

ESCUELA TÉCNICA SUPERIOR DE INGENIEROS  
INDUSTRIALES Y DE TELECOMUNICACIÓN

UNIVERSIDAD DE CANTABRIA



*Trabajo Fin de Grado*

**SELECCIÓN DE SENSORES EN SISTEMAS  
DE FUSIÓN DE DECISIONES  
INALÁMBRICAS**

**(Sensor selection in Wireless Decision Fusion  
Systems)**

Para acceder al Título de

***Graduado en  
Ingeniería de Tecnologías de Telecomunicación***

Autor: Adrián Martín Ruiz

Julio- 2025



**GRADUADO EN INGENIERÍA DE TECNOLOGÍAS DE  
TELECOMUNICACIÓN**

**CALIFICACIÓN DEL TRABAJO FIN DE GRADO**

**Realizado por:** Adrián Martín Ruiz

**Director del TFG:** Jesús Pérez Arriaga

**Título:** “Selección de sensores en sistemas de fusión de decisiones inalámbricas”

**Title:** “Sensor selection in wireless decision fusion systems”

**Presentado a examen el día:** 8 de Julio de 2025

para acceder al Título de

**GRADUADO EN INGENIERÍA DE TECNOLOGÍAS DE  
TELECOMUNICACIÓN**

Composición del Tribunal:

Presidente (Apellidos, Nombre): Santamaría Caballero, Luis Ignacio

Secretario (Apellidos, Nombre): Prieto Torralbo, Pablo

Vocal (Apellidos, Nombre): Quintela Incera, Antonio

Este Tribunal ha resuelto otorgar la calificación de: .....

**Fdo: El Presidente**

**Fdo: El Secretario**

**Fdo: El Vocal**

**Fdo: El Director del TFG**

(sólo si es distinto del Secretario)

Vº Bº del Subdirector

Trabajo Fin de Grado Nº

(a asignar por Secretaría)

## **AGRADECIMIENTOS**

Quisiera expresar mi más sincero agradecimiento a Jesús, tutor de este trabajo, por su dedicación, guía y apoyo a lo largo de todo el desarrollo del proyecto.

También quiero agradecer de manera especial a mi familia por estar siempre a mi lado. A mis padres, por transmitirme el valor del esfuerzo y la constancia, y a mi hermano, por su apoyo incondicional y por acompañarme en este camino.

## **RESUMEN**

Este trabajo explora una técnica de selección de sensores en sistemas de fusión de decisiones inalámbricos, en el cual sensores distribuidos espacialmente transmiten decisiones binarias a un centro de fusión (FC). En escenarios realistas, las probabilidades de detección locales de cada sensor son desconocidas y deben ser estimadas. Para ello, se emplea el algoritmo Expectation-Maximization (EM), que permite estimar estas probabilidades de forma ciega a partir de las decisiones recibidas en distintos periodos de observación. A partir de las estimaciones obtenidas, se propone una métrica de clasificación que combina las probabilidades de detección y de falsa alarma en un único cociente que refleja la fiabilidad de cada sensor. Esta métrica permite descartar selectivamente los sensores de bajo rendimiento que pueden introducir ruido y perjudicar la decisión global.

Mediante simulaciones de Monte Carlo, se analiza el impacto de esta estrategia sobre dos métricas clave: la raíz cuadrada del error cuadrático medio (RMSE) de la estimación y la probabilidad de detección del sistema. Los resultados demuestran que, cuando se utiliza una regla de fusión óptima como la LRT, eliminar sensores no mejora la probabilidad de detección. No obstante, descartar los sensores menos fiables permite mantener una alta precisión en la detección, al tiempo que se reduce la sobrecarga de comunicación y el consumo energético.

### **Palabras clave:**

Redes de Sensores Inalámbricos (WSN), Centro de fusión, Selección de Sensores, Algoritmo de Expectation-Maximization (EM), Radio Cognitiva, Probabilidad de Detección, Falsa Alarma, Probabilidad de error, Simulación de Monte Carlo, Eficiencia Energética.

## **ABSTRACT**

This work explores a sensor selection technique in wireless decision fusion systems, in which spatially distributed sensors transmit binary decisions to a fusion center (FC). In realistic scenarios, the local detection probabilities of each sensor are unknown and must be estimated. To this end, the Expectation-Maximization (EM) algorithm is employed, allowing for blind estimation of these probabilities based on the decisions received over different observation periods. Based on the obtained estimates, a classification metric is proposed that combines the detection and false alarm probabilities into a single ratio reflecting each sensor's reliability. This metric enables the selective exclusion of low-performing sensors that may introduce noise and degrade the overall decision.

Monte Carlo simulations are used to analyze the impact of this strategy on two key metrics: the root mean square error (RMSE) of the estimation and the system's detection probability. The results show that when using an optimal fusion rule, such as the LRT, removing sensors does not improve the detection probability. However, discarding the least reliable sensors allows for maintaining high detection accuracy while reducing communication overhead and energy consumption.

### **Keywords:**

Wireless Sensor Networks (WSN), Fusion Center, Sensor Selection, Expectation-Maximization (EM) Algorithm, Cognitive Radio, Detection Probability, False Alarm, Error Probability, Monte Carlo Simulation, Energy Efficiency.

# ÍNDICE

<b>1. Introducción .....</b>	<b>9</b>
1.1 Contexto y motivación.....	9
1.2 Propósitos del trabajo .....	11
1.3 Estructura del trabajo .....	12
<b>2. Descripción del problema.....</b>	<b>13</b>
<b>3. Regla de fusión .....</b>	<b>15</b>
3.1 Otras reglas de fusión .....	17
<b>4. Criterio de selección de sensores .....</b>	<b>18</b>
<b>5. Simulaciones .....</b>	<b>21</b>
5.1 Configuración de la simulación .....	21
5.2 Implementación Software .....	23
5.3 Análisis del error cuadrático medio de las estimaciones.....	25
5.3.1 Resultados .....	26
5.4 Análisis de la probabilidad de detección y de la probabilidad de error del centro de fusión .....	29
5.4.1 Resultados.....	31
5.5 Comparación con selección aleatoria .....	34
5.6 Comparación con la regla de fusión de voto mayoritario .....	35
5.7 Selección basada en umbral de calidad .....	38
<b>6. Conclusiones .....</b>	<b>40</b>
<b>Anexo .....</b>	<b>42</b>
<b>Referencias .....</b>	<b>44</b>

## ÍNDICE DE FIGURAS

Fig. 1. Sensado cooperativo para radio cognitiva .....	10
Fig. 2. Red de fusión en paralelo .....	13
Fig. 3. Curva ROC de un sensor .....	19
Fig. 4. Representación de sensores en el espacio ROC .....	20
Fig. 5. RMSE en función del número de sensores eliminados (conjunto inicial de 10 sensores) .....	26
Fig. 6. RMSE en función del número de sensores eliminados (conjunto inicial de 15 sensores) .....	27
Fig. 7. RMSE en función del número de sensores eliminados (conjunto inicial de 25 sensores) .....	27
Fig. 8. Probabilidad de detección en función del número de sensores eliminados (conjunto inicial de 10 sensores) .....	31
Fig. 9. Probabilidad de detección en función del número de sensores eliminados (conjunto inicial de 20 sensores) .....	32
Fig. 10. Probabilidad de detección y probabilidad de error en función del número de sensores eliminados (conjunto inicial de 10 sensores) .....	33
Fig. 11. Probabilidad de detección en función del número de sensores eliminados (eliminación aleatoria) .....	34
Fig. 12. Probabilidad de detección en función del número de sensores eliminados aplicando voto mayoritario .....	35
Fig. 13. Comparación de la probabilidad de detección: VM, LRT-EM y LRT ....	36
Fig. 14. Comparación de la probabilidad de error: VM, LRT-EM y LRT .....	37
Fig. 15. Probabilidad de detección en función del umbral establecido .....	38
Fig. 16. Probabilidad de detección en función del umbral establecido manteniendo al menos dos sensores .....	39

# 1. INTRODUCCIÓN

## 1.1 CONTEXTO Y MOTIVACIÓN

En las últimas décadas, el rápido crecimiento en las comunicaciones inalámbricas ha llevado a una demanda cada vez mayor de espectro radioeléctrico. Tradicionalmente, las políticas de asignación fija del espectro han demostrado ser ineficientes, generando espacios de espectro infrutilizados mientras que otras bandas se encuentran saturadas debido al incremento en el número de dispositivos conectados y servicios multimedia. Este problema ha impulsado el desarrollo de tecnologías emergentes capaces de mejorar significativamente la eficiencia espectral, destacando entre ellas la radio cognitiva (CR). [1]

La radio cognitiva es una tecnología inteligente que permite a los sistemas inalámbricos detectar de forma autónoma su entorno electromagnético, identificar bandas de frecuencias no utilizadas (también conocidas como "espacios blancos") y ajustar dinámicamente sus parámetros operativos (frecuencia, potencia, modulación) para aprovechar estas oportunidades espectrales sin interferir con los usuarios primarios (PU), que tienen prioridad o licencia sobre determinadas bandas. Esta capacidad adaptativa y consciente del entorno convierte a la radio cognitiva en una solución prometedora para resolver el problema de la congestión espectral.

Sin embargo, el despliegue efectivo de sistemas basados en radio cognitiva presenta desafíos técnicos importantes. Uno de los principales retos es la detección fiable de los usuarios primarios a través del sensado espectral, tarea esencial para evitar interferencias perjudiciales [2] [3]. El sensado espectral implica medir y analizar periódicamente el espectro en busca de señales primarias débiles que podrían estar afectadas por múltiples fenómenos físicos adversos como el desvanecimiento multipath, el ensombrecimiento (shadowing) y la incertidumbre del receptor. Estos efectos dificultan la detección precisa, exigiendo métodos robustos y eficientes.

Para abordar estos desafíos, se han propuesto diversos métodos de sensado espectral tales como la detección de energía, la detección por filtrado adaptado y la detección cicloestacionaria, entre otros [4]. El sensado cooperativo (cooperative sensing) ha surgido como una estrategia clave para mejorar significativamente la fiabilidad del sensado espectral.

El sensado cooperativo se basa en la colaboración de múltiples nodos cognitivos distribuidos espacialmente, que realizan mediciones individuales del espectro. Estos nodos comparten la información obtenida para alcanzar una decisión conjunta sobre la presencia o ausencia de usuarios primarios. La cooperación permite mejorar limitaciones individuales derivadas del desvanecimiento y la incertidumbre del canal, ofreciendo así una mejora significativa en la fiabilidad y precisión de la detección.

Dentro de este enfoque colaborativo, uno de los métodos destacados es el sensado centralizado [5], en el que una unidad central, denominada centro de fusión (FC), recoge

las informaciones locales de los distintos sensores y toma una decisión final. Este modelo no solo mejora la precisión de detección, sino que también permite mitigar los efectos de canal que afectan a cada nodo de manera individual.

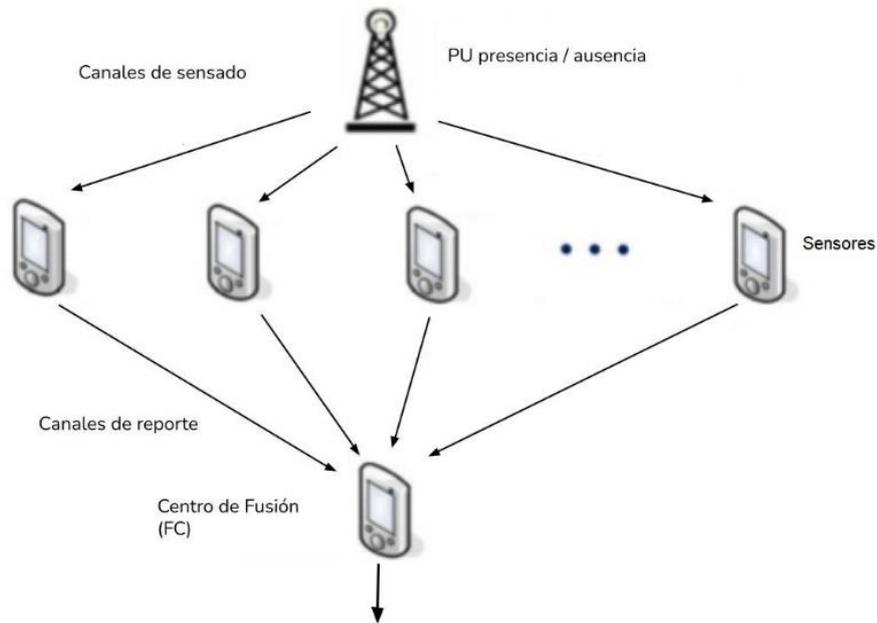


Fig. 1. Sensado cooperativo para radio cognitiva

En cuanto al tipo de información que se comparte con la unidad central, existen dos enfoques principales: *hard combining* y *soft combining*. En el primero, cada nodo envía únicamente una decisión binaria (por ejemplo, “canal libre” o “canal ocupado”), lo cual reduce significativamente el uso de ancho de banda, pero puede conllevar pérdida de información útil [6]. En cambio, el enfoque *soft* implica el envío de métricas más detalladas, como niveles de energía u otras medidas estadísticas, que permiten a la unidad central aplicar técnicas más sofisticadas de fusión, mejorando la sensibilidad del sistema a costa de un mayor consumo de recursos [7].

