2.4. Metodología para la sincronización de la WASN . . . . .	11
2.4.1. Sincronización . . . . .	11
2.4.2. Protocolos de sincronización . . . . .	12
2.4.3. Sincronización de relojes basada en el acceso a Internet . . . . .	13
2.4.4. Metodología de sincronización . . . . .	15
<b>3. Diseño de la Red Inalámbrica de Sensores Acústicos</b>	<b>17</b>
3.1. Selección de sensor acústico . . . . .	17
3.2. Análisis de la sensibilidad del sensor acústico seleccionado . . . . .	19
3.3. Selección de unidad de procesamiento . . . . .	21
3.3.1. Arduino MKR Zero . . . . .	21
3.3.2. Raspberry Pi Zero W y Raspberry Pi 4B+ . . . . .	23
3.4. Software de captura . . . . .	24
<b>4. Modelo matemático para la generación del mapa de energía acústica</b>	<b>27</b>
4.1. Funcionamiento del modelo de mapeo de energía acústica . . . . .	27
4.2. Campo lejano y campo cercano . . . . .	28
4.2.1. Modelo de mapeo de energía acústica para campo lejano . . . . .	29
4.2.2. Modelo de mapeo de energía acústica para campo cercano . . . . .	30
4.3. Formadores de haz para la generación del mapa de energía acústica . . . . .	32
4.3.1. Delay And Sum (DAS) . . . . .	33
4.3.2. Minimum Variance Distortionless Response (MVDR) . . . . .	34
4.3.3. Phase-Based Frequency Masking (PBM) . . . . .	36
4.3.4. Steered Response Power with Phase Transform (SRP-PHAT) . . . . .	37

<b>5. Caracterización del modelo de mapeo acústico</b>	<b>39</b>
5.1. Ejemplo para la generación de un mapa de energía acústica . . . . .	39
5.2. Tiempo de procesamiento . . . . .	43
5.2.1. Dependencia con el tamaño de muestra . . . . .	43
5.2.2. Dependencia con el número de micrófonos . . . . .	44
5.2.3. Dependencia con los puntos de prueba . . . . .	45
5.3. Forma del mapa de energía . . . . .	47
5.3.1. Dependencia del funcionamiento con la presencia de ruido . . . . .	47
5.3.2. Caracterización para la variación de la frecuencia de la fuente . . . . .	51
5.3.3. Caracterización para múltiples fuentes acústicas . . . . .	54
<b>6. Metodologías de sincronización y captura de la WASN</b>	<b>57</b>
6.1. Grabación local con protocolos: NTP y PTP . . . . .	57
6.2. Grabación Remota . . . . .	58
<b>7. Análisis de resultados experimentales y simulados</b>	<b>61</b>
7.1. Simulación de grabación de señales con elementos de sincronización . . . . .	65
7.2. Resultados experimentales y análisis . . . . .	73
7.2.1. Análisis de sincronización . . . . .	76
7.2.2. Metodología NTP-Local . . . . .	77
7.2.3. Metodología PTP-Local . . . . .	79
7.2.4. Metodología PTP-Remota . . . . .	81
<b>8. Discusión</b>	<b>85</b>
<b>9. Conclusiones</b>	<b>89</b>
<b>Bibliografía</b>	<b>91</b>

# Índice de figuras

1.1. Mapas de energía acústica . . . . .	2
1.2. Sistema de captura . . . . .	2
2.1. Elementos presentes en un ambiente acústico y procesamiento de señales . . . . .	8
2.2. Funcionamiento básico de un formador de haz. . . . .	8
2.3. Patrón de radiación obtenido en un formador de haz. . . . .	10
2.4. WASN: sensores acústicos, servidores y red de comunicación. . . . .	10
2.5. Gráfica de sincronización de un reloj ideal, uno adelantado y uno atrasado . . . . .	12
2.6. Protocolos de sincronización basados en Internet. . . . .	13
2.7. Esquema de sincronización tipo cliente-servidor. . . . .	15
3.1. Adafruit I2S MEMS Microphone Breakout. . . . .	18
3.2. Montaje para determinar la sensibilidad del dispositivo de captura. . . . .	19
3.3. Montaje para determinar la respuesta del sistema como función de la frecuencia. . . . .	20
3.4. Frecuencia máxima del espectro en la captura de la señal de la fuente. . . . .	20
3.5. RMS de la señal capturada por el nodo en función de la posición de la fuente. . . . .	21
3.6. Primer prototipo de sistema de captura. . . . .	22
3.7. Nodo de las WASN con Raspberry Pi Zero W. . . . .	23
3.8. Unidades de procesamiento utilizadas para los nodos de la WASN. . . . .	24
4.1. Modelo de mapeo de energía acústica. . . . .	27
4.2. Mapa de energía acústica en el conjunto de puntos de prueba. . . . .	28
4.3. Frente de ondas acústicas para campo lejano y campo cercano. . . . .	28
4.4. Frente de onda para una fuente en el campo lejano. . . . .	29
4.5. Frente de onda para una fuente en el campo cercano. . . . .	31
4.6. Diagrama del funcionamiento del formador de haz <b>DAS</b> . . . . .	33
4.7. Suma de las señales con tiempos de arribo distintos. . . . .	33
5.1. Red de sensores elegida para la simulación. . . . .	40
5.2. Señales de entrada capturada por los nodos de la red. . . . .	40
5.3. Malla de puntos de 625 prueba propuestos. . . . .	41
5.4. Mapas de energía generados con los diferentes formadores de haz. . . . .	42
5.5. Mapas de calor generados con los diferentes formadores de haz. . . . .	42
5.6. Tiempo de procesamiento contra número de muestras. . . . .	43
5.7. Tiempo de ejecución contra el número de micrófonos. . . . .	45
5.8. Mapas de energía acústica obtenidos para distintas cantidades de puntos de prueba. . . . .	46
5.9. Tiempo de ejecución contra número de puntos de prueba. . . . .	46
5.10. Mapa generado sin la presencia de ruido. . . . .	48

5.11. Mapa generado con la presencia de ruido de amplitud 0.02A. . . . .	49
5.12. Mapa generado con la presencia de ruido de amplitud 0.1A. . . . .	49
5.13. Mapa generado con la presencia de ruido de amplitud 0.2A. . . . .	50
5.14. Mapa generado con la presencia de ruido de amplitud 0.4A. . . . .	50
5.15. Mapa generado con una señal de frecuencia de 10 Hz. . . . .	51
5.16. Mapa generado con una señal de frecuencia de 100 Hz. . . . .	52
5.17. Mapa generado con una señal de frecuencia de 250 Hz. . . . .	52
5.18. Mapa generado con una señal de frecuencia de 500 Hz. . . . .	53
5.19. Mapa generado con una señal de frecuencia de 700 Hz. . . . .	53
5.20. Mapa generado con una señal de frecuencia de 900 Hz. . . . .	54
5.21. Mapa generado para dos fuentes acústicas puntuales. . . . .	55
5.22. Mapa generado el formador <b>MVDR</b> . . . . .	55
5.23. Mapa generado el formador <b>PBM</b> . . . . .	56
6.1. Metodología de grabación local. Archivos creados en cada nodo. . . . .	57
6.2. Metodología de grabación remota. . . . .	58
7.1. Amplitud y energía originales de la señal de la fuente. . . . .	62
7.2. Señales de entrada capturadas por cada nodo. . . . .	62
7.3. Mapas de energía acústica para los formadores: DAS, MVDR, SRP-PHAT y PBM. . . . .	63
7.4. Mapas de calor para los formadores: DAS, MVDR, SRP-PHAT y PBM. . . . .	64
7.5. Energía obtenida de la señal reconstruida en el punto de prueba $(x, y) = (0, 0)$ . . . . .	65
7.6. Señales de entrada con un retraso aumentando en 20 muestras en el nodo 1. . . . .	66
7.7. Mapas de calor para el formador de haz DAS con diferentes errores de sincronización. . . . .	67
7.8. Mapas de calor para el formador de haz MVDR con errores de sincronización. . . . .	68
7.9. Mapas de calor para el formador de haz PBM con errores de sincronización. . . . .	69
7.10. Mapas de calor para el formador de haz SRP-PHAT con errores de sincronización. . . . .	70
7.11. Errores de localización de una fuente en $(x, y) = (0, 0)$ y errores de sincronización. . . . .	71
7.12. Errores de localización para todos los formadores de haz. . . . .	71
7.13. Ajuste para los formadores de haz: DAS y MVDR. . . . .	72
7.14. Diseño experimental de la WASN para pruebas. . . . .	73
7.15. Arreglo experimental para la evaluación de las metodologías. . . . .	73
7.16. Procesamiento de archivos capturados por los nodos de la WASN. . . . .	74
7.17. Post-procesamiento de las señales de entrada. . . . .	75
7.18. Mapas de calor generados a partir de las señales capturadas mediante las WASN con la fuente en la posición $(0m, 0m)$ . . . . .	75
7.19. Errores de sincronización para cada nodo y cada metodología. . . . .	76
7.20. Mapas de energía del formador DAS (izquierda: mejor, derecha: peor). . . . .	77
7.21. Mapas de energía del formador MVDR (izquierda: mejor, derecha: peor). . . . .	78
7.22. Mapas de energía del formador PBM (izquierda: mejor, derecha: peor). . . . .	78
7.23. Mapas de energía del formador SRP-PHAT (izquierda: mejor, derecha: peor). . . . .	79
7.24. Mapas de energía del formador DAS (izquierda: mejor, derecha: peor). . . . .	79
7.25. Mapas de energía del formador MVDR (izquierda: mejor, derecha: peor) . . . . .	80
7.26. Mapas de energía del formador PBM (izquierda: mejor, derecha: peor). . . . .	80
7.27. Mapas de energía del formador SRP-PHAT (izquierda: mejor, derecha: peor). . . . .	81
7.28. Mapas de energía del formador DAS (izquierda: mejor, derecha: peor). . . . .	81
7.29. Mapas de energía del formador MVDR (izquierda: mejor, derecha: peor). . . . .	82
7.30. Mapas de energía del formador PBM (izquierda: mejor, derecha: peor). . . . .	82

7.31. Mapas de energía del formador SRP-PHAT (izquierda: mejor, derecha: peor). . . . . 83

# Índice de tablas

2.1. Estampas temporales obtenidas mediante el protocolo NTP . . . . .	14
3.1. Tabla comparativa de micrófonos . . . . .	17
3.2. Tabla de opciones de <i>evaluation boards</i> (E. B.) y <i>beakouts</i> . . . . .	18
3.3. Tabla con las características del sensor de captura . . . . .	18
3.4. Tabla comparativa de prototipos de sistema de captura. . . . .	23
3.5. Tabla con las características de los Raspberry Pi Zero W y 4B+. . . . .	24
5.1. Regresión lineal para la relación tiempo de procesamiento contra el número de muestras de las señales capturadas. . . . .	44
5.2. Regresión lineal para la relación tiempo de procesamiento contra el número de micrófonos de la WASN . . . . .	44
5.3. Regresión lineal para la relación tiempo de procesamiento contra el número de puntos de prueba. . . . .	47
8.1. Análisis de error de localización para cada formador de haz y cada metodología . . .	87

# Capítulo 1

## Introducción

### 1.1. Abstract

El proyecto desarrollado consistió en la propuesta y evaluación de una metodología que permita generar un mapa acústico utilizando señales capturadas por una red de sensores acústicos distribuidas en un espacio físico. El proyecto consiste de las siguientes etapas:

- El diseño de una red de sensores acústicos inalámbrica y una metodología para capturar las señales del ambiente acústico y para su sincronización.
- El desarrollo de un modelo matemático y su implementación algorítmica para la generación de mapas acústicos utilizando formadores de haz para el procesamiento de las señales capturadas por la red de sensores.
- La caracterización del modelo propuesto para los diferentes parámetros presentes en el proceso de la generación del mapa acústico.
- Experimentación con la red de sensores diseñada y implementación del modelo propuesto para su validación.

### 1.2. Resumen del proyecto

La exploración del ambiente acústico consiste en la descripción y localización de las fuentes de energía acústica mediante el análisis y la interpretación de las señales capturadas con sensores acústicos, una herramienta útil en la exploración del ambiente acústico es el mapeo acústico, ya que brinda información de la localización e intensidad de las fuentes acústicas presentes. En este trabajo se desarrolló un modelo matemático y su implementación algorítmica para generar un mapa de energía acústica mediante el procesamiento de señales capturadas por una red de sensores acústicos distribuidos en un espacio físico.

Con el modelo matemático es posible obtener un mapa de energía acústica del ambiente como los que se muestran en las Figuras 1.1a y 1.1b, en donde se determina la cantidad de energía acústica que hay en un conjunto de puntos de interés (ambas Figuras representan el mismo mapa de energía obtenido), para lo cual se realiza un proceso de formador de haz para reconstruir la señal en dichos puntos y posteriormente se determina la cantidad de energía por el tiempo que las señales son procesadas.

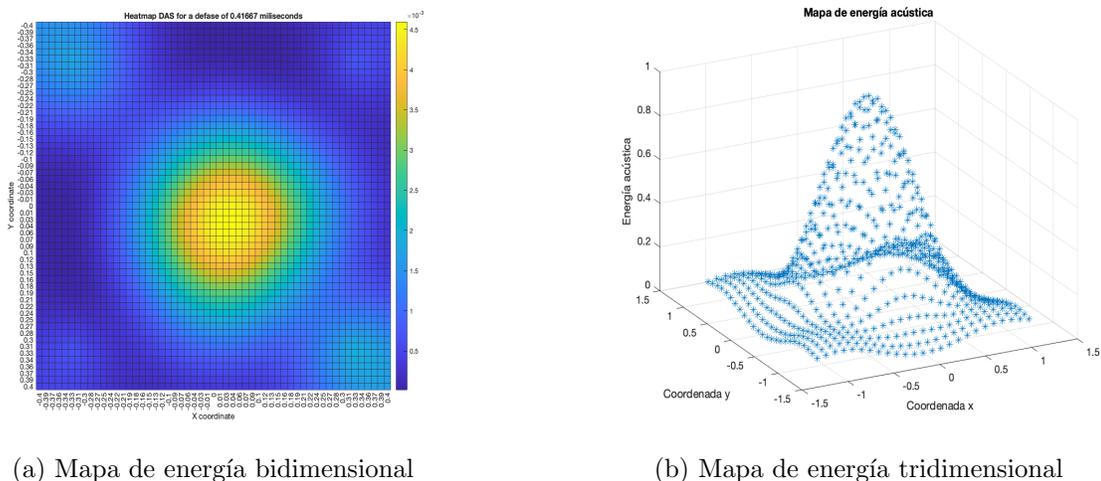


Figura 1.1: Mapas de energía acústica

El modelo matemático para la generación del mapa de energía acústica utiliza como fundamento el concepto de formador de haz. El modelo se caracterizó para cuatro diferentes algoritmos de formadores de haz:

- Delay-and-sum (DAS)
- Minimum Variance Distortionless Response (MVDR)
- Phase-Based Frequency Masking (PBM)
- Steered Response Power with Phase Transforms (SRP-PHAT)

Se diseñó una Red Inalámbrica de Sensores Acústicos (WASN, por sus siglas en inglés). Una WASN debe de contar con nodos de captura de señal acústica, y una unidad central de procesamiento que cuente con características de conexión inalámbrica y almacenamiento interno.

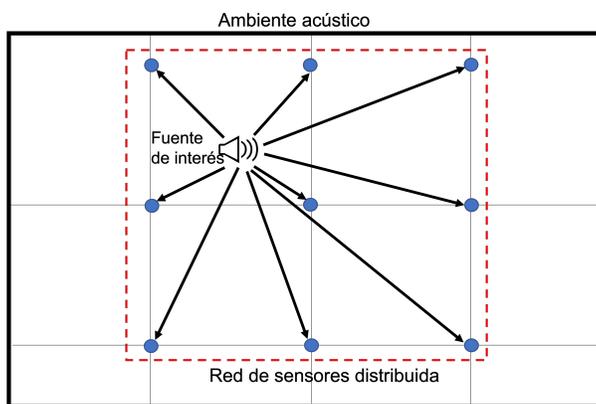


Figura 1.2: Sistema de captura

Al estar distribuida en una superficie, la señal acústica emitida por una fuente deberá viajar diferentes distancias, y como consecuencia, diferentes tiempos hasta cada uno de los nodos de la red, como puede observarse en la Figura 1.2, en donde el ambiente acústico se representa dentro del recuadro negro, el área que cubre la red de sensores como la línea punteada roja, cada nodo de la red como un punto azul y una fuente de interés. Son estas diferencias en los tiempos de arribo de la señal

emitida por la fuente a cada nodo, las bases para el funcionamiento de los algoritmos de formador de haz para la generación del mapa de energía acústica. Sin embargo, pueden producirse variaciones en el reloj interno de la unidad de procesamiento de los nodos de la red debido a diferentes factores como: intermitencia en la red de transmisión de la información, cambios en las variables físicas del entorno, etc. Una red de sensores des-sincronizada, no captura señales que representen el ambiente acústico real, por lo que es de vital importancia asegurar la sincronización dentro de los parámetros necesarios para las aplicaciones deseadas. Se optó por implementar protocolos de sincronización basado en internet, dentro de los que se eligieron el Network Time Protocol (NTP) y el Precision Time Protocol (PTP).

Por esta razón, se desarrollaron y caracterizaron metodologías para la sincronización de la red y la captura de la señal del ambiente acústico para asegurar condiciones en las que las señales capturadas representen al ambiente acústico real. Las metodologías desarrolladas son:

- Grabación local con protocolo de sincronización NTP.
- Grabación local con protocolo de sincronización PTP.
- Grabación remota con protocolo de sincronización PTP.

### **1.3. Objetivo**

El objetivo principal del proyecto doctoral es desarrollar un modelo matemático para generar un mapa de energía acústica en un ambiente determinado, que puede ser doméstico, urbano, de bioacústica, etc. y su validación experimental mediante señales capturadas en los nodos de una red de sensores acústicos.

Como objetivo específico está el desarrollar una WASN contemplando diversas metodologías de sincronización y captura de las señales acústicas del ambiente para la validación experimental del modelo de mapeo acústico desarrollado.

Como objetivos específicos finales del proyecto están la implementación, experimentación y evaluación de los modelos y metodologías desarrollados, para su aplicación en un ambiente real y el estudio de su desempeño.

### **1.4. Motivación**

En la investigación bibliográfica realizada, se encontró que el mapeo de energía acústica se realiza midiendo la energía de la señal acústica en cada uno de los nodos sensores de manera local, esto impone una restricción al número de puntos en donde puede obtenerse información, y esta restricción es el número de nodos con el que cuenta la WASN. La motivación principal del proyecto es desarrollar una metodología que permita obtener información de la energía acústica del ambiente mediante el procesamiento de las señales capturadas por la WASN, pero que permita calcular la energía en puntos diferentes a las posiciones de los nodos, de manera ideal, en cualquier punto del espacio físico, incluso en lugares donde no existen sensores. El resultado obtenido permite generar una WASN con un número de sensores determinado y la capacidad de calcular la energía en la posición de una cantidad puntos de interés sin la necesidad de tener un nodo en cada uno de ellos. El proyecto tiene aplicaciones en diversas áreas, como pueden ser: el monitoreo urbano [1], la bioacústica [2, 3], labores de rescate en situaciones de desastre [4] y puede implementarse como complemento de los asistentes virtuales inteligentes [5].

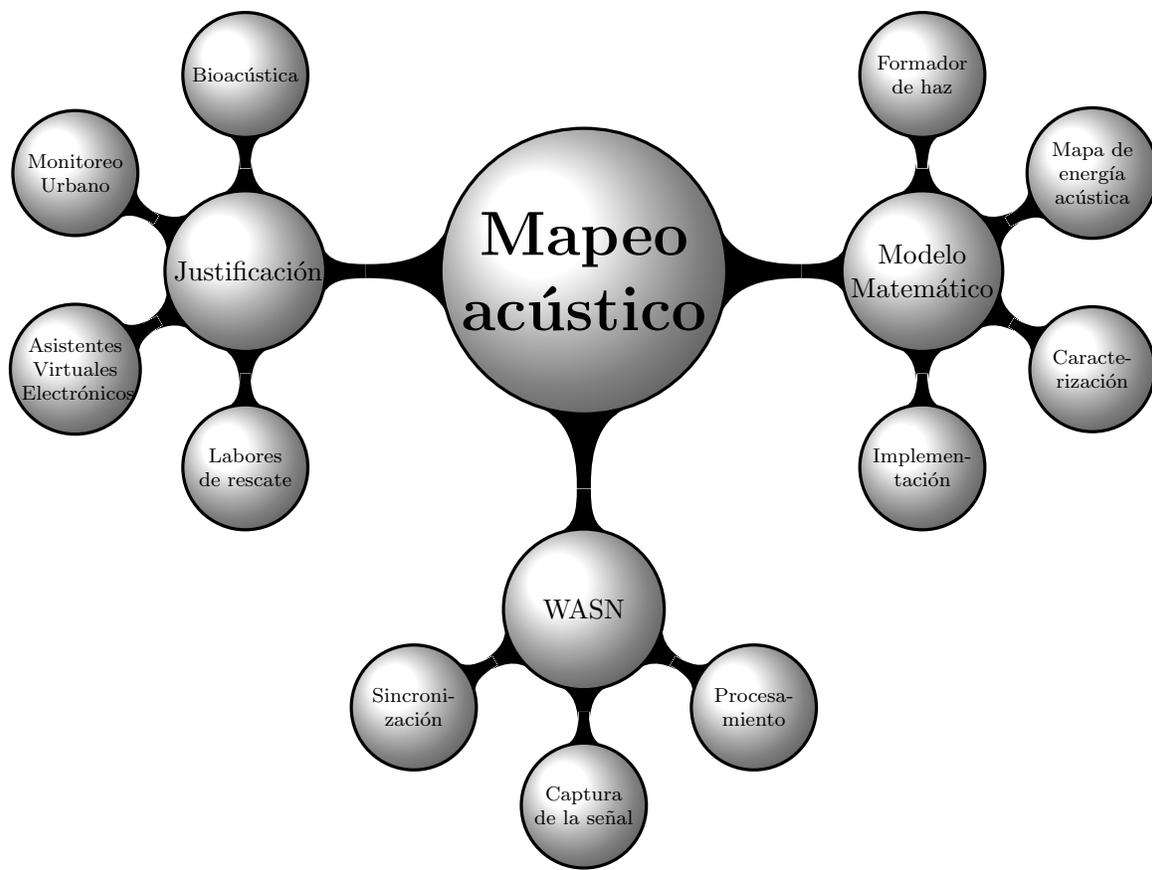


Diagrama 1. Mapa de la estructura del proyecto de mapeo acústico.

En el Diagrama 1 se muestra de manera gráfica la estructura del proyecto, los conceptos, y las técnicas y estrategias que se desarrollaron.

## 1.5. Contribución

El proyecto de doctorado tuvo como resultado el desarrollo de una metodología para la generación de un mapa de energía acústica mediante el procesamiento de señales de un ambiente acústico capturadas utilizando una Red Inalámbrica de Sensores Acústicos. (**WASN**).

La metodología desarrollada permite que el número de sensores de la WASN no determine la cantidad de puntos en donde se calcula la energía, este resultado permite reducir el número de nodos necesarios para generar el mapa de energía en un espacio físico.

Se desarrolló una **WASN** de bajo costo y se propusieron metodologías para la sincronización de sus nodos y la captura de la señal acústica del ambiente.

Con el modelo desarrollado e implementadas, tanto la WASN como las metodologías de sincronización y captura de señal, se obtiene como resultado una mayor portabilidad en redes de sensores empleadas en la localización de fuentes acústicas que beneficia a diferentes áreas de aplicación.

El proyecto desarrollado que tuvo como conclusión la metodología de mapeo acústico puede resumirse en las siguientes etapas:

- Desarrollo de un modelo matemático para el cálculo de energía acústica utilizando los formadores de haz: DAS, MVDR, SRP-PHAT y PBM.
- Simulación de aplicación del modelo para los formadores de haz elegidos con variación en los parámetros de experimentación para fuentes y WASN simuladas.
- Caracterización del modelo para cuatro diferentes formadores de haz, poniendo énfasis en el tiempo de procesamiento y la forma del mapa de energía. Para la caracterización se tomaron en cuenta los siguientes parámetros de variación: tamaño de la muestra, número de nodos de la WASN, número de puntos de prueba, presencia de ruido, frecuencia de la fuente de interés, múltiples fuentes acústicas y errores de sincronización en la WASN.
- Diseño de una WASN y evaluación de los diferentes dispositivos seleccionados; micrófonos y unidades de procesamiento.
- Desarrollo, evaluación y selección de metodologías para la captura de la señal y la sincronización de la WASN.

## 1.6. Estructura de la tesis

La tesis se presenta con la siguiente estructura:

- **Capítulo 2. Marco teórico.** Se definen los conceptos que se utilizarán en las metodologías de mapeo acústico, y de sincronización y captura de la red inalámbrica de sensores acústicos, o WASN, por sus siglas en inglés. Dentro de los conceptos principales están: análisis del ambiente acústico, WASN, sincronización y protocolos de sincronización, formadores de haz y la base teórica de los utilizados en el modelo de mapeo acústico.
- **Capítulo 3. Diseño de la Red Inalámbrica de Sensores Acústicos.** En este capítulo se desarrolla el proceso mediante el cual se diseñó la WASN, tomando en cuenta la evaluación y selección de micrófonos y unidades de procesamiento y el proceso para la captura de la señal.
- **Capítulo 4. Modelo matemático para la generación del mapa de energía acústica.** En este capítulo se desarrolló el modelo para la generación del mapeo acústico tanto para campo lejano como para campo cercano y su implementación para los formadores seleccionados.

- **Capítulo 5. Caracterización del modelo de mapeo acústico.** En este capítulo se analiza el efecto que tiene la variación de parámetros en la implementación simulada de mapeo acústico. Se tomaron en cuenta dos elementos de evaluación; la diferencia en el tiempo de procesamiento y la forma del mapa de energía obtenido. Los parámetros que se variaron fueron: el tamaño de la muestra, el número de nodos de la WASN, el número de puntos de interés para el cálculo de la energía, variación en la frecuencia de la fuente, presencia de ruido y múltiples fuentes de energía acústica.
- **Capítulo 6. Metodologías de sincronización y captura de la WASN.** En este capítulo se definen diferentes maneras de realizar la sincronización de la WASN y la captura de las señales acústicas del ambiente. Se definieron dos maneras diferentes de captura (local y remota) y dos protocolos de sincronización para la WASN (NTP y PTP).
- **Capítulo 7. Análisis de resultados experimentales y simulados.** En este capítulo se aplicó el modelo de mapeo acústico para señales simuladas con elementos de sincronización y se analizaron los resultados. Se realizaron mediciones con la WASN diseñada y las metodologías para su sincronización y captura de la señal para evaluar su desempeño.
- **Capítulo 8. Discusión.** En este capítulo se analizan los resultados obtenidos mediante las simulaciones y las mediciones realizadas con las diferentes metodologías y se discuten los factores que intervienen en el desempeño encontrado.
- **Capítulo 9. Conclusión.** Con base en los resultados discutidos hasta este capítulo, se determinan los elementos dentro de la metodología implementada que tuvieron mejor desempeño en la aplicación de mapeo acústico, se tomaron en cuenta: hardware (micrófonos y unidad de procesamiento), sincronización y captura (metodologías) y desempeño de los diferentes formadores de haz.

## 1.7. Abreviaciones

Las siguientes abreviaciones son utilizadas comúnmente en la tesis presentada:

WASN	Wireless Acoustic Sensor Network
NTP	Network Time Protocol
PTP	Precision Time Protocol
MEMS	Micro–electret microphone
SNR	Signal–to–noise ratio
DAS	Delay and Sum
MVDR	Minimum Variance Distortionless Response
SRP–PHAT	Steered–Response Power Phase Transform
PBM	Phase–based Binary Masking

# Capítulo 2

## Marco teórico

En este capítulo se desarrollan los conceptos utilizados en las diferentes etapas necesarias para la implementación del modelo de mapeo acústico propuesto. Estas etapas son: el diseño de la red de sensores acústicos, el modelo matemático para el mapeo acústico, la caracterización del modelo matemático, el desarrollo de metodologías para la sincronización, experimentación y comparación con los resultados simulados.

### 2.1. Análisis del ambiente acústico

El análisis de señales acústicas consiste en el procesamiento de la información capturada en un ambiente acústico. Las señales obtenidas en sensores acústicos son procesadas para determinar la presencia (detección), ubicación (localización), tipo (clasificación) y trayectoria (seguimiento) de una o más fuentes acústicas [6, 7, 8].

El análisis de señales acústicas tiene como objetivo la caracterización del ambiente acústico. Un ambiente acústico es un espacio físico en donde se propagan ondas acústicas generadas por una o más fuentes de sonido y receptores, que pueden ser una persona que escucha esas ondas o sensores acústicos que capturen la señal; en el caso particular del proyecto, los receptores se encuentran dentro de una red de sensores acústicos.

Un ambiente acústico ideal es aquel en donde no existe ruido (que en el contexto del proyecto se refiere a la presencia de fuentes acústicas que interfieren con la de interés), por lo que, la propagación de las ondas acústicas es directa, es decir, no hay reflexiones dentro del ambiente, no existe la reverberación ni los ecos. Este tipo de ambientes pueden obtenerse hasta cierto grado en una cámara anecoica, sin embargo, cualquier otro ambiente acústico tendrá ruido de fuentes ajenas a la de interés: eco, reflexiones, reverberación, obstáculos que impidan que las ondas acústicas lleguen de manera directa. Además, un ambiente real puede contar con una o más fuentes acústicas fijas, fuentes acústicas móviles, y receptores en movimiento, la Figura 2.1 representa un ambiente con sólo una fuente de interés fija y elementos de ruido e interferencia.

Una vez capturada la información, se procesa mediante la implementación de algoritmos que permitan la caracterización del ambiente acústico que puede centrarse en localizar fuentes sonoras, detectar dirección de arribo de las señales acústicas, separación de fuentes o una combinación de estos. La determinación de la dirección de arribo y la localización de fuentes acústicas dependen de dos factores principalmente; el primero, es la naturaleza de la señal capturada por los sensores acústicos, y el segundo, son la geometría y el rango de captura de la red de los sensores acústicos [9]. El modelo desarrollado utiliza como base de su funcionamiento a un formador de haz, en la siguiente sección se define lo que es.

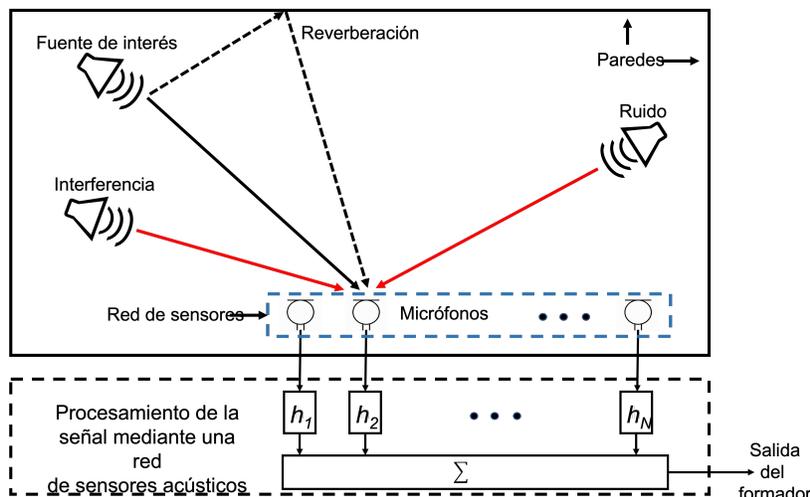


Figura 2.1: Elementos presentes en un ambiente acústico y procesamiento de señales

## 2.2. Formador de haz

Un formador de haz es un procesador que trabajando en conjunto con una red de sensores permite generar un filtrado espacial. La red de sensores captura muestras de ondas acústicas propagándose a través de un medio, en este caso el aire, que después son procesadas por el formador de haz. El objetivo es estimar la señal capturada en una dirección deseada cuando el ambiente acústico tiene elementos de ruido e interferencia de señales.

Un formador de haz funciona de la siguiente manera: se captura la señal del ambiente acústico mediante una red de  $M$  nodos de sensores (en la Figura 2.1 se muestran las señales capturadas por esta red), se realiza el procesamiento por alguno de los diferentes algoritmos que existen y se obtiene a la salida la señal reconstruida en la dirección elegida, dicho procedimiento se ilustra en la Figura 2.2.

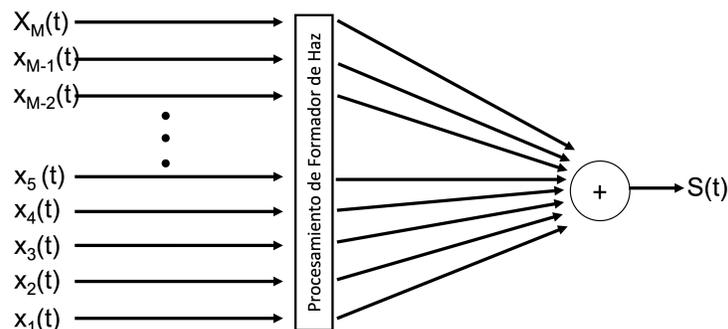


Figura 2.2: Funcionamiento básico de un formador de haz.

Existen diferentes clasificaciones y modelos de formadores de haz, pero el funcionamiento básico es el mismo para todos, la señal acústica del ambiente es capturada mediante una red de sensores, se procesa la información capturada y se reconstruye la señal en una dirección deseada [10]. Las redes de sensores permiten orientar y recibir el patrón en una dirección específica. Un factor importante

para la generación de un formador de haz es el diseño de la red de sensores.

En cada sensor se tiene un patrón de radiación dependiente de la señal emitida por la fuente, el patrón de radiación del sensor mismo, y del ángulo con el que llega esa señal al sensor.

Para obtener el patrón de emisión de toda la red de sensores, es necesario sumar las señales de entrada de cada sensor,  $S_i(\theta)$  desde  $i = 1$ , hasta  $i = M - 1$ . La ecuación 2.1 representa el patrón de radiación de la red [11], y la manera de separar entre los elementos de la misma. La radiación obtenida mediante esta ecuación es una función del ángulo ( $\theta$ ) que se crea a partir de un sistema de referencia determinado por la geometría de la red de sensores.

$$S(\theta) = \sum_{i=1}^{M-1} S_i(\theta) = S_e(\theta)S_a(\theta). \quad (2.1)$$

de donde:

- $S_e(\theta)$ : Factor elemento, patrón de radiación asociado a cada sensor.
- $S_a(\theta)$ : Factor red, patrón de radiación del conjunto de  $M$  sensores.

En cuanto al manejo de las señales para generar este patrón, la salida del formador de haz,  $S(t)$ , para un tiempo  $t$  puede escribirse de la siguiente manera:

$$S(t) = \sum_{i=1}^{M-1} w_i^H x_i(t) \quad (2.2)$$

donde  $w_i^H$  es el complejo conjugado de un concepto con el que se trabaja en los diferentes modelos de formador de haz que es una matriz con elementos por los cuales hay que multiplicar las señales capturadas por cada nodo de micrófonos para recuperar la señal de la fuente en la dirección deseadas, a esta matriz se le denomina *steering vector* [12].

Con esta información es posible obtener un patrón de radiación de la energía obtenida en la dirección analizada y de donde puede obtenerse información analizando los siguientes elementos:

- **Lóbulo principal:** Indica la dirección del haz y la mayor concentración de energía.
- **Lóbulos laterales:** Son producidos por la interferencia constructiva y destructiva de las señales alrededor de la fuente, y de la geometría de la red de sensores.

Estos elementos pueden observarse en la Figura 2.3, en donde se genera un patrón particular para las condiciones particulares de una medición: la geometría de la red de sensores y la posición de la fuente. En la Figura 2.3 se muestra un patrón de amplitud, pero sin escala, ya que nuevamente, su magnitud dependería de las características de la fuente y la red.

Analizando el patrón de radiación y las magnitudes de los elementos antes mencionados, es posible determinar la dirección de la que proviene la energía de la fuente de energía acústica. En el caso del patrón de ejemplo (Figura 2.3), la fuente se encuentra en  $0^\circ$  con respecto a la posición de la red de  $M$  sensores.

## 2.3. Red Inalámbrica de Sensores Acústicos (WASN)

Una red inalámbrica de sensores consiste de un conjunto de nodos distribuidos en un espacio físico que cooperativamente monitorean condiciones físicas o ambientales, para el caso particular

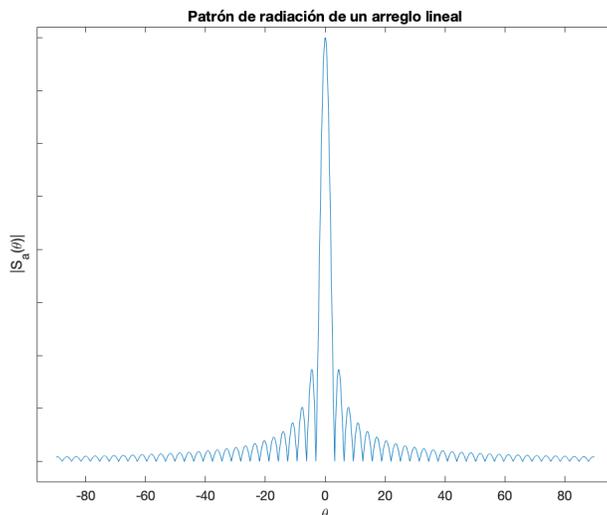


Figura 2.3: Patrón de radiación obtenido en un formador de haz.

de estudio, el parámetro físico monitoreado es el sonido. Cada nodo está equipado con un **sensor** para medir alguna propiedad física, un **transceptor** para comunicarse de manera inalámbrica y las conexiones necesarias para llevar a cabo la **interconexión** de los dispositivos y una **fuentes de energía**, usualmente una batería o conexión a la red eléctrica.