En resumen, la radio cognitiva representa una solución innovadora y eficaz frente a la creciente demanda de espectro radioeléctrico, al permitir una gestión dinámica y consciente del entorno espectral. No obstante, su implementación conlleva retos significativos, especialmente en lo que respecta a la detección fiable de usuarios primarios. En este contexto, el sensado cooperativo, y en particular el modelo centralizado, se presenta como una estrategia robusta para mitigar los efectos adversos del canal y mejorar la precisión del sistema. Así, el estudio y optimización del sensado espectral cooperativo se convierte en una línea de investigación fundamental para el desarrollo de redes cognitivas más eficientes y confiables.

## 1.2 PROPÓSITOS DEL TRABAJO

En el contexto específico de este trabajo, nos centraremos en el estudio de redes de sensores con un esquema de sensado centralizado, en las que cada nodo transmite al centro de fusión (FC) una decisión binaria sobre la presencia o ausencia de un evento. Para ello, se adoptará un enfoque basado en *hard combining*.

El propósito principal de este trabajo es establecer un criterio de selección de sensores y analizar si esta selección puede mejorar las prestaciones globales de la red. Algunos estudios, como [8], han demostrado que colaborar con todos los nodos disponibles no necesariamente garantiza el rendimiento óptimo.

El criterio de selección planteado se basará en las probabilidades de detección y falsa alarma de los sensores. Las probabilidades de falsa alarma se considerarán conocidas, mientras que las probabilidades de detección se asumen desconocidas y deberán ser estimadas a partir de los datos recogidos. A partir de estas estimas de probabilidad de detección y de la información sobre las falsas alarmas, se evaluará si excluir aquellos sensores con peor comportamiento contribuye a una mejora en las prestaciones globales del sistema. La hipótesis de partida es que, al eliminar sensores menos fiables, se reduce la influencia de observaciones ruidosas y se mejora la fiabilidad del proceso de decisión conjunta.

### 1.3 ESTRUCTURA DEL TRABAJO

El presente Trabajo Fin de Grado se estructura en seis capítulos que desarrollan de forma progresiva los distintos aspectos abordados:

- **Capítulo 1 – Introducción:** Incluye el contexto y la motivación del trabajo, los objetivos principales y la descripción general de la estructura del documento.
- **Capítulo 2 – Descripción del problema:** Se presenta el modelo de red considerado, incluyendo los sensores, los canales de comunicación y las hipótesis sobre sus parámetros. Además, se introduce el algoritmo *Expectation-Maximization* (EM) como herramienta para estimar las probabilidades de detección de los sensores.
- **Capítulo 3 – Regla de fusión:** Se desarrolla la regla de decisión global empleada por el centro de fusión, basada en el riesgo bayesiano y la razón de verosimilitud. Se adapta esta regla al caso en que las probabilidades de detección se estiman mediante el algoritmo EM.
- **Capítulo 4 – Criterio de selección de sensores:** Se define un ratio que combina la probabilidad de detección y de falsa alarma para evaluar la calidad de cada sensor. Este índice permite priorizar sensores más fiables y descartar aquellos que podrían degradar el rendimiento del sistema.
- **Capítulo 5 – Simulaciones:** Se presentan los experimentos realizados en MATLAB para evaluar el impacto del criterio de selección sobre distintas métricas: error cuadrático medio (RMSE), probabilidad de detección y probabilidad de error. Se exploran diferentes configuraciones de red y se comparan los resultados con técnicas alternativas como la selección aleatoria o el uso de reglas de fusión simples como el voto mayoritario.
- **Capítulo 6 – Conclusiones:** Se resumen los hallazgos más relevantes del trabajo, destacando el valor de la selección de sensores como herramienta para optimizar el rendimiento y la eficiencia energética de redes de sensores, especialmente en contextos con recursos limitados.

## 2. DESCRIPCIÓN DEL PROBLEMA

En este trabajo abordamos el problema de la fusión de decisiones en una red de sensores inalámbricos, en la cual cada sensor transmite su decisión binaria,  $y_k$  (donde  $k$  es el índice del sensor), sobre la ausencia ( $z = 0$ ) o presencia ( $z = 1$ ) de un evento al FC. Denotaremos estas hipótesis como  $H_0$  y  $H_1$ .

La decisión de cada sensor se transmitirá a través de canales inalámbricos independientes y no ideales. El FC recibirá las decisiones ruidosas de los sensores a la salida de los canales, las cuales denotaremos como  $x_k$ .

En particular, modelamos los enlaces de comunicación mediante canales binarios simétricos (BSC) [9], los cuales tendrán una probabilidad de error de bit  $e_k$ , definida como:

$$e_k = P(x_k = 1 | y_k = 0) = P(x_k = 0 | y_k = 1) \quad (1)$$

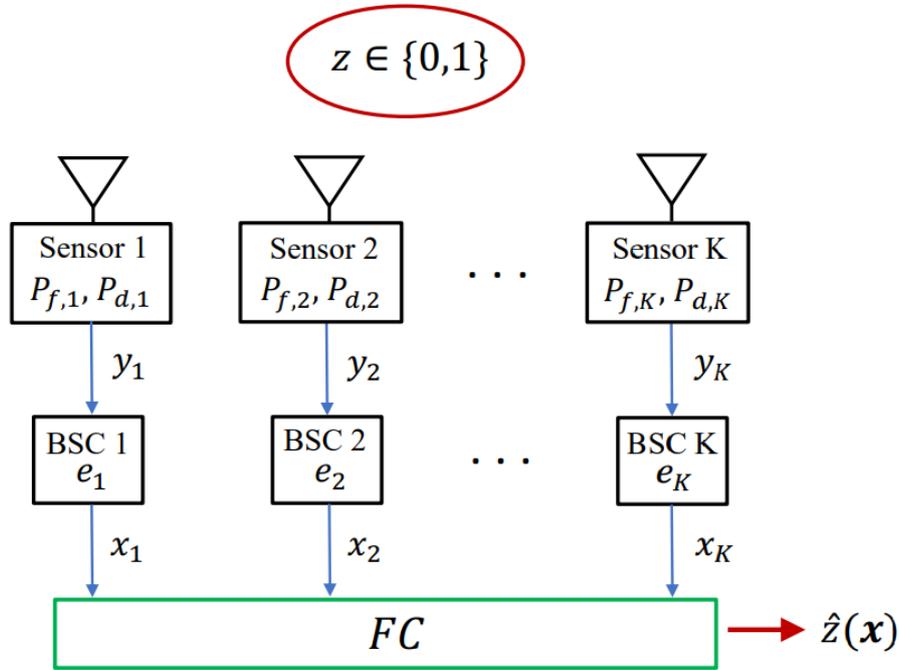


Fig. 2. Red de fusión en paralelo

Nos centraremos en un escenario heterogéneo, en el que los sensores presentan distintas probabilidades de detección  $P_{d,k} = P(y_k = 1 | H_1)$  y probabilidades de falsa alarma  $P_{f,k} = P(y_k = 1 | H_0)$ . Las probabilidades de falsa alarma y de detección en el FC (incluyendo el efecto de los canales binarios simétricos BSCs) son:

$$f_k = P(x_k = 1 | \mathcal{H}_0) = e_k + P_{f,k}(1 - 2e_k) \quad (2)$$

$$d_k = P(x_k = 1 | \mathcal{H}_1) = e_k + P_{d,k}(1 - 2e_k) \quad (3)$$

Suponemos que las probabilidades de error de los BSC están en el intervalo  $0 < e_k < \frac{1}{2}$ . Esta restricción asegura que la probabilidad de recibir correctamente una decisión es mayor que la de recibirla de forma errónea. Si  $e_k = 0.5$ , el canal sería completamente aleatorio. Asimismo, si  $e_k > 0.5$ , sería más eficiente invertir las decisiones recibidas. Por ello, restringimos el valor de  $e_k$  a este intervalo para garantizar que el canal es informativo.

Además, suponemos que  $P_{f,k} \leq P_{d,k}$ . Esta condición refleja un comportamiento coherente y esperado del sistema de detección, ya que suponemos que los sensores no son maliciosos ni defectuosos intencionadamente, por lo que es razonable esperar que tengan mayor capacidad para detectar correctamente un evento que para generar falsas alarmas. Así, las probabilidades efectivas están restringidas de la siguiente manera:

$$0 < e_k \leq f_k \leq d_k \leq 1 - e_k < 1 \quad (4)$$

Además, asumiremos que el centro de fusión desconoce las probabilidades de detección de los sensores, mientras que conoce sus probabilidades de falsa alarma [10],[11], así como las probabilidades de error de los BSC  $\mathbf{e} = (e_1, e_2, \dots, e_K)$  y las probabilidades a priori  $P(H_1) = u$  y  $P(H_0) = 1-u$ .

Esta situación refleja de manera realista muchos escenarios prácticos, donde las condiciones dinámicas del entorno (como variaciones en la intensidad de un fenómeno, cambios en la calidad de los sensores o degradación de sus capacidades) hacen que las probabilidades de detección sean inciertas y difíciles de conocer previamente. Estas probabilidades dependen de las características del evento a detectar, las cuales normalmente son desconocidas.

Por el contrario, las probabilidades de falsa alarma pueden ser calibradas previamente en condiciones controladas, como se justifica en [10], ya que estas dependen únicamente del comportamiento del sensor en ausencia del evento a detectar, es decir, están determinadas por el ruido interno del sensor.

Dado un conjunto de sensores, la regla de fusión óptima es el test del cociente de verosimilitudes (Likelihood Ratio Test: LRT) [12], [14], [15]. No obstante, la implementación del LRT requiere que el centro de fusión conozca las probabilidades de detección de los sensores, información que en nuestro caso no está disponible.

Para superar esta limitación, se propone el uso del algoritmo Expectation-Maximization (EM) [16], [17], el cual permite estimar de manera ciega las probabilidades de detección de los sensores a partir de las decisiones binarias recibidas durante un número finito de periodos de observación, tal como se desarrolla en [13].

Una vez estimadas las probabilidades de detección el FC toma una decisión final, a partir del vector de decisiones binarias  $\mathbf{x} = (x_1, x_2, \dots, x_K)$ , de acuerdo con una cierta regla de fusión  $\hat{z}(\mathbf{x}) \in \{0,1\}$ .

### 3. REGLA DE FUSIÓN

Como se ha explicado anteriormente, en un sistema de fusión de decisiones, el FC tiene como objetivo tomar una decisión global sobre la presencia o ausencia de un evento, a partir del vector de observaciones  $\mathbf{x}$ . Por tanto, es necesario establecer una regla de fusión que permita tomar dicha decisión final.

$$\Lambda(\mathbf{x} | \mathbf{d}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \gamma \quad (5)$$

Donde  $\Lambda(\mathbf{x} | \mathbf{d})$  denota el estadístico de decisión, que en general depende de las probabilidades de detección  $\mathbf{d} = (d_1, d_2, \dots, d_K)$  y  $\gamma$  es el umbral de decisión.

En este trabajo se ha optado por utilizar la regla de fusión que minimiza la probabilidad de error, lo que resulta adecuado cuando se desea penalizar los errores de detección y de falsa alarma de forma equitativa.