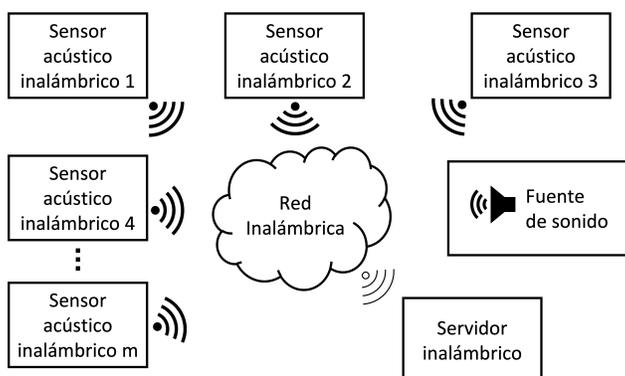


Figura 2.4: WASN: sensores acústicos, servidores y red de comunicación.

Los nodos inalámbricos de la red trabajan en conjunto para capturar, recibir y transmitir información, y/o actuar como repetidores dentro de la misma. Los nodos que conforman la red suelen tener recursos limitados y mantienen su posición dentro de ella. Cada nodo captura la señal de su ambiente acústico local y la transmite a una unidad central de procesamiento. En el caso de la red diseñada, se emplean sensores acústicos, por lo que se denomina una **Red Inalámbrica de Sensores Acústicos** o **WASN** [13, 14], por sus siglas en inglés, un esquema del funcionamiento de este tipo de WASN se muestra en la Figura 2.4. La propagación de una señal acústica se produce por la vibración del medio de transmisión, en el caso del estudio, el medio es el aire. El sensor responsable de capturar la información de este tipo de vibración es el micrófono. Se definen los conceptos que forman una WASN:

- **Nodo.** Unidad de procesamiento que junto con un sensor es capaz de capturar la señal acústica

y transmitirla a otros nodos de la WASN o almacenarla internamente.

- **Unidad de procesamiento.** Componente encargado de servir como procesador y transmisor de la información capturada por sensores del nodo.
- **Sensor.** Componente encargado de sensar alguna propiedad física del ambiente y convertirla a señales eléctricas.

## 2.4. Metodología para la sincronización de la WASN

La necesidad de la sincronización en una WASN surge porque los relojes internos de cada dispositivo dentro de la Red están expuestos a cambios producidos por diferentes interacciones físicas con el medio. El análisis de la escena acústica tiene como principal elemento de procesamiento la diferencia entre los tiempos de arribo a cada uno de los nodos de medición dentro de la WASN. Es por esta razón, que es de vital importancia implementar estrategias que permitan que los relojes se mantengan sincronizados, es decir, que los tiempos internos de cada nodo no presenten desviaciones, dentro de los límites que permita el *hardware*. En esta sección se definirán las metodologías implementadas para la sincronización de la WASN diseñada.

### 2.4.1. Sincronización

La sincronización se refiere al proceso mediante el cual, el reloj interno de un dispositivo se modifica para ajustarse al reloj interno de algún dispositivo externo, este proceso se realiza normalmente en cualquier dispositivo que asocie alguna clase de medición con las marcas temporales en las que se realizó la medición. Es necesario definir y conocer cómo funciona el reloj interno de un dispositivo y los factores que pueden afectarlo.

Los circuitos dentro de un reloj de computadora consisten en un oscilador y un contador. Tomando como referencia la frecuencia angular del oscilador, el contador incrementa su valor para representar un reloj local  $C(t)$  en la WASN. El tiempo real se define como la magnitud del tiempo que se mide en la naturaleza. En una situación ideal, la frecuencia angular del oscilador no varía con el tiempo, si se interpreta la relación  $\frac{dC(t)}{dt}$  como la variación del reloj interno con respecto al tiempo real, el reloj ideal se representaría por la condición  $\frac{dC(t)}{dt} = 1$ . Sin embargo, la frecuencia angular del oscilador está expuesta a variaciones debido a cambios físicos en el ambiente como: temperatura, vibración y presión, cuando esto sucede, la frecuencia angular presenta una desviación o *drift*, en inglés.

El reloj interno de un nodo en la WASN puede relacionarse con el tiempo real de la siguiente manera:

$$C_i(t) = a_i t + b_i \quad (2.3)$$

en donde,  $t$  es el tiempo real,  $C_i(t)$  representa al tiempo interno dentro del  $i$ -ésimo nodo,  $a_i$  es el *drift* que presenta el reloj interno con respecto al tiempo real y  $b_i$  es la diferencia en segundos entre el reloj interno y el tiempo real, *offset*, en inglés.

Cuando el reloj interno de uno de los nodos se encuentra adelantado, se tiene que  $\frac{dC(t)}{dt} > 1$ , mientras que cuando un reloj se atrasa con respecto al tiempo real, la condición que se cumple es:  $\frac{dC(t)}{dt} < 1$ [15].

Del mismo modo, puede compararse el tiempo interno de dos nodos ( $i$  y  $j$ ) de la WASN de la siguiente manera:

$$C_i(t) = a_{ij} C_j(t) + b_{ij} \quad (2.4)$$

$a_{ij}$  es el *drift* relativo y  $b_{ij}$  es el *offset* relativo entre los relojes de los nodos  $i$  y  $j$ . Si dos relojes están perfectamente sincronizados, su *drift* = 1, lo que implica que sus relojes tienen la misma razón de cambio. Si su *offset* = 0, significa que en ese instante, ambos relojes tienen el mismo valor.

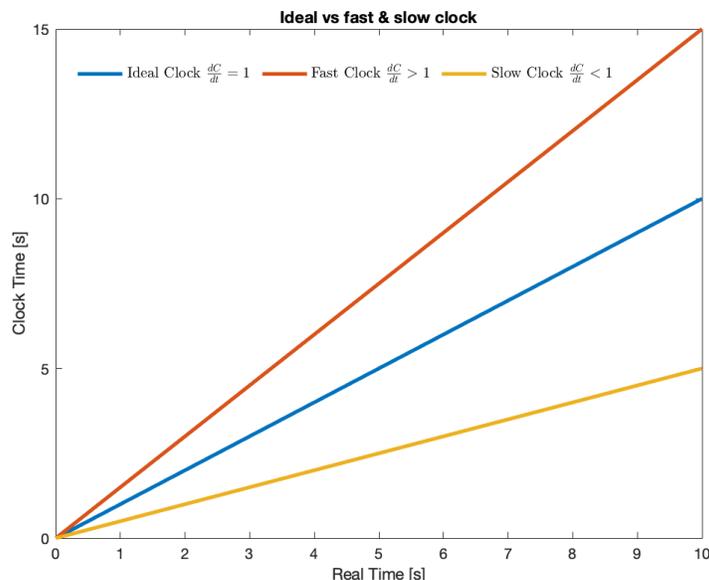


Figura 2.5: Gráfica de sincronización de un reloj ideal, uno adelantado y uno atrasado

En la Figura 2.5, puede observarse una representación gráfica del comportamiento de los diferentes tipos de relojes (ideal, adelantado y atrasado). Después de cierto periodo de tiempo, en un reloj con *drift* > 1, el tiempo en el reloj interno del nodo se adelantará con respecto al tiempo real, mientras que sucederá lo opuesto para un reloj con *drift* < 1.

### 2.4.2. Protocolos de sincronización

Existen diferentes estrategias para lograr que cada uno de los nodos dentro de una WASN se encuentre sincronizado. Una manera directa de solucionar este problema es que cada uno de los nodos de la WASN cuente con *hardware* de control de tiempo externo; sin embargo, este método resulta costoso y complejo en su implementación, por lo que usualmente se busca resolver el problema desde un punto de vista de *software* [16, 17, ?, 18].

Existen varios métodos de sincronización que tienen como fundamento el acceso a Internet. Aunque este tipo de sincronización sirve para algunas aplicaciones en redes de sensores, los métodos más utilizados para la sincronización de WASN toman como referencia el tiempo de nodos dentro de la Red misma o realizan la sincronización en cada nodo de la Red.

Independientemente del método que se emplee para resolver el problema de la sincronización, la mayoría de ellos basan su funcionamiento en el intercambio de mensajes. El intercambio de mensajes tiene asociados errores que pueden producirse en el proceso de sincronización, ya sea cuando se emite el mensaje, cuando se recibe o en el tránsito entre nodos de la Red. A continuación se listan las diferentes etapas en este proceso en donde pueden producirse errores.

- **Tiempo de envío.** El tiempo total que toma la construcción del mensaje de sincronización y su transferencia a la interfaz de comunicación. Este tiempo varía principalmente dependiendo

del sistema operativo en uso.

- **Tiempo de acceso.** Es el tiempo de espera requerido para que el mensaje de sincronización pueda acceder al canal de transmisión.
- **Tiempo de transmisión.** Es el tiempo que le toma al mensaje de sincronización en ser transmitido bit por bit, a través de la antena. Este tiempo puede determinarse conociendo el tamaño del mensaje y la velocidad de transmisión de la antena.
- **Tiempo de propagación.** El tiempo real que le toma al mensaje en transmitirse desde el emisor hasta el receptor en el medio.
- **Tiempo de recepción.** Tiempo que le toma a la capa física del nodo en recibir el mensaje de sincronización a través de la antena.
- **Tiempo de procesamiento.** Tiempo que le toma al nodo receptor procesar el mensaje.

La sincronización de una WASN tiene que tomar en cuenta todas estas posibles fuentes de error y contrarrestarlas mediante los diferentes protocolos existentes para que el tiempo interno de cada nodo se aleje lo menos posible del tiempo real [19, 20, 21, 22].

### 2.4.3. Sincronización de relojes basada en el acceso a Internet

El desarrollo del internet y su fácil acceso hace que se planteen protocolos de sincronización utilizando esta herramienta. Estos protocolos tienen una combinación de flexibilidad y bajo costo en su implementación. Este tipo de protocolos logran la sincronización mediante el envío de mensajes a diferentes nodos dentro de una red local. Entre los protocolos más comunes de este tipo de sincronización se encuentran los protocolos: *Time Protocol*, *Digital Service Time (DTS)*, *Network Time Protocol (NTP)* y *Precision Time Protocol (PTP)*. A continuación, se explican brevemente el funcionamiento de los protocolos NTP y PTP[23].

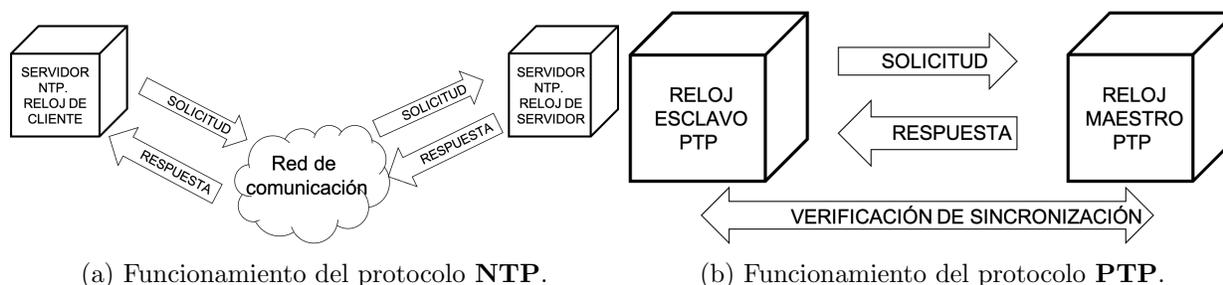


Figura 2.6: Protocolos de sincronización basados en Internet.

**NTP.** Existen tres maneras comunes de implementar este protocolo: el modo *cliente-servidor*, el modo *maestro-esclavo*, y el modo *unicast-multicast*. El más comúnmente utilizado es el *cliente-servidor*, en donde el cliente envía periódicamente información sobre su reloj interno, y el servidor responde al cliente con su tiempo actual para que este se sincronice con él, el funcionamiento de este protocolo puede observarse en la Figura 2.6a. Este protocolo es ampliamente utilizado, ya que para que el cliente pueda sincronizar su reloj interno, lo único que tiene que asegurarse es que el servidor reciba la petición por parte del cliente. La precisión de este protocolo de sincronización está en el rango de los milisegundos. Es por esta razón que se optó por este modo de implementación, pero se podría estudiar como trabajo futuro la viabilidad de los demás modos.

**PTP.** Cuando las necesidades de precisión son mayores, se emplea este protocolo, ya que se encuentra en rangos menores que los microsegundos. La arquitectura de este protocolo es similar al NTP, pero toma medidas adicionales para mejorar la precisión de la sincronización. El esquema de

funcionamiento se muestra en la Figura 2.6b, en donde los relojes de los nodos de la red se dividen entre *reloj maestro* y *reloj esclavo*. Los relojes esclavos de la red ajustan su tiempo de acuerdo a los mensajes emitidos por los relojes maestros. Los mensajes de sincronización entre relojes maestros y relojes esclavos siguen el mismo mecanismo de solicitud-respuesta del protocolo NTP, pero se someten a una etapa de verificación con los demás relojes maestros.

La implementación para ambos protocolos en el modo *cliente-servidor* funciona de la siguiente manera:

- El cliente genera un paquete de información que contiene campos a llenar de: estampas de tiempo, direcciones IP, entre otros.
- El cliente envía el paquete de información al servidor y en el paquete se guarda la información del tiempo cliente **t1** y dirección IP del cliente.
- El servidor recibe el mensaje y lo analiza, sustituye los campos de direcciones IP y algunos otros campos dentro del paquete de información que le corresponden, como el tiempo en el que se recibe el paquete de información **t2**.
- Después de analizar y procesar el paquete de información, el servidor lo envía de vuelta a la dirección IP especificada en el paquete de información recibido. El paquete se envía con la marca de tiempo de transmisión en el servidor **t3**.
- El cliente recibe el paquete de información y graba la última estampa de tiempo, el de recepción del paquete en el cliente **t4**.

En la Tabla 2.1 se observan los procesos en los que se generan las diferentes estampas de tiempo.

Cliente	Conexión	Servidor
Estampa de tiempo 1 ( <b>t1</b> )	⇒ solicitud	Estampa de tiempo 2 ( <b>t2</b> )
Estampa de tiempo 4 ( <b>t4</b> )	⇐ respuesta	Estampa de tiempo 3 ( <b>t3</b> )

Tabla 2.1: Estampas temporales obtenidas mediante el protocolo NTP

Con los valores de estas marcas temporales es posible determinar: el desfase temporal que existe entre el cliente y el servidor (**offset**), y el tiempo que tomó al paquete realizar todo su recorrido (**delay**). Estos valores se calculan de la siguiente manera:

$$offset = \frac{(t2 - t1) + (t3 - t4)}{2}, \quad (2.5)$$

$$delay = (t4 - t1) - (t3 - t2). \quad (2.6)$$

Con estos valores, cuando se implementa alguno de los protocolos, se ajustan los relojes internos de cada nodo cliente con el reloj interno del servidor.

El protocolo **NTP** es el método más común para sincronizar el reloj del *software* de un sistema GNU/Linux con los servidores horarios de Internet. Está diseñado para mitigar los efectos de la variable de latencia de la red. La precisión en las redes de área local es aún mejor, hasta un milisegundo. El protocolo **PTP** realiza el mismo proceso de solicitud y respuesta de mensajes para determinar el desfase y corregir los relojes internos, pero realiza un proceso de verificación que, según la literatura, puede hacer que la precisión en redes locales sea de décimas de microsegundo[24, 25].

#### 2.4.4. Metodología de sincronización

La sincronización es uno de los ejes principales del proyecto, ya que el funcionamiento de los modelos tanto de detección de arribo como de formadores de haz dependen principalmente de la diferencia en los tiempos de arribo de la señal de la fuente con los nodos de la WASN, por lo que sin una WASN sincronizada, no es posible generar un mapa de energía acústica.

El esquema para la sincronización propuesto para la WASN diseñada en el proyecto de investigación se muestra en la Figura 2.7. Los dispositivos seleccionados después de realizar el diseño de la WASN son Raspberry, por lo que el esquema contempla estos dispositivos en cada uno de los nodos de la red. El nodo servidor de tiempo también es un Raspberry como el de los nodos de captura, este nodo sincroniza su reloj interno con un servidor de tiempo mediante el acceso al internet. Una vez sincronizado el Raspberry servidor de tiempo, los nodos de captura (Raspberry cliente) sincronizan su reloj con el del servidor. El protocolo de sincronización se muestra en la Figura 2.7.

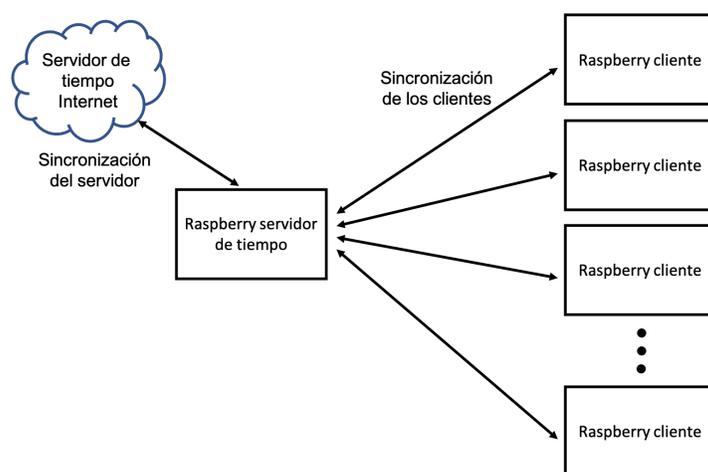


Figura 2.7: Esquema de sincronización tipo cliente-servidor.

Cada uno de los procesos de sincronización, tanto **cliente-servidor** y **servidor-servidor de tiempo en internet**, se realizaron mediante protocolos basados en el internet.

Además de las marcas temporales solicitadas al servidor de tiempo, se realizó un proceso de sincronización mediante el *daemon*, que es un proceso en segundo plano que sincroniza de manera periódica el reloj interno del cliente con el reloj del servidor.



## Capítulo 3

# Diseño de la Red Inalámbrica de Sensores Acústicos

En este capítulo se detallan los elementos que se tomaron en cuenta para el diseño de una Red Inalámbrica de Sensores Acústicos (WASN) que tenga la capacidad de capturar y procesar señales del ambiente para la generación de un mapa de energía acústica adecuados para la implementación del modelo de mapeo acústico propuesto. Se analizó el desempeño de los siguientes dispositivos para el diseño de la WASN: micrófonos para la captura de la señal y unidades de procesamiento con la capacidad de conexión entre nodos de la red.

### 3.1. Selección de sensor acústico

El primer componente del cual se realizó una investigación de mercado fue el micrófono, ya que este resulta una pieza fundamental en la captura de señal acústica y debía cumplir los requerimientos de medición. Dentro de las diferentes variedades que existen en el mercado se optó por un tipo de micrófonos conocido como MEMS (microelectromechanical systems, por sus siglas en inglés) que son dispositivos con partes microscópicas móviles.

En la tabla 3.1 se presentan algunos de los micrófonos analizados y sus características.

Marca	Modelo	Sensibilidad [dB]	SNR [dB]	Ancho de Banda [Hz]
Analog Devices	AMP401	-42	62.0	60-15,000
Vesper	VM3000	-26	62.5	—
TDK Invensense	ICS-43432	-26	65.0	5-20,000
Knowles	SPH0641	-26	64.3	45-20,000
STM	IMP34DT05	-26	64.0	—

Tabla 3.1: Tabla comparativa de micrófonos

Además de la diferencia en el tamaño, un micrófono MEMS tiene, generalmente, una mayor sensibilidad y también una mejor razón de señal a ruido (SNR). Las características que tiene el micrófono que se eligió son: una sensibilidad (sensitivity)  $\sim -26$  dB y una relación de señal a ruido (SNR)  $\sim 65$  dB. Se optó por un micrófono con estas características para determinar si era viable utilizarlo en la aplicación de mapeo acústico utilizando una WASN, pero para trabajos futuros se puede analizar el desempeño del modelo propuesto para dispositivos con mejores características.

Otro parámetro que influyó en la selección del micrófono fue la facilidad de utilizarse para la adquisición de la señal por parte del sistema de procesamiento. En la investigación que se realizó,

se encontraron diversos dispositivos que ya implementan un amplificador y diferentes tipos de salida de la señal. Se eligió entre dos tipos de dispositivos, el primero se llama **Breakout**, que consiste en un circuito pequeño que sólo entrega las conexiones de salida del micrófono. El segundo se llama **Evaluation Board**, que usualmente tiene un sistema de procesamiento para entregar a la salida la señal en voltaje. En la Tabla 3.2 se muestran los dispositivos que implementan un micrófono con las características de sensibilidad y SNR establecidas.

Marca	Modelo	Tipo	Tipo de salida	Precio [Dólares]
Invensense	EV-INMP441	E. B.	Digital	94.25
ST	STEVAL-MIC003V1	Breakout.	Salida single-bit	35.00
Knowles	KAS-33100-004	E. B.	Digital, PDM	373.75
Vesper	S-VM1001-S	E. B.	PDM	81.25
Adafruit	3421	Breakout	I <sup>2</sup> S	9.04
Spakfun	BOB-09868	Breakout	Analógica	14.24

Tabla 3.2: Tabla de opciones de *evaluation boards* (E. B.) y *breakouts*.

El factor más importante por el que se escogió el dispositivo final fue el tipo de salida, ya que como puede observarse en la tabla 3.2, algunos dispositivos cuentan con un diferente protocolo de procesamiento. El protocolo que se escogió fue *I<sup>2</sup>S* (Inter-IC Sound), que permite separar las señales de datos y de reloj, lo que resulta en una fluctuación menor de la señal.

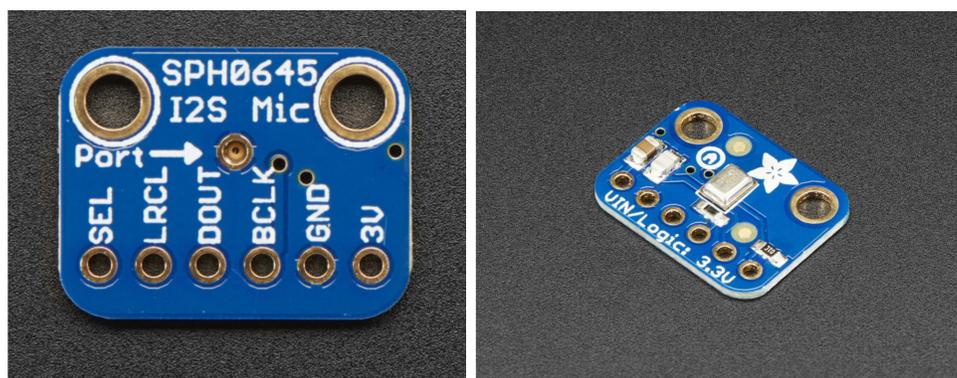


Figura 3.1: Adafruit I2S MEMS Microphone Breakout.

Tomando en consideración todos los factores, se determinó que el dispositivo que actúa como sensor en la WASN es el **Adafruit I2S MEMS Microphone Breakout** (Figura 3.1) y con las características que se muestran en la Tabla 3.3.

Breakout-marca y modelo	Protocolo de comunicación	MEMS-marca modelo	Sensibilidad @ 96 dB	SNR @ 96 dB
Adafruit SPH0645LM4H	I <sup>2</sup> S Inter-IC Sound	Knowles SPH0641LM4H-1	-26dB ±3dB	65 dB

Tabla 3.3: Tabla con las características del sensor de captura

### 3.2. Análisis de la sensibilidad del sensor acústico seleccionado

Se realizó un análisis de la sensibilidad de los sensores utilizados en la WASN. Conociendo la sensibilidad de los nodos de captura, es posible determinar el alcance que tienen, para poder así definir las características mínimas de funcionamiento de la red de sensores acústicos.

Para poder caracterizar la sensibilidad del dispositivo de captura se generó una señal y se realizó una medición de la intensidad capturada como función de la distancia con la fuente de sonido. Para determinar la relación que existe entre la intensidad y la distancia, se varió la distancia de una fuente de sonido con el dispositivo de captura y se midió la intensidad registrada en el dispositivo de captura. El análisis se realizó con un sólo micrófono conectado al Raspberry como se muestra en la Figura 3.2. En la Figura se muestra un arreglo de Raspberry y MEMS, que es el nodo de captura de la WASN, y las mediciones se realizaron haciendo que la fuente se aleje del nodo.

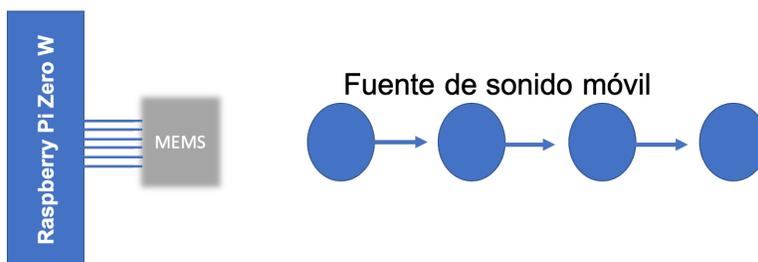


Figura 3.2: Montaje para determinar la sensibilidad del dispositivo de captura.

La amplitud de la señal varía con respecto a su frecuencia debido a condiciones del ambiente acústico como el medio de transmisión o la naturaleza de la fuente, por lo que primero se determinó la frecuencia que tenía mejor respuesta para el nodo que con el que se caracterizó el sistema de captura.

Para caracterizar la respuesta, se generó una función chirp que contiene frecuencias que varían entre los 20 Hz. y los 600 Hz. Con la señal simulada se creó un archivo de audio tipo WAV que se transfirió a un dispositivo móvil para poder reproducirla mediante un altavoz.

Una vez generada la señal, se realizó un montaje para observar la respuesta de un altavoz y los micrófonos conectados a manera de grabación estereofónica para saber cuál es la frecuencia de mejor respuesta y el resultado se muestra en la Figura 3.3.

Con este montaje, se capturó la señal para poder determinar de cuál frecuencia se obtiene la mejor respuesta para el sistema de captura.

En la Figura 3.4 se observa el espectro de la señal capturada por los micrófonos obtenido mediante la aplicación de la transformada de Fourier. Observando el espectro de Fourier, es posible determinar cuáles frecuencias fueron capturadas por los micrófonos con una mayor amplitud.

Se tomó la decisión de utilizar la frecuencia correspondiente al valor máximo del espectro. Puede observarse en la Figura 3.4 que este máximo se encuentra alrededor de los 134 Hz. La amplitud de la señal capturada por el nodo es  $X(t)$ , que es diferente que la señal de todo el arreglo  $S(t)$ .



Figura 3.3: Montaje para determinar la respuesta del sistema como función de la frecuencia.

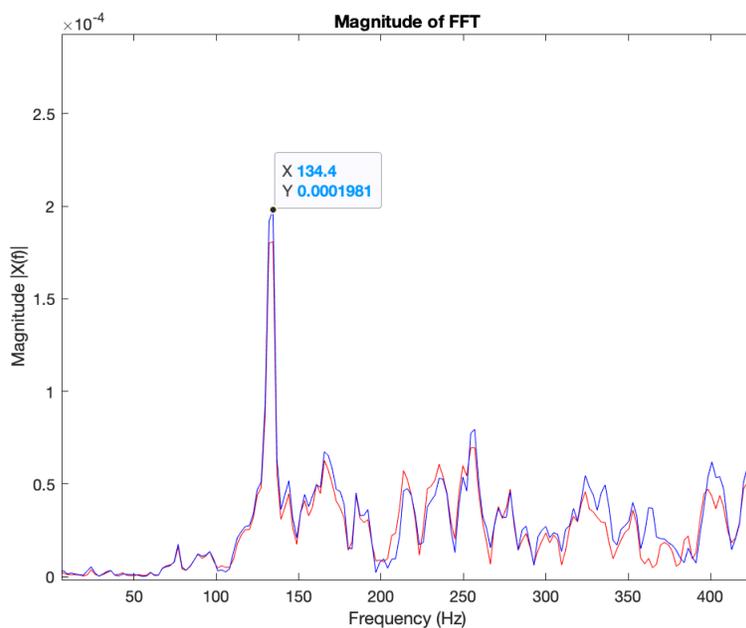


Figura 3.4: Frecuencia máxima del espectro en la captura de la señal de la fuente.

Emitiendo una señal sinusoidal con la frecuencia máxima, se realizó una captura de señal variando la distancia entre la fuente y un micrófono. Se calculó el valor RMS (*root mean square*, por sus siglas en inglés) de la señal capturada como indicador de la energía capturada por el nodo, y se graficó con respecto a la distancia para obtener una gráfica que reflejara la sensibilidad del sistema de captura.

Cuando se varía la posición de la fuente, también cambia su fase, esto debido al cambio en el

tiempo de arribo de la señal de la fuente al micrófono y la intensidad, ya que la amplitud de la onda acústica decrece con la distancia.

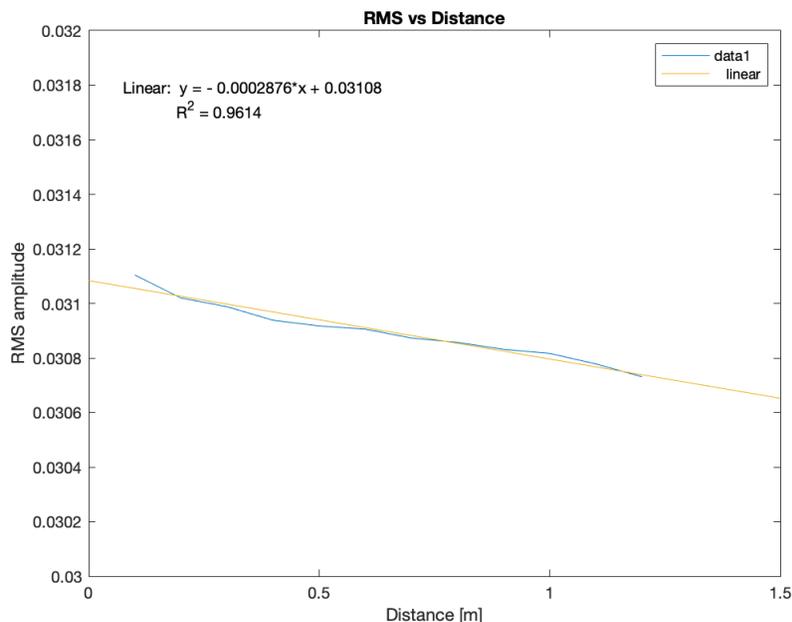


Figura 3.5: RMS de la señal capturada por el nodo en función de la posición de la fuente.

Por último, se calculó el valor RMS para cada una de las mediciones y se graficó contra la posición para observar el desempeño que tiene el sistema de captura. Los resultados se muestran en la Figura 3.5, en donde también se realizó un ajuste por mínimos cuadrados para interpolar la señal y se obtuvo la ecuación mostrada. La ecuación es:

$$y = -0.0002876x + 0.03108. \quad (3.1)$$

Con esta ecuación puede establecerse un umbral de separación entre las fuentes y los nodos de la WASN según sea necesaria la aplicación. Dependiendo del ambiente acústico en el que se desea generar un mapeo acústico, se establecerá la distribución de los nodos de la WASN, y tomando en cuenta este análisis de sensibilidad será posible determinar el alcance máximo para realizar la captura de la señal de una fuente acústica.

### 3.3. Selección de unidad de procesamiento

Se analizaron tres unidades de procesamiento que son capaces de conectarse con el micrófono seleccionado (Figura 3.1) y que tienen la capacidad de conexión con otros nodos dentro de la WASN: Arduino MKR Zero, Raspberry Pi Zero W y Raspberry Pi 4B+.

#### 3.3.1. Arduino MKR Zero

Se diseñó el primer prototipo del nodo de la WASN en donde cada nodo consiste en un Arduino MKR Zero, al cual se conecta un Adafruit I<sup>2</sup>S MEMS Microphone Breakout y un tranceptor de radio nRF24L0 con un ancho de banda de 2.4-2.5 GHz. como se muestra en la Figura 3.6.

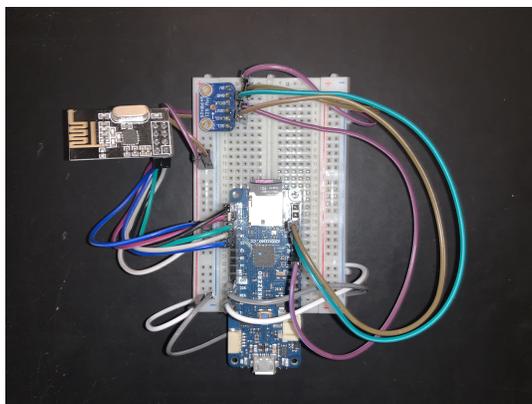


Figura 3.6: Primer prototipo de sistema de captura.

Este primer prototipo tiene la capacidad de capturar la señal y posteriormente transmitirla mediante la antena de radiofrecuencia a otro Arduino para la recolección de la información. La señal que desea capturarse corresponde al espectro de frecuencia de audición humana, es decir, desde los 20 Hz. hasta los 20 kHz.

La razón por la que este dispositivo no resultó ser funcional para el análisis de la señal capturada es que el Arduino tienen una capacidad limitada de almacenamiento de valores. La memoria SRAM, que es en donde se almacenan los valores medidos por el micrófono, es de 32 kB. Con esta capacidad, y las características que debe tener la señal capturada (48 kHz de frecuencia) sólo es posible obtener fragmentos de información en bloques de 1500 muestras aproximadamente.

Este número de muestras, a una frecuencia de 48 kHz, equivale 0.03125 segundos de información capturada de manera constante, después de este intervalo de captura, el dispositivo requiere de un tiempo para ejecutar alguna instrucción con la información almacenada (puede imprimirse en el monitor serial, transmitirse a otro dispositivo mediante la antena de radio o guardarse en la memoria microSD), y después vuelve al modo de captura de información.

Esta limitante en el funcionamiento del dispositivo hace que la información con la que se debe realizar el análisis acústico tenga intervalos de tiempo en donde el micrófono no está capturando señal acústica.

Se buscaron maneras de hacer que el tamaño del bloque de muestras fuera más grande o que el intervalo en el que el Arduino pudiera procesar la información almacenada se redujera. Aunque se logró reducir el tiempo en el que el micrófono no captura señal mientras espera a que el Arduino procese la información, nunca fue posible eliminar estos espacios en “blanco”. Al no lograr optimizar el procesamiento de la información, se buscó alguna alternativa para aumentar el tamaño de la información que es capaz de almacenar el Arduino antes de tener que procesarla.

Se realizó una investigación sobre los diferentes modelos de Arduino para determinar si existe alguna opción que cumpla los requerimientos. Ningún dispositivo Arduino cuenta con más de 32 kB, incluso, muchos de ellos tiene un tamaño de memoria inferior, por lo que se descartó cambiar de modelo. Se investigó la posibilidad de aumentar de manera externa la memoria SRAM y se encontró que existen chips para expandir la memoria. Se aumentó la memoria de un Arduino con un chip modelo 23LC1024 de la marca Microchip que permitió aumentar la memoria SRAM de 32 kB a 1024 kB, y aumentó el tiempo de captura 32 veces y permitiendo realizar capturas de aproximadamente 3 segundos. La expansión de la memoria SRAM es posible y se logró implementar, pero sólo para algunos modelos de Arduino, se aumentó la SRAM de un Arduino Mega, pero no pudo implementarse para los MKR Zero, por lo que esta opción tampoco es viable.

Para poder realizar el análisis de ambiente acústico que se pretende, es necesario tener más muestras, y por un intervalo mayor de tiempo. Por esta razón se propuso una nueva unidad de procesamiento. La segunda opción de unidad de procesamiento analizada es el Raspberry Pi Zero W.

### 3.3.2. Raspberry Pi Zero W y Raspberry Pi 4B+

Para solucionar los problemas observados del nodo de captura con unidad de procesamiento Arduino MKR Zero, se utilizó la unidad de procesamiento más sencilla de la familia de Raspberry; el Pi Zero W. El Raspberry Pi Zero W cuenta con conectividad inalámbrica 802.11 b/g/n wireless LAN, Bluetooth 4.1, procesador de 1 GHz. y 512 MB de RAM. Con este nodo es posible capturar señal acústica con una frecuencia de muestreo que va desde los 2 hasta los 192 kHz. Para los requerimientos del experimento, sólo es necesario grabar con una frecuencia de muestreo de 48000 Hz. Utilizando este prototipo es posible capturar señal en un archivo WAV que tiene un tamaño máximo de 2 GB. La cantidad de información que es posible capturar de manera continua con este prototipo es mucho mayor que la que se logró llegar a capturar con el nodo de Arduino. El montaje del nodo de captura se muestra en la Figura 3.7.

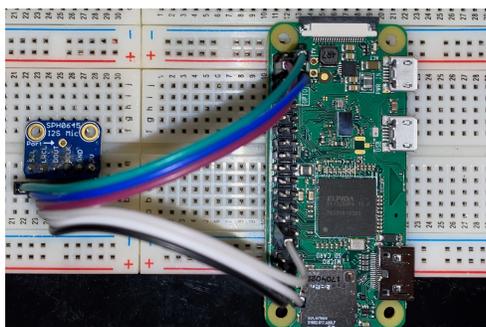


Figura 3.7: Nodo de las WASN con Raspberry Pi Zero W.

Esta unidad de procesamiento mejora la capacidad de captura de señal porque no presenta los espacios vacíos que no fue posible eliminar en el nodo con Arduino.

En la Tabla 3.4 se presenta la comparativa de los aspectos más importantes de las unidades de procesamiento que se tomaron en cuenta para tomar la decisión del trabajar con el Raspberry Pi Zero W.

Unidad de Procesamiento	Sensor acústico	Intervalo de captura continua	Conectividad inalámbrica	Unidad de almacenamiento	Costo [pesos]
Arduino MKR Zero	I <sup>2</sup> S MEMS Microphone	31.25 ms [espacios vacíos]	Transceptor nRF24L0	On board SD connector	950 aprox.
Raspberry Pi Zero W	I <sup>2</sup> S MEMS Microphone	Determinada por la capacidad de la SD	Wireless LAN 802.11 b/g/n	MicroSD intera	750 aprox.

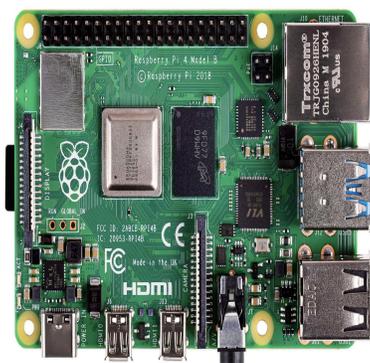
Tabla 3.4: Tabla comparativa de prototipos de sistema de captura.

El Raspberry Pi Zero W es una unidad adecuada para nuestra aplicación, ya que su precio permite la generación de una **WASN** de bajo costo y funcional. Conforme avanzó el proyecto, se realizó una mejora en la unidad de procesamiento con la que se genera la **WASN**, esto para determinar cómo afecta la capacidad de procesamiento del nodo en la captura de la señal y por consiguiente a la generación del mapa de energía acústica.

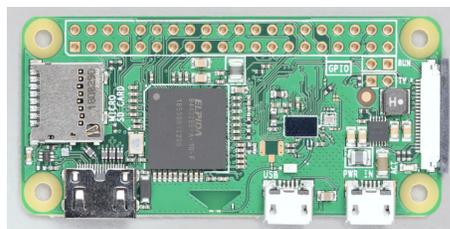
En este sentido, la mejora consistió en sustituir en los nodos de captura, el Raspberry Pi Zero W por un Raspberry Pi 4B+, ambos dispositivos se muestran en las Figuras 3.8a y 3.8b. Las ventajas que se obtuvieron al sustituir este dispositivo, no sólo tuvieron efecto en el procesamiento de la información, sino también en una mejora considerable en la comunicación entre los nodos de la red y los dispositivos que ejecutan las instrucciones, las características de ambas se muestran en la Tabla 3.5.

Modelo	Procesador	Memoria RAM	Conectividad	Precio
Pi Zero W	1 núcleo de 1 Ghz	512 MB	802.11 b/g/n 2.5 Ghz Bluetooth 4.0	540
Pi 4 B+	4 núcleos de 1.5 Ghz	1 GB	802.11 b/g/n/ac 5.0 GHz Bluetooth 5.0	1300

Tabla 3.5: Tabla con las características de los Raspberry Pi Zero W y 4B+.



(a) Raspberry Pi 4 B+



(b) Raspberry Pi Zero W

Figura 3.8: Unidades de procesamiento utilizadas para los nodos de la WASN.

Tomando en cuenta las características de las unidades de procesamiento, el nodo seleccionado para el diseño de la WASN tiene como unidad de procesamiento el Raspberry Pi 4 B+ y como sensor el Adafruit I2S MEMS Microphone Breakout.

### 3.4. Software de captura

El *software* utilizado para la captura de la señal es **JACK Audio Connection Kit** [26], software libre que permite capturar señal acústica en tiempo real. Este *software* debe instalarse en cada uno de los dispositivos Raspberry para poder utilizarlos como nodos dentro de la WASN.

Un ejemplo de comando para iniciar el servicio de JACK y poder comenzar a grabar con uno de sus agentes es:

```
jackd -dalsa -r48000 -p1024 -n4 -Chw:0
```

En donde `jackd` indica que se desea iniciar el servicio, `-dalsa` que se trabajará con el driver de ALSA, `-r48000` indica la frecuencia de muestreo (en este caso 48000 Hz.), `-p1024` el número de muestras por ventana, `-n4` el número de ventanas que se graban en el buffer antes de capturarse y `-Chw:0` el número de tarjeta de audio que realiza la captura.

Es importante explicar cómo funciona el *software* de **JACK**, ya que las metodologías desarrolladas se dividen en dos categorías:

La primera realizando grabaciones en donde un agente de JACK se ejecuta en cada dispositivo y graba la información localmente, y posteriormente se recopila la información en una unidad central para su procesamiento.

La segunda metodología realiza el proceso de grabación mediante la transmisión de la información capturada en cada nodo a través de internet para su almacenamiento en un nodo central, de esta manera, sólo es un nodo el que realiza la grabación de información, mientras que cada nodo sólo se encarga de la transmisión de su propia información.

Se utiliza un paquete llamado netjack2 [27], en donde se configura la red para que se tenga un nodo maestro y varios nodos esclavos. En el nodo maestro se ejecuta el cliente de JACK **netmanager** que comienza la comunicación, escuchando y recibiendo, a través del internet, la información capturada por cada uno de los esclavos. A su vez, cada esclavo inicia un cliente de JACK en donde establece la comunicación entre este nodo y el maestro, enviando, cada vez que se tenga información suficiente, paquetes de información con la señal capturada.



## Capítulo 4

# Modelo matemático para la generación del mapa de energía acústica

En este capítulo se desarrolla el modelo que permite la generación del mapa de energía acústico mediante el procesamiento de señales capturadas utilizando una WASN. Se desarrollan por separado el modelo para el campo cercano y para el campo lejano y se presentan los resultados obtenidos con los formadores de haz seleccionados.

### 4.1. Funcionamiento del modelo de mapeo de energía acústica

El modelo de mapeo de energía acústica propuesto toma en cuenta los elementos básicos del funcionamiento de un formador de haz para determinar la cantidad de energía como función de la posición.

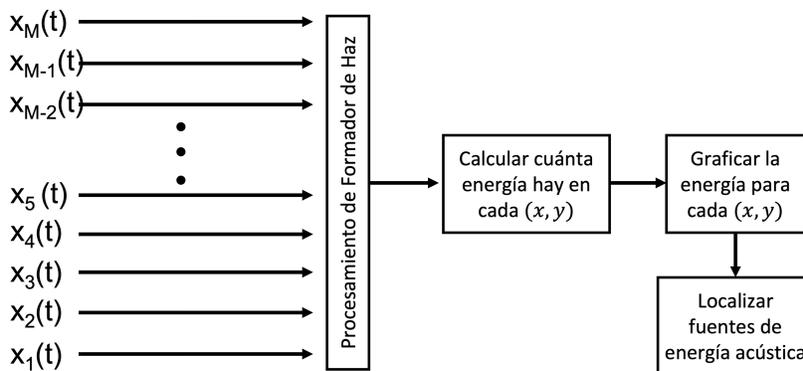


Figura 4.1: Modelo de mapeo de energía acústica.

Como se muestra en la Figura 4.1, el modelo consiste en calcular la energía en puntos  $(x, y)$  dentro del ambiente acústico para lo cual se propone generar una colección de “puntos de prueba”, en donde para cada uno de estos puntos, se realice un procesamiento de formador de haz para reconstruir la señal que proviene de ese punto, suponiendo que existe una fuente acústica, y después calcular la cantidad de energía que hay en ese punto. Por último, se grafica la cantidad de energía correspondiente a cada uno de los puntos de prueba. El resultado de este desarrollo produce una

gráfica como la que se muestra en la Figura 4.2 en donde sólo se representa la energía correspondiente a 5 puntos de prueba, pero en la implementación del modelo se generará toda una malla de puntos de prueba.

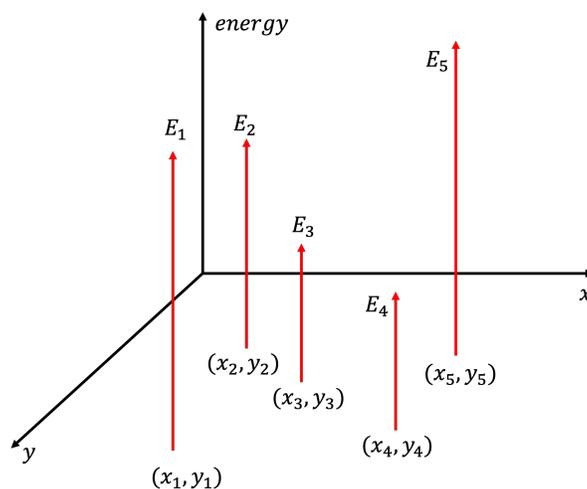


Figura 4.2: Mapa de energía acústica en el conjunto de puntos de prueba.

Una vez generado el mapa de energía acústica, es posible determinar la localización de fuentes acústicas, observando los puntos donde exista una mayor cantidad de energía acústica.

## 4.2. Campo lejano y campo cercano

Para generar el patrón de radiación de un formador de haz, lo primero que hay que considerar es la diferencia entre **campo lejano** y **campo cercano**, ya que el análisis de las señales es diferente para cada caso.

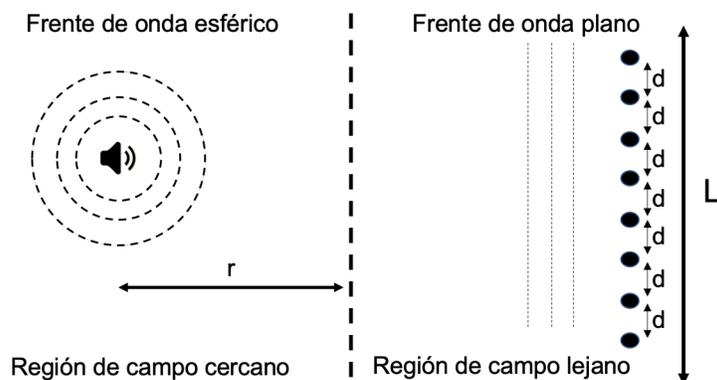


Figura 4.3: Frente de ondas acústicas para campo lejano y campo cercano.

En la Figura 4.3 se ejemplifican ambos casos. Se supone una fuente acústica, como puede observarse en la figura, en la región cercana a la fuente el frente de ondas acústicas es esférico, mientras el comportamiento del frente de ondas sea así, se considera **campo cercano**. A cierta distancia y sobre un área finita, los frentes de ondas pueden considerarse planos, y esto es conocido como **campo lejano**.

El valor de la distancia que separa ambos casos depende de la frecuencia de la onda sonora emitida por la fuente, de la velocidad de propagación de las ondas acústicas y del tamaño total de la WASN. De esta manera pueden definirse dos regiones, en la ecuación 4.1 está definida la condición que hace posible considerar el comportamiento como de campo cercano y en la ecuación 4.2 está definida la condición para considerar el caso del campo lejano[28].

$$r_c < \frac{L^2 f}{v}, \quad (4.1)$$

$$r_l > \frac{L^2 f}{v}. \quad (4.2)$$

donde,  $r_c$  representa todos los valores para los cuales el frente de onda se considera esférico,  $r_l$  representa la región donde los frentes de onda pueden considerarse planos,  $L$  es el tamaño total de la red de sensores,  $f$  es la frecuencia y  $v$  es la velocidad de propagación de las ondas, que en el caso del aire es  $v = 340 \frac{m}{s}$ .

#### 4.2.1. Modelo de mapeo de energía acústica para campo lejano

Suponiendo que se conoce la posición de la fuente,  $(x, y)$ , es posible determinar el ángulo  $\theta$  con el que el frente de ondas planas emitidas por la fuente llegan a la red de sensores mediante la expresión trigonométrica para la tangente:

$$\tan(\theta) = \frac{\text{cateto opuesto}}{\text{cateto adyacente}} = \frac{y}{x}, \quad (4.3)$$

por lo que el ángulo  $\theta$

$$\theta = \arctan\left(\frac{y}{x}\right). \quad (4.4)$$

El modelo para generar un mapa de energía acústica supone que en ciertos puntos en el espacio  $(x, y)$  se encuentra una fuente y se realiza el proceso de formador de haz para reconstruir la señal que proviene de esa dirección.

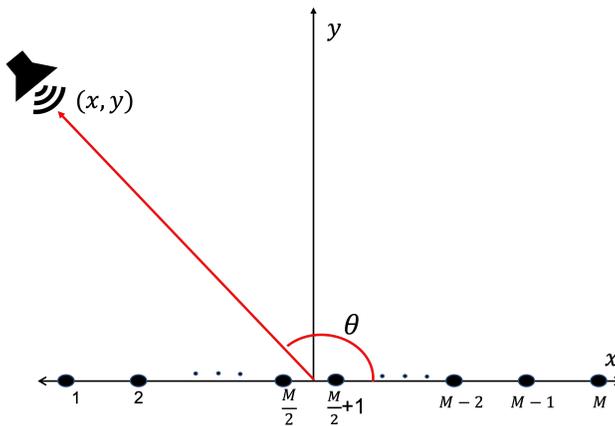


Figura 4.4: Frente de onda para una fuente en el campo lejano.

La geometría para una red de sensores en el campo lejano se presenta en la Figura 4.4. Suponiendo una red de sensores lineal que tiene una separación uniforme de  $d$  metros entre cada uno de sus nodos, la señal llegará primero a uno de los nodos, para este caso en particular el nodo que se

designa como el de referencia es el correspondiente a  $M = 1$ , con un tiempo de arribo de  $T_1 = 0$ , el tiempo que demorará en llegar al segundo nodo  $M = 2$  sería:

$$T_2 = \frac{d}{c} \sin \theta \quad (4.5)$$

en donde  $c = 343 \frac{m}{s}$  es la velocidad del sonido en el aire. Como los nodos están separados uniformemente, el vector de tiempos de arribo  $T$  quedaría escrito como:

$$T = \left[ 0 \quad \frac{d}{c} \sin \theta \quad \frac{2d}{c} \sin \theta \quad \frac{3d}{c} \sin \theta \quad \dots \quad \frac{(M-1)d}{c} \sin \theta \right]. \quad (4.6)$$

y como es posible determinar el ángulo  $\theta$ , conociendo las coordenadas  $(x, y)$ , el vector de tiempos de retardo sería:

$$T = \left[ 0 \quad \frac{d}{c} \sin \left( \arctan \left( \frac{y}{x} \right) \right) \quad \frac{2d}{c} \sin \left( \arctan \left( \frac{y}{x} \right) \right) \quad \dots \quad \frac{(M-1)d}{c} \sin \left( \arctan \left( \frac{y}{x} \right) \right) \right]. \quad (4.7)$$

Este es el concepto más importante para los formadores de haz, el vector que contiene las diferencias en los tiempos de arribo de la señal de la fuente a cada uno de los nodos de la WASN.

Para formar la señal, asumiendo que se conoce la posición  $(x, y)$  de la fuente, se calcula el desfase apropiado para cada micrófono y se suman para cada frecuencia.

$$S(f, x, y) = \sum_{i=1}^{M-1} \frac{W(f, x, y)^H X(f, x, y)}{M}. \quad (4.8)$$

en donde  $f$  es la frecuencia,  $W^H$  es el complejo conjugado del *steering vector*, y  $X$  es la matriz que contiene todas las señales capturadas por el arreglo de sensores. La señal reconstruida está en el espacio de las frecuencias, por lo que para obtener la señal en el espacio del tiempo, se aplica la transformada de Fourier inversa, denotada por  $\mathcal{F}\{\}^{-1}$ :

$$\mathcal{F}\{S(f, x, y)\}^{-1} = S(t, x, y). \quad (4.9)$$

Una vez obtenida la señal, la energía es una magnitud proporcional al promedio del cuadrado de la señal:

$$E \approx \|S(t, x, y)\|^2. \quad (4.10)$$

De esta manera puede generarse un mapa de energía acústica con el plano de los puntos de prueba y la cantidad de energía calculada para cada uno de ellos. Analíticamente, si existiera sólo una fuente de energía acústica, su posición se encontraría donde se concentra la mayor cantidad de energía. Para el caso de más de una fuente, tendría que analizarse cada uno de los máximos locales para determinar si se trata de una fuente o de algún efecto del ambiente acústico. El procedimiento para cada uno de los casos consiste en comparar el valor de la energía de un punto con las de sus alrededores, y considerar como un máximo local a aquel que tiene una energía superior a la de los puntos en su vecindad.

#### 4.2.2. Modelo de mapeo de energía acústica para campo cercano

El modelo anterior se propuso para solucionar el problema del cálculo de la energía acústica cuando la fuente se encuentra alejada de la red de sensores lo suficiente como para que el frente de ondas acústicas se considere como un frente de ondas plano.

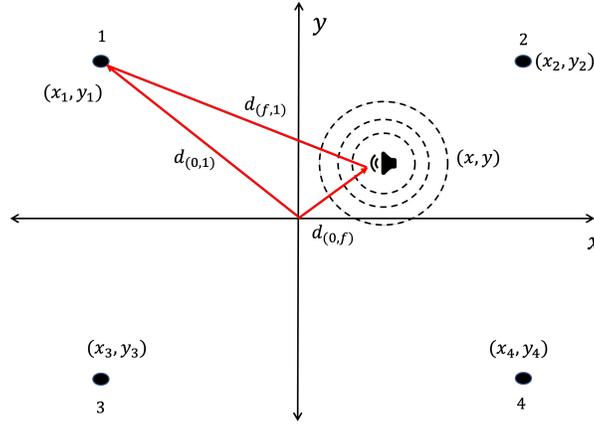


Figura 4.5: Frente de onda para una fuente en el campo cercano.

En este segundo desarrollo, se plantea el caso en el que la fuente de energía acústica se encuentra dentro de una red de sensores distribuida sobre una superficie. Para este caso, la propagación de las ondas no pueden considerarse como frente de ondas plano, por lo que el modelo tiene que modificarse, pero algunos conceptos que se utilizaron en el modelo anterior conservan su validez. La propagación de ondas en este modelo son esféricas como se muestra en la Figura 4.5, por lo que se tiene que calcular la diferencia en los tiempos de arribo de la señal a los nodos de la red de sensores, cada nodo se representa como un punto negro y tiene asociado su número dentro de la red y su posición.

En la Figura 4.5 se muestra una red de cuatro sensores, pero el análisis puede generalizarse para  $M$  sensores. Como puede observarse, una fuente en la posición  $(x, y)$ , tiene diferentes tiempos de arribo para cada uno de los sensores de la red. Los sensores tendrán posiciones:

$$X = [(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_M, y_M)]. \quad (4.11)$$

La distancia entre la fuente y cada uno de los  $M$  sensores ( $d_{f,i}$ ) puede calcularse de la siguiente manera:

$$d_{f,i} = \sqrt{(x_f - x_i)^2 + (y_f - y_i)^2}. \quad (4.12)$$

El tiempo que le toma a la señal en llegar a cada uno de los sensores, dependerá de la distancia de la fuente al sensor y la velocidad del sonido de la siguiente manera:

$$T_i = \frac{d_{f,i}}{c}. \quad (4.13)$$

con  $c = 344 \frac{m}{s}$ . De esta manera podemos obtener el vector de tiempos de arribo  $T$  de la fuente a cada uno de los sensores. El procedimiento para la generación del formador de haz es la misma que para el caso del campo lejano, teniendo como factor más importante el vector de tiempos de arribo de la señal a cada uno de los  $M$  nodos de la red de sensores:

$$T = [T_1 \quad T_2 \quad T_3 \quad \dots \quad T_M]. \quad (4.14)$$

o ya con los tiempos de arribo calculados:

$$T = \left[ \frac{\sqrt{(x_f - x_1)^2 + (y_f - y_1)^2}}{c} \frac{\sqrt{(x_f - x_2)^2 + (y_f - y_2)^2}}{c} \dots \frac{\sqrt{(x_f - x_M)^2 + (y_f - y_M)^2}}{c} \right]. \quad (4.15)$$

Entonces, volviendo al modelo planteado anteriormente, se propone una red de puntos  $(x, y)$  en donde se supone que existe una fuente, y para cada uno se calcula el desfase apropiado con respecto a cada nodo de la red. Para cada punto propuesto, se calcula la señal reconstruida, mediante la suma:

$$S(f, x, y) = \sum_{i=1}^{M-1} \frac{W(f, x, y)^H X(f, x, y)}{M}. \quad (4.16)$$

con:

$$W(f, x, y) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{-i2\pi f_1 T_2} & e^{-i2\pi f_2 T_2} & \dots & e^{-i2\pi f_N T_2} \\ e^{-i2\pi f_1 T_3} & e^{-i2\pi f_2 T_3} & \dots & e^{-i2\pi f_N T_3} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-i2\pi f_1 T_M} & e^{-i2\pi f_2 T_M} & \dots & e^{-i2\pi f_N T_M} \end{bmatrix}. \quad (4.17)$$

Para obtener la señal en el espacio del tiempo, se aplica la transformada inversa:

$$\mathcal{F}\{S(f, x, y)\}^{-1} = S(t, x, y). \quad (4.18)$$

Una vez obtenida la señal, la energía es una magnitud proporcional al promedio del cuadrado de la señal:

$$E \approx \|S(t, x, y)\|^2 \quad (4.19)$$

Puede generarse un mapa de energía acústica relacionando el plano de puntos de prueba con la energía acústica calculada para cada uno de ellos.

### 4.3. Formadores de haz para la generación del mapa de energía acústica

En esta sección se desarrollan los fundamentos conceptuales para los siguientes formadores de haz:

- Delay and sum (**DAS**).
- Minimum Variance Distortionless Response (**MVDR**).
- Phase-Based Frequency Masking (**PBM**).
- Steered Response Power with Phase Transform (**SRP-PHAT**).

Se analizó y caracterizó el funcionamiento de cada uno de estos formadores de haz para la generación del mapa de energía acústica.

### 4.3.1. Delay And Sum (DAS)

El formador de haz **Delay and sum** [29] tiene un funcionamiento como se muestra en la Figura 4.6. Para generar el formador de haz, se siguen los siguientes pasos:

- Se reciben señales con diferentes tiempos de arribo a los distintos nodos de sensores acústicos, tomando uno de los nodos como referencia, la señal llega con un retardo temporal ( $t_1, t_2, \dots, t_M$ ) a los demás nodos dentro de la red.
- El formador de haz calcula una matriz que compensa los retardos  $W^H$  (*steering vector*) para que hacer que las señales de cada uno de los nodos tengan el mismo retardo que la señal del primer nodo.
- Cuando las señales tienen el mismo retardo temporal, se suman para generar el patrón de radiación de la señal amplificada en la dirección deseada.

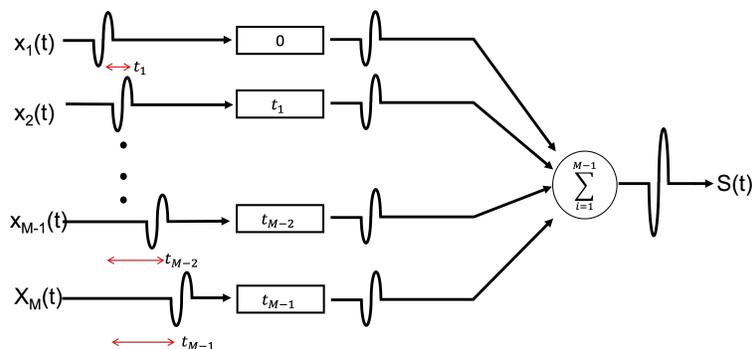


Figura 4.6: Diagrama del funcionamiento del formador de haz **DAS**.

Hay que tener en cuenta que aplicación del *steering vector* sólo hace que las señales en cada nodo se amplifiquen cuando la dirección seleccionada sea la de la fuente, para cualquier otra dirección, que generaría una matriz diferente, se obtendría un resultado como el que se muestra en la Figura 4.7.

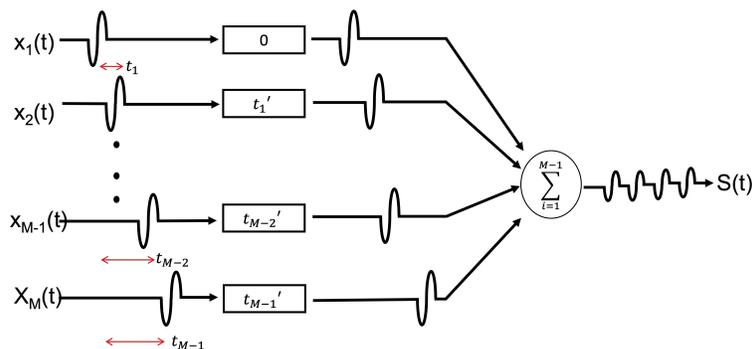


Figura 4.7: Suma de las señales con tiempos de arribo distintos.

El modelo genera la señal de salida en la dirección deseada, multiplicando el llamado *steering vector*, que contiene los retardos para los nodos de la red de acuerdo a la dirección de arribo de la señal de la fuente, por la matriz de señales capturadas  $\mathbf{X}$ , que contiene las muestras capturadas por los  $M$  nodos de la red de sensores.

$$\hat{S} = W^H X, \quad (4.20)$$

de donde:

$$X(f) = \begin{bmatrix} x_1(f_1) & x_1(f_2) & \cdots & x_1(f_N) \\ x_2(f_1) & x_2(f_2) & \cdots & x_2(f_N) \\ \vdots & \vdots & \ddots & \vdots \\ x_M(f_1) & x_M(f_2) & \cdots & x_M(f_N) \end{bmatrix}. \quad (4.21)$$

es la matriz de señales capturadas en el espacio de las frecuencias.

El vector de tiempos de arribo de la señal de la fuente con cada uno de los nodos de la red, tomando un nodo como el de referencia,  $T_1 = 0$ , puede escribirse de la siguiente manera:

$$T = [T_1 \ T_2 \ T_3 \ \cdots \ T_M] = [0 \ T_2 \ T_3 \ \cdots \ T_M] \quad (4.22)$$

De esta manera, el *steering vector*, también en el espacio de las frecuencias, tiene en su primer renglón al nodo de referencia.

$$W(f, x, y) = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-i2\pi f_1 T_2} & e^{-i2\pi f_2 T_2} & \cdots & e^{-i2\pi f_N T_2} \\ e^{-i2\pi f_1 T_3} & e^{-i2\pi f_2 T_3} & \cdots & e^{-i2\pi f_N T_3} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-i2\pi f_1 T_M} & e^{-i2\pi f_2 T_M} & \cdots & e^{-i2\pi f_N T_M} \end{bmatrix}. \quad (4.23)$$

Debido a que, tanto la matriz de las señales capturadas como la matriz de *steering vectors* existen en el espacio de frecuencias, se obtendrá la señal a la salida en este mismo espacio frecuencias, por lo que hará falta realizar una transformada inversa de Fourier para obtener la salida en el espacio del tiempo.

La salida del formador de haz puede escribirse como:

$$S(f, x, y) = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-i2\pi f_1 T_2} & e^{-i2\pi f_2 T_2} & \cdots & e^{-i2\pi f_N T_2} \\ e^{-i2\pi f_1 T_3} & e^{-i2\pi f_2 T_3} & \cdots & e^{-i2\pi f_N T_3} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-i2\pi f_1 T_M} & e^{-i2\pi f_2 T_M} & \cdots & e^{-i2\pi f_N T_M} \end{bmatrix} \begin{bmatrix} x_1(f_1) & x_1(f_2) & \cdots & x_1(f_N) \\ x_2(f_1) & x_2(f_2) & \cdots & x_2(f_N) \\ \vdots & \vdots & \ddots & \vdots \\ x_M(f_1) & x_M(f_2) & \cdots & x_M(f_N) \end{bmatrix}. \quad (4.24)$$

Cabe recordar que la señal obtenida  $S(f, x, y)$ , corresponde a la fuente que se encuentra en una posición previamente conocida. En el modelo que se propone para generar el mapa de energía, se va a calcular la energía para  $p$  puntos de prueba, por lo que para graficar se relacionará la energía de los  $p$  puntos de prueba  $\{x_1, x_2, \dots, x_p\}$ .

### 4.3.2. Minimum Variance Distortionless Response (MVDR)

Antes de analizar el formador de haz **MVDR** [30, 31, 32, 33], se necesita definir la matriz de covarianza.

Considerando una red de  $M$  micrófonos y una señal capturada de  $n$  muestras, la señal de entrada del arreglo de sensores puede modelarse como  $x(n) = A(\theta)S(n) + \eta(n)$  de donde:  $A(\theta) = [\vec{a}(\theta_1) \ \vec{a}(\theta_2) \ \cdots \ \vec{a}(\theta_M)]$  es la matriz de direcciones, y  $\vec{a}(\theta)$  es el vector de direcciones.

Al definir la señal de esta manera se separa la señal del ruido, utilizando la propiedad de ortogonalidad a través de la eigendescomposición de la matriz de covarianza. Se supone que el ruido adquirido por los micrófonos no está correlacionado con la señal fuente:  $r_{sn}(l) = 0 = r_{ns}(l)$ .

La matriz de covarianza de la salida de la red se define como la esperanza matemática  $\mathcal{E} \{ \cdot \}$  de la energía de la señal.

$$\hat{R}_{xx} = \mathcal{E} \{ x(n)x^H(n) \}. \quad (4.25)$$

Si la señal es real, se tiene que:  $x(n)^H = x(n)^T$ , donde  $H$  es hermitiano y  $T$  es traspuesto. Por lo que la expresión anterior queda como:

$$\mathcal{E} \{ x(n)x^T(n) \} = \mathcal{E} \{ (A(\theta)S(n) + \eta(n)) (A(\theta)S(n) + \eta(n))^T \} = A(\theta)\vec{R}_{ss}A(\theta)^T + \vec{R}_{nn}. \quad (4.26)$$

con  $\hat{R}_{ss} = \{ s(n)s^T(n) \}$  y  $\vec{R}_{nn} = \{ \eta(n)\eta^T(n) \}$ . En la práctica,  $R_{xx}$  puede estimarse como:

$$\hat{R}_{xx} = \frac{1}{N} \sum_{n=1}^N x(n)x^H(n). \quad (4.27)$$

si  $N \rightarrow \infty$ ,  $\hat{R}_{xx} \rightarrow R_{xx}$ , para la aplicación deseada, el número de sensores no puede hacerse infinito, pero se considera suficiente para la implementación de este formador de haz, por lo que la ecuación 4.27 puede reescribirse como:

$$\hat{R}_{ss} = \hat{A}(\theta)\hat{R}(ss)\hat{A}^H(\theta) + \sigma_\eta^2 I = \sigma_s A(\theta)A^H(\theta) + \sigma_\eta I. \quad (4.28)$$

donde,  $\hat{A}(\theta)$  es la matriz de direcciones,  $\vec{a}(\theta)$  es el vector de dirección de cada señal,  $\sigma_s$  es la varianza de la señal,  $\sigma_\eta$  es la varianza del ruido e  $I$  es la matriz identidad de  $MXM$ . Para la aplicación del formador de haz con señales capturadas por la WASN, la matriz que contiene estas señales puede escribirse como:  $x(n) = X(f, x, y)$ , ya que representan el mismo conjunto de elementos.

Este algoritmo optimiza la energía de la señal de salida de un formador de haz, minimizando el ruido de la señal sin distorsionar la señal deseada. Como ya se había planteado anteriormente, la salida de un formador de haz es:  $S(f, x, y) = W(f, x, y)^T X(f, x, y)$ .

La energía  $\|S(f, x, y)\|^2 = (W^T X(f, x, y)) (W^T X(f, x, y)) = W^T X(f, x, y)X^T(f, x, y)W$ , de esta expresión, puede calcularse el valor esperado como:

$$E \{ y^2(f) \} = E \{ W^T x(f)x^T(f)W \} = W^T E \{ x(f)x^T(f) \} W = W^T R_{xx} W. \quad (4.29)$$

señales capturadas por una WANS se consideran señales reales, por lo que  $H = T$  y se cumple que:  $E \{ x(f)x^T(f) \} = R_{xx}$ .

Se minimiza la potencia de las señales de interferencia provenientes de direcciones no deseadas  $\theta$ , pero simultáneamente se mantiene una ganancia fija en la dirección de interés.

Lo que se busca es minimizar  $W^T R_{xx} W$  con la condición de que  $W^H \vec{a}(\theta) = 1$ , de donde se tiene que  $W^H \vec{a}(\theta) - 1 = 0$ . Para resolver problemas de optimización, uno de los métodos que pueden emplearse es el de los **multiplicadores de Lagrange**, que es un procedimiento para encontrar

los máximos y mínimos relativos de funciones de multivariantes que se encuentran sujetas a ciertas restricciones. Este método reduce el problema restringido con  $n$  variables a uno sin restricciones de  $n + k$  variables, donde  $k$  es igual al número de restricciones, y cuyas ecuaciones pueden ser resueltas más fácilmente. Se plantea la ecuación 4.30.

$$H(W) = \frac{1}{2}W^T R_{xx}W + \lambda (W^H a(\theta) - 1). \quad (4.30)$$

Para calcular los coeficientes óptimos, se calcula:

$$\nabla_w H(W) = \frac{\partial H(W)}{\partial W_i} = 0. \quad (4.31)$$

Los coeficientes óptimos  $W_{op}$  están asociados con la restricción de la salida óptima, dada por:  $S_{op}(n) = W_{op}^T x(n)$  en la dirección  $\theta$  deseada. Desarrollando la ecuación 4.31, se obtiene:

$$\frac{\partial H(W)}{\partial W_i} = R_{xx}W + \lambda a(\theta) = 0 \implies W_{op} = -\lambda R_{xx}^{-1} a(\theta). \quad (4.32)$$

Sustituyendo en la restricción, se obtiene una expresión para  $\lambda$ :

$$(-\lambda R_{xx}^{-1} a(\theta))^T a(\theta) = 1 \implies -\lambda a(\theta)^T R_{xx}^{-1} a(\theta) = 1 \implies \lambda = \frac{1}{a(\theta)^T R_{xx}^{-1} a(\theta)}. \quad (4.33)$$

De esta manera se obtiene que:

$$W_{op} = \frac{R_{xx}^{-1} a(\theta)}{a(\theta)^T R_{xx}^{-1} a(\theta)} \implies S_{op}(n) = W_{op}^T x(n) = \left( \frac{R_{xx}^{-1} a(\theta)}{a(\theta)^T R_{xx}^{-1} a(\theta)} \right)^T X(f). \quad (4.34)$$

La salida del formador de haz para una posición previamente conocida de la fuente es:

$$S(f, x, y) = W_{op}^T x(f) = \left( \frac{R_{xx}^{-1} a(\theta)}{a(\theta)^T R_{xx}^{-1} a(\theta)} \right)^T X(f). \quad (4.35)$$

De igual modo que en el formador DAS, se proponen una colección de puntos de prueba en donde se calcula la energía y se realiza la gráfica para cada uno de ellos.

### 4.3.3. Phase-Based Frequency Masking (PBM)

Para este formador se propone una matriz de  $M \times N$  que contiene las señales capturadas por una red de micrófonos como se muestra en ecuación 4.21. Suponiendo que la fuente se encuentra en una posición definida por las coordenadas  $(x, y)$ , la primera etapa del procesamiento de este formador de haz es alinear las señales de entrada utilizando el *steering vector* ( $W(f, x, y)$ ).

$$W(f, x, y) = \begin{bmatrix} e^{-i2\pi f_1 T_1} & e^{-i2\pi f_2 T_1} & \dots & e^{-i2\pi f_N T_1} \\ e^{-i2\pi f_1 T_2} & e^{-i2\pi f_2 T_2} & \dots & e^{-i2\pi f_N T_2} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-i2\pi f_1 T_M} & e^{-i2\pi f_2 T_M} & \dots & e^{-i2\pi f_N T_M} \end{bmatrix}. \quad (4.36)$$

El resultado de la multiplicación genera una nueva matriz de señales alineadas en la dirección de la fuente de interés:

$$X_a(f, x, y) = X(f, x, y)W(f, x, y). \quad (4.37)$$

Como puede observarse, en el *steering vector* hay un vector que contiene los diferentes tiempos de arribo con los que llega a cada micrófono de la red la señal emitida por la fuente de interés:

$$T = [T_1 \quad T_2 \quad \cdots \quad T_M]. \quad (4.38)$$

La diferencia de fase promedio calculada para cada frecuencia  $f$  es:

$$|\varphi|_f = \frac{2}{M(M-1)} \sum_{i=1}^{M-1} \sum_{j=i+1}^M |\varphi_{i,f} - \varphi_{j,f}|. \quad (4.39)$$

donde  $M$  es el número de micrófonos y la fase  $\varphi_i$  en el  $i$ -ésimo micrófono.

Esta suma determinará si la frecuencia que se está analizando después de alinearla proviene de la fuente de interés o del componente de interferencia-ruido del ambiente acústico.