El punto de partida es el riesgo bayesiano, que representa el valor esperado del coste asociado a tomar una decisión  $\hat{z}(\mathbf{x})$  bajo incertidumbre [12], [14], [15]. Se define como:

$$C = \sum_{i=0}^1 \sum_{j=0}^1 C_{ij} P(H_j) P(\hat{z} = i | H_j) \quad (6)$$

Donde  $C_{ij}$  es el coste de decidir  $\hat{z} = i$  cuando la hipótesis verdadera es  $H_j$  y  $P(H_j)$  es la probabilidad a priori de la hipótesis  $H_j$  siendo  $P(H_1) = u$  y  $P(H_0) = 1 - u$ .

Cuando no se penalizan los aciertos ( $C_{11} = C_{00} = 0$ ) y los fallos se penalizan por igual ( $C_{10} = C_{01} = 1$ ), el riesgo bayesiano se reduce a la probabilidad de error.

$$P_e = (1 - u) \cdot P(\hat{z} = 1 | \mathcal{H}_0) + u \cdot P(\hat{z} = 0 | \mathcal{H}_1) \quad (7)$$

El objetivo es minimizar esta expresión para cada posible vector de observaciones  $\mathbf{x}$ , lo que nos lleva a la siguiente regla de fusión óptima basada en el cociente de verosimilitud [14], [15]:

$$\frac{p(\mathbf{x} | H_1, \mathbf{d})}{p(\mathbf{x} | H_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \frac{1 - u}{u} \quad (8)$$

donde  $p(\mathbf{x} | H_1, \mathbf{d})$  y  $p(\mathbf{x} | H_0)$ , son las funciones de probabilidad condicionada que representan la probabilidad de recibir el vector de observaciones  $\mathbf{x}$  bajo las hipótesis  $H_1$  (presencia de evento) y  $H_0$  (ausencia de evento), respectivamente.

$$p(\mathbf{x} | H_0) = \prod_{k=1}^K f_k^{x_k} (1 - f_k)^{1-x_k} \quad (9)$$

$$p(\mathbf{x} | H_1, \mathbf{d}) = \prod_{k=1}^K d_k^{x_k} (1 - d_k)^{1-x_k} \quad (10)$$

En la regla de fusión (8), se muestra explícitamente la dependencia con  $\mathbf{d}$ , ya que estas probabilidades de detección son desconocidas. Por otro lado, no se muestra la dependencia de  $p(\mathbf{x} | H_0)$  con  $\mathbf{f}$ , ya que las probabilidades de falsa alarma se consideran conocidas.

Tomando logaritmos en ambos lados para facilitar los cálculos, nos queda el siguiente estadístico y umbral de decisión:

$$\Delta(\mathbf{x} | \mathbf{d}) = \log \left( \frac{p(\mathbf{x} | H_1, \mathbf{d})}{p(\mathbf{x} | H_0)} \right) \quad (11)$$

$$\gamma = \log \left( \frac{1 - u}{u} \right) \quad (12)$$

Como podemos observar el estadístico de decisión (9) se basa en el cociente de verosimilitudes (LR: Likelihood Ratio) de las observaciones  $\mathbf{x}$ .

Por tanto,

$$\hat{z}(\mathbf{x}) = \begin{cases} 1, & \text{si } \Lambda(\mathbf{x} | \mathbf{d}) > \gamma \\ 0, & \text{si } \Lambda(\mathbf{x} | \mathbf{d}) \leq \gamma \end{cases} \quad (13)$$

No obstante, en el caso práctico considerado, las probabilidades de detección  $\mathbf{d}$  no son conocidas a priori, por lo que trabajaremos con las estimas,  $\hat{\mathbf{d}}$ , que obtenemos al aplicar el algoritmo EM. La regla de fusión empleada será:

$$\hat{z}(\mathbf{x}) = \begin{cases} 1, & \text{si } \Lambda(\mathbf{x} | \hat{\mathbf{d}}) > \gamma \\ 0, & \text{si } \Lambda(\mathbf{x} | \hat{\mathbf{d}}) \leq \gamma \end{cases} \quad (14)$$

### 3.1 OTRAS REGLAS DE FUSIÓN

Existen otras reglas de fusión denominadas *ciegas*, en las cuales no es necesario conocer ningún parámetro de los sensores. Un ejemplo de este tipo son las reglas de conteo, que se basan únicamente en las decisiones locales recibidas de los sensores.

En este enfoque, la decisión final se toma a partir del número total de decisiones positivas recibidas, mediante la siguiente expresión:

$$\Delta(\mathbf{x}) = \mathbf{x}^T \cdot \mathbf{1}_K = \sum_{k=1}^K x_k \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \gamma \quad (15)$$

Dentro de las reglas de conteo, una de las más utilizadas es la regla de voto mayoritario (*majority voting*). La decisión final se toma por mayoría simple, en la cual el umbral de decisión  $\gamma$  se fija en:

$$\frac{K}{2}$$

donde  $K$  es el número total de sensores.

Esto significa que se decide por la hipótesis  $H_1$  si la mitad o más de los sensores reportan una detección positiva ( $x_k = 1$ ). Formalmente, se expresa como:

$$\Delta(\mathbf{x}) = \sum_{k=1}^K x_k \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \frac{K}{2} \quad (16)$$

#### 4. CRITERIO DE SELECCIÓN DE SENSORES

El criterio utilizado para valorar el desempeño de los sensores y clasificarlos según su calidad de detección se basa en el estadístico utilizado, el cual se basa a su vez en el cociente de verosimilitudes de las observaciones  $\mathbf{x}$ .

Sustituyendo en (11), las expresiones (9) y (10), obtenemos:

$$\begin{aligned}\Lambda(\mathbf{x} | \mathbf{d}) &= \sum_{k=1}^K \log \left( \frac{d_k^{x_k} (1 - d_k)^{1-x_k}}{f_k^{x_k} (1 - f_k)^{1-x_k}} \right) \\ &= \sum_{k=1}^K x_k \log \left( \frac{d_k (1 - f_k)}{f_k (1 - d_k)} \right) + \sum_{k=1}^K \log \left( \frac{1 - d_k}{1 - f_k} \right)\end{aligned}\quad (17)$$

Por tanto, podemos observar que el estadístico de decisión es una combinación lineal de las decisiones de los sensores más un término que no depende de las observaciones. La decisión del sensor  $k$  está ponderada por el factor:

$$R_k = \log \left( \frac{d_k (1 - f_k)}{f_k (1 - d_k)} \right) \quad (18)$$

Este factor combina en un solo número las dos cualidades esenciales de cualquier detector: su capacidad para advertir la presencia del fenómeno que buscamos y su propensión a lanzar falsas alarmas.

En el numerador del cociente aparece el término  $d_k(1 - f_k)$ , que puede interpretarse como el producto de dos probabilidades: por un lado, la probabilidad de obtener un verdadero positivo, es decir, detectar correctamente la presencia del fenómeno  $d_k = TP = P(x_k = 1 | H_1)$ , y, por otro lado, la probabilidad de obtener un verdadero negativo  $(1 - f_k) = TN = P(x_k = 0 | H_0)$ , lo que equivale a que no se produzca una falsa alarma.

En el denominador aparece el término  $f_k(1 - d_k)$ , que representa también el producto de dos probabilidades, pero en este caso asociadas a errores: la probabilidad de un falso positivo  $f_k = FP = P(x_k = 1 | H_0)$ , correspondiente a una falsa alarma, y la probabilidad de un falso negativo  $(1 - d_k) = FN = P(x_k = 0 | H_1)$ , correspondiente a no detectar un evento real.

El ratio propuesto compara directamente el producto de las probabilidades de acertar (verdaderos positivos y verdaderos negativos) con el producto de las probabilidades de fallar (falsos positivos y falsos negativos).

Este ratio ofrece varias ventajas. Para empezar, equilibra los dos errores posibles: un sensor que alcance una gran sensibilidad a costa de disparar el número de falsas alarmas no tendrá un peso elevado, porque el denominador crecerá en la misma medida. De manera similar, un sensor que reduzca las falsas alarmas a costa de bajar su probabilidad de detección tampoco saldrá beneficiado, porque el numerador se verá mermado.

El valor de  $R_k$  se interpreta de forma muy directa. Cuando es mayor que cero indica que los aciertos predominan sobre los errores y, cuanto mayor sea, mayor será la distancia entre el comportamiento del sensor y el simple azar. El caso límite,  $R_k = 0$ , equivale a lanzar una moneda al aire: el sensor no aporta información útil. Por debajo de cero, el sensor resulta claramente contraproducente. Si  $R_k < 0$  es probable que el sensor  $k$  sea malicioso.

Una forma de ilustrar cómo un sensor varía la probabilidad de detección a costa de la probabilidad de falsa alarma es mediante la curva ROC (Receiver Operating Characteristic) [14] a medida que se modifica el umbral de decisión. La curva ROC representa gráficamente el compromiso inherente entre estas dos probabilidades clave en el desempeño de un sensor.

La curva ROC de un sensor se construye variando el umbral de decisión y calculando, en cada caso, la probabilidad de detección y la probabilidad de falsa alarma. A umbrales bajos, la probabilidad de detección aumenta a costa de incrementar la tasa de falsas alarmas, mientras que umbrales más altos reducen tanto la detección como las falsas alarmas.

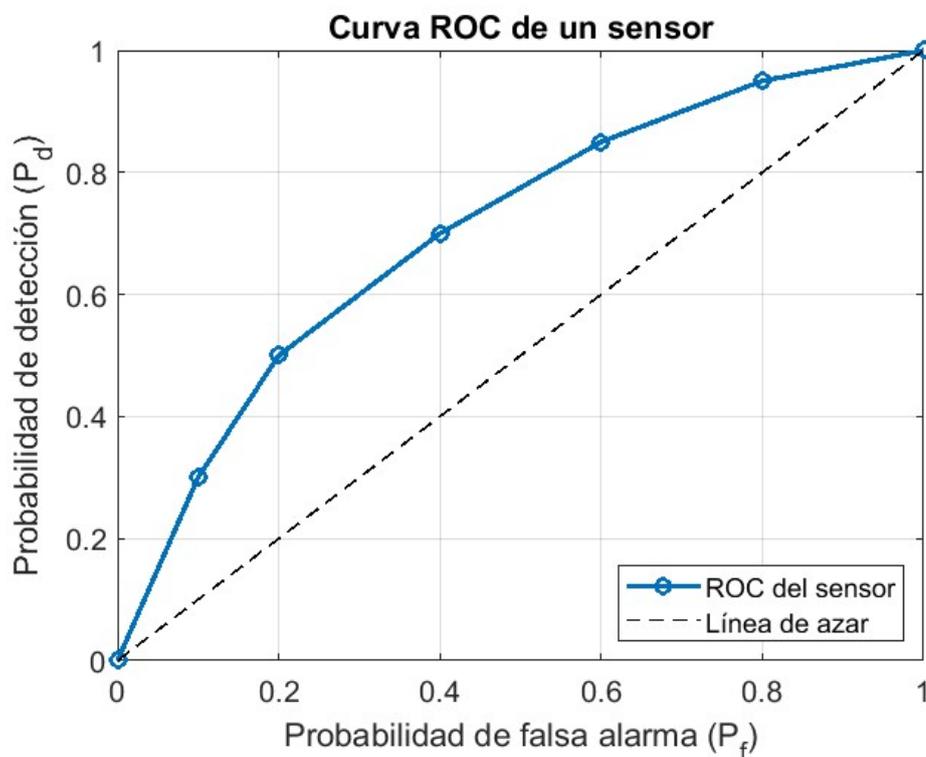


Fig. 3. Cruva ROC de un sensor

La línea oblicua que une los puntos (0,0) y (1,1) en la figura 3 representa un sensor totalmente aleatorio. Los sensores que se encuentren por debajo de esta línea podrían considerarse como maliciosos o perjudiciales para el sistema.

Además, el espacio ROC permite representar y comparar distintos sensores simultáneamente. Cada sensor se puede visualizar como un punto en este espacio, lo que facilita la comparación directa de su desempeño. Nuestro objetivo en este trabajo será seleccionar sensores que se sitúen lo más cerca posible de la esquina superior izquierda, ya que esta posición indica una alta probabilidad de detección y una baja tasa de falsas alarmas.