Para la implementación de este algoritmo, es necesario definir un límite  $\varphi_{max}$  en la magnitud de  $\varphi_f$  con lo cual puede generarse un par de filtros, llamados “*máscaras de frecuencia*” [34, 35] en donde se separarán las señales que provienen de la fuente y las señales que provienen de la interferencia-ruido del ambiente. El valor del límite  $\varphi_{max}$  depende de la aplicación deseada y provee un umbral que podrá amplificar las señales provenientes de la fuente de interés y reducir las que provengan de otros elementos acústicos como la interferencia o el ruido.

$$P_{fuente}[f] = \begin{cases} 1, & \text{if } \varphi_f \geq \varphi_{max} \\ 0, & \text{cualquier otro valor} \end{cases} \quad (4.40)$$

$$P_{int-ruido}[f] = \begin{cases} 0, & \text{if } \varphi_f \geq \varphi_{max} \\ 1, & \text{cualquier otro valor} \end{cases} \quad (4.41)$$

Por último, se recupera la señal realizando la multiplicación matricial de la máscara obtenida con la señal de entrada en el micrófono  $m = 1$ . La ecuación para la señal recuperada se muestra en la siguiente ecuación:

$$S(f, x, y) = P_{fuente}[f] * X[1, f]. \quad (4.42)$$

#### 4.3.4. Steered Response Power with Phase Transform (SRP-PHAT)

Este formador de haz busca una posición propuesta que maximice la salida de un formador de haz **DAS**, pero aplicando un *steering vector* para alinear las señales de entrada [36, 37].

Considerando una red de  $M$  micrófonos, donde cada micrófono tiene una señal discreta temporal de la forma  $x_n(t)$ . La salida de un formador de haz **DAS** para las coordenadas  $(\theta, r)$ (polares) y  $(x, y)$  (cartesianas) en el dominio de la frecuencia es:

$$S(f, \theta) = W(\theta)^H X(f). \quad (4.43)$$

La potencia de esta señal está dada por:

$$P(\theta) = \int_{-\infty}^{\infty} |S(f, \theta)|^2 df. \quad (4.44)$$

Esta potencia, recibe el nombre de **Steered Response Power** y puede escribirse de la siguiente manera:

$$P(\theta) = \sum_{n=1}^{M-1} \sum_{m=n+1}^{M-1} R_{n,m}(\tau_{n,m}). \quad (4.45)$$

El factor  $R_{n,m}$  corresponde a la correlación cruzada entre el  $n$ -ésimo y el  $m$ -ésimo micrófonos. La expresión que calcula esta correlación cruzada es:

$$R_{n,m}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi_{n,m}(f) X_n(f) X_m^*(f) e^{ift} df, \quad (4.46)$$

de donde  $\Psi_{n,m}$  corresponde a la función de pesos, y aplicando la transformada de fase, que permite homogeneizar los pesos de las frecuencias, se obtiene:

$$\Psi_{n,m}^{PHAT}(f) = \frac{1}{|X_n(f) X_m^*(f)|}. \quad (4.47)$$

Para el procesamiento, en lugar de realizar una integral en todo el espacio, se realiza suma para todos los valores posibles de frecuencia, que escrito de manera discreta,  $R_{n,m}(\tau)$  es:

$$R_{n,m}(\tau) = \sum_{l=1}^L \frac{X_n(f) X_m^*(f)}{|X_n(f) X_m^*(f)|} e^{if(l)\tau_{n,m}}. \quad (4.48)$$

Es en este punto donde se calcula la diferencia en los tiempos de arribo de la señal de la fuente con respecto a cada uno de los puntos de prueba en donde se quiere calcular la respuesta de la potencia, este retraso temporal puede escribirse como:

$$\tau_{n,m} = \frac{r_n - r_m}{c} = \frac{\sqrt{(x_q - x_n)^2 + (y_q - y_n)^2} - \sqrt{(x_q - x_m)^2 + (y_q - y_m)^2}}{c}, \quad (4.49)$$

donde  $c$  es la velocidad del sonido,  $(x_n, y_n)$ ,  $(x_m, y_m)$ ,  $(x_q, y_q)$  son las coordenadas cartesianas del micrófono  $n$ , del micrófono  $m$  y del punto de prueba en donde quiere determinarse la potencia.

Entonces, la salida de este formador de haz para una posición específica de coordenadas  $(x, y)$  puede escribirse simplemente como:

$$S(f, x, y) = \sum_{n=1}^{M-1} \sum_{m=n+1}^{M-1} R_{n,m}(x, y)(\tau_{n,m}(x, y)). \quad (4.50)$$

## Capítulo 5

# Caracterización del modelo de mapeo acústico

Este capítulo se enfoca en mostrar los resultados al aplicar los cuatro modelos de formadores de haz seleccionados para la generación del mapa de energía acústica. Se probaron estos modelos con señales simuladas, a fin de caracterizar el desempeño de cada uno de ellos.

Los elementos que se consideran para la caracterización de los diferentes formadores de haz son:

- Tamaño de muestra.
- Número de micrófonos.
- Puntos de prueba.
- Ruido.
- Variación de la frecuencia.
- Múltiples fuentes sonoras.

Para el tamaño de muestra, número de micrófonos y puntos de prueba se considera el efecto que tienen estos parámetros con el tiempo de procesamiento y la forma del mapa de energía obtenido. Mientras que para el ruido, la frecuencia, y la presencia de múltiples fuentes sonoras en el ambiente acústico, se analiza cómo afecta su presencia a la forma del mapa de energía obtenido.

### 5.1. Ejemplo para la generación de un mapa de energía acústica

Para la generación del mapa de energía, utilizando los cuatro formadores de haz, se simuló una red de sensores cuadrada de 25 micrófonos, 5 filas y 5 columnas de micrófonos. Las coordenadas de las posiciones de los micrófonos ubican al centro de la red el punto  $(0, 0)$  como se muestra en la Figura 5.1.

Se simuló una fuente puntual en las coordenadas  $(0, 0)$ . Esta fuente emite una señal cosenoidal con dependencia temporal de la forma  $s(t) = \cos(2\pi\omega t)$ , en donde  $\omega$  es la frecuencia angular de la señal y  $t$  es el tiempo. La señal se simuló un segundo con una frecuencia de muestreo de 1000 muestras por segundo. Las señales de entrada del algoritmo se muestran en la Figura 5.2.

Con esta señal, se procede a proponer los puntos de prueba en donde se aplican los diferentes formadores de haz para recuperar la señal en ese punto y posteriormente se calcula la energía para cada uno de estos puntos.

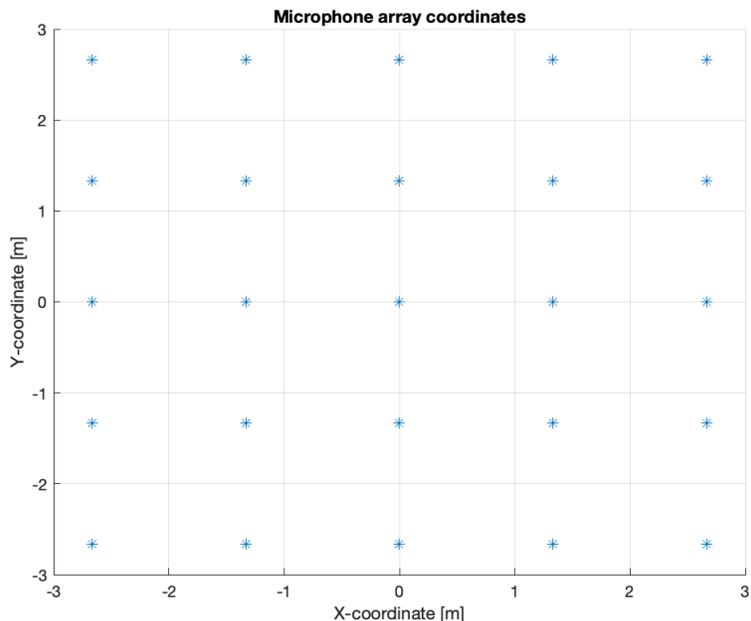


Figura 5.1: Red de sensores elegida para la simulación.

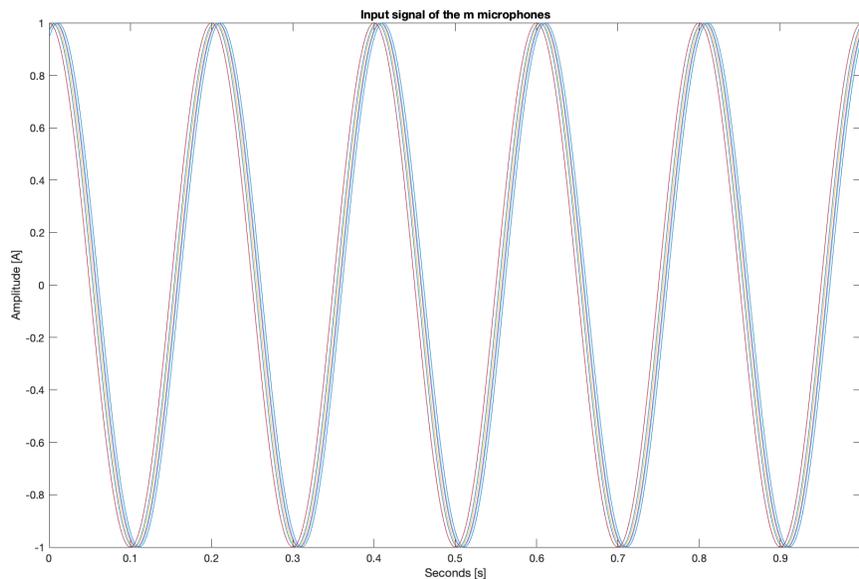


Figura 5.2: Señales de entrada capturada por los nodos de la red.

Para los resultados mostrados, se propusieron 625 puntos de prueba, distribuidos en 25 filas y 25 columnas de puntos. La malla de puntos de prueba tiene su centro coincidente con el de la red de los micrófonos y los puntos están separados de manera uniforme como se muestra en la Figura 5.3. Aplicando el modelo, se supone que en cada uno de estos  $n$  puntos de prueba hay una “fuente”, el vector de tiempos de arribo de la señal emitida por esta “fuente” al  $i$ -ésimo micrófono de la red se

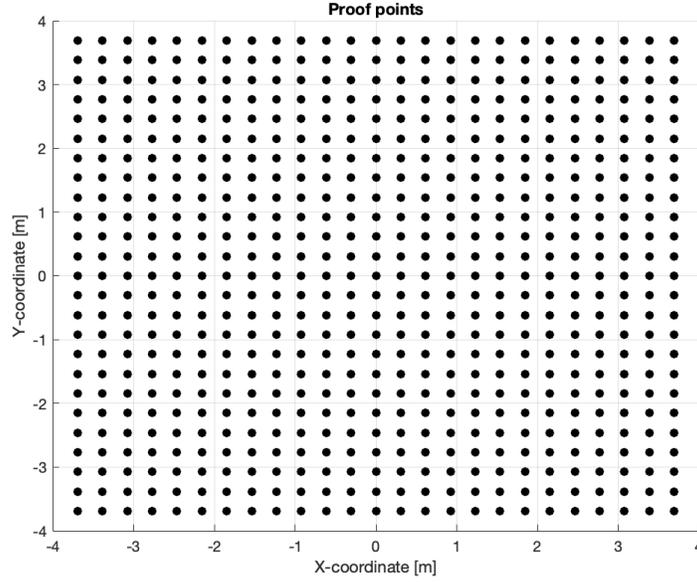


Figura 5.3: Malla de puntos de 625 prueba propuestos.

calcula de la siguiente manera:

$$T(i) = \left[ \frac{\sqrt{(x_i - x_1)^2 + (y_i - y_1)^2}}{c} \frac{\sqrt{(x_i - x_2)^2 + (y_i - y_2)^2}}{c} \dots \frac{\sqrt{(x_i - x_M)^2 + (y_i - y_M)^2}}{c} \right]. \quad (5.1)$$

Recordando el principio bajo el cual los formadores de haz funcionan (sección 4), cada uno de los puntos de prueba tendrá su correspondiente *steering vector*.

$$W(i) = \begin{bmatrix} e^{-i2\pi f_1(i)T_1(i)} & e^{-i2\pi f_2(i)T_1(i)} & \dots & e^{-i2\pi f_N(i)T_1(i)} \\ e^{-i2\pi f_1(i)T_2(i)} & e^{-i2\pi f_2(i)T_2(i)} & \dots & e^{-i2\pi f_N(i)T_2(i)} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-i2\pi f_1(i)T_M(i)} & e^{-i2\pi f_2(i)T_M(i)} & \dots & e^{-i2\pi f_N(i)T_M(i)} \end{bmatrix}. \quad (5.2)$$

El mapa de energía acústica obtenido mediante cada uno de los formadores de haz se obtendrá graficando la energía de la señal reconstruida en cada uno de los puntos de prueba, a los que se denominarán  $n_p$ , suponiendo que una hay una fuente en cada uno de ellos. La energía y la señal reconstruida en los puntos de prueba son:  $E(n_p)$ , y  $S(n_p)$ , respectivamente y la energía puede obtenerse como:

$$E(n_p) \approx \|S(n_p)\|^2. \quad (5.3)$$

Los resultados se muestran en la Figura 5.4, en donde cada uno de los mapas obtenidos grafican la energía calculada para cada uno de los puntos de prueba en un mapa tridimensional y en la Figura 5.5 en donde una escala de color indica la cantidad de energía asociada con ese punto de prueba. La caracterización del modelo de mapeo acústico evalúa aspectos como el tiempo de procesamiento y la forma del mapa de energía acústica en la implementación de los diferentes formadores de haz. Para lograr esto, se realizó la simulación variando uno de los elementos y se midió algún aspecto en la generación del mapa de energía acústica.

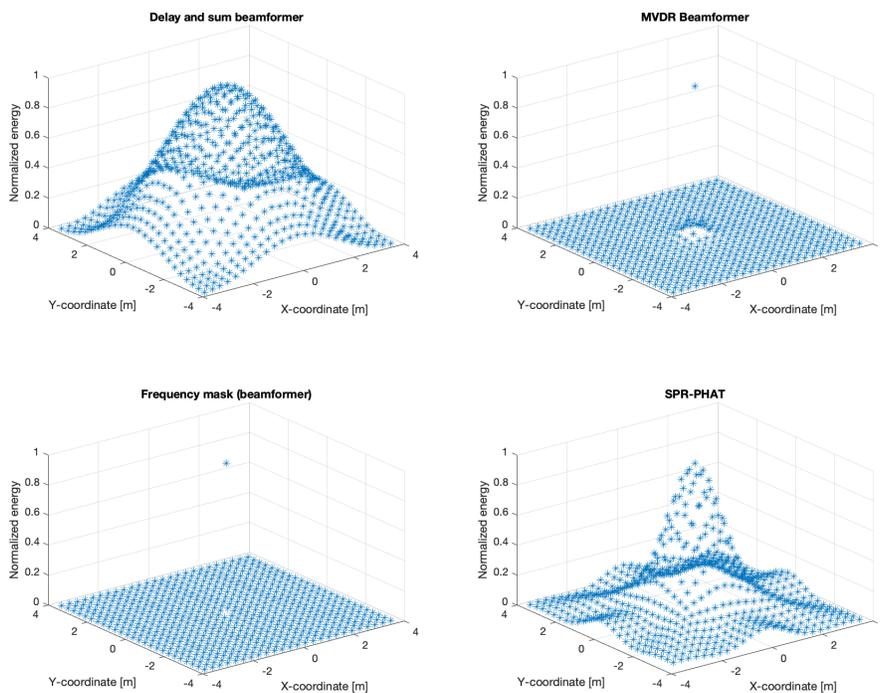


Figura 5.4: Mapas de energía generados con los diferentes formadores de haz.

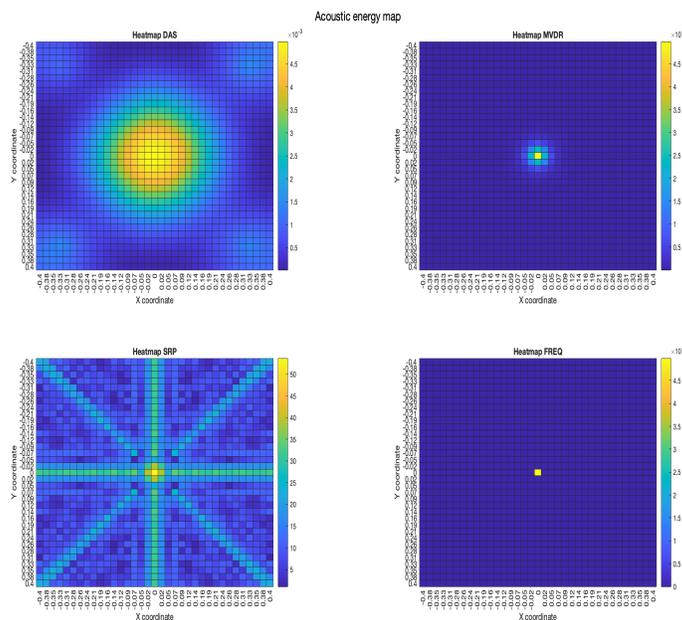


Figura 5.5: Mapas de calor generados con los diferentes formadores de haz.

Para las dependencias del tiempo de procesamiento contra el tamaño de muestra, el número de micrófonos y la cantidad de puntos de prueba, se varió el valor de cada uno de estos y se midió el tiempo que tomaba a cada uno de los modelos generar el mapa de energía acústica.

Para observar el funcionamiento de los diferentes formadores de haz se analizó la forma del mapa de energía acústica al variar los siguientes parámetros: ruido, frecuencia, número de fuentes.

## 5.2. Tiempo de procesamiento

### 5.2.1. Dependencia con el tamaño de muestra

La dependencia del tiempo de procesamiento y el tamaño de muestra que se desea analizar es importante, ya que esto determinará el tiempo necesario que se requiere para obtener un mapa de energía acústica. La caracterización se realizó con una señal simulada y con los siguientes parámetros: 4 micrófonos, una fuente acústica puntual localizada en el centro del arreglo, frecuencia de la señal cosenoidal de 134 Hz y 625 puntos de prueba, las características de la WASN consideraron un número menor de micrófonos que en el ejemplo para poder hacerla de bajo costo. Para la caracterización del tiempo de procesamiento y su dependencia con el tamaño de muestra que se analiza, se tomaron valores de 2000 muestras y aumentando en 2000 muestras se llegó a 48000 muestras con una frecuencia de muestreo de 48000 Hz. Los resultados se observan en la Figura 5.6.

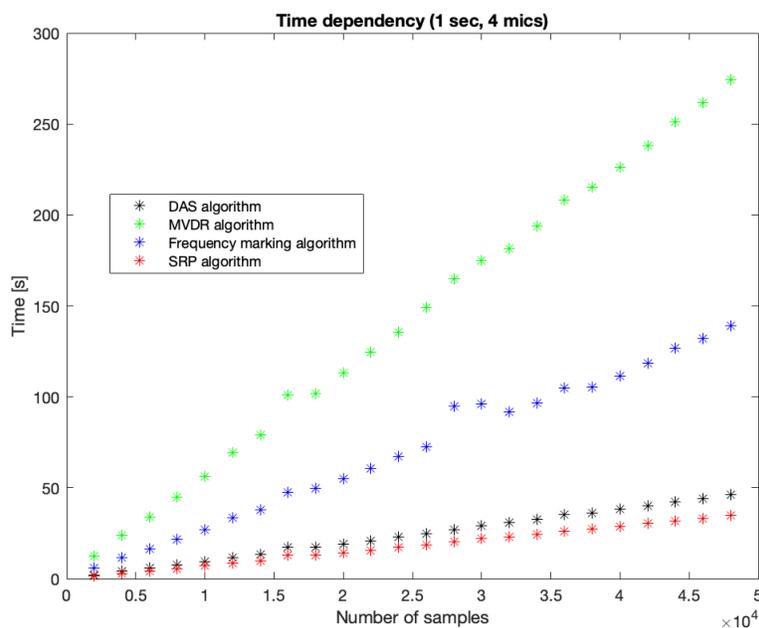


Figura 5.6: Tiempo de procesamiento contra número de muestras.

Como puede observarse, el algoritmo que presenta una razón de crecimiento mayor en cuanto a su tiempo de procesamiento en relación con el número de muestras es el formador de haz MVDR, seguido del PBM, mientras que los algoritmos DAS y SRP-PHAT tienen una razón de crecimiento similar, pero con un orden de magnitud menor.

Con los datos obtenidos puede realizarse un análisis estadístico. Como la relación del tiempo de procesamiento y el número de muestras parece lineal, se realizó un ajuste mediante mínimos cuadrados, y los resultados para todos los algoritmos se muestran en la tabla 5.1. Al realizar el

Formador de haz	Ecuación lineal	Coficiente de determinación ( $R^2$ )	Incremento normalizado
DAS	$y = 0.000955x + 0.201$	0.9989	1.3197
MVDR	$y = 0.005679x + 1.136$	0.9989	7.8482
PBM	$y = 0.002898x - 0.6289$	0.9904	4.0049
SRP-PHAT	$y = 0.0007236x - 0.01442$	0.9989	1

Tabla 5.1: Regresión lineal para la relación tiempo de procesamiento contra el número de muestras de las señales capturadas.

análisis se normalizó el incremento en el tiempo de procesamiento al menor tiempo de ejecución que corresponde al algoritmo SRP-PHAT y se nota que nuevamente que la diferencia entre el tiempo de procesamiento entre SRP-PHAT y DAS no es tan considerable, ya que sólo aumenta en 1.3197 veces el tiempo de procesamiento. El algoritmo PBM aumenta en una proporción de 4.0049 veces el tiempo de procesamiento y el algoritmo MVDR aumenta en 7.8482 veces el tiempo de procesamiento con respecto al tiempo más bajo que es SRP-PHAT.

### 5.2.2. Dependencia con el número de micrófonos

El número de micrófonos también es un factor muy importante, ya que si se desea generar una WASN con más nodos de captura de señal, el tiempo de procesamiento no debe aumentar como para que no sea viable la implementación del modelo, este parámetro se define dependiendo de la aplicación.

La red de sensores consiste en un arreglo cuadrado con el mismo número de micrófonos en columnas y renglones. Para realizar la caracterización, se aumentó el número de micrófonos de renglón y columna, comenzando en 2 micrófonos y terminando en 10 micrófonos. Cada mapa se generó utilizando 1 segundo de señal cosenoidal de 134 Hz con una frecuencia de muestreo de 1000 Hz, con 625 puntos de prueba. El resultado de los tiempos de procesamiento como función del número de micrófonos dentro de la red se muestra en la Figura 5.7.

Puede observarse inmediatamente que la relación que tiene el tiempo de procesamiento contra el número de micrófonos no es lineal. Se realizó un análisis de regresión y se ajustaron los datos a una ecuación polinomial para identificar la ecuación que determina la razón de aumento del tiempo de procesamiento con respecto al número de micrófonos. Los resultados se muestran en la tabla 5.2.

Formador de haz	Ecuación obtenida del análisis de regresión	Coficiente de determinación ( $R^2$ )	Incremento normalizado
DAS	$y = 0.008766x + 0.8505$	0.4299	—
MVDR	$y = 0.02554x^2 + 0.4058x + 0.2502$	0.9989	18.5610
PBM	$y = 0.001376x^2 + 0.9556x - 1.448$	0.998	1
SRP-PHAT	$y = 0.06438x^2 - 0.08877x - 0.9746$	0.999	46.7877

Tabla 5.2: Regresión lineal para la relación tiempo de procesamiento contra el número de micrófonos de la WASN

La ecuación que mejor define la relación entre el tiempo de procesamiento y el número de micrófonos es una cuadrática, para casi todos los algoritmos. El menor tiempo de procesamiento lo tiene el algoritmo DAS, pero no se encuentra una ecuación cuadrática que se ajuste a esta relación. Los demás algoritmos sí presentan una relación cuadrática en el incremento, y de igual manera que en

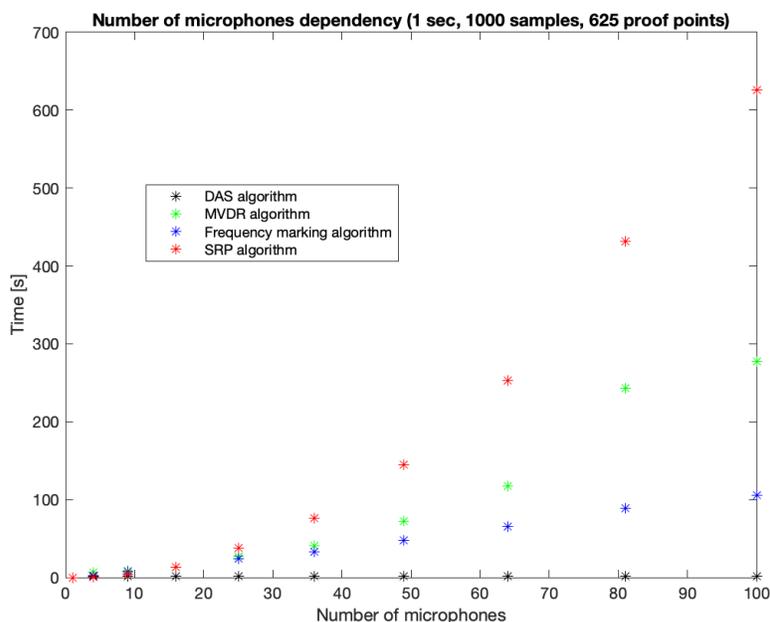


Figura 5.7: Tiempo de ejecución contra el número de micrófonos.

el análisis del tiempo de procesamiento contra el número de muestras, se normalizó para determinar la proporción en la que aumenta el tiempo de procesamiento con relación al número de micrófonos para el algoritmo con un menor tiempo de procesamiento.

El aumento del tiempo de procesamiento tiene una dependencia más fuerte con el término cuadrático de la ecuación, por lo que para realizar la comparativa, el incremento normalizado, sólo toma en cuenta la diferencia entre los coeficientes de los términos cuadráticos. El algoritmo con el menor coeficiente es el PBM y contra ese se realiza la normalización de los demás.

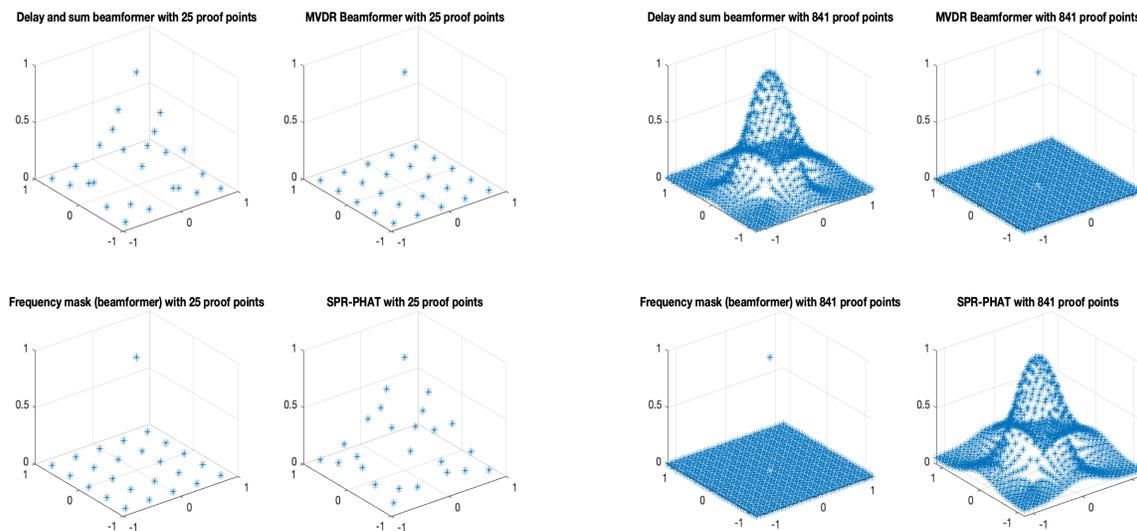
Como puede observarse, el algoritmo que tiene un menor aumento en el tiempo de procesamiento al aumentar el número de micrófonos es el PBM, por lo que si se deseara aumentar el tamaño de la red de sensores, sería la mejor opción.

### 5.2.3. Dependencia con los puntos de prueba

El análisis del tiempo de procesamiento contra los puntos de prueba en los que se calcula la energía determinará qué tan viable es la implementación para generar mapas con suficiente resolución, pero que no comprometa el tiempo de procesamiento de una manera excesiva. Estos parámetros de suficiente resolución y tiempo excesivamente largo se determinarán cuando se realicen mediciones en ambientes reales.

En las Figuras 5.8a y 5.8b se muestra claramente el objetivo de este análisis y es que ambas figuras representan el procesamiento de la señal de la forma  $s(t) = A\cos(2\pi\omega t)$ , en donde  $\omega = 143$  es la frecuencia angular de la señal y  $t$  es el tiempo, de una fuente puntual localizada en el punto  $(0, 0)$  capturada por 4 micrófonos, pero en la Figura 5.8a sólo se calcula la energía en 25 puntos de prueba, mientras que en la Figura 5.8b se calcula en 841. La diferencia entre ambos consiste en la definición que tiene el mapa creado, ya que en el primero, la forma del mapa es apenas distinguible, mientras que para el segundo, la forma queda claramente representada.

Para realizar el análisis, se generaron puntos de prueba de manera cuadrada homogénea, en donde se varió el número de puntos de prueba de filas y columnas desde calcular la energía en una red de puntos de prueba de  $5 \times 5$ , 25 puntos de prueba, hasta tener una red de puntos de prueba de  $30 \times 30$ ,



(a) Mapa generado para 5X5 puntos de prueba. (b) Mapa generado para 29X29 puntos de prueba.

Figura 5.8: Mapas de energía acústica obtenidos para distintas cantidades de puntos de prueba.

900 puntos de prueba. El resultado se muestra en la Figura 5.9. En la figura puede observarse que para pocos puntos de prueba, el tiempo de procesamiento presenta un comportamiento en el que primero baja y luego sube, pero a partir de 100 puntos de prueba, el comportamiento parece ser lineal, por lo que para hacer el análisis, se tomarán en cuenta los tiempos de procesamiento para más de 100 puntos de prueba.

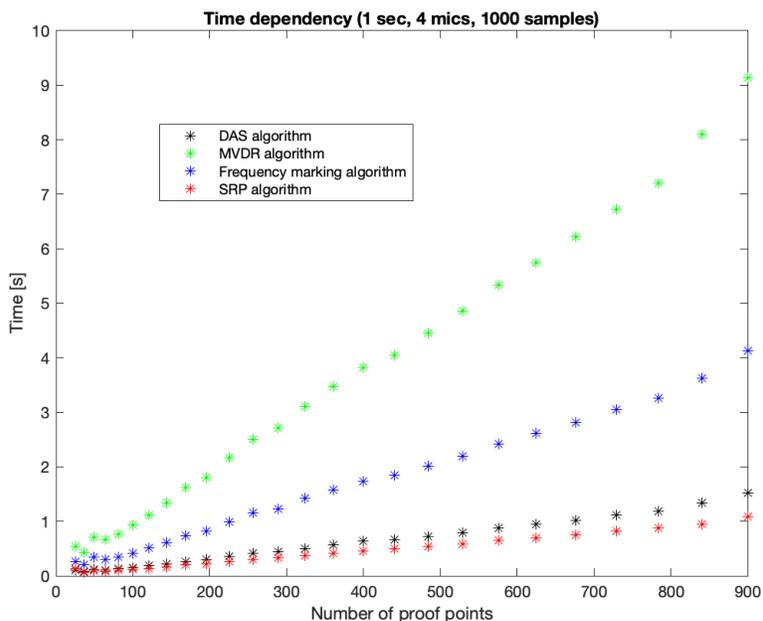


Figura 5.9: Tiempo de ejecución contra número de puntos de prueba.

Con los resultados obtenidos mediante este análisis, se notaron diferencias en los tiempos de

Formador de haz	Ecuación lineal	Coefficiente de determinación ( $R^2$ )	Incremento normalizado
DAS	$y = 0.002083x + 0.02091$	0.9995	1.6962
MVDR	$y = 0.01182x + 0.1131$	0.9997	9.6254
PBM	$y = 0.005446x + 0.0618$	0.9997	4.4348
SRP-PHAT	$y = 0.001228 + 0.01$	0.9995	1

Tabla 5.3: Regresión lineal para la relación tiempo de procesamiento contra el número de puntos de prueba.

procesamiento que le lleva los diferentes formadores de haz la generación del mapa de energía acústica. Si bien, todos los formadores de haz son capaces de generar un mapa de energía acústica y todos localizan posición de la fuente simulada en el punto exacto donde fue simulada, la diferencia entre ellos radica en la forma del mapa de energía acústica obtenido. Por esta razón, se realizará un análisis de la forma del mapa de energía acústica, pero una observación muy importante de esta sección es que la cantidad de puntos de prueba no afecta la forma del mapa de energía, pero sí aumenta la definición en la forma del mapa obtenido, pero tiene que tenerse en consideración que el tiempo de procesamiento aumenta.

### 5.3. Forma del mapa de energía

#### 5.3.1. Dependencia del funcionamiento con la presencia de ruido

Un ambiente real, al contrario que en las simulaciones que se han presentado hasta este punto, estará sujeto a la presencia de ruido, este puede ser ambiental, de otras fuentes o interferencias de otras señales. Un análisis que tiene que realizarse para la evaluación del desempeño de los algoritmos para generar el mapa de energía acústica es cómo afecta el ruido a su funcionamiento.

El que un modelo no funcione de manera adecuada ante la presencia de ruido no lo descartará para poder ser utilizado en una etapa del procesamiento de la información, ya que, probablemente, la implementación final del modelo tendrá consideradas varias etapas en donde pueden utilizarse diferentes tipos de filtros, algoritmos de procesamiento, etc.

En la comparativa que se realizará, se analiza cómo afecta la presencia de ruido al modelo de mapa de energía acústica que se obtiene para los diferentes algoritmos. En la simulación se propone una fuente puntual, localizada en el centro del arreglo, que emite una señal cosenoidal a una frecuencia de muestreo de 4800 Hz. de la forma:  $s(t) = A \cos(2\pi\omega t)$ , en donde  $\omega$  es la frecuencia angular de la señal y  $t$  es el tiempo. Se generaron mapas de energía acústica variando la presencia de ruido para las señales simuladas mediante la inclusión de ruido blanco de la forma:  $[0 \ 0.02A \ 0.1A \ 0.2A \ 0.4A]$ , en donde  $A$  es la amplitud de la señal y  $0.4A$  representa un ambiente muy ruidoso. A continuación, se presentan los mapas obtenidos para estas cantidades de ruido, y en la sección de conclusiones se realizará un análisis de estos resultados. La simulación se realizó para un tamaño de muestra de 1024 muestras que representan 0.2133 segundos y 625 puntos de prueba. En la Figura 5.10 se muestra el mapa sin presencia de ruido en un tamaño grande para que se aprecien los detalles del mapa, para las siguiente Figuras, los mapas se presentan más pequeños, con el objetivo de que se observe su comportamiento.

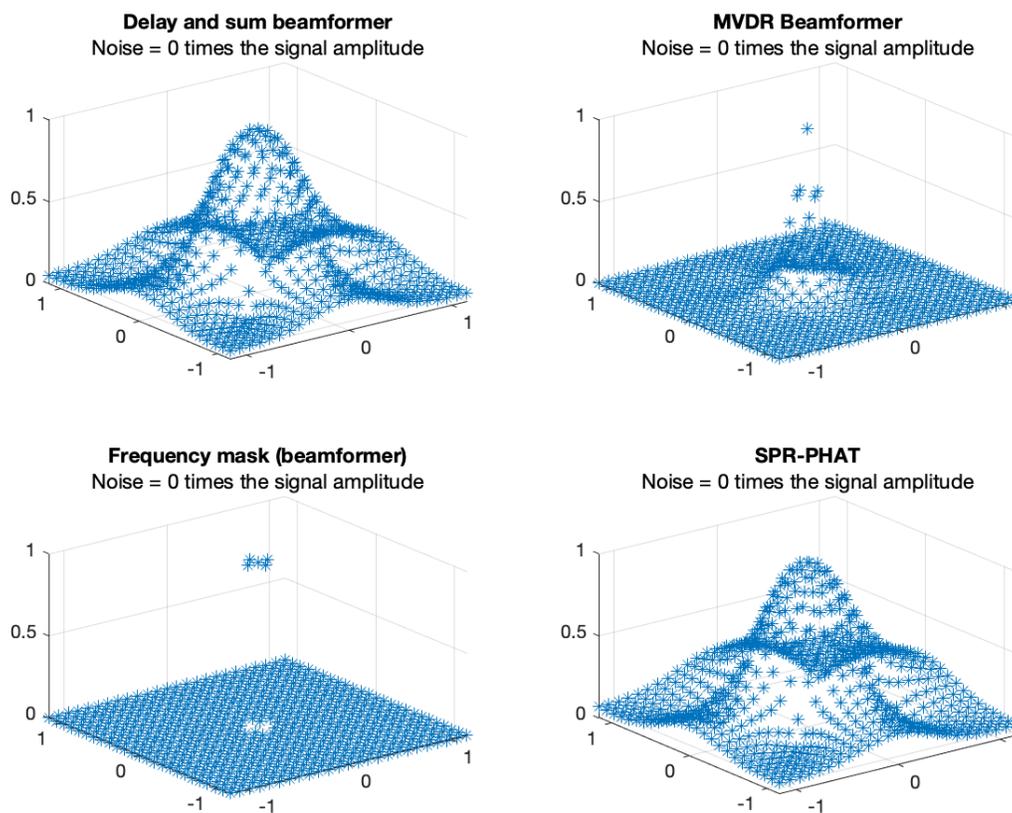


Figura 5.10: Mapa generado sin la presencia de ruido.