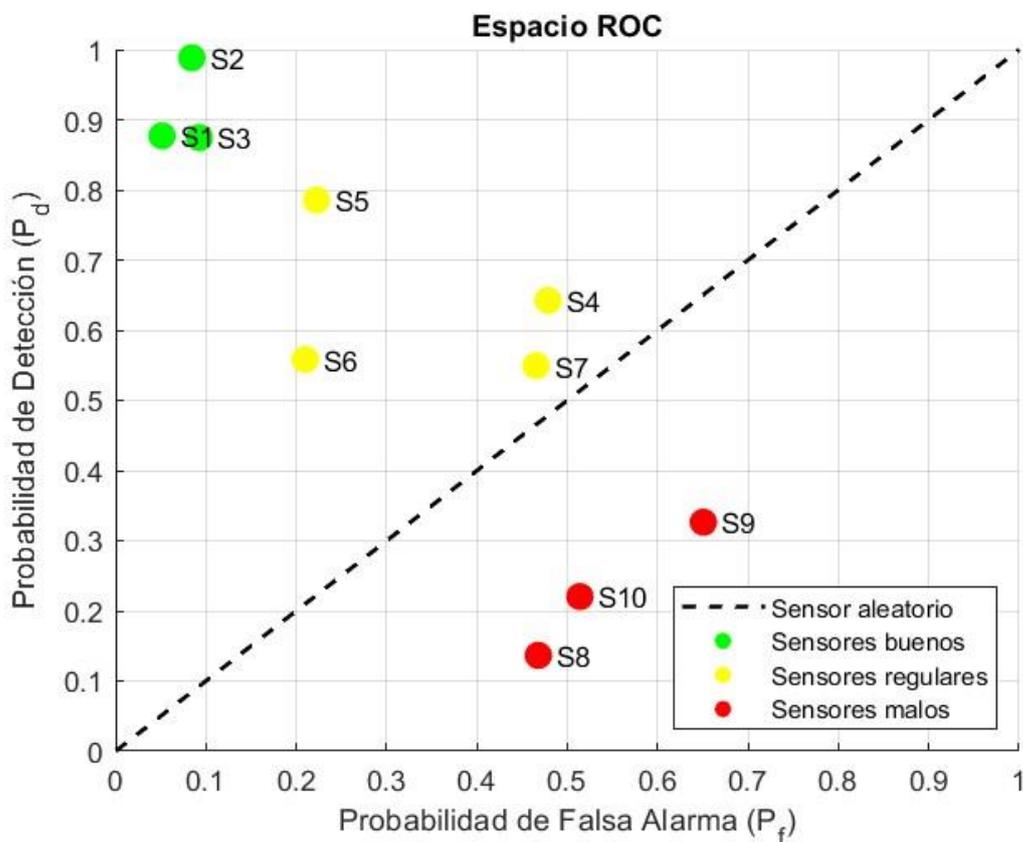


Fig. 4. Representación de sensores en el espacio ROC

A lo largo de este trabajo se seguirán dos criterios a la hora de eliminar sensores. Primero, se eliminarán sistemáticamente los  $K$  peores sensores en función de su valor de  $R_k$ , y posteriormente realizaremos simulaciones en las cuales nos quedaremos únicamente con aquellos sensores cuyo valor de  $R_k$  supere un umbral determinado.

## 5. SIMULACIONES

Con el objetivo de evaluar el impacto del criterio de selección de sensores propuesto, en esta sección se presentan una serie de simulaciones realizadas en MATLAB. Estas simulaciones permiten analizar el rendimiento del sistema en términos de estimación de parámetros y capacidad de detección, considerando distintos escenarios y configuraciones de red.

### 5.1 CONFIGURACIÓN DE LA SIMULACIÓN

A la hora de simular siguiendo [10], [11], [18], asumimos que los sensores emplean una modulación ASK con símbolos  $Y_{k,n} \in \{0,1\}$  para transmitir sus decisiones al FC a través de los canales de reporte, los cuales experimentan desvanecimiento plano lento.

Entonces, en el instante  $n$ , la señal en banda base en tiempo discreto recibida en el FC proveniente del sensor  $k$  está dada por:

$$r_{k,n} = h_k y_{k,n} + w_{k,n} \quad (19)$$

donde  $h_k$  representa la respuesta del canal y  $w_{k,n}$  es el ruido en el FC el cual consideramos que es independiente, idénticamente distribuido, gaussiano complejo y circularmente simétrico con varianza  $\sigma_w^2$ . Asumimos que el FC conoce las respuestas de los canales,  $h_k$ , por lo que emplea detección coherente para obtener  $x_{k,n}$  a partir de  $r_{k,n}$  usando el siguiente umbral de decisión óptimo  $\frac{|h_k|}{2} \cdot \frac{1-u}{u}$ . Este umbral minimiza la probabilidad de error para el sensor  $k$  si se conoce la probabilidad a priori  $u$ . Entonces, la tasa de error de bit es:

$$e_k = Q\left(\frac{|h_k|}{2\sigma_w} \cdot \frac{1-u}{u}\right) \quad (20)$$

Se ha elegido este modelo de canal BSC en las simulaciones debido a su amplia utilización en la literatura técnica [10], [11].

Analizamos el rendimiento de estimación del algoritmo EM y la probabilidad de detección del FC a medida que se eliminan progresivamente los sensores con peor desempeño, de acuerdo con el criterio de selección establecido, mediante simulaciones de Monte Carlo. Cada experimento promedia el rendimiento de  $R = 5 \cdot 10^4$  ejecuciones independientes. En cada experimento, fijamos la probabilidad a priori  $u$ , el número de sensores  $K$  y el número de vectores de decisión  $N$ . También elegimos  $\sigma_w^2$  de forma que la probabilidad de error promedio de los canales de reporte tome un valor dado  $\bar{e} = E[e_k] = E\left[Q\left(\frac{|h_k|}{2\sigma_w} \cdot \frac{1-u}{u}\right)\right]$ . Estos parámetros no cambian en las  $R$  ejecuciones.

En cada ejecución, las probabilidades de los sensores se extraen de forma independiente de acuerdo con las siguientes distribuciones uniformes:

$$P_{f,k} \sim U(0,0.1)$$

$$P_{d,k} - P_{f,k} \sim U(0,0.9)$$

También generamos de forma independiente las respuestas de canal  $h_k$ , asumimos que los canales de reporte experimentan desvanecimiento Rayleigh independiente  $h_k \sim \mathcal{N}_c(0,1)$ , y las correspondientes tasas de error  $e_k$ . Finalmente, se genera una secuencia de estados  $z$  a partir de  $u$ , y una matriz de observaciones  $\mathbf{X}$  de dimensiones  $K \times N$  se genera a partir de  $z$ ,  $\mathbf{f}$  y  $\mathbf{d}$ .

## 5.2 IMPLEMENTACIÓN SOFTWARE

Para llevar a cabo las simulaciones se han desarrollado dos funciones principales en MATLAB: `inicializacion_parametros` y `EM_algorithm`. Estas funciones constituyen la base del entorno de simulación y serán utilizadas de forma recurrente en los distintos experimentos del trabajo. A continuación, se describen sus funcionalidades, así como sus entradas y salidas.

### **Función `inicializacion_parametros`:**

Esta función se encarga de generar de forma automática los parámetros necesarios para simular un sistema de detección distribuida.

Entradas:

- K: número total de sensores.
- N: número de instantes de observación.
- u: probabilidad a priori de que se cumpla la hipótesis H1, es decir, la presencia del evento observado.

Salidas:

- f: vector de probabilidades de falsa alarma de cada sensor ( $K \times 1$ ).
- d\_reales: vector con las probabilidades reales de detección de cada sensor ( $K \times 1$ ).
- e: vector con las probabilidades de error asociadas al canal de transmisión para cada sensor ( $K \times 1$ ).
- X: matriz binaria de observaciones recogidas por los sensores ( $K \times N$ ), en función del estado real y los parámetros anteriores.

La matriz de observaciones X se genera utilizando los estados reales del fenómeno (z) y los parámetros f y d\_reales. De este modo, se obtiene una representación simulada de las decisiones binarias emitidas por los sensores a lo largo del tiempo.

### **Función `EM_algorithm`:**

Esta función implementa el algoritmo de Expectation-Maximization (EM), empleado para estimar las probabilidades de detección de los sensores a partir de las observaciones recogidas y los parámetros conocidos del sistema.

Entradas:

- X: matriz binaria de observaciones ( $K \times N$ ).
- f: vector de probabilidades de falsa alarma ( $K \times 1$ ).

- $e$ : vector de errores del canal ( $K \times 1$ ).
- $u$ : probabilidad a priori de la hipótesis  $H_1$ .

Salida:

- $d$ : vector estimado de probabilidades de detección ( $K \times 1$ ).

La estimación se realiza de forma iterativa mediante los pasos clásicos del algoritmo EM:

1. **E-step**: se calcula, para cada observación, la probabilidad de que el evento observado corresponda a la hipótesis  $H_1$ , en función de las observaciones y de los parámetros actuales.
2. **M-step**: se actualizan las estimaciones de las probabilidades de detección de cada sensor en función de los valores esperados calculados en la etapa anterior, respetando las restricciones impuestas por el modelo.

El proceso se repite hasta que se alcanza la convergencia.

Este procedimiento permite obtener una estimación robusta de las verdaderas capacidades de detección de cada sensor, incluso en presencia de errores de canal y decisiones ruidosas.

### 5.3 ANÁLISIS DEL ERROR CUADRÁTICO MEDIO DE LAS ESTIMACIONES

Para evaluar el rendimiento del algoritmo EM en la estimación de  $\mathbf{d}$ , utilizamos la raíz cuadrada del error cuadrático medio (RMSE: Root Mean Squared Error). Este indicador nos permite cuantificar cómo varía la precisión de las estimaciones a medida que vamos eliminando los sensores que presentan un peor desempeño. La expresión del RMSE es:

$$\text{RMSE} = \sqrt{\frac{1}{KR} \sum_{r=1}^R \|\widehat{\mathbf{d}}^{(r)} - \mathbf{d}^{(r)}\|^2} \quad (21)$$

donde  $r$  es el índice de las ejecuciones,  $\mathbf{d}^{(r)}$  representa las probabilidades de detección reales de los sensores y  $\widehat{\mathbf{d}}^{(r)}$  son las estimaciones obtenidas mediante el algoritmo EM en la simulación  $r$ .

Para llevar a cabo este análisis, se ha desarrollado un programa en MATLAB que simula  $R$  ejecuciones independientes del sistema. En cada iteración, los sensores se ordenan de menor a mayor calidad según el ratio propuesto, y se eliminan progresivamente aquellos con peor desempeño. Este procedimiento se repite aumentando el número de sensores descartados en cada paso, con el objetivo de analizar cómo varía el valor del RMSE en función del subconjunto de sensores seleccionados. De esta manera, se evalúa empíricamente el impacto que tienen los sensores de baja calidad en la estimación global, así como la eficacia del criterio utilizado para identificarlos y excluirlos.

### 5.3.1 Resultados

En primer lugar, hemos simulado cómo evoluciona el RMSE a medida que eliminamos progresivamente los sensores con peor rendimiento. Comenzamos eliminando un único sensor y vamos incrementando el número de sensores descartados en cada iteración hasta quedarnos con únicamente un sensor. Tras cada eliminación, se vuelve a aplicar el algoritmo EM para obtener nuevas estimas de  $d$ .

Los parámetros utilizados para esta simulación son los siguientes: partimos de un total de  $K=10$  sensores, el error medio de los canales BSC es de  $\bar{\epsilon} = 0.05$ , la probabilidad *a priori* es  $u = 0.5$  y el número de observaciones disponibles para estimar el parámetro  $d$  es de  $N=40$ .

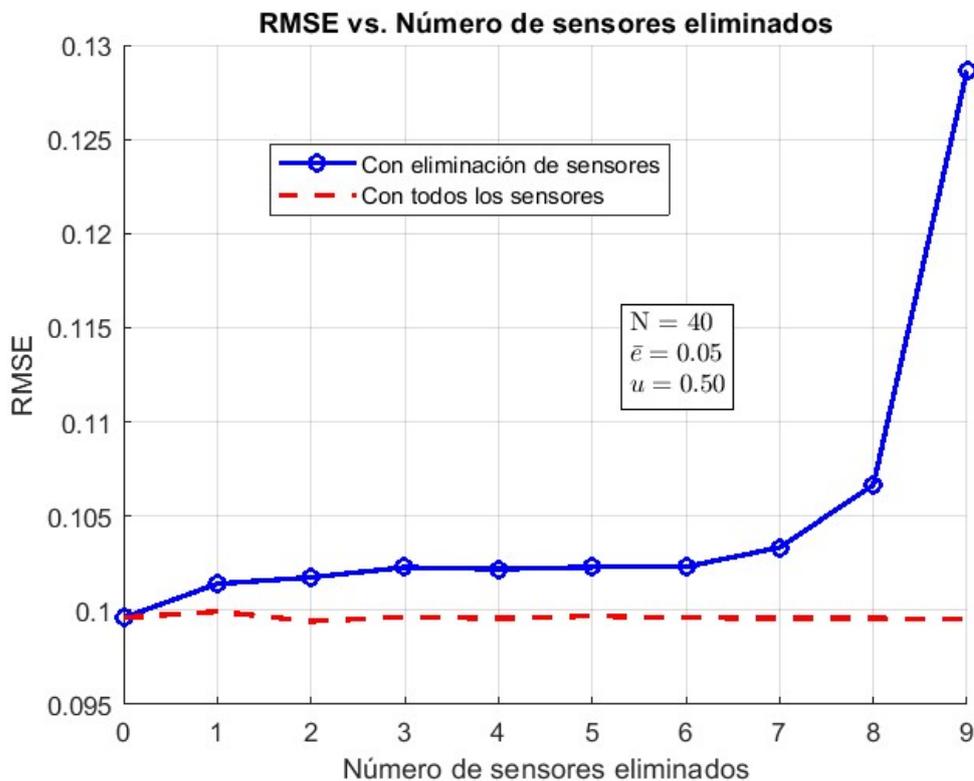


Fig. 5. RMSE en función del número de sensores eliminados (conjunto inicial de  $K=10$  sensores)

Al simular con estos 10 sensores, observamos que la estimación no mejora pese a eliminar sensores con bajo rendimiento, que podrían estar introduciendo ruido en el proceso de estimación. Esto sugiere que, en configuraciones con un número limitado de sensores, la eliminación de algunos de ellos no siempre es beneficiosa, ya que se reduce la cantidad total de observaciones disponibles.

A continuación, se ha repetido la simulación incrementando el número inicial de sensores a 15. Los resultados obtenidos son los siguientes:

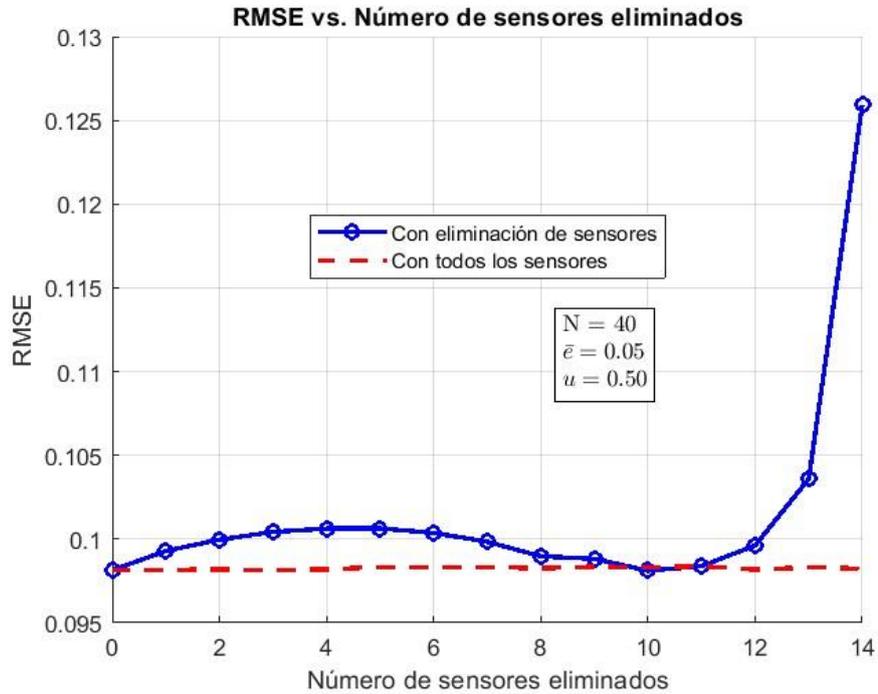


Fig. 6. RMSE en función del número de sensores eliminados (conjunto inicial de K=15 sensores)

En la figura 6 se puede observar que se consigue mejorar ligeramente la estimación al elegir los cinco mejores sensores. Por ello, se va a volver a simular con un número mayor de sensores para comprobar si se puede conseguir un RMSE aún menor.

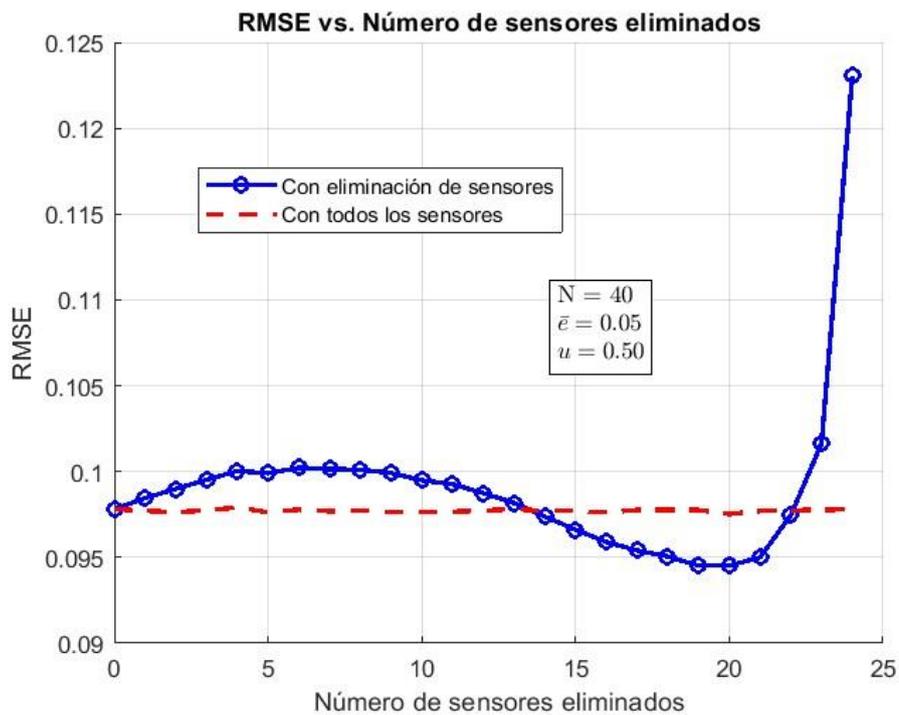


Fig. 7. RMSE en función del número de sensores eliminados (conjunto inicial de K = 25 sensores)

En este caso, observamos que sí se consigue mejorar la estimación. Este comportamiento puede explicarse por el hecho de que, al disponer de un conjunto inicial mayor, tenemos la posibilidad de seleccionar sensores con características significativamente mejores. Sus observaciones son más fiables, lo que permite obtener estimaciones más precisas que al usar todos los sensores indiscriminadamente.

No obstante, también se aprecia que cuando el número de sensores seleccionados es demasiado reducido, las estimaciones empeoran considerablemente. Esto se debe a que, aunque los sensores escogidos sean de alta calidad, la cantidad total de observaciones disminuye y no compensa la ganancia en fiabilidad individual. Por tanto, existe un equilibrio entre la calidad y la cantidad de sensores seleccionados que permite optimizar la estimación.

## 5.4 ANALISIS DE LA PROBABILIDAD DE DETECCIÓN Y DE LA PROBABILIDAD DE ERROR DEL CENTRO DE FUSIÓN

En esta sección se analiza el impacto del criterio de selección de sensores sobre el rendimiento global del sistema, evaluando dos métricas clave: la probabilidad de detección y la probabilidad de error del centro de fusión. Estas métricas permiten valorar la eficacia de la detección conjunta y la robustez del sistema ante la eliminación de sensores poco fiables.

### Probabilidad de detección del FC

Para calcular la probabilidad de detección del centro de fusión, se considera  $\mathbf{x}$  como un vector aleatorio. Su espacio muestral es el conjunto de  $K$ -tuplas binarias, denotado como  $\Omega_x$ , con una cardinalidad de  $|\Omega_x| = 2^K$ .

El FC toma la decisión global para cada  $\mathbf{x}$ , de acuerdo con cierta regla de fusión  $\hat{z}(\mathbf{x}) \in \{0,1\}$ . Su probabilidad de detección se expresa como:

$$P_d = P(\hat{z}(\mathbf{x}) = 1 | \mathcal{H}_1, \mathbf{d}) = \sum_{\mathbf{x} \in \Omega_x} \hat{z}(\mathbf{x}) p(\mathbf{x} | \mathcal{H}_1, \mathbf{d}) \quad (22)$$

No obstante, dado que en las simulaciones se emplea un número elevado de sensores, resulta computacionalmente muy costoso calcular esta expresión de forma exacta, ya que el conjunto de posibles tuplas crece exponencialmente con  $K$ .

Por tanto, se ha optado por emplear simulaciones de Monte Carlo para estimar la probabilidad de detección del FC. Para ello, se genera una matriz con mil vectores de observaciones  $\mathbf{x}$  bajo la hipótesis  $H_1$ . A continuación, se aplica la regla de fusión para cada uno de estos vectores. Finalmente, se cuenta el número de veces que el centro de fusión decide  $\hat{z}(\mathbf{x}) = 1$ . Así, la probabilidad de detección se estima como:

$$P_d = \frac{R_1}{R} \quad (23)$$

donde  $R$  es el número total de simulaciones realizadas y  $R_1$  es el número de veces que se ha decidido  $\hat{z}(\mathbf{x}) = 1$ .

### Probabilidad de error del FC

La probabilidad de error ( $P_e$ ) se define como la probabilidad de que la decisión global  $\hat{z}(\mathbf{x})$  difiera de la verdadera hipótesis  $z$ , y se expresa como:

$$P_e = P(\hat{z}(\mathbf{x}) \neq z) = P_f(1 - u) + (1 - P_d)u = \sum_{\mathbf{x} \in \Omega_x} \hat{z}(\mathbf{x}) p(\mathbf{x}|\mathbf{d}) \quad (24)$$

Al igual que ocurre con la probabilidad de detección, el cálculo exacto de  $P_e$  se vuelve computacionalmente inviable cuando se dispone de un número elevado de sensores, debido a la complejidad combinatoria de las observaciones posibles. Por ello, se opta por una estimación mediante simulaciones Monte Carlo.

Para ello, se ha desarrollado una función en MATLAB que genera numerosas observaciones tanto bajo la hipótesis  $H_0$  como bajo  $H_1$ , y se contabiliza cuántas veces el centro de fusión comete un error en la decisión global. La estimación de la  $P_e$  mediante simulación se obtiene como:

$$P_e = \frac{\text{errores}_{H_0} \cdot (1-u) + \text{errores}_{H_1} \cdot u}{\text{total de observaciones}} \quad (25)$$

donde  $\text{errores}_{H_0}$  denota el número de falsa alarmas y  $\text{errores}_{H_1}$  el número de no detecciones.

### 5.4.1 Resultados

Para las simulaciones emplearemos los mismos parámetros utilizados previamente en las simulaciones del RMSE. Primero, analizaremos cómo evoluciona la probabilidad de detección a medida que se eliminan sensores. Cada vez que se descartan sensores, se vuelve a aplicar el algoritmo EM para obtener nuevas estimaciones de  $\mathbf{d}$ , y posteriormente se aplica de nuevo la regla de fusión.