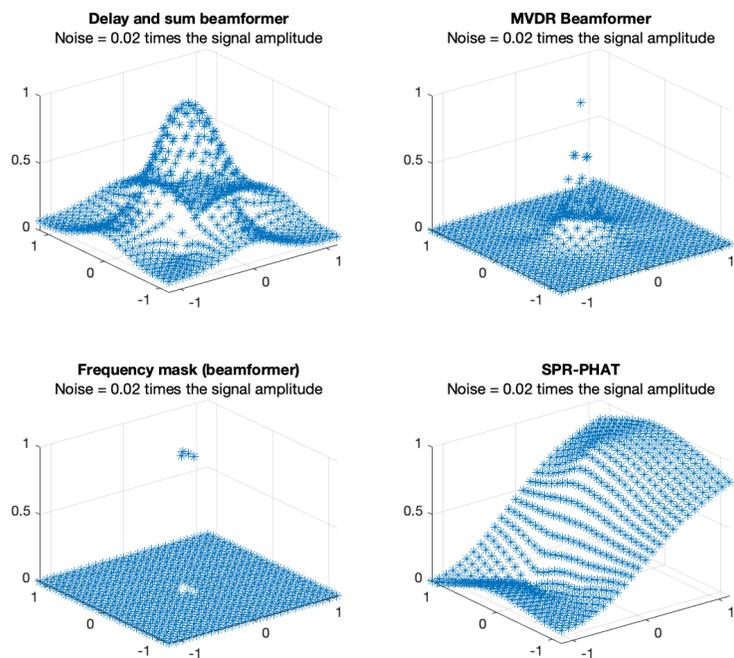


Figura 5.11: Mapa generado con la presencia de ruido de amplitud 0.02A.

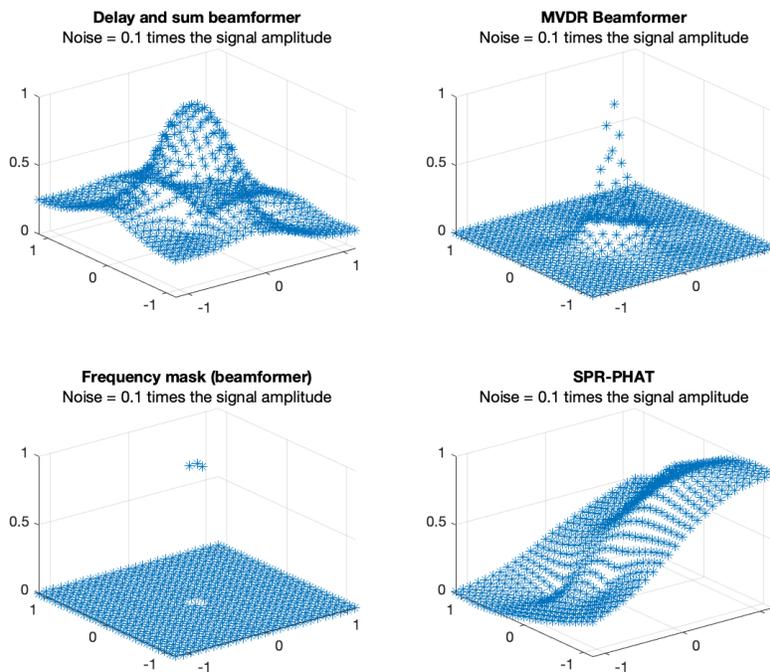


Figura 5.12: Mapa generado con la presencia de ruido de amplitud 0.1A.

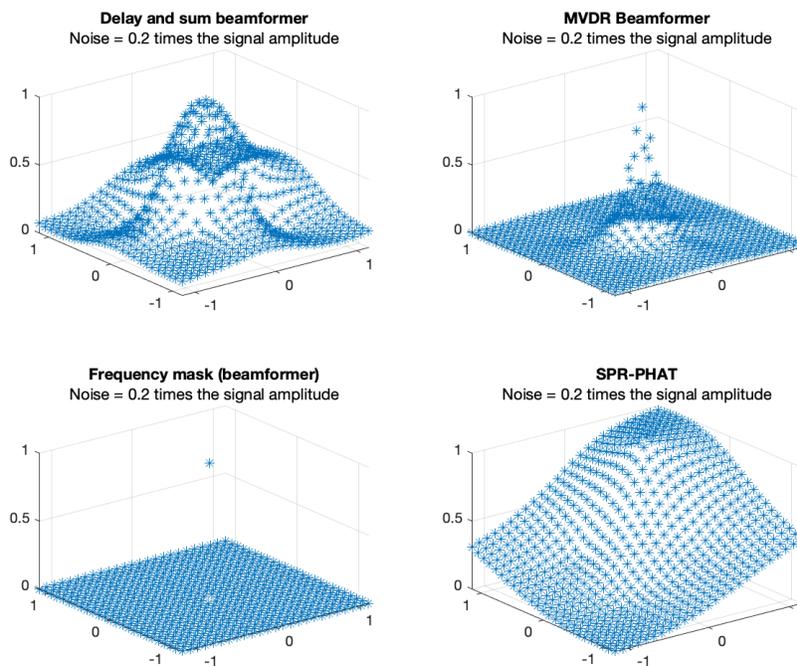


Figura 5.13: Mapa generado con la presencia de ruido de amplitud 0.2A.

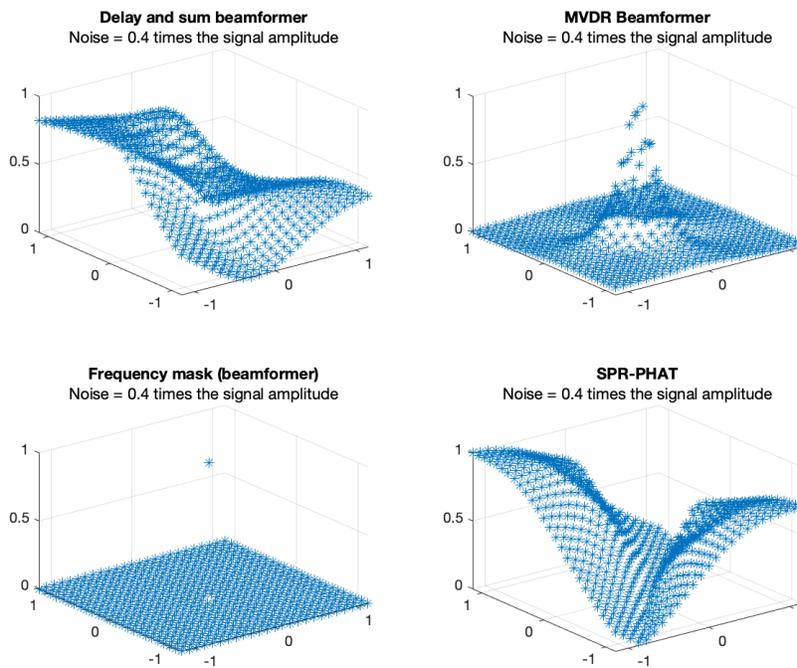


Figura 5.14: Mapa generado con la presencia de ruido de amplitud 0.4A.

### 5.3.2. Caracterización para la variación de la frecuencia de la fuente

En esta sección se analiza el efecto que tiene la variación de la frecuencia de emisión de la fuente en el mapa de energía obtenido. Se generaron mapas de energía con las siguientes características: frecuencia de muestreo: 48000 Hz, separación entre los sensores: 0.45 m, número de sensores: 4, posición de la fuente: (0,0), 75X75 puntos de prueba para lograr una mejor definición y poder apreciar los detalles obtenidos. Los mapas obtenidos se obtuvieron con los parámetros descritos y variando la frecuencia de la fuente en un intervalo de 10 Hz, comenzando en 10 Hz. y llegando hasta 900 Hz. estos valores fueron utilizados, ya que están relacionados con intervalo de frecuencias de la voz humana, que es una de las posibles aplicaciones del modelo. Los resultados mostrados incluyen las frecuencias que se consideran importantes para el análisis, como frecuencias bajas(10, 100, y 250), y se muestran también frecuencias altas, (500,700, y 900) en donde el fenómeno de *spatial aliasing* se presenta, que es un fenómenos que distorsiona las señales reconstruidas y que depende de la frecuencia de la señal y la geometría del arreglo. Los resultados obtenidos se muestran en las siguientes figuras.

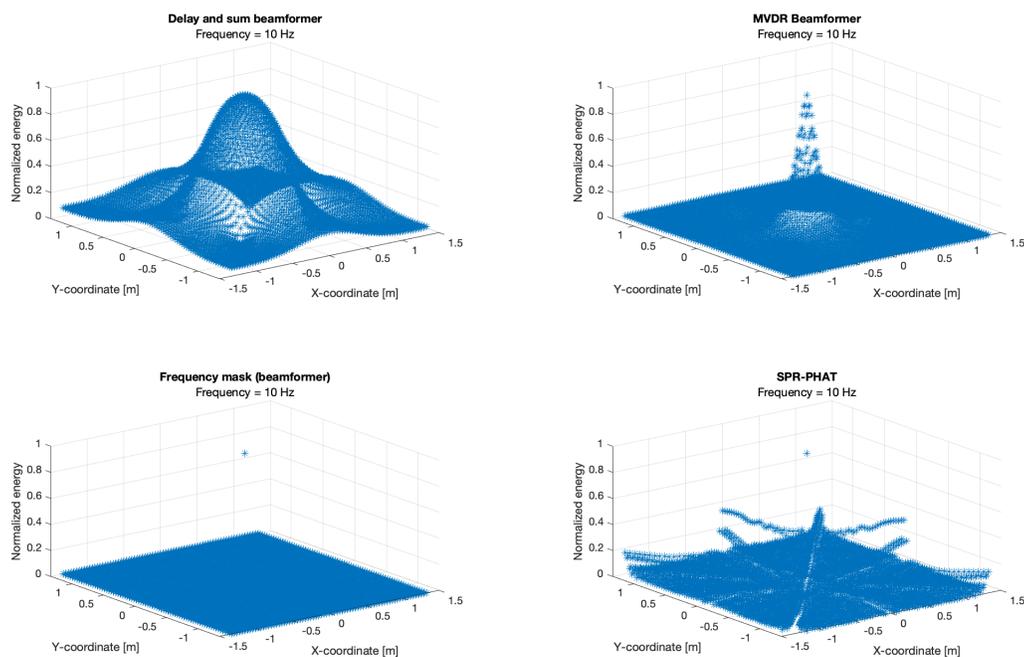


Figura 5.15: Mapa generado con una señal de frecuencia de 10 Hz.

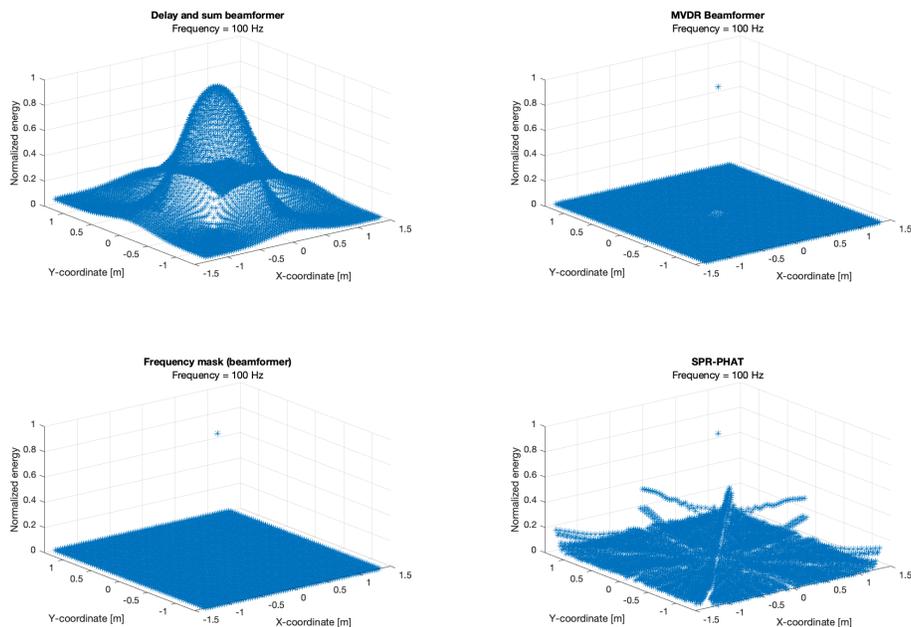


Figura 5.16: Mapa generado con una señal de frecuencia de 100 Hz.

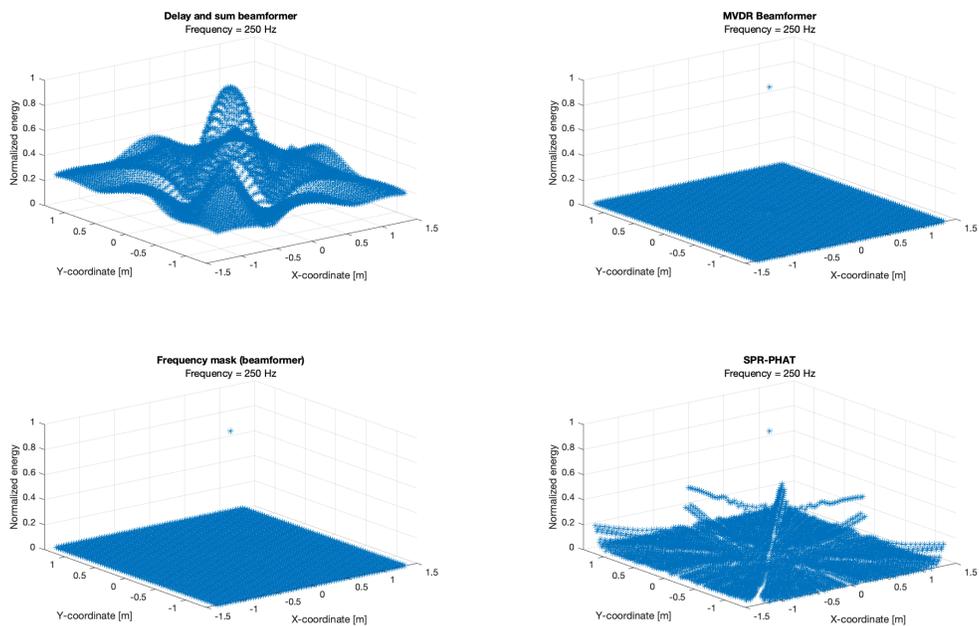


Figura 5.17: Mapa generado con una señal de frecuencia de 250 Hz.

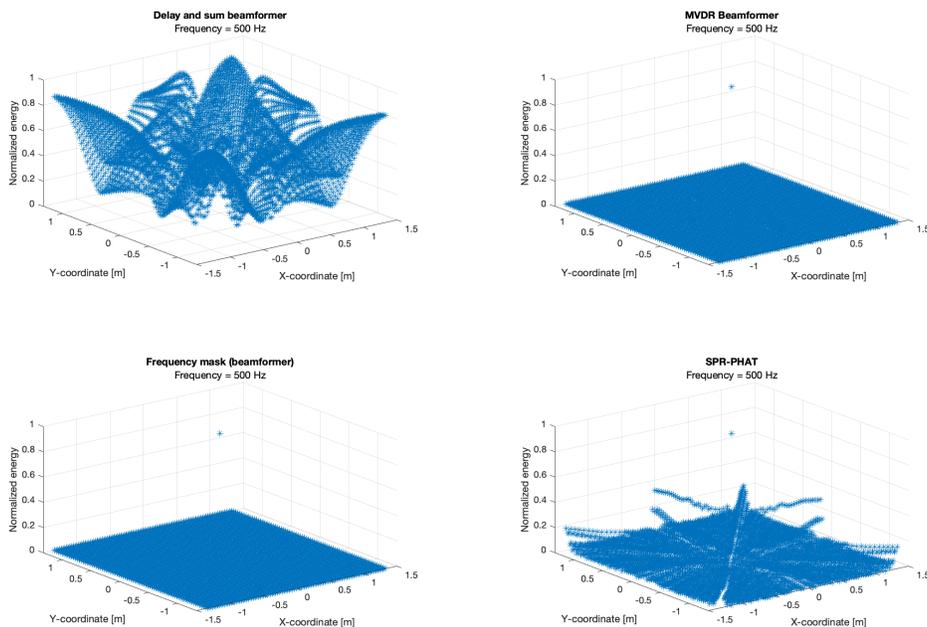


Figura 5.18: Mapa generado con una señal de frecuencia de 500 Hz.

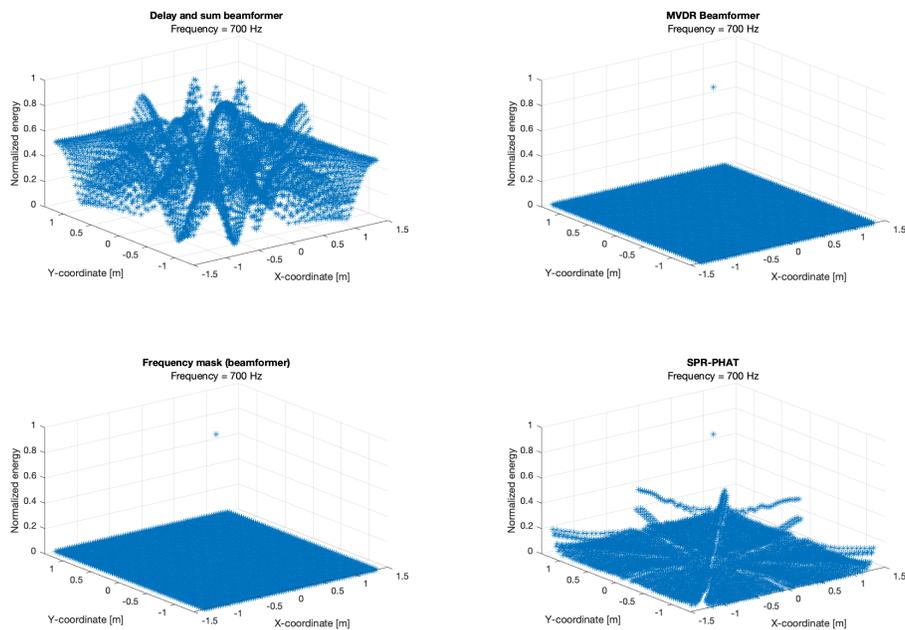


Figura 5.19: Mapa generado con una señal de frecuencia de 700 Hz.

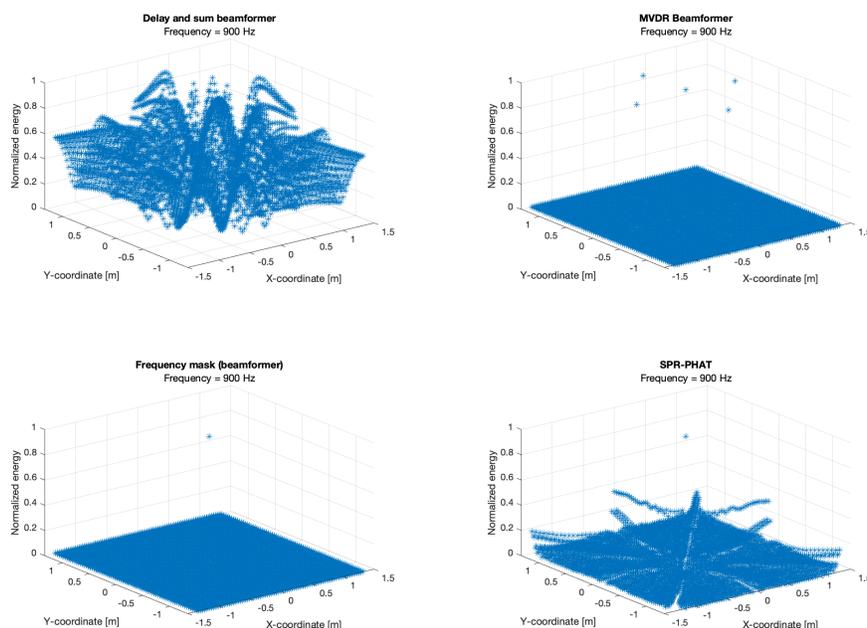


Figura 5.20: Mapa generado con una señal de frecuencia de 900 Hz.

### 5.3.3. Caracterización para múltiples fuentes acústicas

Se probó el funcionamiento para generar un mapa de energía acústica en la presencia de dos fuentes de energía acústica puntuales. Las frecuencias de excitación de las fuentes son: 100 y 200 Hz y las posiciones de las fuentes son:  $(0.2, -0.2)$  y  $(0.1, 0.1)$ . La simulación tiene las siguientes características: 4 micrófonos,  $75 \times 75$  puntos de prueba, sin presencia de ruido. Se presentan estas posiciones y frecuencias para las fuentes acústicas, pero se demostró con otras simulaciones que el modelo es capaz de generar mapas similares para otras posiciones y frecuencias.

En la Figura 5.21 se observa el funcionamiento de los diferentes formadores de haz para la generación del mapa de energía acústica en la presencia de dos fuentes acústicas puntuales. Como puede observarse, no es posible determinar a simple vista la localización de las fuentes para los formadores de haz DAS y SRP-PHAT, pero tanto el algoritmo MVDR como el PBM entregan mapas de energía acústica que localizan los máximos de energía en la misma posición en la que se simularon las fuentes. En las Figuras 5.22 y 5.23 se muestran los mapas de energía acústica obtenidos mediante los algoritmos MVDR y PBM, respectivamente. Como puede observarse, en ambos casos la posición de la fuente simulada coincide con el máximo de energía del mapa de energía acústica obtenido mediante el modelo de mapeo acústico propuesto.

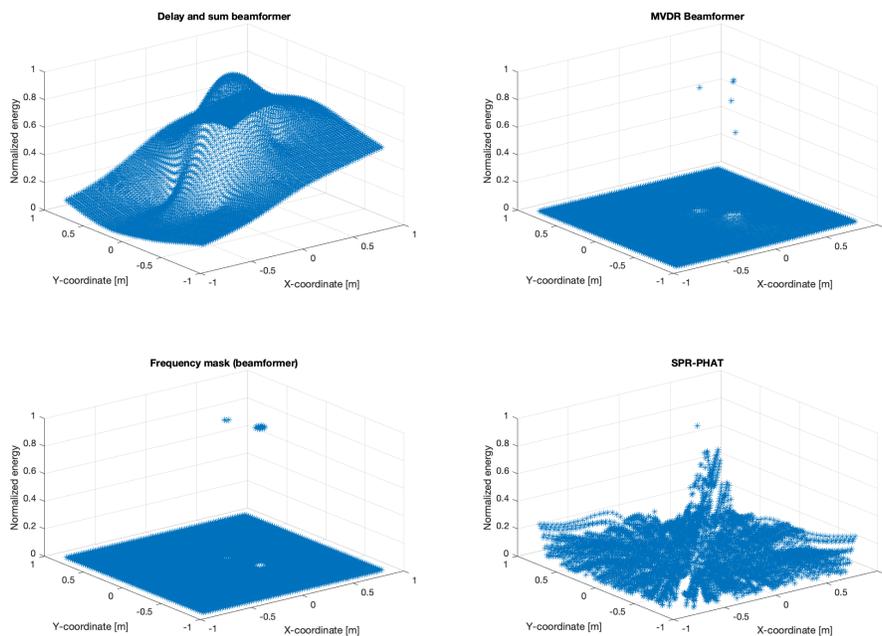


Figura 5.21: Mapa generado para dos fuentes acústicas puntuales.

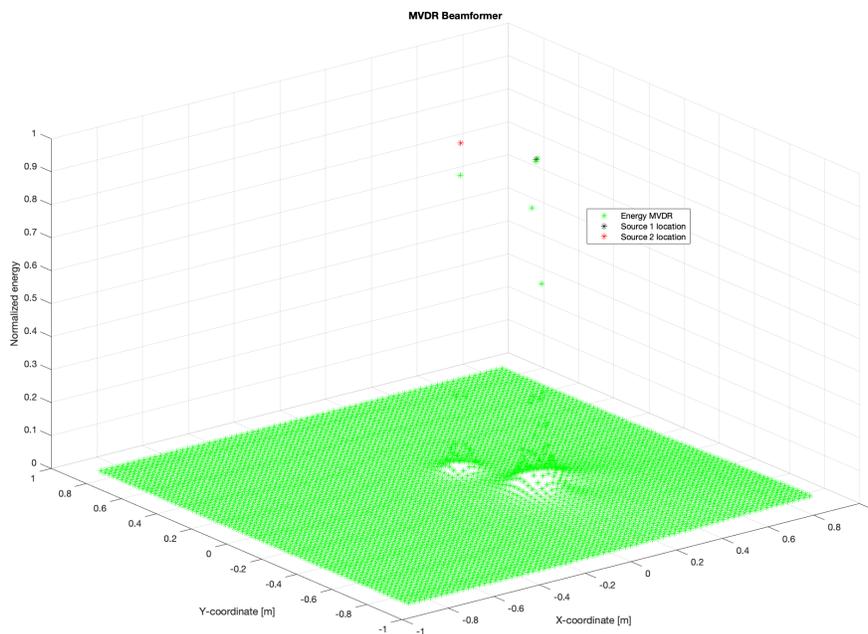


Figura 5.22: Mapa generado el formador MVDR

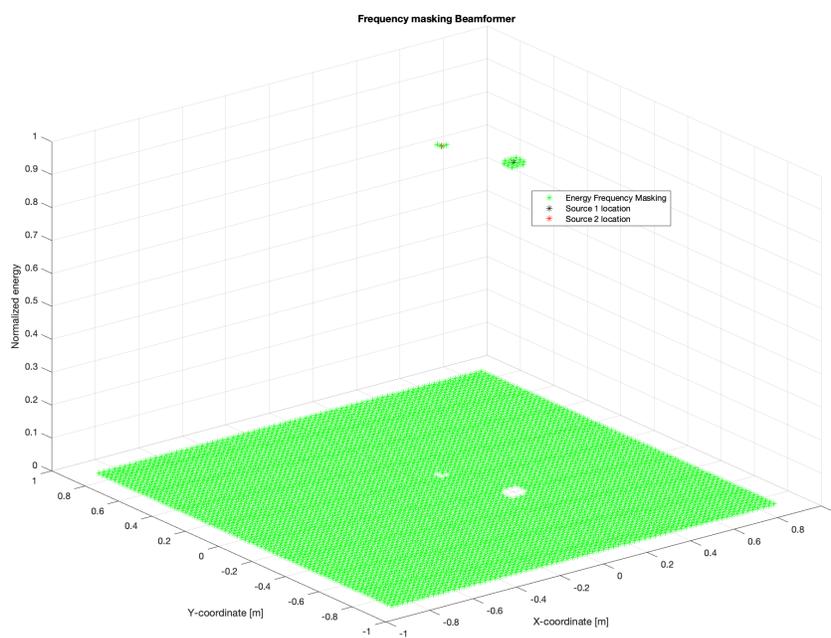


Figura 5.23: Mapa generado el formador **PBM**

## Capítulo 6

# Metodologías de sincronización y captura de la WASN

En esta sección se proponen tres metodologías para la captura de la señal acústica mediante la WASN. El objetivo de desarrollar diferentes metodologías es para analizar el desempeño que cada una de ellas tiene para el procesamiento de la información capturada en la generación del mapa de energía acústica. Las metodologías desarrolladas pueden clasificarse por sus características en dos grupos: grabación local y grabación a través de internet. En las siguientes subsecciones se explica el funcionamiento de cada una de estas metodologías.

### 6.1. Grabación local con protocolos: NTP y PTP

Se plantearon dos metodologías para la grabación local, en donde cada uno de los nodos captura información del ambiente acústico mediante su sensor y graba la información en su unidad de procesamiento. Posteriormente, se recupera la información capturada y se analiza la información en post-procesamiento para la generación del mapa de energía acústica.

En esta metodología, cada uno de los nodos genera su propio conjunto de marcas temporales dadas por el reloj interno de cada unidad de procesamiento. Esquemáticamente, la manera de grabar con esta metodología se representa en la Figura 6.1, en donde se muestran los archivos creados durante la grabación, indicados en cada nodo y se indica también la labor que realiza el servidor dentro de la WASN.

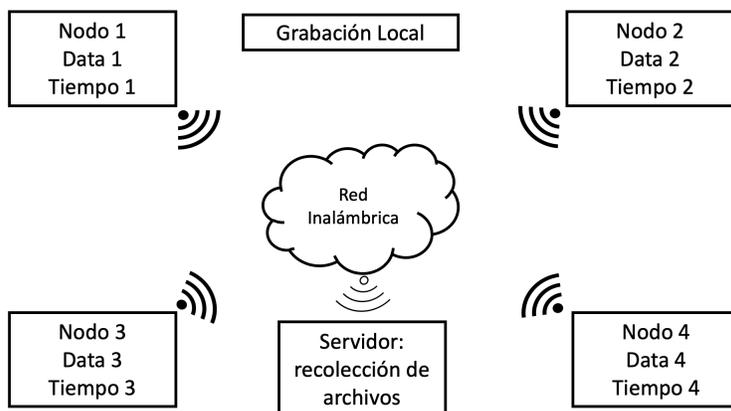


Figura 6.1: Metodología de grabación local. Archivos creados en cada nodo.

Se propusieron dos metodologías para la grabación local, y la diferencia entre ellas, es el protocolo con el que se sincronizó la **WASN**. La primera metodología de grabación local, sincroniza a la **WASN** mediante el protocolo **NTP**, en donde en cada uno de los nodos sincroniza su reloj interno con el de un nodo servidor de tiempo, éste a su vez, sincroniza su reloj con un servidor de tiempo de Internet. Para lograr que este protocolo de sincronización funcione de manera adecuada, se requieren eliminar todas las fuentes de sincronización en los archivos de configuración del protocolo **NTP**.

Existe un proceso en segundo plano *daemon* que actualiza el reloj interno de cada uno de los nodos clientes con el servidor de tiempo y es con este proceso que se sincroniza la **WASN** periódicamente [38].

La segunda metodología de grabación local sincroniza a la **WASN** mediante el protocolo **PTP** y utilizando su propio proceso en segundo plano (*daemon*), actualiza el reloj de los clientes con el del servidor de tiempo [39] de manera periódica.

Cualquiera de estas metodologías, plantea que cada uno de los nodos de captura se encargue de sensar el ambiente acústico, monitorear sus tiempos locales y generar archivos de audio y marcas temporales de manera local. Este proceso, como se analizará en la sección de resultados, genera que se tenga que realizar un proceso de ajuste en las marcas temporales para cada uno de los conjuntos tiempo-información.

## 6.2. Grabación Remota

Después de realizar mediciones empleando las metodologías de grabación local, se estudió la posibilidad de realizar las grabaciones a través de internet, de esta manera, se evita generar marcas temporales particulares para cada uno de los nodos y se genera en su lugar sólo una marca temporal global, dependiente del reloj interno del servidor de JACK que recibe la información del ambiente acústico capturada por cada uno de los nodos. El funcionamiento de esta metodología se representa en la Figura 6.2 en donde se muestran los archivos creados durante la grabación, y en donde se observa la ventaja. de que sólo se crea un archivo con las marcas temporales en el servidor y los cuatro archivos con la información acústica capturada por cada uno de los nodos.

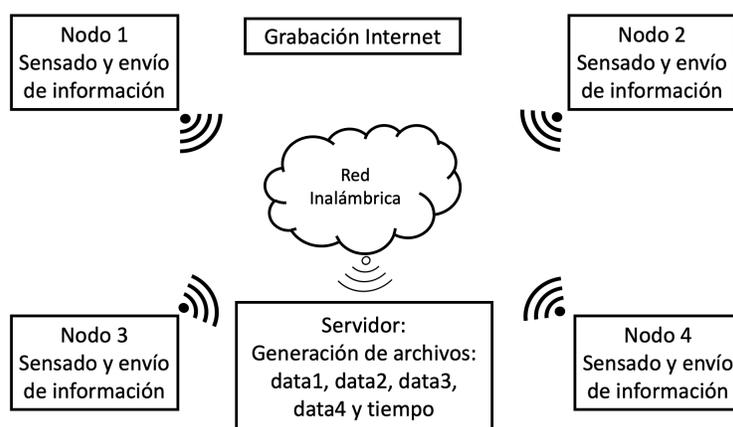


Figura 6.2: Metodología de grabación remota.

Esta metodología ofrece ciertas ventajas con respecto a la grabación local, por ejemplo, cada uno de los nodos de grabación sólo se encarga de sensar y transmitir la información del ambiente acústico, lo que le quita las labores de generar un archivo de audio y el monitoreo de su reloj interno para

después escribir un archivo de marcas temporales. También, en el procesamiento de la información para la generación del mapa de energía acústica, esta metodología tiene la ventaja de sólo generar una marca temporal para la información capturada por cada uno de los nodos de medición, lo que simplifica mucho el algoritmo, tanto en tiempo de procesamiento como en complejidad.

Por último, tanto en las metodologías de grabación local como en la de grabación a través de Internet, el *software* de captura (JACK audio connection kit) genera errores en su ejecución. Estos errores se llaman **xrun** y su existencia se debe al funcionamiento interno de JACK. JACK acumula información en ventanas de tiempo (definidas por el usuario) y después realiza lo que se le indique, pero cuando JACK, por alguna razón no cuenta con la información necesaria en esa ventana de tiempo, genera un xrun y la información de esa ventana no se genera de manera correcta.

En las metodologías de grabación local, al generar archivos tanto de audio (.wav) como de marcas temporales (.txt), es muy común obtener xrun en una grabación, ya que cada unidad de procesamiento está realizando muchas actividades simultáneamente. Por esta razón es que se propuso la metodología de grabación a través de Internet, ya que libera de algunas actividades a las unidades de procesamiento de cada nodo y esto hace que JACK no genere tantos xruns como en la grabación local.

Sin embargo, la metodología de grabación a través de Internet no está exenta de xruns, ya que como la transmisión de la información se hace a través de un módem en una red local, la latencia en la transmisión es la responsable de que, en ocasiones, JACK no tenga la información necesaria para una ventana en específico.

Otra consideración que se tuvo que tomar en cuenta cuando se propuso la metodología de grabación a través de Internet es que los dispositivos **Raspberry Pi Zero W** no se desempeñan correctamente. Los dispositivos **Raspberry Pi 4 B+** sí funcionan de manera correcta y los resultados se presentarán en la sección de resultados.



## Capítulo 7

# Análisis de resultados experimentales y simulados

En esta sección se analizará el impacto que tienen errores de sincronización en la generación del mapa de energía acústica para determinar los niveles de precisión necesarios para esta aplicación. La hipótesis presentada es que si existe un error de sincronización, la forma del mapa de energía acústica se modificará y resultará en un error en la posición de las fuentes acústicas.

Para este análisis, se simula una fuente acústica en un punto fijo conocido  $(x_f, y_f)$ , también se simula una WASN con cuatro nodos en una distribución rectangular con la fuente acústica localizada en el mismo plano. Para simplificar el proceso, la posición de la fuente se coloca en  $(x, y) = (0, 0)$ , de tal modo que la distancia entre la fuente y cada nodo de la red sea la misma.

Posteriormente, los formadores de haz descritos en la sección 2 son usados para reconstruir la señal en un dado “punto de prueba”. El mapa acústico obtenido puede ser representado ya sea en una gráfica tridimensional, en donde el plano  $XY$  representa el plano físico donde se encuentran tanto la WASN como la fuente y el eje  $Z$  representa la energía acústica asociada con cada punto. Otra manera de representar el mapa acústico es mediante un mapa de calor, en donde cada punto de prueba es presentado con un determinado color asociado con su energía.

Las características del ambiente acústico simulado son:

- Frecuencia de muestreo para cada uno de los nodos:  $f_s = 48,000$  Hz.
- Velocidad del sonido en el aire:  $c = 343 \frac{m}{s}$ .
- Amplitud de la señal en la fuente =  $\cos(ft)$ .
- Frecuencia de excitación:  $f = 300$  Hz.
- Tiempo entre dos muestras capturadas:  $\tau_s = \frac{1}{f_s} = 0.00625$  s.
- Coordenadas de la fuente:  $(x, y) = (0m, 0m)$ .
- Coordenadas de los nodos:
  - $(x_1, y_1) = (-0.32m, 0.32m)$ .
  - $(x_2, y_2) = (0.32m, 0.32m)$ .
  - $(x_3, y_3) = (-0.32m, -0.32m)$ .
  - $(x_4, y_4) = (0.32m, -0.32m)$ .