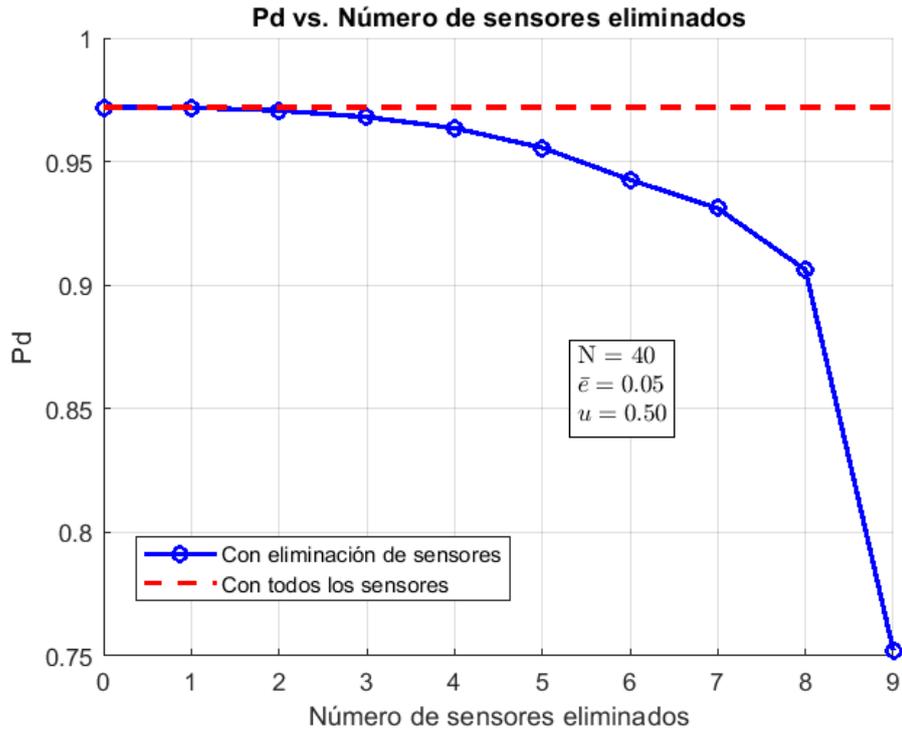


Fig. 8. Probabilidad de detección en función del número de sensores eliminados (conjunto inicial de  $K=10$  sensores)

A la vista de los resultados (Figura 8), observamos que la probabilidad de detección no mejora a pesar de eliminar sensores. Aunque se podrían estar eliminando sensores que introducen ruido en la estima de  $\mathbf{d}$ , no se aprecia una mejora en el rendimiento del detector global.

A continuación, se repiten las simulaciones aumentando el número de sensores para comprobar si, al disponer de un conjunto más amplio, es posible realizar una mejor selección y, de esta forma, mejorar las prestaciones, tal como ocurrió en el caso del RMSE.

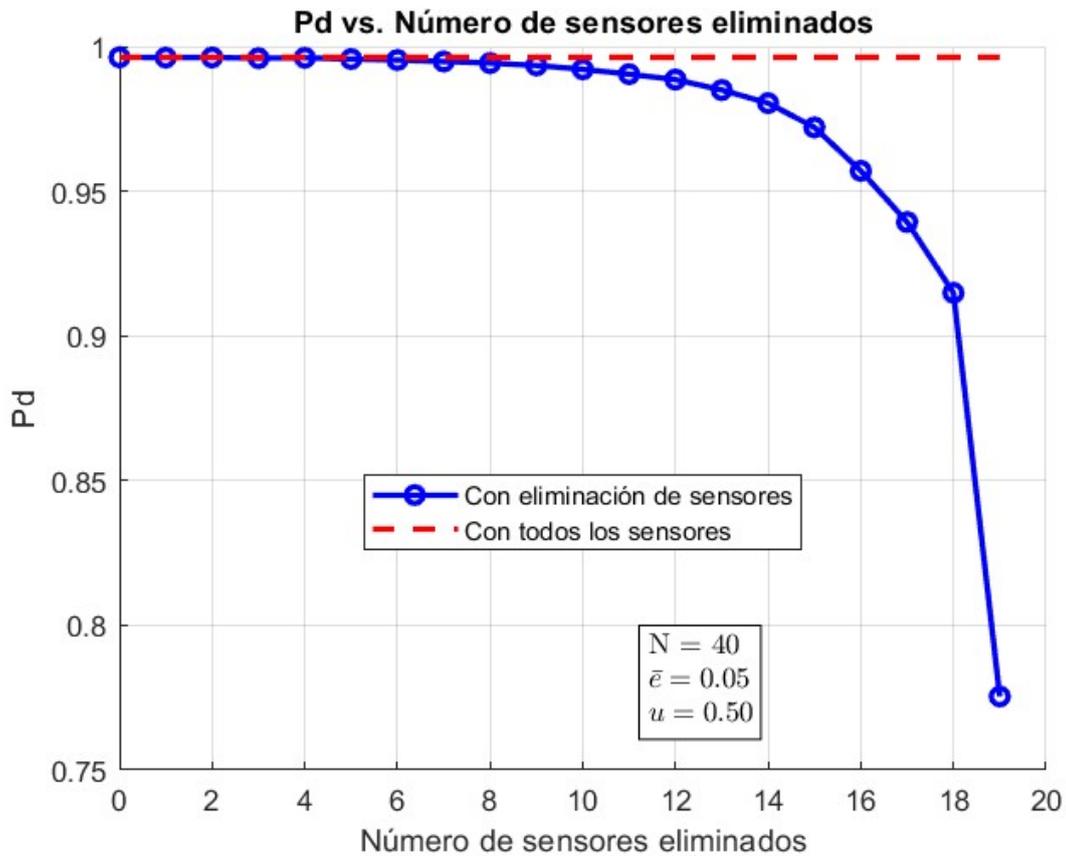


Fig. 9. Probabilidad de detección en función del número de sensores eliminados (conjunto inicial de  $K=20$  sensores)

En la figura 9 vemos que aun aumentado el conjunto inicial de sensores no se consigue mejorar la probabilidad de detección en comparación con el caso en que se utilizan todos los sensores. Este comportamiento puede explicarse por el funcionamiento de la propia regla de fusión empleada. Tal como se discutió anteriormente en la expresión (17), dicha regla pondera las decisiones individuales de los sensores en función del factor elegido para clasificar los sensores el cual refleja su fiabilidad. Como consecuencia, los sensores con peor desempeño tienen un peso reducido en la decisión final del centro de fusión, lo que atenúa su influencia en el proceso global de detección. Por este motivo, eliminarlos no supone una mejora en la probabilidad de detección.

No obstante, se puede observar que, aunque la probabilidad de detección no mejora, tampoco empeora de forma significativa al eliminar un pequeño número de sensores. Esto indica que es posible prescindir de los sensores con peor desempeño sin afectar de manera notable el rendimiento del sistema.

Además, al disponer de un conjunto inicial más amplio de sensores, se constata que la mayor variedad de sensores permite una mejor selección de los que presentan peores prestaciones según el ratio de calidad. Así, se pueden eliminar un mayor número de ellos sin que la probabilidad de detección se vea comprometida de forma significativa.

Por otro lado, se han analizado los resultados de la probabilidad de error, obteniéndose un comportamiento similar al de la probabilidad de detección.

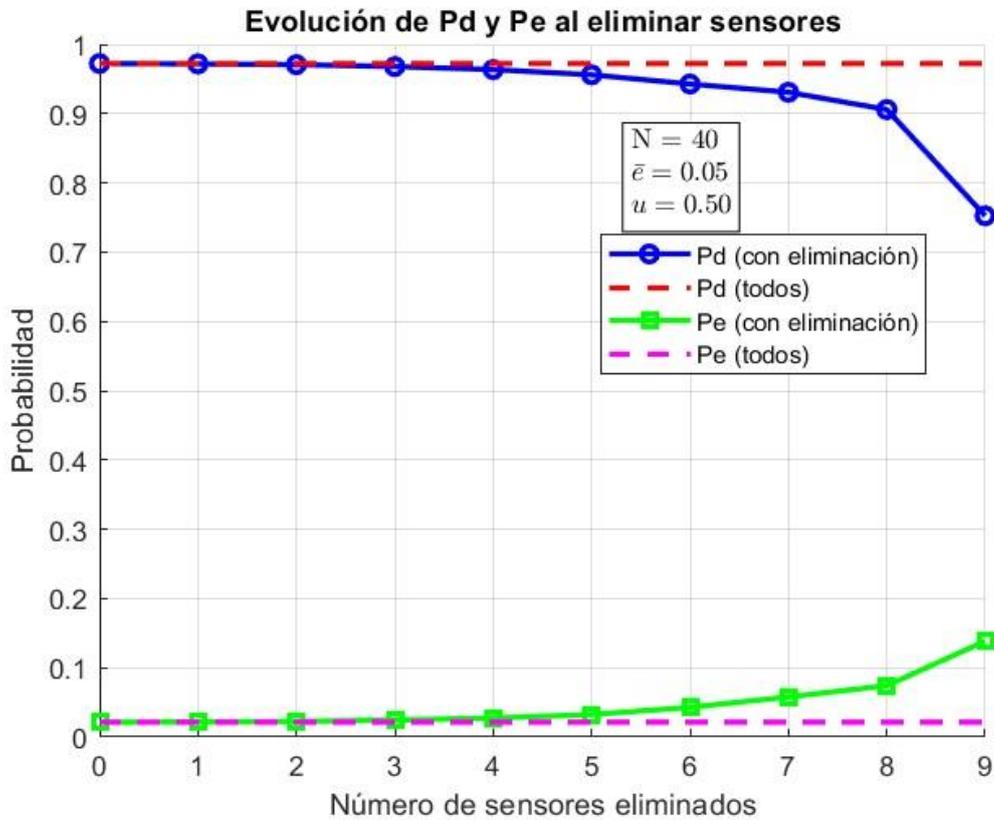


Fig.10. Probabilidad de detección y probabilidad de error en función del número de sensores eliminados (conjunto inicial de  $K=10$  sensores)

Si observamos la figura 10 apreciamos que a medida que se eliminan sensores, la probabilidad de error no presenta un deterioro relevante, lo que confirma la robustez del criterio de selección propuesto. Este resultado destaca que, al eliminar los sensores menos fiables, no solo se mantiene la probabilidad de detección, sino que tampoco se incrementa la probabilidad de error de manera sustancial. De este modo, se refuerza la conclusión de que es posible prescindir de sensores con bajo rendimiento sin afectar significativamente el desempeño global del sistema.

## 5.5 COMPARACIÓN CON SELECCIÓN ALATORIA

Para analizar la robustez del criterio de selección propuesto, vamos a comparar cómo evoluciona la probabilidad de detección en el caso en que los sensores se eliminen de forma aleatoria.