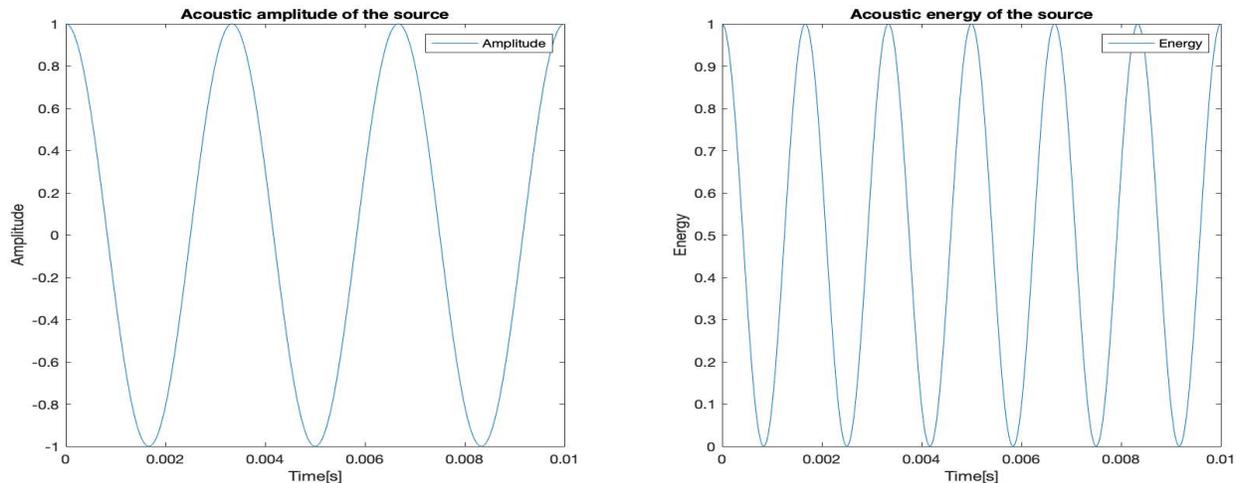


Figura 7.1: Amplitud y energía originales de la señal de la fuente.

La amplitud original y la energía de la fuente se ven representadas en la Figura 7.1. Como se mencionó previamente, la función dependiente del tiempo de la energía es proporcional al cuadrado de la señal emitida por la fuente, por simplicidad y sin pérdida de generalidad, la constante de proporcionalidad se propone que sea 1, de modo que la función de la energía pueda ser escrita como:

$$E(x, y, t) = \|S(x, y, t)\|^2. \quad (7.1)$$

Las señales capturadas por los cuatro nodos de la WASN simulada se muestran en la Figura 7.2. En esta imagen, todas las señales se encuentran sobrepuestas, ya que todas las señales tienen la misma frecuencia de muestreo y el mismo tiempo de arribo. En el análisis de sincronización, uno de estos nodos será des-sincronizado artificialmente, de modo que la señal de ese nodo está retrasada una cierta cantidad de muestras.

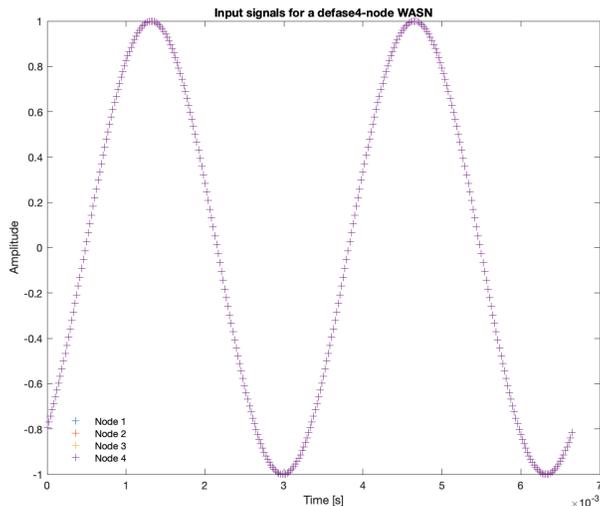


Figura 7.2: Señales de entrada capturadas por cada nodo.

Los mapas de energía acústica obtenidos al aplicar cada uno de los formadores de haz para las señales de entradas simuladas se muestran en la Figura 7.3, y sus respectivos mapas de calor en la Figura 7.4.

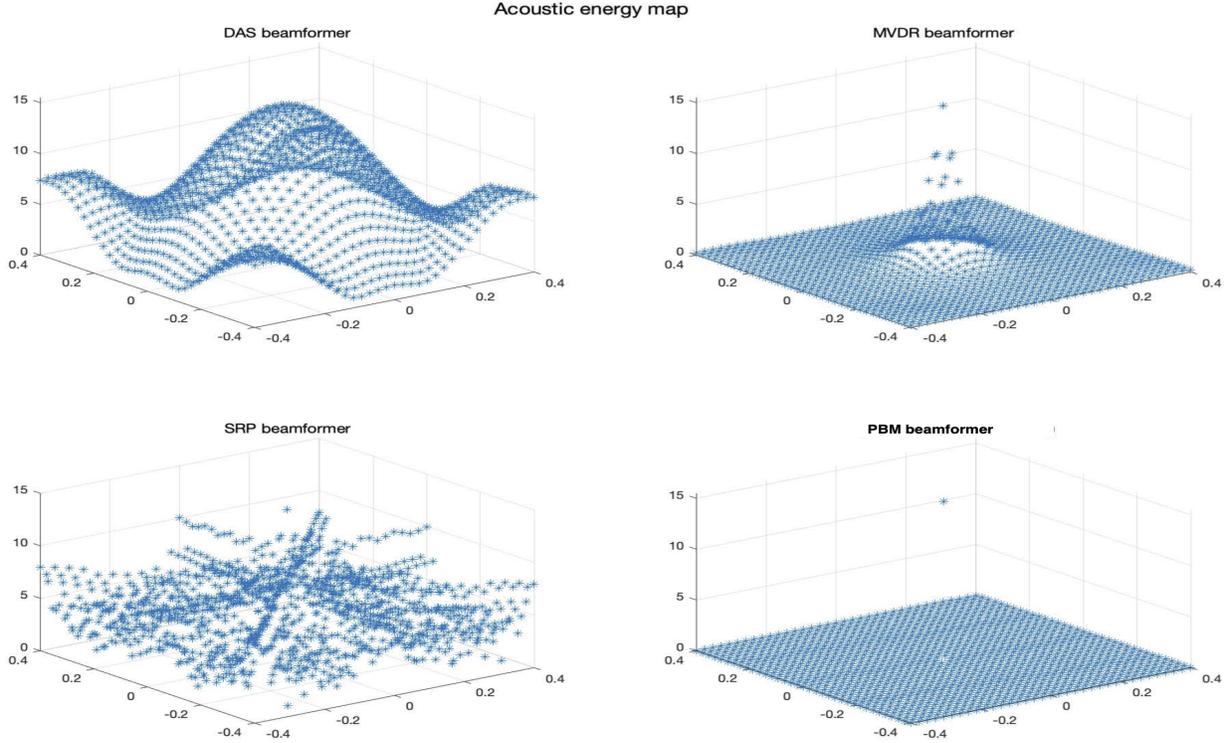


Figura 7.3: Mapas de energía acústica para los formadores: DAS, MVDR, SRP-PHAT y PBM.

El punto de prueba con la mayor cantidad de energía es  $(x, y) = (0, 0)$ . La energía obtenida con la señal reconstruida en este punto de prueba con los formadores de haz DAS, MVDR y PBM se muestran en la Figura 7.5. Adicionalmente, también se muestra el mapa bidimensional de la potencia obtenida con el formador SRP-PHAT.

Para analizar el desempeño en la generación del mapa de energía acústica, se calcula la energía total en un periodo de tiempo. En la señal simulada, la energía en un periodo de tiempo puede ser calculada mediante la integral y su valor puede aproximarse utilizando el método trapezoidal, explicado a profundidad en [40], mediante la siguiente expresión:

$$\int_{t=0}^{t=t_f} S(x, y, t) dt \approx \frac{t_f}{2N} \sum_{n=1}^N (S(x, y, t_n) + S(x, y, t_{n+1})) := \sum_{n=1}^N S(x, y, t), \quad (7.2)$$

donde  $S(x, y, t)$  es la señal simulada de la fuente,  $t_f$  es la longitud en tiempo de la muestra de la señal y  $N$  es el tamaño de la señal en número de muestras. La suma calculada para la señal simulada es:

$$\frac{t_f}{2N} \sum_{n=1}^N (\cos^2(2\pi\omega t_n) + \cos^2(2\pi\omega t_{n+1})) = 0.0049791A. \quad (7.3)$$

Mientras que la suma calculada para la señal reconstruida en el punto de prueba  $(0, 0)$  es:

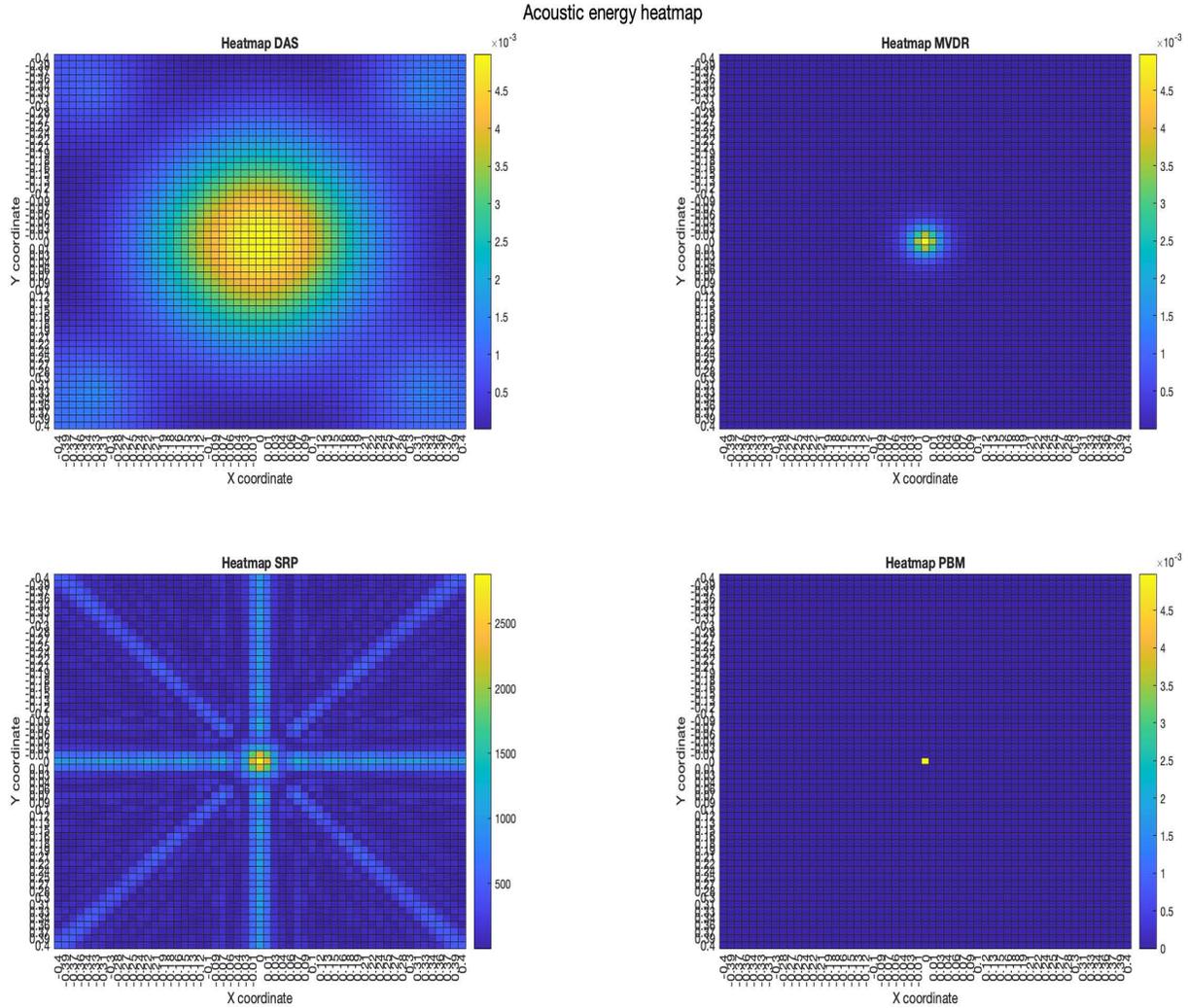


Figura 7.4: Mapas de calor para los formadores: DAS, MVDR, SRP-PHAT y PBM.

$$\sum_{n=1}^N S_{DAS}(0, 0, t) = \sum_{n=1}^N S_{MVDR}(0, 0, t) = \sum_{n=1}^N S_{PBM}(0, 0, t) = 0.0049861A. \quad (7.4)$$

Este resultado muestra que la energía obtenida por el mapa de energía acústica generado tiene un valor muy cercano a la señal simulada original.

Los resultados también muestran que el mapa de energía acústica obtenido mediante los formadores de haz tiene una mayor concentración de energía alrededor del punto donde la señal fue simulada. Sin embargo, hay que tener ciertas consideraciones:

- Existen diferencias en la forma del mapa de energía acústica obtenido con cada formador de haz. En este sentido, algunos son más robustos frente al ruido y la reverberación (como MVDR), mientras que otros son más eficientes en términos de tiempo de procesamiento y simplicidad en sus implementaciones (como DAS y PBM). El principal objetivo de este análisis es caracterizar el desempeño de los formadores de haz ante la presencia de errores en la sincronización, los resultados mencionados ya se describieron en la caracterización del

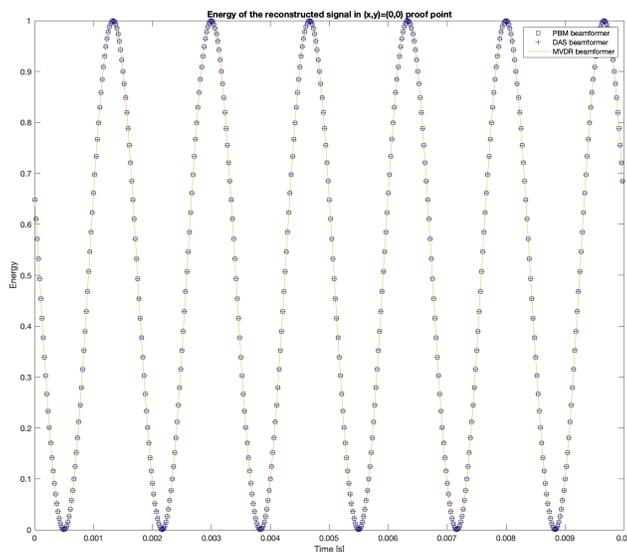


Figura 7.5: Energía obtenida de la señal reconstruida en el punto de prueba  $(x, y) = (0, 0)$ .

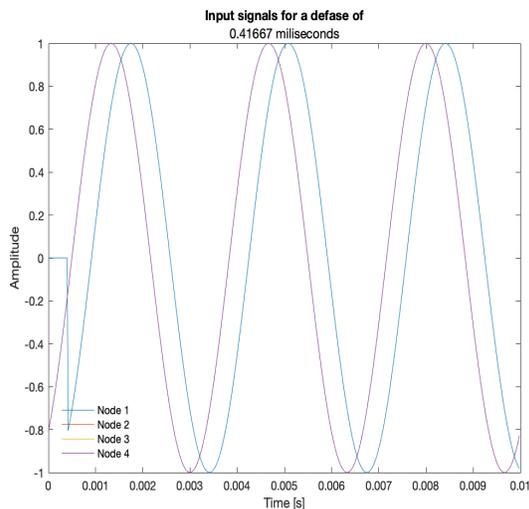
modelo matemático.

- El valor de la energía calculada a partir de la señal reconstruida en el punto de prueba  $(x, y) = (0, 0)$  para DAS, MVDR y PBM tiene un pequeño error de 0.14 % en comparación con la energía de la entrada original señal que se considera un valor pequeño como diferencia, ya que una reconstrucción ideal tendría un error de 0.00 %. Esto se puede considerar como evidencia del buen desempeño de estos formadores de haz cuando se trata de reconstruir la fuente de interés en el punto de prueba que coincide con la posición simulada de la fuente.
- DAS, MVDR y SRP-PHAT proporcionan valores de energía moderadamente altos en puntos de prueba que no están en la posición de la fuente. El formador de haz PBM encuentra energía solo en la posición simulada de la fuente.
- El formador de haz PBM parece ser el más preciso para localizar la posición de la fuente. Sin embargo, esta técnica requiere que se establezca un parámetro *a-priori*, por lo que existe una etapa adicional de calibración para la implementación de esta técnica.
- Tanto el formador de haz SRP-PHAT, como los formadores DAS y MVDR, encuentra energía en puntos donde no hay energía acústica de la fuente, pero también encuentra un patrón que no representa un efecto esperado del formador de haz.

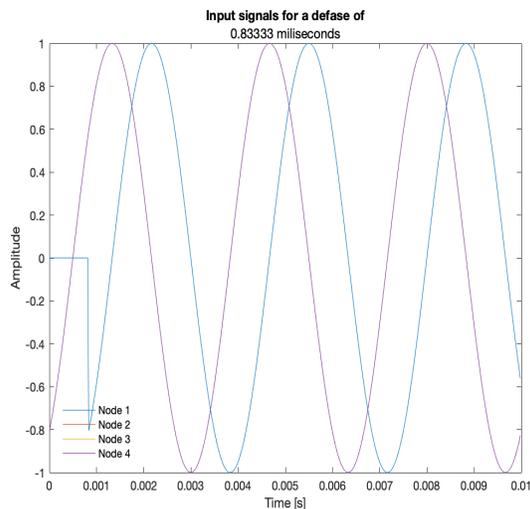
## 7.1. Simulación de grabación de señales con elementos de sincronización

El modelo de mapa de energía acústica para los cuatro formadores de haz da como resultado la localización correcta de la fuente acústica simulada. Se analizará su comportamiento ante la presencia de error en la sincronización.

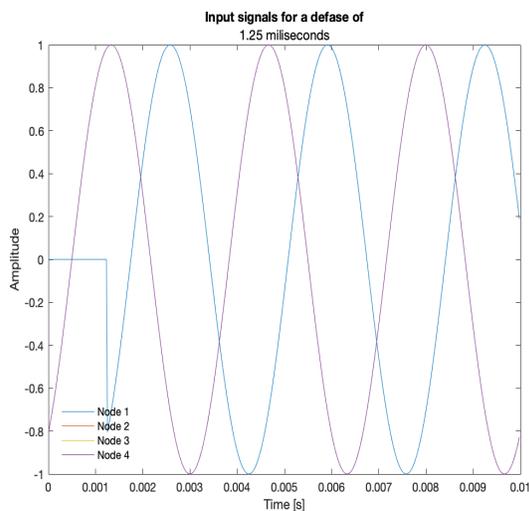
Es importante recordar que los formadores de haz requieren una sincronización adecuada de la red de nodos para su funcionamiento correcto, lo que no siempre ocurre con los WASN. Para simular esto, se inserta artificialmente un retardo (establecido como un número determinado de muestras) en el primer nodo simulado, ubicado en  $(x_1, y_1) = (-0.32m, 0.32m)$ , en la posición superior izquierda de la red, referido aquí como “nodo con error de sincronización”. El número de 20 muestras fue seleccionado como incremento para establecer una relación más precisa entre errores de sincronización y errores en la creación del mapa de energía acústica.



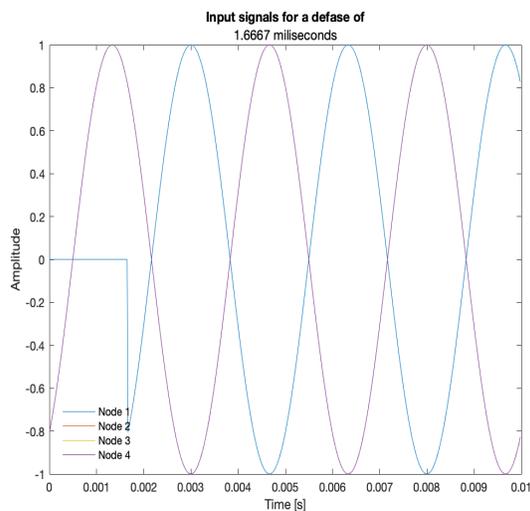
(a) 20 muestras de retraso



(b) 40 muestras de retraso



(c) 60 muestras de retraso

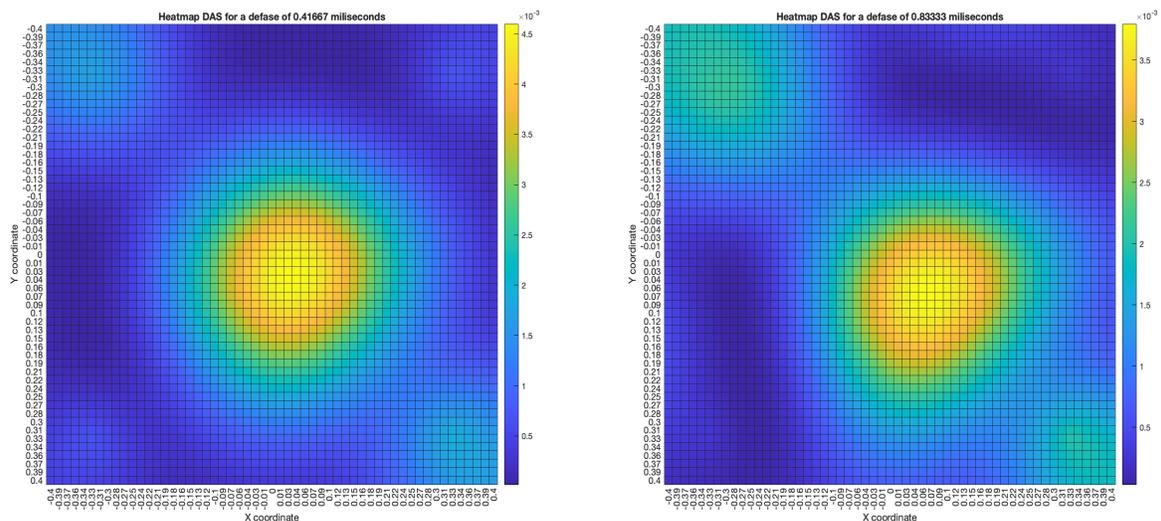


(d) 80 muestras de retraso

Figura 7.6: Señales de entrada con un retraso aumentando en 20 muestras en el nodo 1.

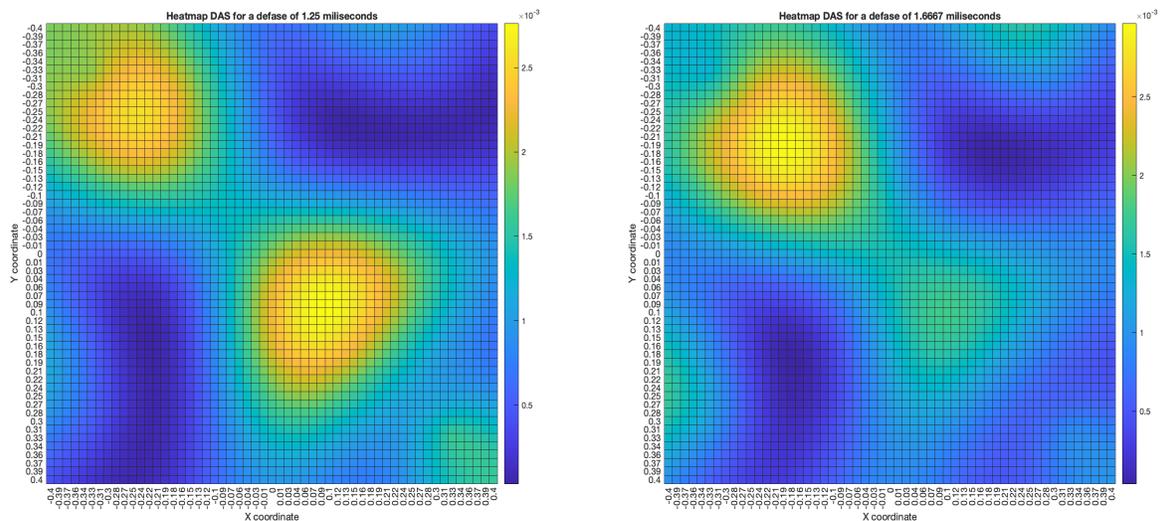
Un error de sincronización implica que antes de que la señal llegue al nodo retrasado, no hay señal acústica, como se muestra en la Figura 7.6: la señal de entrada retrasada tiene un valor de cero para tiempos previos al retraso simulado. Los formadores de haz pueden tener problemas con esta

discontinuidad en el dominio del tiempo, y podrían resolverse aplicando una ventana (como Hann o Hamming) para reducir sus efectos, pero este análisis se deja para un trabajo futuro. Este error de sincronización simulado es alimentado luego a los formadores de haz, y su impacto se observa en las siguientes figuras y se cuantifica en el cambio de forma del mapa de energía acústica, en la posición de la fuente y en el valor de su energía reconstruida.



(a) Mapa de calor con 20 muestras de retraso.

(b) Mapa de calor con 40 muestras de retraso.



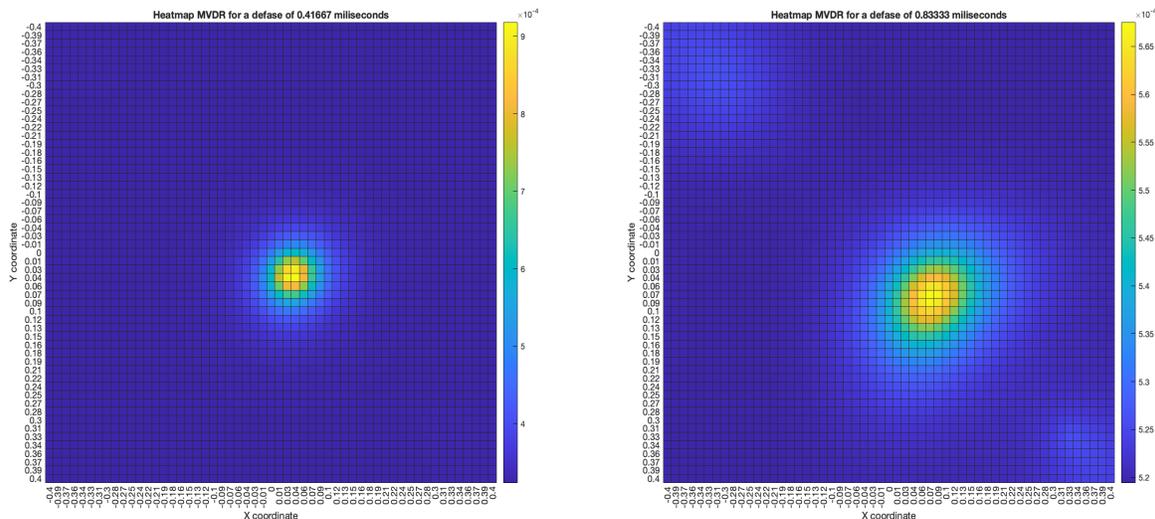
(c) Mapa de calor con 60 muestras de retraso.

(d) Mapa de calor con 80 muestras de retraso.

Figura 7.7: Mapas de calor para el formador de haz DAS con diferentes errores de sincronización.

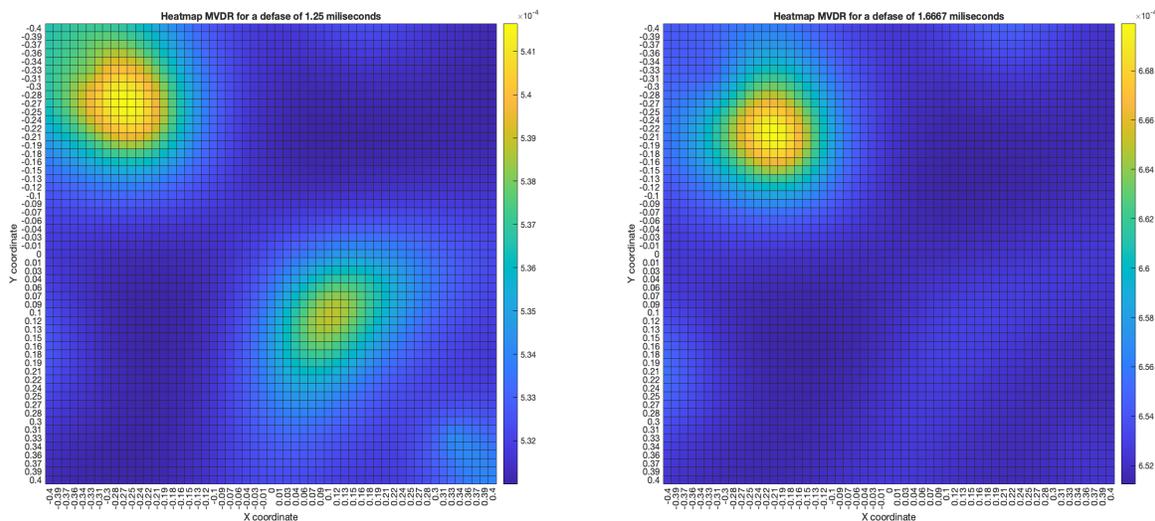
Los mapas de energía del formador de haz DAS se muestran en la Figura 7.7. Como puede verse, la posición del valor máximo de energía (y su “colina” o lóbulo circundante) aparece en una posición diferente a medida que aumentan las muestras de retraso, alejándose del nodo de error de sincronización. En las Figuras 7.7a,b, este comportamiento es consistente hasta el retraso de 60

muestras; a partir de ese momento, aparece otro lóbulo cerca del nodo de error de sincronización. Las tendencias “móviles” del formador de haz DAS también están presentes en el formador de haz MVDR, como se muestra en la Figura 7.8. Sin embargo, una diferencia importante es que el tamaño de los lóbulos es mucho más pequeño, lo que proporciona una mejor precisión.



(a) Mapa de calor con 20 muestras de retraso.

(b) Mapa de calor con 40 muestras de retraso.



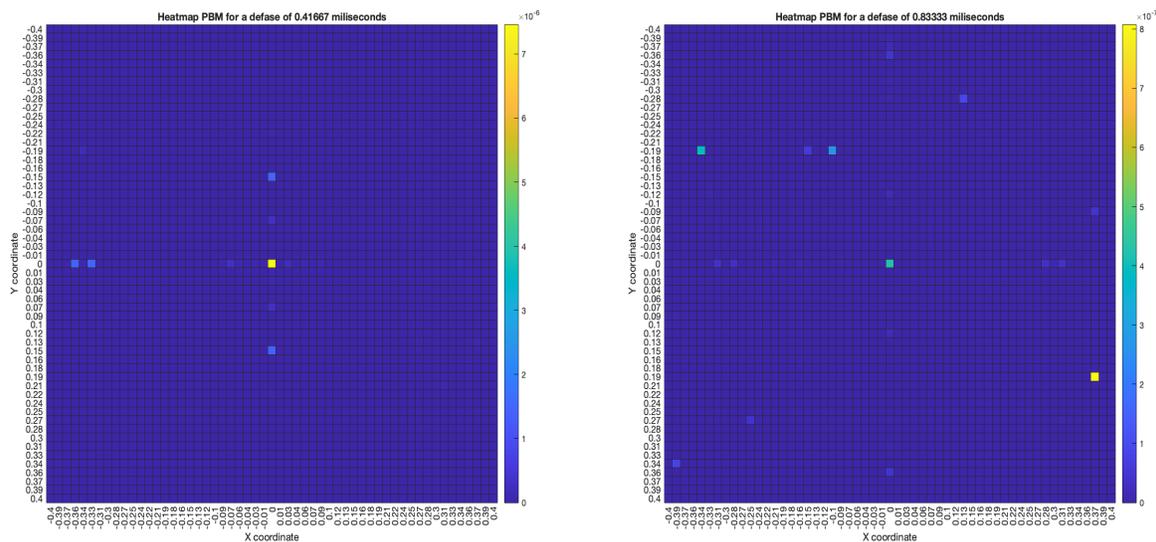
(c) Mapa de calor con 60 muestras de retraso.

(d) Mapa de calor con 80 muestras de retraso.

Figura 7.8: Mapas de calor para el formador de haz MVDR con errores de sincronización.

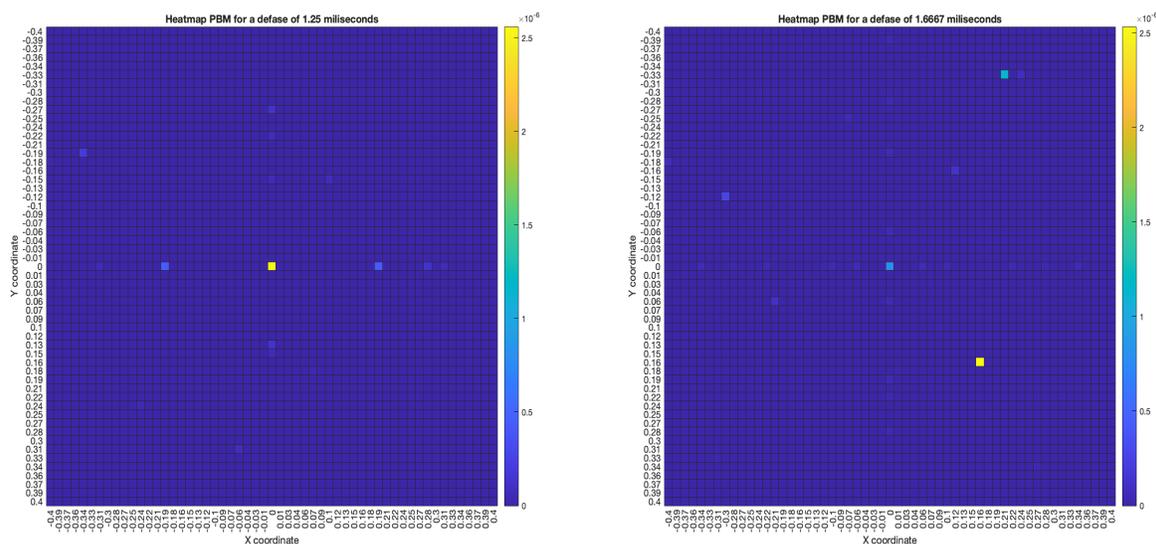
Las razones por las que ocurren estas tendencias son diversas, estos formadores de haz basan su funcionamiento en la diferencia de tiempo de llegada entre la fuente y los nodos, y al imponerle un retraso externo controlado, resulta en el efecto de que la fuente se “mueva”; los formadores de haz compensan estas diferencias de tiempo cambiando artificialmente las señales de entrada, lo que muestra una naturaleza cíclica, causando el “retorno” al nodo de error de sincronización. En

cualquier caso, puede observarse que para DAS y MVDR la forma del mapa de energía acústica cambia si los nodos de la red no están sincronizados. Esto da como resultado un cambio de la posición estimada de la fuente, determinada al ubicar el valor máximo de energía en el mapa. En cuanto al formador de haz PBM, como se puede ver en la Figura 7.9, el valor máximo de energía para algunos de los retrasos simulados, se ubica en la posición  $(x, y) = (0, 0)$ , sin embargo, hay algunos retrasos simulados para los cuales la fuente se localiza en otra posición completamente.



(a) Mapa de calor con 20 muestras de retraso.

(b) Mapa de calor con 40 muestras de retraso.



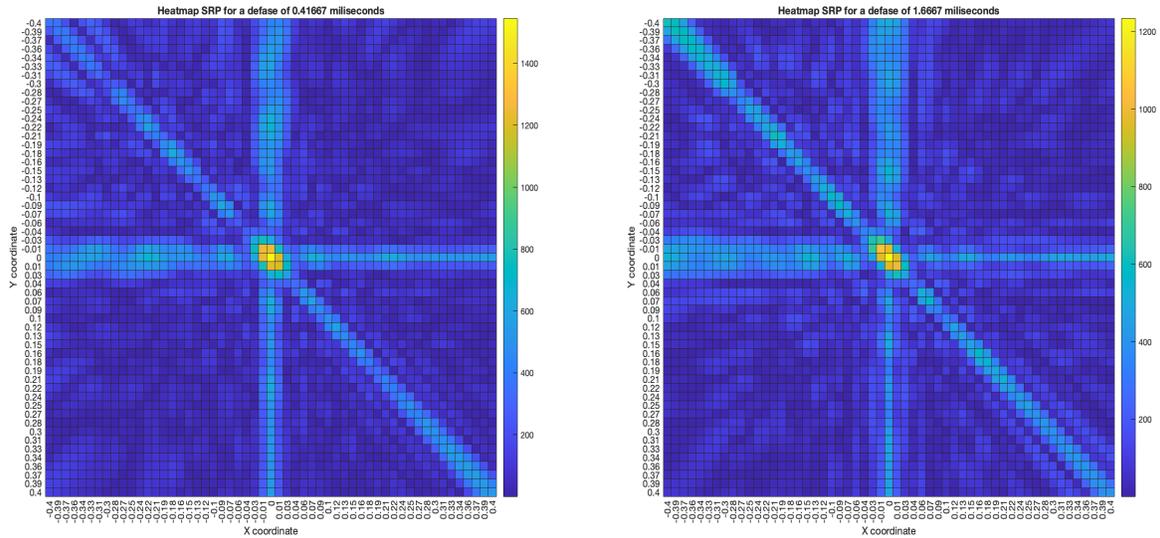
(c) Mapa de calor con 60 muestras de retraso.

(d) Mapa de calor con 80 muestras de retraso.

Figura 7.9: Mapas de calor para el formador de haz PBM con errores de sincronización.

Los mapas de calor que se muestran en la Figura 7.10 no son técnicamente mapas de energía acústica dado que el formador de haz SRP-PHAT no calcula la energía real (solo la potencia dirigida). Sin

embargo, todavía se pueden utilizar para localizar fuentes acústicas. Se puede ver que la forma del mapa de respuesta de potencia dirigida no cambia mucho para errores de sincronización que varían entre 20 y 80 muestras de retardo (razón por la cual no se muestran otros mapas de calor en la Figura 7.10). Por lo que, esta técnica se puede utilizar principalmente para determinar solo el punto de prueba más probable donde se puede ubicar una fuente única. Las líneas diagonales producidas en el mapa de energía acústica se dejarán para estudios posteriores, pero se tiene la intuición de que son producidas por una combinación entre la geometría del arreglo y la frecuencia de la señal.