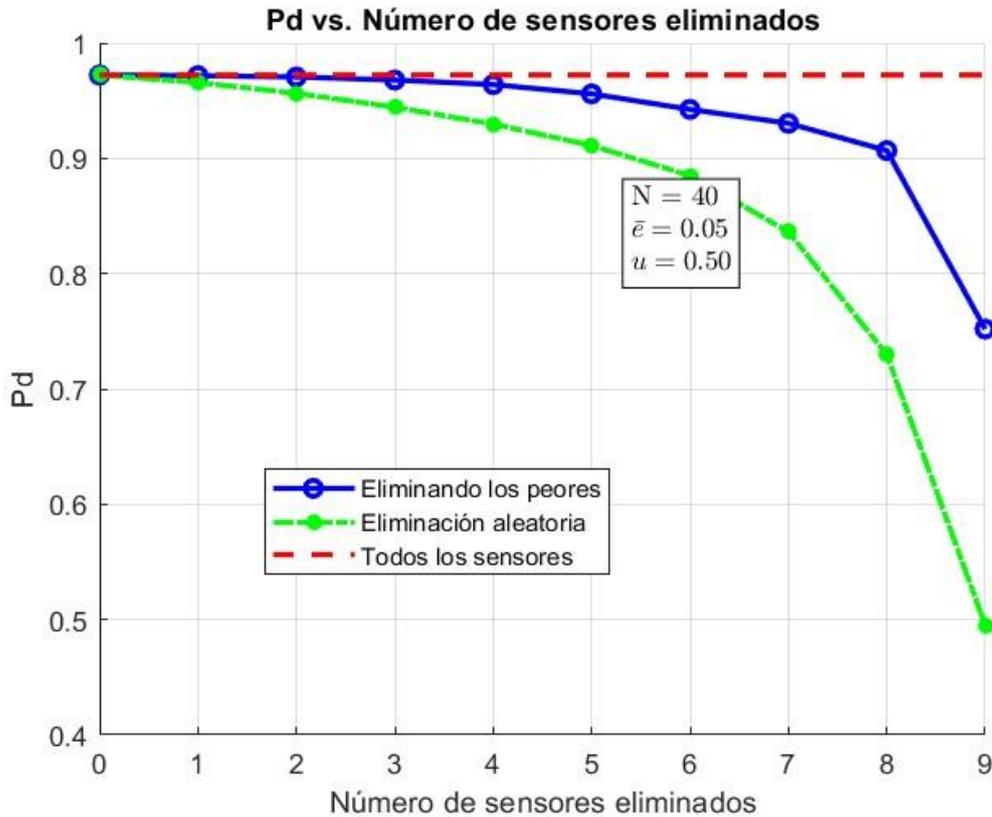


Fig.11. Probabilidad de detección en función del número de sensores eliminados (eliminación aleatoria)

Se puede observar (Figura 11) que la probabilidad de detección disminuye considerablemente más rápido cuando la eliminación de sensores se realiza aleatoriamente. Por ejemplo, al eliminar únicamente dos sensores, en el caso aleatorio la probabilidad de detección cae en torno a 0.95, mientras que aplicando el criterio de selección propuesto se mantiene aproximadamente en 0.97, es decir, en la misma probabilidad de detección que usando todos.

Esta diferencia se debe a que, en el caso de la eliminación aleatoria, se están eliminando sensores que aportan información útil al FC, lo que degrada el rendimiento del detector. En cambio, el criterio propuesto prioriza la eliminación de aquellos sensores con menor calidad, preservando así los más informativos y manteniendo un buen desempeño del sistema global.

## 5.6 COMPARACIÓN CON LA REGLA DE FUSIÓN DE VOTO MAYORITARIO

Con el objetivo de analizar la robustez del criterio de selección de sensores propuesto, se ha comparado su efecto al aplicarlo junto con una regla de fusión distinta: la regla de voto mayoritario (16).

Al simular cómo evoluciona la probabilidad de detección y compararla con el caso en el que se usen todos los sensores obtenemos los siguientes resultados:

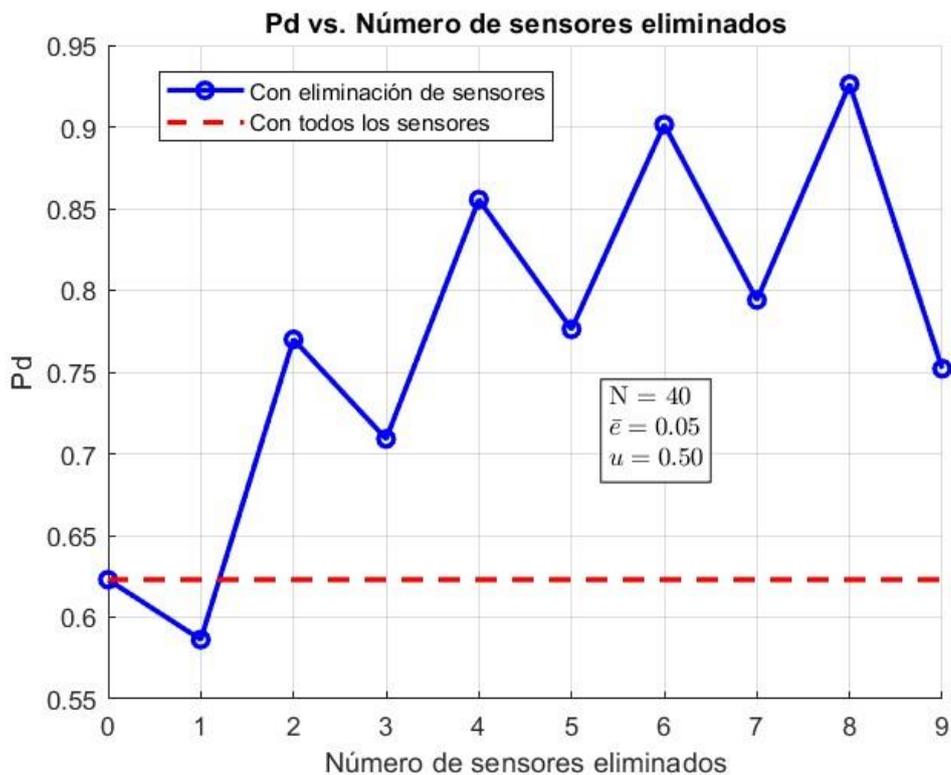


Fig.12. Probabilidad de detección en función del número de sensores eliminados aplicando voto mayoritario

Al aplicar el proceso de selección de sensores sobre esta regla, se observa una mejora sustancial en la probabilidad de detección en comparación con el uso de todos los sensores disponibles. Esto contrasta con los resultados obtenidos con la regla de fusión de mínima probabilidad de error, donde eliminar sensores malos no producía mejoras, ya que la propia regla ya penalizaba internamente a los sensores menos fiables.

En cambio, en la regla de voto mayoritario, donde todos los sensores tienen el mismo peso, eliminar sensores de baja calidad sí tiene un impacto directo y positivo sobre la Pd, al reducir el ruido en la decisión final y aumentar el peso relativo de los sensores más fiables.

Además, se observa que la gráfica presenta un patrón en zigzag, lo cual se debe a la asimetría inherente a esta regla de decisión. En particular, cuando el número de sensores es par, la probabilidad de detección tiende a ser mayor, ya que al definir la regla de voto mayoritario se estableció que, en caso de empate, se decide a favor de la hipótesis  $H_1$ .

Llama la atención que la mayor probabilidad de detección se alcanza cuando únicamente se conservan 2 sensores, lo que indica que, en esta configuración concreta, eliminar sensores de baja calidad tiene un impacto positivo muy significativo sobre el rendimiento del sistema.

También cabe destacar que pese a ser una regla muy sencilla, aplicando el criterio de selección de sensores conseguimos unas probabilidades de detección bastante buenas.

A continuación, se va a comparar la regla de fusión del voto mayoritario con las reglas de fusión de mínima probabilidad de error: una basada en el algoritmo EM, y otra en la que se supone que conocemos todos los parámetros reales de los sensores (la regla de fusión óptima con información completa).

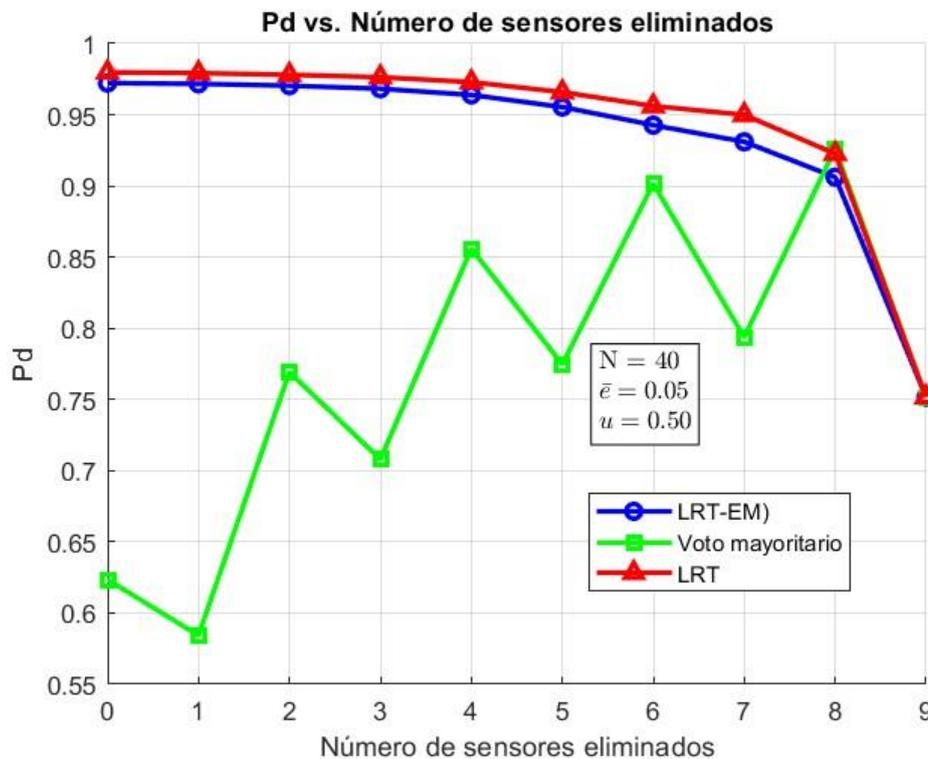


Fig.13. Comparación de la probabilidad de detección cuando se aplica VM, LRT-EM y LRT

Los resultados obtenidos son interesantes y aportan información valiosa sobre el comportamiento de los distintos métodos de fusión. Cuando únicamente nos quedamos con dos sensores, la regla de decisión basada en el voto mayoritario logra superar la regla de fusión basada en el LRT con estimaciones de las probabilidades de detección ( $d$ ) a partir del algoritmo EM. Esto se debe a que, al trabajar con tan solo dos sensores, la simplicidad y robustez del voto mayoritario permite obtener una mejor decisión conjunta, especialmente cuando las estimaciones de los parámetros pueden no ser muy precisas.

Sin embargo, a medida que aumenta el número de sensores disponibles, la ventaja del voto mayoritario se va perdiendo y la regla de fusión basada en el LRT muestra un rendimiento superior. Este resultado es lógico, ya que el LRT aprovecha toda la

información estadística disponible y, con un mayor número de sensores, la agregación de datos mejora la precisión de las decisiones conjuntas. La regla de fusión óptima basada en el conocimiento completo de los parámetros proporciona la mejor referencia posible y sirve como límite teórico para comparar los demás métodos.

Lo interesante de este análisis es que, en situaciones en las que no contamos con muchos sensores y hemos realizado una buena selección de sensores el voto mayoritario puede ofrecer un rendimiento competitivo.

Por otro lado, cuando la red cuenta con muchos sensores y podemos aprovechar el mayor volumen de información, las reglas de fusión más sofisticadas, como la del LRT con estimaciones de los parámetros mediante EM, empiezan a ofrecer resultados significativamente mejores. La precisión en la estimación de las probabilidades de detección de los sensores se traduce en decisiones más fiables y robustas frente a las incertidumbres.

Además, vamos a simular también la evolución de la probabilidad de error ( $Pe$ ) en los mismos casos. Como se esperaba, observamos un comportamiento similar al de la probabilidad de detección ( $Pd$ ). Con dos sensores, la regla de voto mayoritario obtiene menores valores de  $Pe$ , superando incluso a la regla de fusión basada en el LRT con estimaciones. Sin embargo, para configuraciones con más sensores, la regla de fusión basada en el LRT ofrece una probabilidad de error significativamente menor.