(a) Mapa de calor con 20 muestras de retraso. (b) Mapa de calor con 80 muestras de retraso.

Figura 7.10: Mapas de calor para el formador de haz SRP-PHAT con errores de sincronización.

Con estos resultados es posible encontrar la posición de la fuente acústica localizando el valor máximo de energía. Por lo tanto, la siguiente parte del análisis es determinar, si la hay, una relación entre los errores de sincronización y el cambio de posición de la fuente acústica. Después de encontrar la posición de la fuente acústica, se calculó la diferencia entre la posición conocida de la fuente  $r_s$  y la posición obtenida de la fuente  $r_m$  de la siguiente manera:

$$\Delta \vec{r} = |\vec{r}_s - \vec{r}_m|. \quad (7.5)$$

Los resultados de la simulación en intervalos de 10 muestras se muestran en la Figura 7.11. Puede observarse que existe una tendencia en los formadores de haz DAS y MVDR a aumentar el error de localización a medida que aumenta el error de sincronización. Los formadores de haz PBM y SRP-PHAT presentan los errores de localización más bajos, pero el formador de haz PBM genera errores de localización para valores particulares de retardo (no con la tendencia observada en DAS y MVDR). El formador de haz SRP-PHAT tiene el rendimiento más consistente, pero dado el “ruido” de la forma de su mapa de potencia dirigida, solo debe usarse para la localización de fuentes acústicas, no para la generación de un mapa de energía acústica.

Como también se puede ver en la Figura 7.11, hay un “salto” considerable en el error de localización para retrasos simulados mayores a 60 muestras. Por lo tanto, este es el valor máximo de retardo que usamos para establecer una relación entre el error de localización y el error de sincronización. En la Figura 7.12, esta relación se muestra utilizando todos los retrasos posibles en el rango de 1

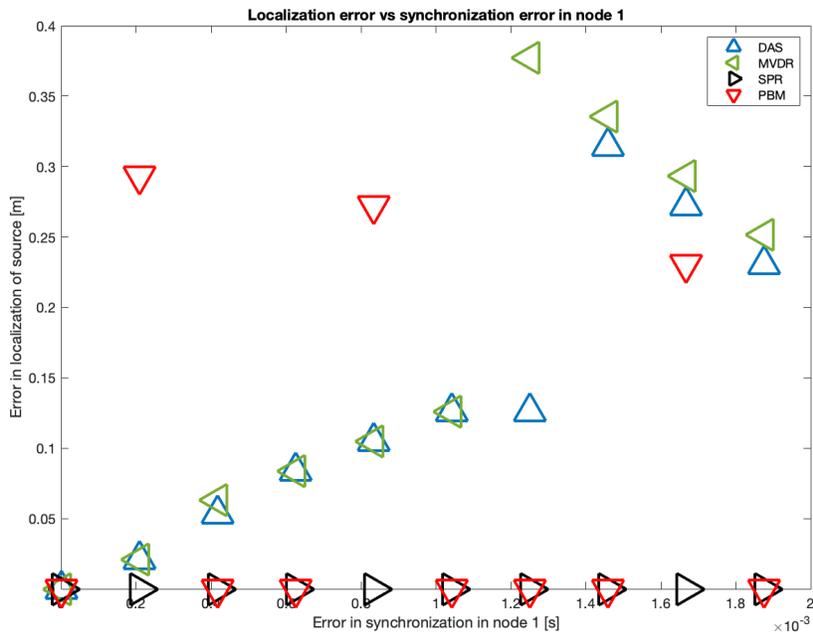


Figura 7.11: Errores de localización de una fuente en  $(x, y) = (0, 0)$  y errores de sincronización.

a 60 muestras:

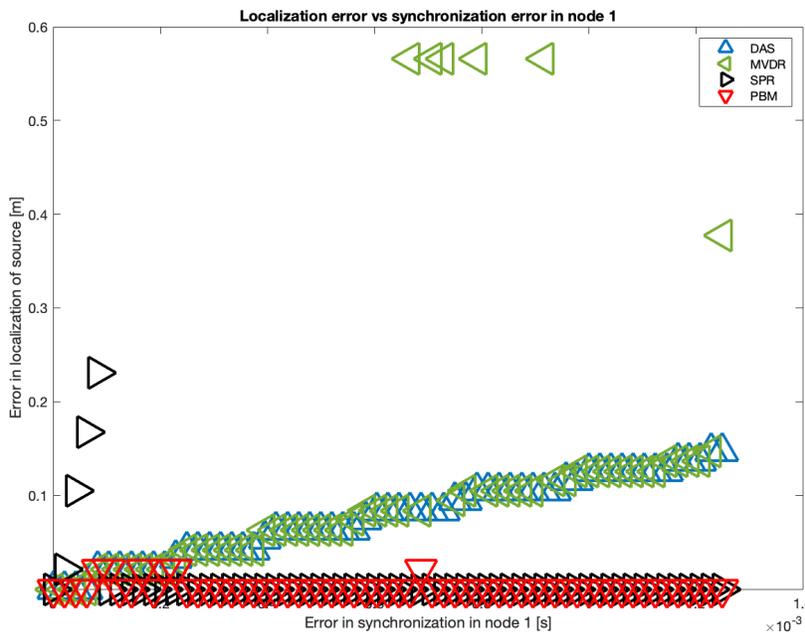


Figura 7.12: Errores de localización para todos los formadores de haz.

Como puede verse, existe una relación casi lineal entre el error de localización y la sincronización para DAS y MVDR. Aplicando un ajuste estadístico básico, estas son las ecuaciones que mejor

describen esta relación:

- $\Delta \vec{r}_{DAS} = 124.8\Delta t - 0.0002519$ , con  $R^2 = 0.9933$ .
- $\Delta \vec{r}_{MVDR} = 123.6\Delta t - 0.0019950$ , con  $R^2 = 0.9844$ .

donde  $\Delta t$  es el error de sincronización en el nodo 1. Es importante tener en cuenta que MVDR tiene algunos valores para  $\Delta t$  donde  $\Delta \vec{r}$  es mucho mayor que el resto, por lo que se trataron como valores atípicos para el proceso de ajuste. Los valores para el ajuste resultante para DAS y MVDR se muestra en la Figura 7.13:

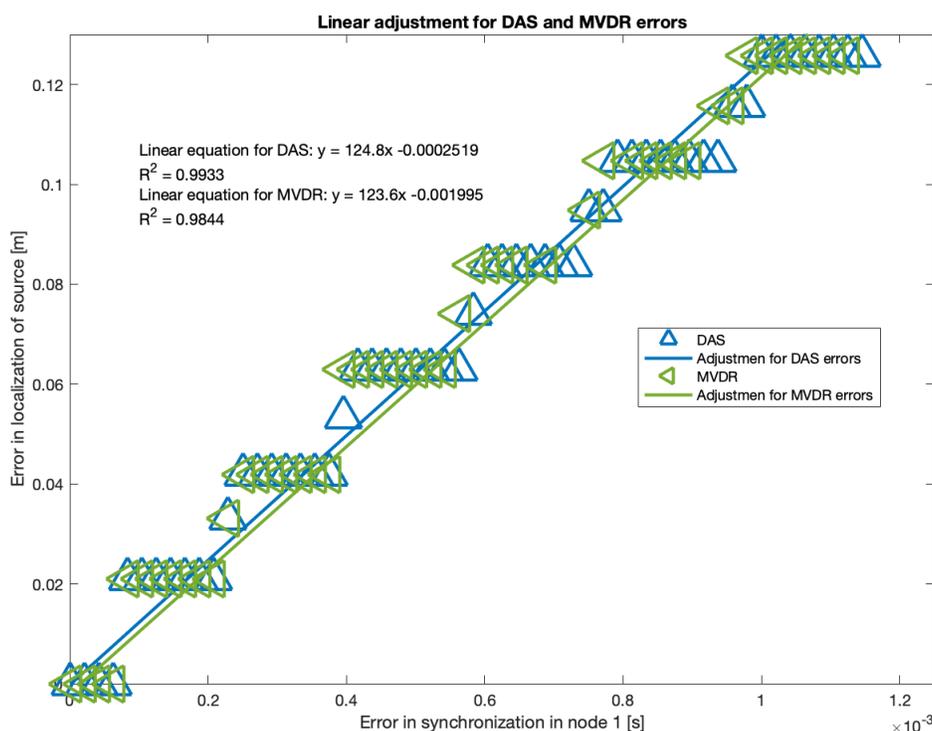


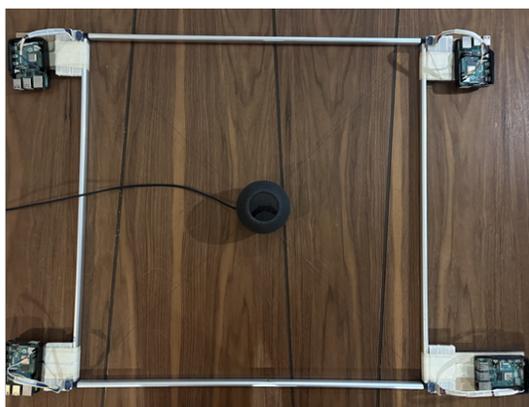
Figura 7.13: Ajuste para los formadores de haz: DAS y MVDR.

Hay un comportamiento de “escalón” en la Figura 7.13, que puede explicarse por la naturaleza de la retícula de puntos de prueba en la que se basa el mapa de energía acústica. Se realizaron simulaciones adicionales con un mayor número de puntos de prueba, y se observó que estos “escalones” estaban más cerca entre sí, es decir, la retícula se refina. Los valores máximos de errores de sincronización no cambiaron, pero el tiempo de procesamiento sí aumentó considerablemente. Por lo tanto, se debe hacer un compromiso: un comportamiento de “escalón” más refinado a cambio de un mayor tiempo de procesamiento.

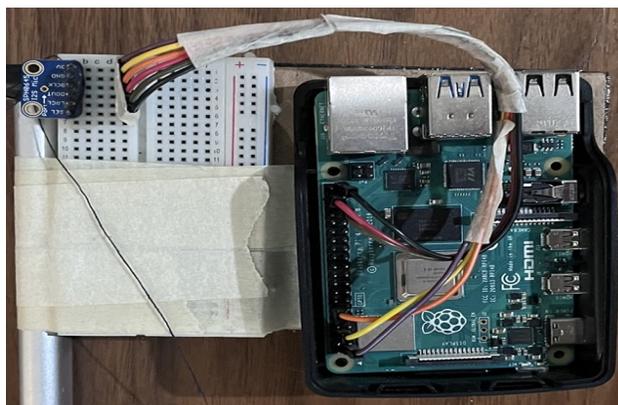
En cuanto a los formadores de haz PBM y SRP-PHAT, tienden a ubicar la fuente acústica exactamente en la posición donde se simuló de una manera casi consistente a lo largo de los retardos de tiempo simulados, con un error cercano a cero para el mismo límite de 60 muestras de retraso.

## 7.2. Resultados experimentales y análisis

En esta sección se analizarán las metodologías desarrolladas para la captura de señal acústica y su aplicación en la generación del mapa de energía acústica. Se utilizó una WASN compuesta por los dispositivos descritos en la Sección 3 para capturar la señal de una fuente acústica (un altavoz que emite una onda sinusoidal acústica de 300 Hz) colocada en el centro de la WASN. La red se construyó en un arreglo rectangular con las dimensiones presentadas en la Sección 7. La WASN y el nodo de captura se muestran en la Figura 7.14. La representación esquemática de la WASN y la fuente se muestra en la Figura 7.15.



(a) WASN y altavoz para pruebas.



(b) Nodo de captura.

Figura 7.14: Diseño experimental de la WASN para pruebas.

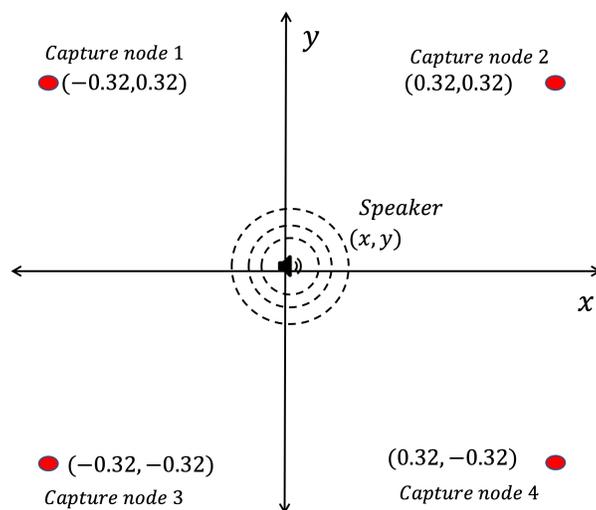
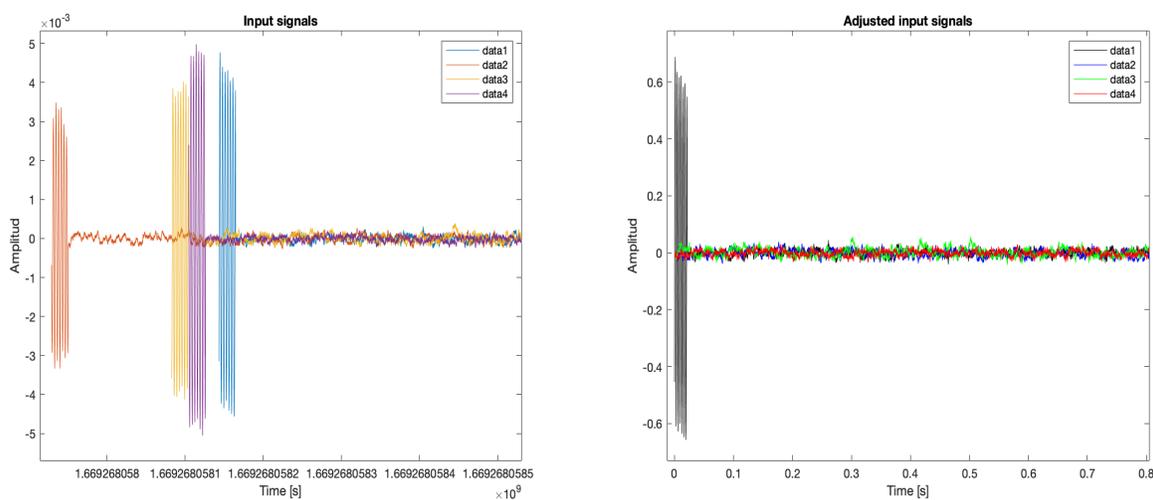


Figura 7.15: Arreglo experimental para la evaluación de las metodologías.

Cada uno de los nodos capturó una señal con los micrófonos MEMS. Luego, se aplicó un ajuste de post-procesamiento para la generación del modelo de energía acústica, de la siguiente manera:

1. **Recuperación de información y ajuste de marcas de tiempo.** Cada nodo genera archivos de audio y de marcas de tiempo para las metodologías de grabación local: un archivo de audio (en formato WAV) para la señal capturada y un archivo de texto con las marcas de tiempo generadas durante la grabación. Posteriormente, estos archivos son recuperados por una unidad de procesamiento. Las cuatro señales capturadas se muestran en la w 7.16a, sin ningún procesamiento posterior, con sus propias marcas de tiempo y valores de amplitud. Como se verá más adelante, cuando se implementan las metodologías locales para la captura de señales, cada nodo comienza a grabar en un momento diferente. Las marcas de tiempo de cada una de las grabaciones de los nodos se ajustan para comenzar en el momento en el que todas las marcas de tiempo están más próximas. Cada grabación tiene asociada la hora del reloj local, por lo que también es necesario establecer un marco de referencia global para la hora. Todas las señales se normalizan y ajustan para comenzar en  $t = 0$  como muestra la Figura 7.16b.

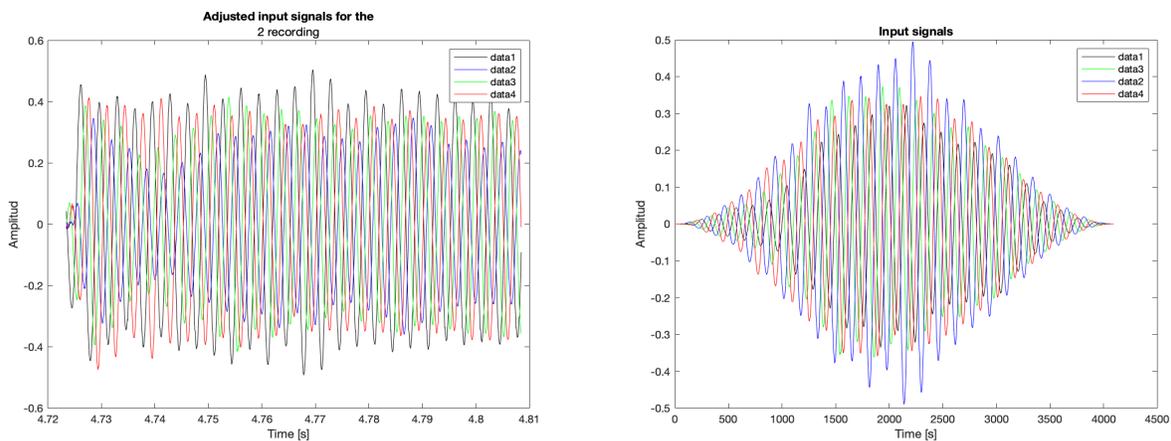


(a) Señales de entrada sin procesamiento.

(b) Señales de entrada ajustadas.

Figura 7.16: Procesamiento de archivos capturados por los nodos de la WASN.

2. **Selección de la muestra.** Con las marcas de tiempo ajustadas, se seleccionó un segmento de la señal completa para la generación del mapa de energía acústica. Se aplicó una ventana de Hann a este segmento para evitar efectos de *bleed-over* de frecuencia cuando las señales son procesadas por todos los formadores de haz que se muestran en la Figura 7.17a,b. La posible utilización y el efecto que tienen otras ventanas se deja para trabajos futuros.
3. **Generación del mapa de energía acústica.** Las señales capturadas ajustadas fueron procesadas por los formadores de haz para generar el mapa de energía acústica con las técnicas descritas en la Sección 4. En esta etapa del procesamiento también es posible determinar la posición de la fuente acústica encontrando el valor máximo en el mapa de energía acústica. Un ejemplo del mapa de calor generado a partir de las señales capturadas se muestra en la Figura 7.18.



(a) Muestra de señal acústica.

(b) Señal con una ventana de Hann.

Figura 7.17: Post-procesamiento de las señales de entrada.

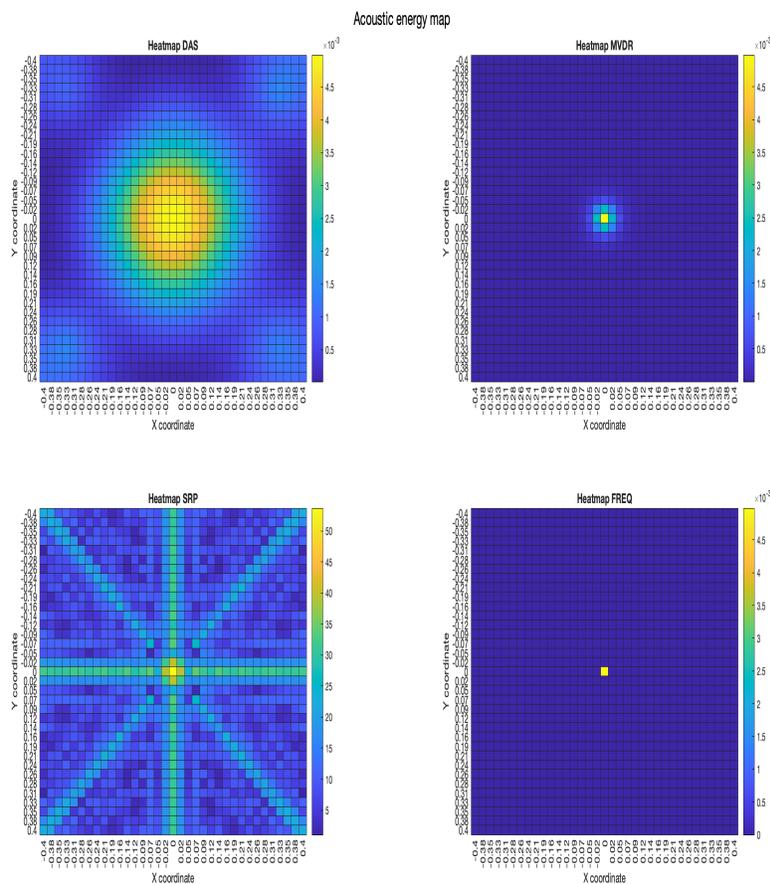


Figura 7.18: Mapas de calor generados a partir de las señales capturadas mediante las WASN con la fuente en la posición  $(0m, 0m)$ .

### 7.2.1. Análisis de sincronización

Dada la ubicación de la fuente acústica  $(x_f, y_f) = (0m, 0m)$ , su señal debería llegar a cada uno de los nodos de la WASN al mismo tiempo. De esta manera, el análisis de sincronización puede basarse en la diferencia de los tiempos de llegada de la señal acústica de la fuente a cada nodo de la WASN. Adicionalmente, también puede tomarse en cuenta la forma del mapa de energía acústica generado por las señales captadas por la WASN en comparación con las obtenidas en la simulación sin errores de sincronización.

Para el análisis de diferencia de tiempos de llegada para cada metodología de sincronización, se calculó el tiempo de llegada de la señal de la fuente a cada nodo y, tomando como referencia el nodo 1, se calcularon las diferencias de tiempos de llegada. En una WASN perfectamente sincronizada, con una fuente acústica puntual perfectamente situada en el centro, la diferencia de tiempos de llegada entre el nodo de referencia y todos los demás del WASN debería ser cero. Es importante mencionar que pueden ocurrir errores de sincronización debido a interferencias inalámbricas. Por lo tanto, cada metodología tendrá un desempeño diferente únicamente en función de sus características e implementación. Los resultados se muestran en Figura 7.19.

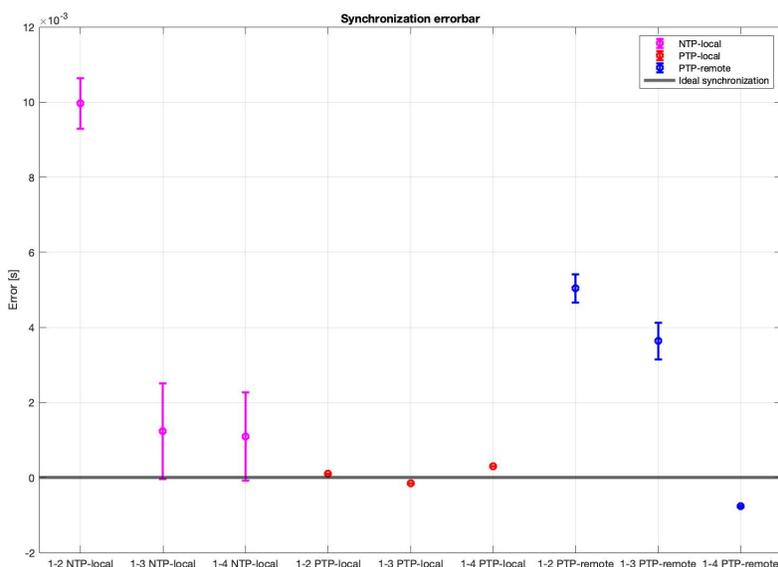


Figura 7.19: Errores de sincronización para cada nodo y cada metodología.

Estos resultados muestran que la metodología PTP-Local tiene los errores de sincronización más bajos, mientras que la metodología NTP-Local tiene los errores de sincronización más grandes. En cuanto al análisis de la forma del mapa, en las siguientes sub-secciones (una para cada metodología de sincronización) se presentarán dos mapas para cada formador de haz. El mapa de la izquierda es representativo de los mejores mapas obtenidos del formador de haz dada la metodología de sincronización, y el mapa de la derecha representa los peores mapas. Los mejores y peores mapas de energía se obtuvieron capturando la señal de una fuente de energía acústica con una WASN como se muestra en las Figuras 7.14 y 7.15. El proceso de evaluación implicó ejecutar cada metodología para capturar 30 muestras de 60 segundos cada una. Con estas señales, en post-procesamiento, se aplicaron los formadores de haz para generar un mapa de energía acústica a partir de cada muestra. El mapa que fue más similar al entorno acústico real se presenta aquí como el mejor mapa. El peor mapa es el que presentó la mayor discrepancia con el entorno acústico real.

### 7.2.2. Metodología NTP-Local

La metodología NTP-Local generó los peores mapas de energía acústica, lo cual es de esperar, ya que obtuvo los errores de sincronización más grandes. Por lo tanto, en los peores casos, los formadores de haz no generaron un mapa de energía acústica que representara la naturaleza del entorno acústico. Sin embargo, en los mejores casos donde los errores de sincronización fueron bajos, los mapas de energía acústica generados utilizando esta metodología pudieron ubicar un máximo de energía en forma de fuente acústica.

Para el formador de haz DAS, Figura 7.20, la forma del mapa de energía acústica no presenta un valor máximo como fuente puntual, sino que se distribuye en una región (o lóbulo). Para MVDR, Figura 7.21, el mapa de energía presenta dos regiones donde parecen existir dos fuentes acústicas, ninguna de las cuales corresponde a la posición real de la fuente acústica. Para PBM, Figura 7.22, el mejor de los casos presentó una fuente acústica casi puntual, pero no en la posición de la fuente acústica; en el peor de los casos, varios puntos de prueba presentaron una cantidad considerable de energía acústica, ninguno de los cuales representaba la posición de la fuente acústica. Finalmente, para SRP-PHAT, Figura 7.23, en el mejor mapa, el punto más probable para la localización de la fuente está cerca de la posición real; en el peor mapa, hay muchos puntos posibles donde se puede ubicar la fuente.

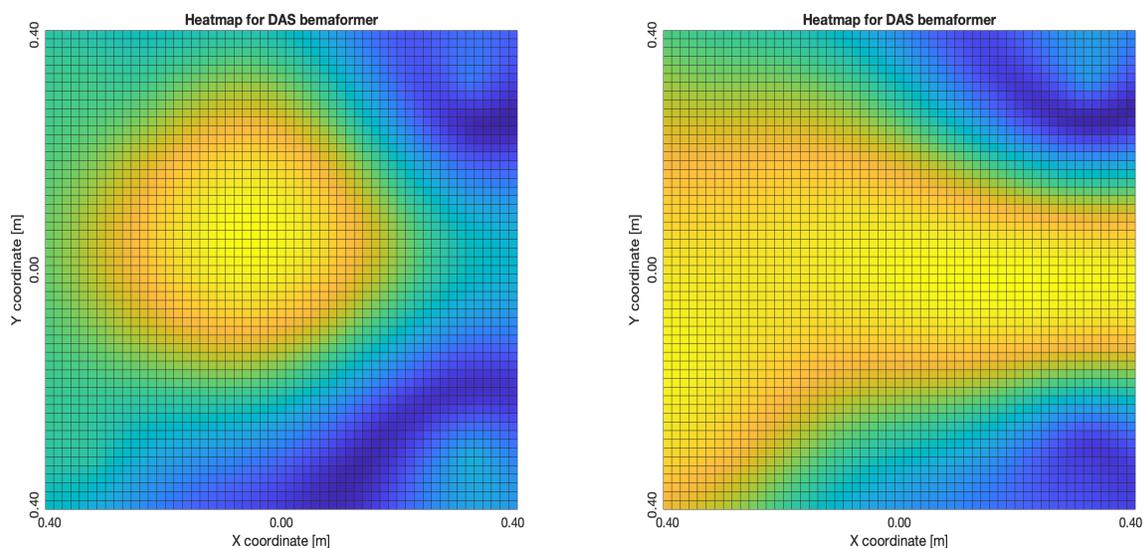


Figura 7.20: Mapas de energía del formador DAS (izquierda: mejor, derecha: peor).

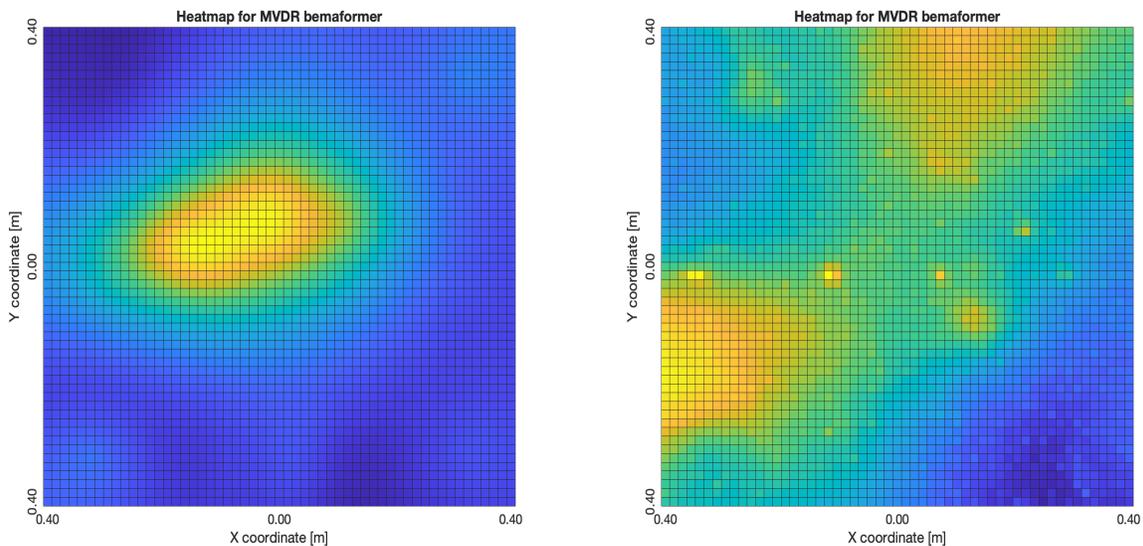


Figura 7.21: Mapas de energía del formador MVDR (izquierda: mejor, derecha: peor).

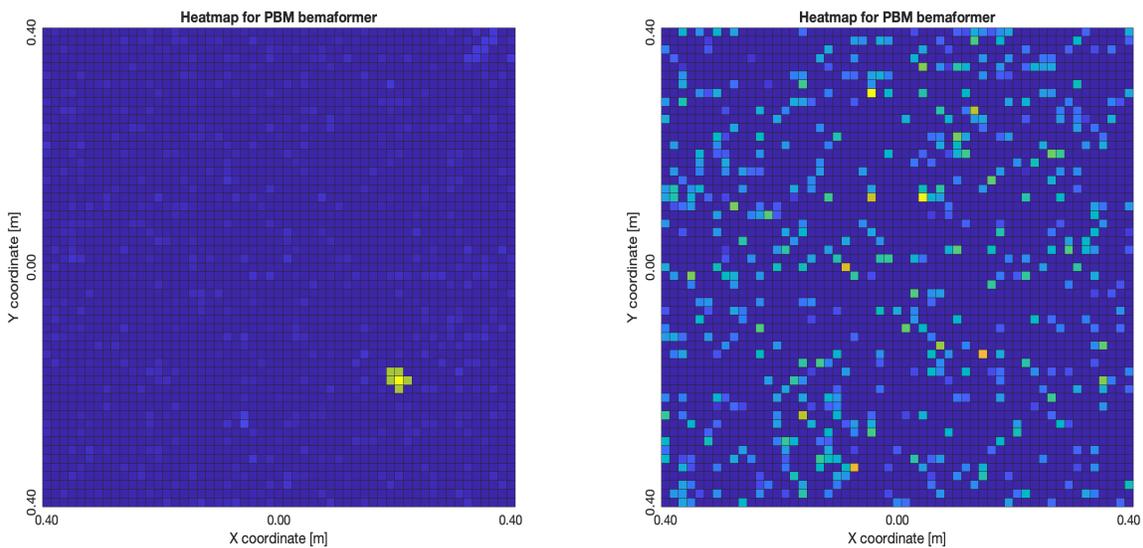


Figura 7.22: Mapas de energía del formador PBM (izquierda: mejor, derecha: peor).

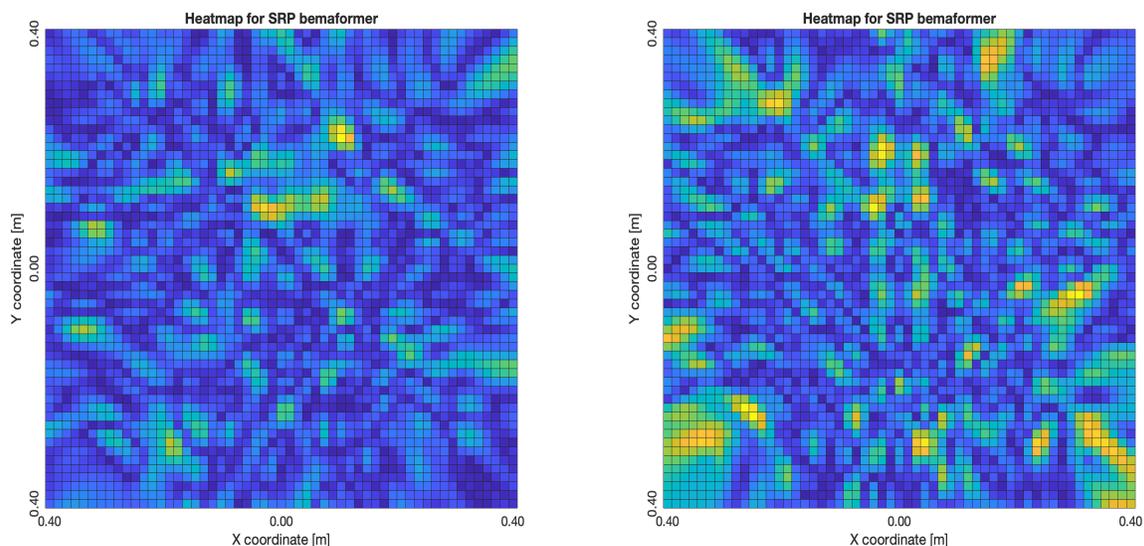


Figura 7.23: Mapas de energía del formador SRP-PHAT (izquierda: mejor, derecha: peor).

### 7.2.3. Metodología PTP-Local

La metodología PTP-Local muestra una mejora con respecto a la metodología NTP-Local en cuanto a la forma y posición de los lóbulos en el mapa de energía acústica. Fue la más consistente de las metodologías, para todos los formadores de haz, ubicando la fuente acústica cerca de su posición real.

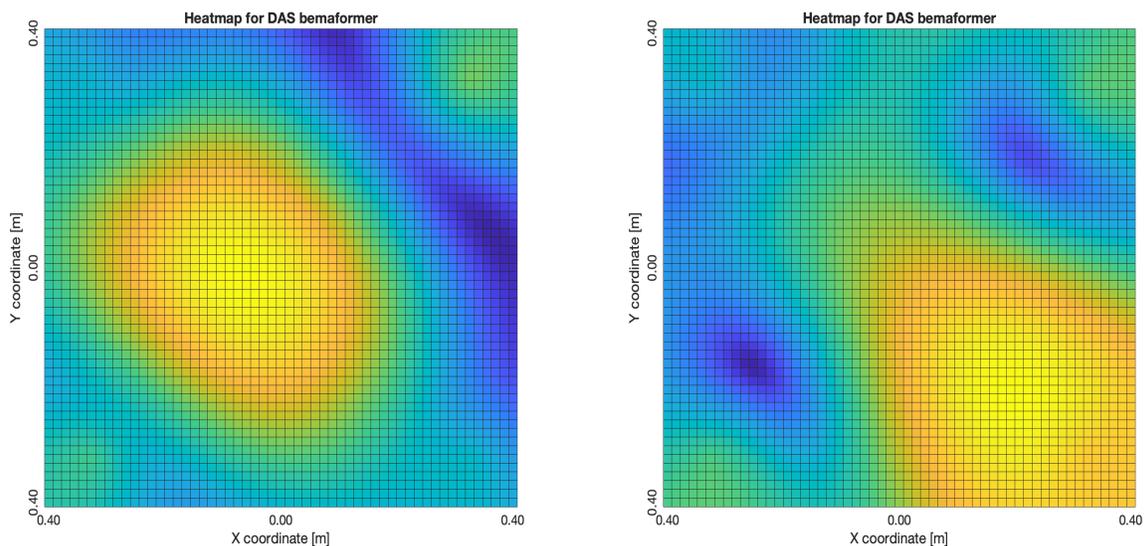


Figura 7.24: Mapas de energía del formador DAS (izquierda: mejor, derecha: peor).

Para el formador de haz DAS, Figura 7.24, incluso en el peor de los casos, el mapa de energía acústica siempre presentó un lóbulo con una energía acústica máxima localizada en la posición conocida de

la fuente. Para MVDR, la Figura 7.25, presenta lóbulos más pequeños (que son deseables ya que implican una mayor precisión en la localización de fuente acústicas), centrados alrededor de una ubicación cercana a la ubicación real de la fuente. Para PBM, Figura 7.26, presenta el mismo problema que presentaba con la metodología NTP-Local: cuando hay un error de sincronización, el formador de haz encuentra energía en muchos puntos de prueba del mapa de energía acústica. Para SRP-PHAT, Figura 7.27, también funciona de manera similar a la metodología NTP-Local.

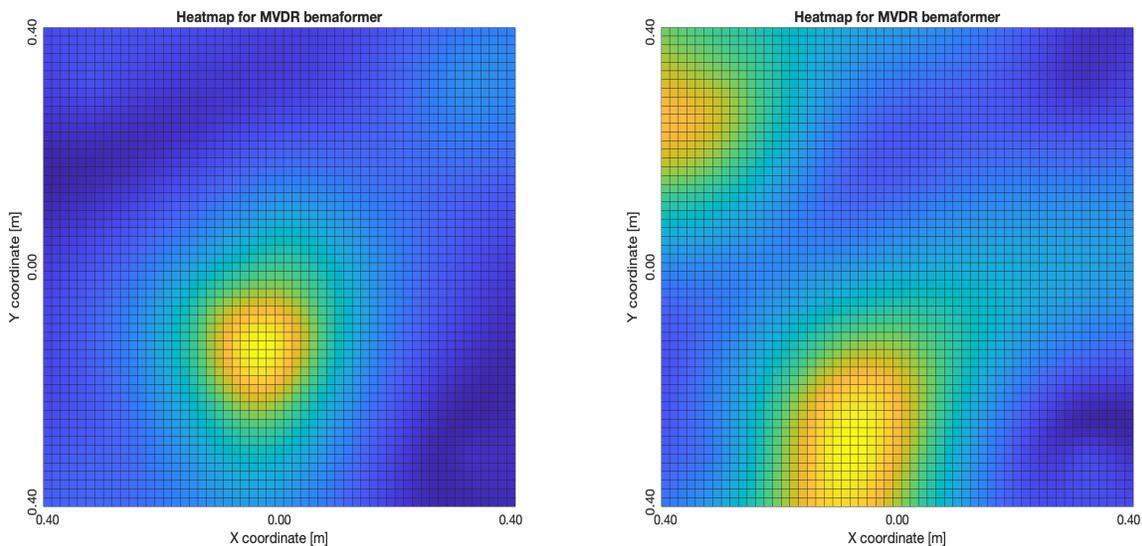


Figura 7.25: Mapas de energía del formador MVDR (izquierda: mejor, derecha: peor)

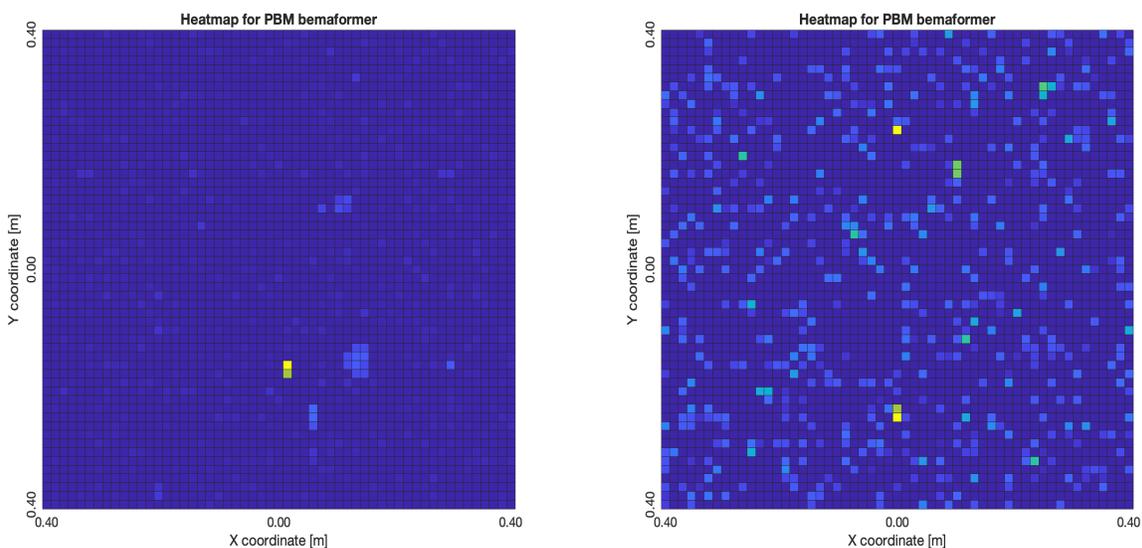


Figura 7.26: Mapas de energía del formador PBM (izquierda: mejor, derecha: peor).

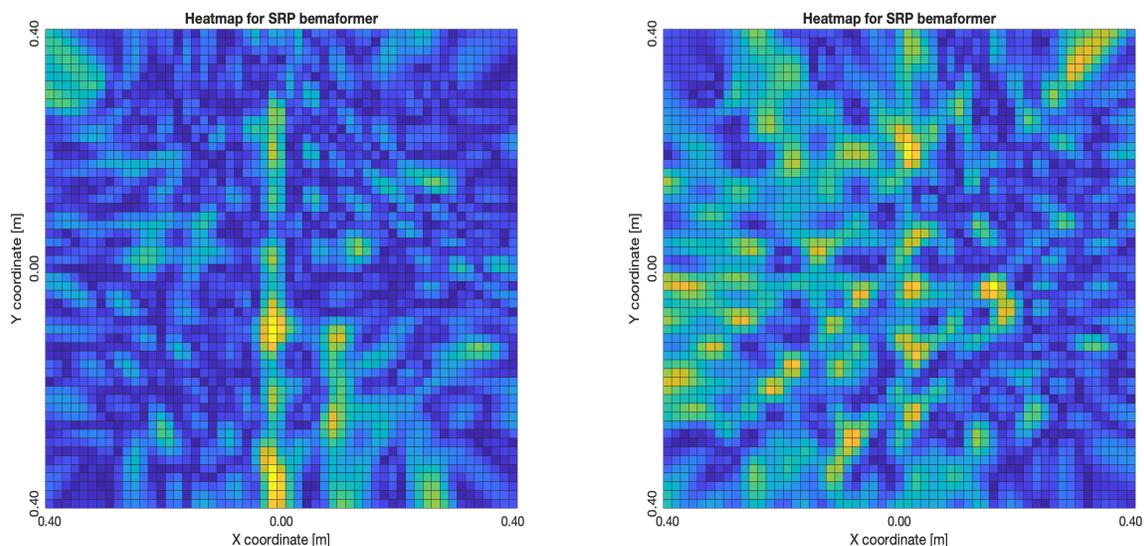


Figura 7.27: Mapas de energía del formador SRP-PHAT (izquierda: mejor, derecha: peor).

#### 7.2.4. Metodología PTP-Remota

La metodología PTP-Remota muestra que cuando la WASN está correctamente sincronizado, los mapas de energía acústica obtenidos se aproximan en precisión a los obtenidos con la metodología PTP-Local. Sin embargo, cuando la WASN presenta errores de sincronización, los mapas obtenidos tienen un desempeño con los mismos problemas que los observados en la metodología NTP-Local.

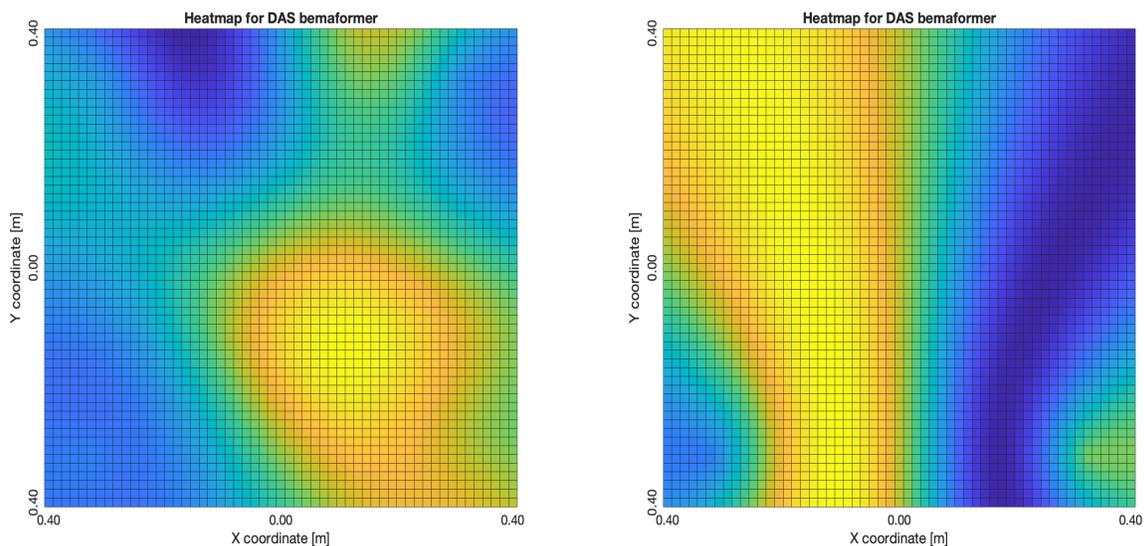


Figura 7.28: Mapas de energía del formador DAS (izquierda: mejor, derecha: peor).

Es importante mencionar que, si bien los peores mapas son similares a los peores mapas obtenidos utilizando la metodología NTP-Local para DAS, Figura 7.28, en el caso de MVDR, Figura 7.29 y

PBM, Figura 7.30, los mejores mapas están localizados de manera similar en comparación con los mejores mapas obtenidos usando la metodología PTP-Local, excepto SRP-PHAT, Figura 7.31. Además, adquirir los datos de audio es más práctico usando la metodología PTP-Remota en comparación con las otras dos metodologías (que requieren extracción manual de cada nodo).

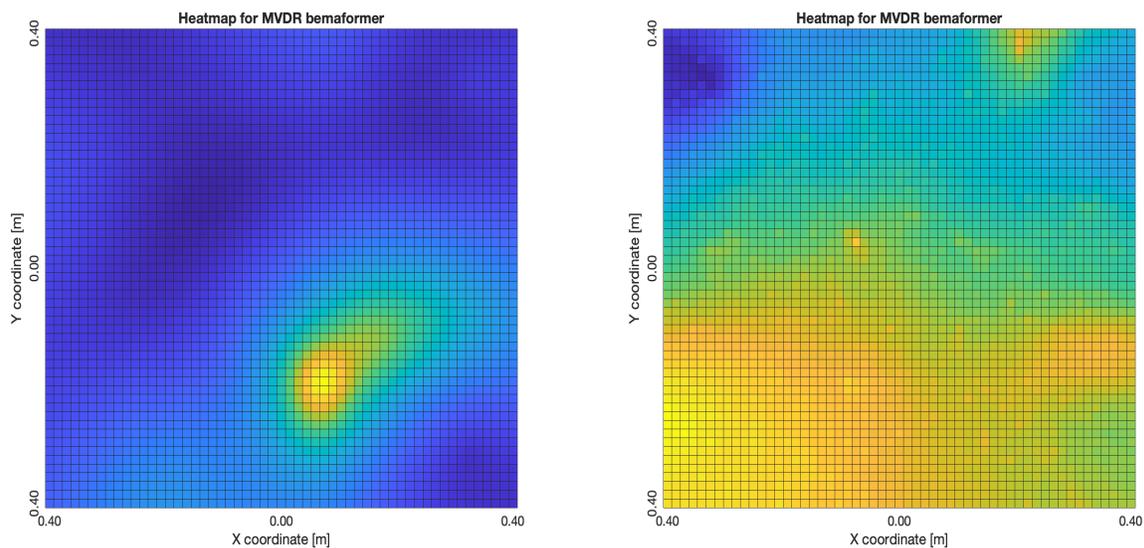


Figura 7.29: Mapas de energía del formador MVDR (izquierda: mejor, derecha: peor).

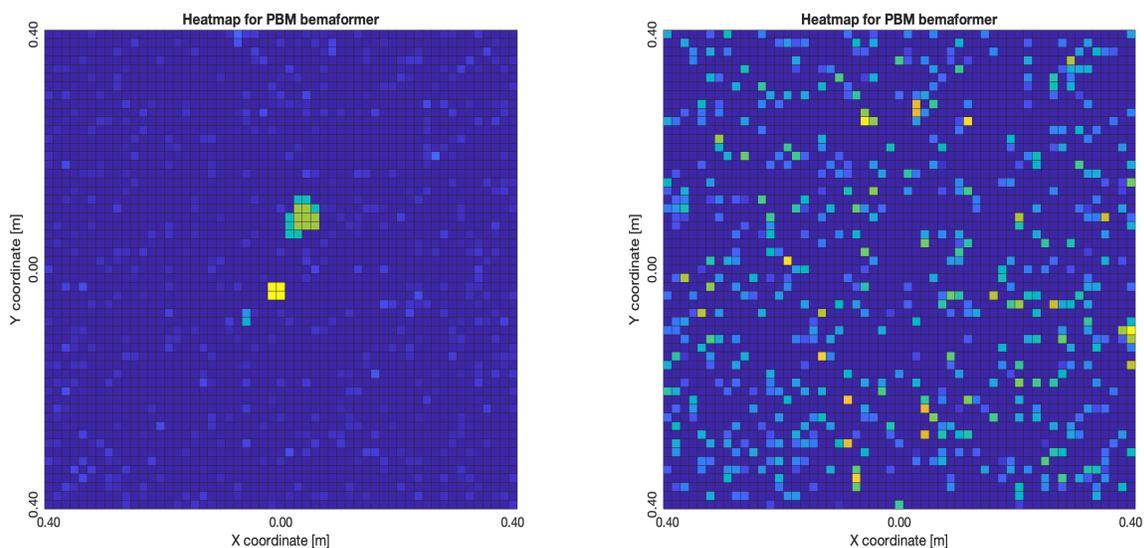


Figura 7.30: Mapas de energía del formador PBM (izquierda: mejor, derecha: peor).

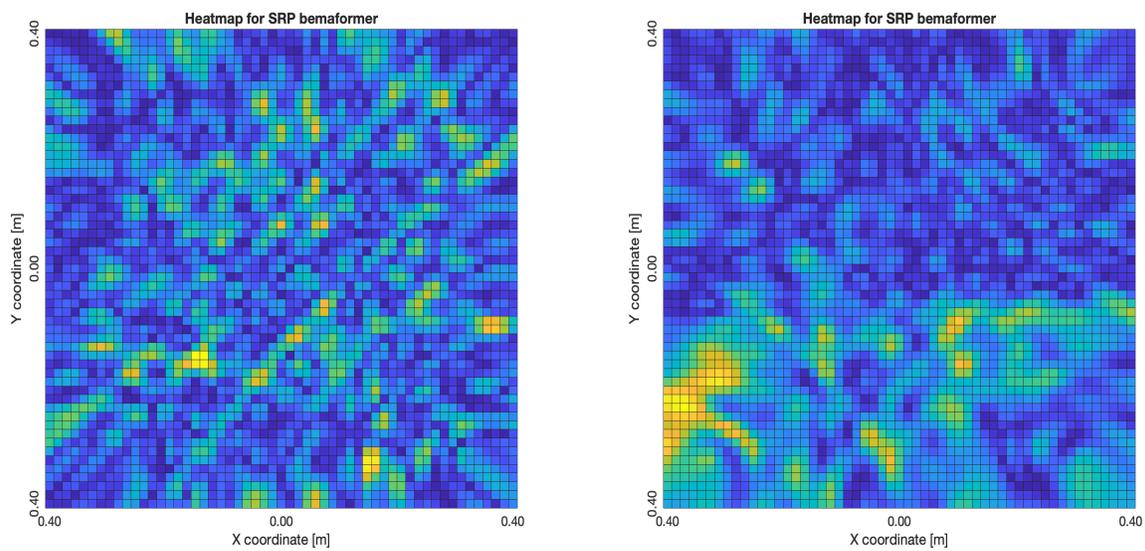


Figura 7.31: Mapas de energía del formador SRP-PHAT (izquierda: mejor, derecha: peor).



# Capítulo 8

## Discusión

El proyecto de investigación del doctorado tuvo tres etapas importantes de desarrollo, a continuación se discuten los resultados obtenidos en cada una de ellas.

1. **Diseño y caracterización de la WASN.** En esta etapa se propusieron y analizaron diferentes dispositivos para el diseño de la WASN. En cuanto al micrófono, se eligió el **Adafruit I2S MEMS Microphone Breakout** debido a su sensibilidad, *SNR* y al tipo de señal que se obtenía a la salida y su facilidad al conectarse con la unidad de procesamiento. Se probaron diferentes unidades de procesamiento para los nodos de la WASN y después de probar con un Arduino MKR Zero y un Raspberry Pi Zero W, se decidió utilizar el Raspberry Pi 4B+, con 1 GB de memoria RAM como unidad de procesamiento. Fue elegido porque permitía la captura de manera eficiente de la señal y mejora el rendimiento en términos de transmisión de información con respecto a las otras unidades. Con estos dos dispositivos se generaron nodos capaces de capturar información del ambiente acústico y transmitir-almacenar información.

Con este nodo diseñado, se procedió a la generación de la WASN, que consistió en conectar estos nodos para que simultáneamente fueran capaces de capturar información, el hecho de que la WASN se encuentra distribuida requirió que se realizara un estudio acerca de las estrategias y técnicas utilizadas para su sincronización. Se analizaron los diferentes protocolos de sincronización disponibles para ser implementados en WASN y se optó por utilizar y probar dos de ellos: el **NTP** y el **PTP**.

2. **Desarrollo y caracterización del modelo de mapeo acústico.** En esta etapa se desarrolló un modelo que a partir del procesamiento de señales capturadas mediante una WASN genere un mapa de energía acústica, y analizando su forma, poder determinar la localización de fuentes acústicas. El modelo basa su funcionamiento en el concepto de **formador de haz**, por lo que la primera labor de esta etapa consistió en el estudio de los diferentes formadores de haz desarrollados en la literatura y su aplicación en el modelo matemático desarrollado. Con el modelo definido, se procedió a analizar los elementos que afectan su funcionamiento, tanto en tiempo de procesamiento como en forma del mapa de energía obtenido. La discusión de estos resultados se encuentra en el Capítulo 7.

Posteriormente, se realizó la caracterización en función de la sincronización de la WASN para los formadores de haz: DAS, MVDR, PBM y SPR-PHAT. En términos de precisión en la localización de una fuente acústica simulada, PBM tiende a encontrar valores de energía altos solo en la posición simulada de la fuente, que es el comportamiento esperado ideal. Sin embargo, es importante tener en cuenta que PBM requiere que se establezca un parámetro

(el umbral de diferencia de fase) antes de la generación del mapa de energía acústica; el resto de los formadores de haz no requieren ninguna calibración *a priori*. En cuanto a SRP-PHAT, aunque técnicamente no crea un mapa de energía acústica, la ubicación con la potencia dirigida más alta coincide con la ubicación de la fuente acústica simulada, incluso cuando se insertan errores de sincronización. En cuanto a DAS y MVDR, aunque presentan mayores errores de localización cuando se simulan errores de sincronización, la relación entre estos errores es casi lineal. Esta información puede ser muy valiosa, ya que esta relación puede utilizarse para establecer un tipo de umbral de precisión dado un nivel esperado de errores de sincronización que, a su vez, puede estimarse a partir de la naturaleza de un escenario de aplicación determinado. Por lo tanto, a partir de los resultados del análisis experimental, un usuario puede saber de antemano qué tipo de errores de localización pueden ocurrir utilizando una WASN para crear un mapa de energía acústica.

Resulta interesante que DAS y MVDR proporcionan mapas de energía acústica de forma muy similar y comparten tendencias similares de cambios en la ubicación del lóbulo sobre la fuente acústica simulada, desplazándolo lejos del nodo donde se simula el error de sincronización hasta un error de 60 muestras. Esto puede explicarse por las características geométricas de la WASN y su relación con la longitud de onda de la señal. El análisis de la naturaleza de los formadores de haz y sus implicaciones en la generación de mapas de energía acústica quedan para estudios posteriores.

### 3. Metodologías para la captura y sincronización de la WASN y experimentación.

Se desarrollaron tres diferentes metodologías para la captura y sincronización de la WASN: **NTP-Local**, **PTP-Local** y **PTP-Remota**. La diferencia consistió en que en las metodologías locales, los nodos eran responsables de capturar y almacenar la información del ambiente acústico internamente, mientras que la grabación remota permite que los nodos sólo se encarguen de capturar la información y transmitirla a través de una red local. También cambia el protocolo de sincronización implementado en la WASN, teniendo dos posibilidades: NTP y PTP.

Los análisis experimentales muestran que puede generarse un mapa de energía acústica utilizando cualquiera de las metodologías de sincronización desarrolladas. Las metodologías NTP-Local y PTP-Remota proporcionan los mayores errores de sincronización, mientras que la metodología PTP-Local presentó los errores de sincronización más bajos, así como un comportamiento de dispersión más bajo. Además, cuando se ubicó una fuente acústica en el centro de la red, donde la diferencia en los tiempos de llegada entre los nodos debería acercarse a 0, la metodología PTP-Local proporcionó los valores y la dispersión más bajos. En el análisis de simulación, se encontró que 60 muestras eran un tipo de umbral superior de error de sincronización; con valores más altos, se observó un gran “salto” en los errores de localización con DAS y MVDR. Un error de sincronización de 60 muestras, muestreadas a 48 kHz, equivale a 1.25 ms. Conectando esto con la diferencia media observada de tiempos de llegada al nodo de captura de referencia ( $media_{1-2} = 0.0958$  ms,  $media_{1-3} = -0.1555$  ms, y  $media_{1-4} = 0,3035$  ms, en donde los subíndices indican entre qué nodos se calculó la media), al utilizar la metodología PTP-Local, puede deducirse que proporciona errores de sincronización muy por debajo del umbral mencionado y, por tanto, resulta adecuado para generar mapas de energía acústica con cualquiera de los formadores de haz descritos.

Habiendo generado el mapa de energía acústica utilizando todos los formadores de haz descritos junto con las metodologías de sincronización descritas, se encontró la posición de su valor máximo y se consideró como la posición estimada de la fuente acústica. Los errores de

localización en el escenario experimental se muestran en la Tabla 8.1. Las medidas estadísticas presentadas son la media y la desviación absoluta mediana (MAD)[41], que es una medida de dispersión no dependiente de su distribución. Estas cantidades se utilizan como factor de comparación entre metodologías.

Tabla 8.1: Análisis de error de localización para cada formador de haz y cada metodología

Metodología	Medida	DAS	MVDR	SRP-PHAT	PBM
NTP-Local	Media [m]	0.3429	0.3224	0.2814	0.3198
	MAD [m]	0.1013	0.1182	0.0823	0.1359
PTP-Local	Media [m]	0.2585	0.2533	0.3046	0.2151
	MAD [m]	0.1069	0.1126	0.0949	0.1247
PTP-Remota	Media [m]	0.2461	0.2922	0.3067	0.2803
	MAD [m]	0.0930	0.1005	0.1015	0.1021

Como puede observarse, estos errores de localización son considerablemente mayores que los encontrados en el análisis de simulación. Esto es de esperar, ya que un escenario acústico real tiene más variables que no fueron simuladas (reverberación, ruido, otras interferencias, etc.). Además, los errores de sincronización simulados se insertaron en un solo nodo (el nodo de error de sincronización), mientras que en el escenario del experimento todos los nodos se ven afectados por tales problemas. Finalmente, el altavoz utilizado en los experimentos no es una fuente acústica puntual, como ocurría en la simulación. Sin embargo, dados todos estos factores, es importante señalar que dado que estos mapas acústicos se generaron a partir de señales capturadas mediante una WASN, la ubicación de los nodos, al ser inalámbrica, no está fija, por lo que se pueden ubicar fácilmente en cualquier lugar del escenario acústico, como en una sala de estar. Teniendo esto en cuenta, los errores de localización observados son razonablemente bajos, y PBM que utiliza la metodología PTP-Local proporciona el error medio más bajo. SRP-PHAT proporciona la varianza más baja (también utilizando la metodología PTP-Local), sin embargo, proporciona el mayor error medio, por lo que puede no ser tan adecuado como PBM. Para trabajos posteriores se utilizará la metodología PTP-Local en conjunto con los formadores de haz MVDR y PBM, y se buscarán estrategias para mejorar su desempeño en la creación de mapas de energía acústica.



# Capítulo 9

## Conclusiones

Se enumeran a continuación las conclusiones del trabajo de investigación desarrollado a lo largo del proyecto.

1. Se propusieron diferentes dispositivos y después de analizar su funcionamiento, se concluye que un micrófono **Adafruit I2S MEMS Microphone Breakout**, un **Raspberry Pi 4B+** en conjunto con el protocolo de sincronización **PTP** y la metodología de grabación local permiten obtener señales del ambiente acústico.
2. Se propuso y caracterizó un modelo matemático para la generación de un mapa de energía acústico aplicando cuatro técnicas populares de formador de haz: Delay-and-Sum (DAS), Minimum Variance Distortionless Response (MVDR), Steered-Response Power with Phase Transform (SRP-PHAT) y Phase-based Binary Masking (PBM). Resultado del análisis, se propone utilizar los formadores de haz **MVDR** y **PBM** como los más adecuados para la aplicación de mapeo acústico. Cada uno de los formadores ofrece ventajas y se propone como trabajo futuro analizarlas y una posible combinación de ambos para mejorar el mapa de energía acústico.
3. Se propusieron tres metodologías para la captura y sincronización de la WASN: **NTP-Local**, **PTP-Local** y **PTP-Remota**. Las metodologías de sincronización están conformadas por: (1) un protocolo de captura WASN (ya sea local o remoto) utilizado para capturar una señal de audio multicanal del entorno acústico; y (2) un protocolo de sincronización, que podría ser Network Time Protocol (NTP) o Precision Time Protocol (PTP). Luego se aplicaron las técnicas de formador de haz a las señales capturadas, para medir la energía proveniente de una colección de puntos de prueba (dispuestos en una cuadrícula) para generar el mapa de energía acústica. De acuerdo con los resultados obtenidos, la metodología de captura de señal y sincronización de la WASN más adecuada para el mapeo acústico es la **PTP-Local**, pero se propone para trabajos futuros desarrollar estrategias para que la metodología **PTP-Remota** sea igual de viable.
4. Se realizó una simulación para, en primer lugar, validar esta prueba de concepto y caracterizar el comportamiento de cada técnica de formador de haz frente a un error de sincronización simulado. Además, se utilizó un arreglo experimental para probar las metodologías de sincronización. Con los resultados obtenidos por simulación y experimentación, se pueden establecer las siguientes conclusiones:
  - En la simulación, SRP-PHAT y PBM mostraron mayor robustez frente a errores de sincronización. DAS y MVDR fueron más sensibles, pero se encontró una relación casi

lineal entre los errores de localización y los errores de sincronización que puede usarse para facilitar la expectativa del usuario.

- En el ámbito experimental, desde un punto de vista subjetivo, MVDR y PBM proporcionaron un mapa de energía acústica cercano a lo esperado, ya sea con un lóbulo o un punto sobre la ubicación de la fuente acústica; SRP-PHAT y DAS proporcionaron mapas inesperados. Si bien PBM requiere que se calibre un parámetro de antemano, genera un mapa de energía acústica más preciso. Y, si bien MVDR genera un mapa de energía acústica menos preciso, es resistente a los elementos ambientales (como el ruido y las interferencias) y no requiere ninguna calibración de parámetros *a-priori*.
- Las metodologías PTP-Local y PTP-Remota son las más adecuadas para el mapeo de energía acústica utilizando una WASN. La metodología PTP-Local tiene el mejor desempeño en cuanto a errores de sincronización, pero el proceso de captura, recuperación y procesamiento de los datos recolectados es más tedioso que la metodología PTP-Remota, ya que requiere que el usuario extraiga manualmente la grabación de cada nodo de captura. La metodología PTP-Remota es menos confiable en términos de sincronización (con un valor medio de error más alto). Sin embargo, las señales capturadas aún se pueden usar para proporcionar un mapa de energía acústica casi preciso (ya sea con PBM o MVDR), mientras que son más fáciles de implementar (ya que no se requiere que un agente de sincronización se ejecute localmente en los nodos de captura), y las grabaciones se transmiten directamente al servidor (no es necesaria la extracción manual).

Para trabajos futuros, se explorarán otros protocolos de sincronización, el efecto de diferentes ventanas (distintas a Hann, utilizadas en este trabajo) para reducir las discontinuidades en el dominio del tiempo y análisis de diferentes formadores haz y combinaciones de ellos.

# Bibliografía

- [1] Arce, P.; Salvo, D.; Piñero, G.; Gonzalez, A. FIWARE based low-cost wireless acoustic sensor network for monitoring and classification of urban soundscape. *Comput. Netw.* 2021, *196*, 108199.
- [2] Verreycken, E.; Simon, R.; Quirk-Royal, B.; Daems, W.; Barber, J.; Steckel, J. Bio-acoustic tracking and localization using heterogeneous, scalable microphone arrays. *Commun. Biol.* 2021, *4*, 1275.
- [3] Xiong, H.; Agcayazi, T.; Latif, T.; Bozkurt, A.; Sichertiu, M.L. Towards acoustic localization for biobotic sensor networks. In Proceedings of the 2017 IEEE Sensors, Glasgow, UK, 29 October–1 November 2017; pp. 1–3.
- [4] Whitmire, E.; Latif, T.; Bozkurt, A. Acoustic sensors for biobotic search and rescue. In Proceedings of the SENSORS, 2014 IEEE, Valencia, Spain, 2–5 November 2014; pp. 2195–2198.
- [5] Hollosi, D.; Schröder, J.; Goetze, S.; Appell, J.–E. Voice activity detection driven acoustic event classification for monitoring in smart homes. In Proceedings of the 2010 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies, Rome, Italy, 7–10 November 2010; pp. 1–5.
- [6] Cobos Maximo, Antonacci Fabio, Alexandridis Anastasios, Mouchtaris Athanasios, Lee Bowon. *A Survey of Sound Source Localization Methods in Wireless Acoustic Sensor Networks*. Wireless Communications and Mobile Computing, Volume 2017.
- [7] Z. Liu; Z. Zhang; L. He.; P. Chou. Energy-based sound source localization and gain normalization for ad hoc microphone arrays. In Proceedings of the 2007 IEEE International Conference on Acoustics, Speech and Signal Processing–ICASSP '07, Honolulu, HI, EUA, 15–20 April 2007.
- [8] T.H. de Groot. *Localization and Classification using an Acoustic Sensor Network*. Tesis para obtener el grado de Master of Science in Electrical Engineering en la Delft University of Technology, Delft, 2010.
- [9] A. Griffin; A. Alexandridis; D. Pavlidi; Y. Mastorakis; A. Mouchtaris Localizing multiple audio sources in a wireless acoustic sensor network. *Signal Process.* **2015**, *107*, 54–67.
- [10] H. Adel, M. Souad, A. Alaqueeli, and A. Hamid, *Beamforming Techniques for Multichannel audio Signal Separation*. JDCTA: International Journal of Digital Content Technology and its Applications, Vol. 6, No. 20, pp. 659–667, 2012
- [11] Jacob Benesty, Jingdong Chen, Yiteng Huang. *Microphone Array Signal Processing*. Springer Topics in Signal Processing, Berlin, 2008, pp. 39–45.

- [12] Wang, R.; Wang, Y.; Han, C.; Gong, Y.; Wang, L. Robust Adaptive Beamforming Based on Interference Covariance Matrix Reconstruction and Steering Vector Estimation. In Proceedings of the 2021 IEEE International Conference on Signal Processing, Communications and Computing, 117–19 August, Xi'an, China, 2021.
- [13] Rastko R. Selmic, Vir V. Phoha, Abdul Serwadda. *Wireless Sensor Networks. Security, Coverage, and Localization*. Springer USA, 2016.
- [14] Ginovart–Panisello, G.J.; Vidaña–Vila, E.; Caro–Via, S.; Martínez–Suquía, C.; Freixes, M.; Alsina–Pagès, R.M. Low–Cost WASN for Real–Time Soundmap Generation. *Eng. Proc.* 2021, 6, 57.
- [15] Yang, Y. Time Synchronization in Wireless Sensor Networks: A Survey. Tesis de Licenciatura, Mittuniversitetet, Sundsvall, Sweden, 2012.
- [16] Shoubin Wang Mingxing Shi, Donghui Li, Tao Du. *A Survey of Time Synchronization Algorithms for Wireless Sensor Networks*. Proceedings of the 38th Chinese Control Conference, Guangzhou, China, 2019, pp. 6392–6397.
- [17] Fengyuan Ren, Chuang Lin. *Self-correcting time synchronization using references broadcast in wireless sensor network*. IEEE Wireless Communications, vol. 15, no. 4, pp. 79–85, 2008.
- [18] Wang Jingchao, Zhao Ruohan, Gu Weiwen. *Time Synchronization in Networks: A Survey*, In Proceedings of the 2nd International Conference on Control and Computer Vision (ICCCV '19). Association for Computing Machinery, New York, NY, EUA, pp 121–126. 2019
- [19] Fikret Sivrikaya and Bulent Yener. *Time Synchronization in Sensor Networks: A Survey*, IEEE Network, vol. 18, no. 4, pp. 45–50, Julio-Ago. 2004
- [20] Sami M. Lasassmeh and James M. Conrad. *Time Synchronization in Wireless Sensor Networks: A Survey*. Proceedings of the IEEE SoutheastCon 2010 (SoutheastCon), Concord, NC, EUA, pp. 242–245, 2010.
- [21] Saurabh Ganeriwal, Ram Kumar, Mani B. Srivastava *Timing-sync Protocol for Sensor Networks*, In Proceedings of the 1st international conference on Embedded networked sensor systems (SenSys '03). Association for Computing Machinery, New York, EUA, pp. 138–149, 2003.
- [22] Darryl Veitch, Julein Ridoux, Satish Babu Korada, *Robust Synchronization of Absolute and Difference Clocks over Networks*. in IEEE/ACM Transactions on Networking, vol. 17, no. 2, pp. 417–430, Abril 2009.
- [23] Neagoe, T.; Cristea, V.; Banica, L. NTP versus PTP in Computer Networks Clock Synchronization. In Proceedings of the 2006 IEEE International Symposium on Industrial Electronics, Montreal, QC, Canada, 9–13 July 2006; pp. 317–362.
- [24] Jingchao, W.; Ruohan, Z.; Weiwen, G. Time Synchronization in Networks: A Survey. In Proceedings of the 2nd International Conference on Control and Computer Vision, Jeju Republic of Korea, June 15–18, 2019; pp. 121–126.
- [25] Ferrari, P.; Flammini, A.; Rinaldi, S.; Bondavalli, A.; Brancati, F. Experimental Characterization of Uncertainty Sources in a Software–Only Synchronization System. *IEEE Trans. Instrum. Meas.* 2012, 61, 1512–1521.

- [26] <https://jackaudio.org>
- [27] [https://github.com/jackaudio/jackaudio.github.com/wiki/WalkThrough\\\_User\\\_NetJack2](https://github.com/jackaudio/jackaudio.github.com/wiki/WalkThrough\_User\_NetJack2)
- [28] Kino, Gordon S. *Acoustic waves : devices, imaging, and analog signal processing*. 1987.
- [29] Perrot, V.; Polichetti, M.; Varray, F.; Garcia, D. So you think you can DAS? A viewpoint on delay–and–sum beamforming. *Ultrasonics* 2021, *111*, 106309.
- [30] Capon, J. High–resolution frequency–wavenumber spectrum analysis. *Proc. IEEE* 1969, *57*, 1408–1418.
- [31] Henry Cox, Robert M. Zeskind, Mark M. Owen. *Robust Adaptive Beamforming*. IEEE Transactions on acoustics, speech, and signal processing, vol. ASSP-35N, O. 10, 1987.
- [32] Ronggui Wang, Yuexian Wang, Chuang Han, Yanyun Gong, Ling Wang. *Robust Adaptive Beamforming Based on Interference Covariance Matrix Reconstruction and Steering Vector Estimation*. 2021 IEEE International Conference on Signal Processing, Communications and Computing, 2021.
- [33] Yuteng Xiao, Jihang Yin, Honggang Qi, Hongsheng Yin, Gang Hua. *MVDR Algorithm Based on Estimated Diagonal Loading for Beamforming*. Mathematical Problems in Engineering Volume 2017, 2017.
- [34] He, L.; Zhou, Y.; Liu, H. Phase Time–Frequency Masking Based Speech Enhancement Algorithm Using Circular Microphone Array. In Proceedings of the 2019 IEEE International Conference on Multimedia and Expo (ICME), Shanghai, China, 8–12 Julio 2019; pp. 808–813.
- [35] Rascon, C. A Corpus–Based Evaluation of Beamforming Techniques and Phase–Based Frequency Masking. *Sensors* 2021, *21*, 5005.
- [36] Diaz–Guerra, D.; Beltran, J.R. Direction of Arrival Estimation with Microphone Arrays Using SRP–PHAT and Neural Networks. In Proceedings of the 2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM), Sheffield, UK, 8–11 Julio 2018; pp. 617–621.
- [37] Zhuo, D.–B.; Cao, H. Fast Sound Source Localization Based on SRP–PHAT Using Density Peaks Clustering. *Appl. Sci.* 2021, *11*, 445.
- [38] <https://manpages.ubuntu.com/manpages/xenial/man8/ntpd.8.html>
- [39] <https://manpages.debian.org/testing/ptpd/ptpd.conf.5.en.html>
- [40] J. A. C. Weideman, Numerical Integration of Periodic Functions: A Few Example; The American Mathematical Monthly 109, Taylor and Francis; pp 21–36; 2002.
- [41] Sachs, L. *Applied Statistics: A Handbook of Techniques*; Springer: New York, NY, USA, 1984; p. 253.