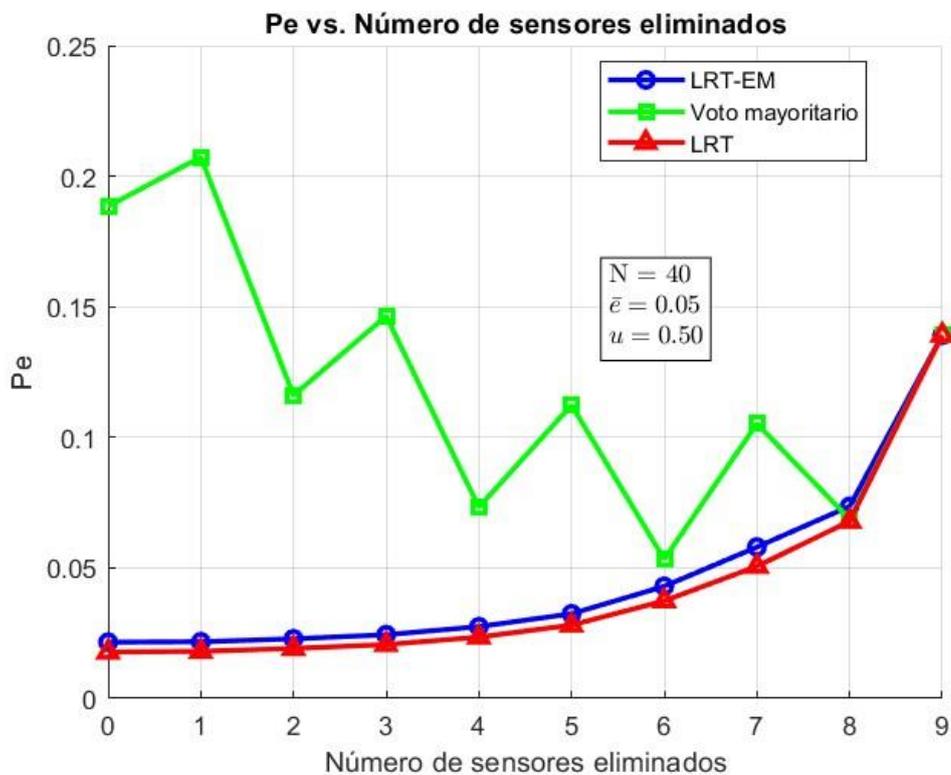


Fig.14. Comparación de la probabilidad de error cuando se aplica VM, LRT-EM y LRT

Este análisis muestra que no siempre una regla de fusión más compleja garantiza mejores resultados, especialmente en escenarios donde el número de sensores es reducido. Además, si se logra realizar una buena selección de sensores, es posible mejorar significativamente el rendimiento del sistema.

## 5.7 SELECCIÓN BASADA EN UMBRAL DE CALIDAD

Otro enfoque consiste en establecer un umbral basado en el ratio propuesto en este trabajo para descartar aquellos sensores cuya calidad no supere un determinado valor, en lugar de eliminar sistemáticamente un número fijo de sensores.

Para ello, se ha simulado la probabilidad de detección en función del umbral seleccionado. El umbral tomará valores desde 0, este caso sería un sensor incapaz de distinguir entre las hipótesis  $H_1$  y  $H_0$ , hasta 5 que sería un sensor de muy altas prestaciones. En esta simulación se ha partido de un conjunto inicial de 10 sensores.

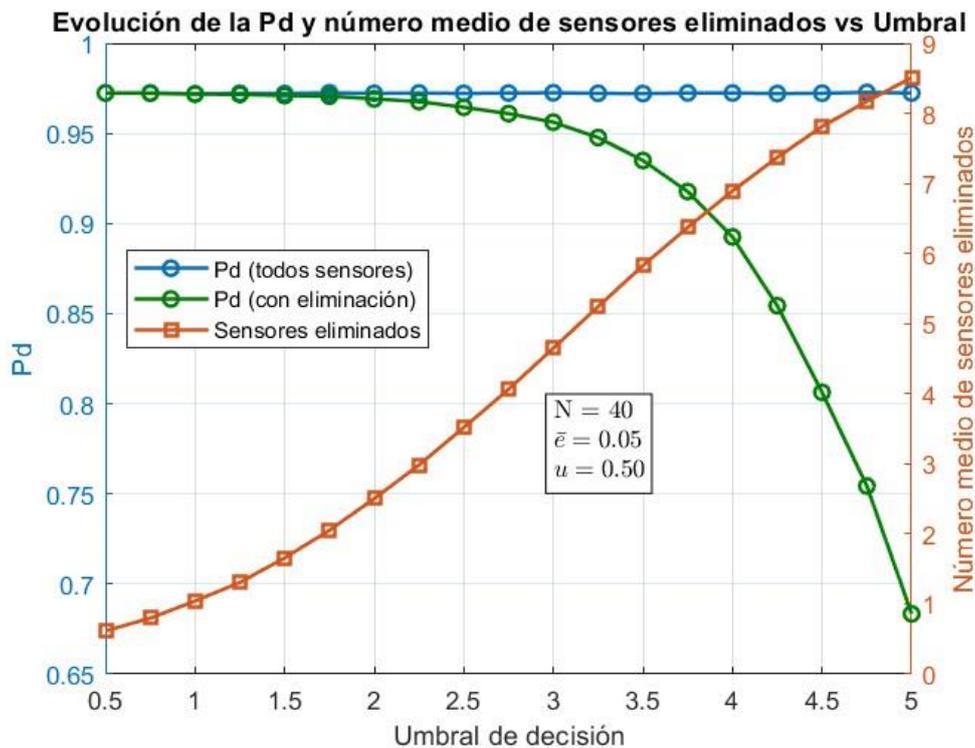


Fig.15. Comparación de la probabilidad de detección en función del umbral establecido

Este enfoque ofrece mayor flexibilidad, ya que permite ajustar dinámicamente el número de sensores descartados en función de su calidad, en lugar de fijar de antemano una cantidad determinada.

Al observar la gráfica correspondiente, se aprecia cómo, a medida que se incrementa el umbral, el número promedio de sensores eliminados también aumenta progresivamente. Esto es esperable, ya que a medida que se exige una mayor calidad (un ratio más alto), disminuye la cantidad de sensores que cumplen ese criterio.

Asimismo, se observa un comportamiento similar al analizado en gráficos anteriores: a partir de un umbral de aproximadamente 1.75, la probabilidad de detección comienza a decrecer de forma más notable. No obstante, incluso en este punto, se sigue manteniendo una Pd relativamente alta. Por tanto, se evidencia el compromiso (trade-off) entre el número de sensores que se desea descartar y el nivel de probabilidad de detección que se busca preservar.

Cabe destacar que, al observar la gráfica, se aprecia que la caída de la Pd es bastante brusca para umbrales elevados. Esto se debe a que, en algunos casos, ninguno de los sensores alcanza el valor umbral y se eliminan todos.

Por ello, se ha repetido la simulación manteniendo al menos los dos mejores sensores.

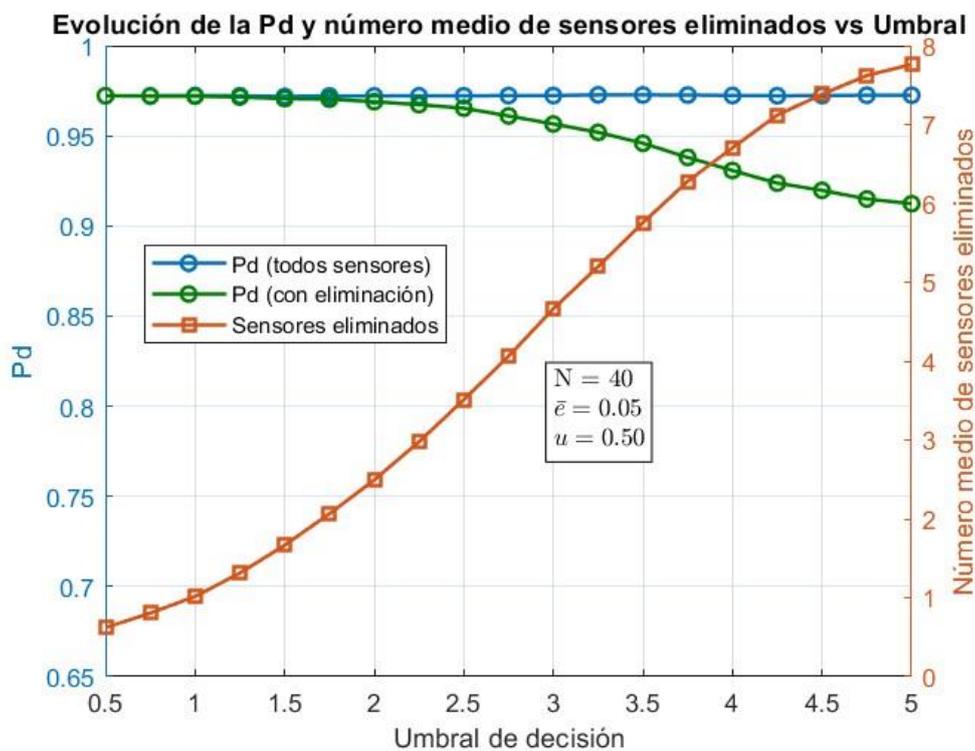


Fig.16. Comparación de la probabilidad de detección en función del umbral establecido manteniendo al menos dos sensores

Se observa en la figura que, al mantener siempre al menos los dos mejores sensores, se consigue preservar una probabilidad de detección superior al 90%, a pesar de eliminar en promedio un número elevado de sensores.

Al no estar limitados a eliminar un número fijo de sensores, este enfoque permite descartar un mayor número de ellos sin comprometer significativamente la probabilidad de detección. Esto resulta especialmente relevante desde el punto de vista de la eficiencia y el consumo energético, ya que se logra mantener el rendimiento del sistema mientras se reduce la cantidad de sensores en funcionamiento, lo que conlleva un ahorro energético considerable.

## 6. CONCLUSIONES

A lo largo de este trabajo se ha abordado el problema de la selección de sensores en una red de sensores inalámbricos cooperativa, con el objetivo de analizar su impacto sobre las prestaciones del sistema. El estudio se ha centrado en redes con combinación dura, en las que los sensores transmiten decisiones binarias al centro de fusión, y donde las probabilidades de detección son desconocidas y deben ser estimadas a partir de las observaciones mediante el algoritmo EM.

Los resultados obtenidos a partir de simulaciones de Monte Carlo permiten extraer varias conclusiones relevantes. En primer lugar, se ha comprobado que, al aplicar una regla de fusión óptima basada en el cociente de verosimilitudes con parámetros estimados, eliminar sensores con bajo rendimiento no mejora la probabilidad de detección, pero tampoco la empeora de manera notable. Esto demuestra que la propia regla de fusión, al ponderar menos a los sensores menos fiables, ya mitiga su impacto negativo en la decisión global.

Aunque los resultados obtenidos no muestran una mejora en dicha probabilidad mediante la eliminación selectiva de sensores, cuando aplicamos una regla de fusión óptima se ha verificado que prescindir de aquellos clasificados como menos eficaces no deteriora de forma sustancial el rendimiento global del sistema. Este resultado tiene importantes implicaciones prácticas, ya que permite reducir el número de sensores activos sin comprometer la capacidad del sistema para realizar tareas de detección fiables.

Por otro lado, se ha observado que cuando se utiliza una regla de fusión no ponderada, como el voto mayoritario, la selección de sensores sí mejora notablemente la probabilidad de detección. En este caso, al no penalizar internamente a los sensores menos fiables, eliminarlos permite reducir el ruido que introducen y obtener un mejor rendimiento global.

Esto es especialmente relevante en el contexto de redes de sensores inalámbricos con recursos energéticos limitados. Tal como se menciona en [19], la mayoría de estos sensores funcionan con baterías, lo que convierte la eficiencia energética en una prioridad de diseño [20]. Reemplazar o recargar sensores puede ser costoso, por lo que maximizar la duración de la red requiere un uso eficiente de los nodos disponibles. En este sentido, eliminar sensores que aportan escasa información permite ahorrar energía y reducir la complejidad del sistema sin degradar su desempeño.

Por otro lado, estudios recientes advierten que aumentar el número de nodos cooperativos puede tener efectos contraproducentes. Aunque una mayor cooperación reduce el tiempo necesario para la detección y los requisitos de sensibilidad de los detectores, también incrementa notablemente la sobrecarga de cooperación en el sistema [20], [21]. Este aumento de la sobrecarga implica más mensajes intercambiados, mayor carga de procesamiento en el centro de fusión y un consumo energético superior, lo que puede comprometer tanto la escalabilidad como la eficiencia del sistema. Además, si la mejora

en la probabilidad de detección es marginal, estos costes adicionales podrían no estar justificados.

Desde este punto de vista, la selección de sensores no solo debe ser entendida como una herramienta para intentar mejorar la precisión de detección, sino como un mecanismo para optimizar el uso de los recursos disponibles en la red. Al identificar y activar únicamente aquellos sensores que aportan información valiosa, se puede reducir la carga de comunicación y procesamiento, minimizando al mismo tiempo el consumo energético. Además, el ratio utilizado en este trabajo para clasificar los sensores podría utilizarse como indicador del deterioro en el rendimiento de los sensores individuales. Esto permitiría implementar mecanismos de mantenimiento preventivo o reemplazo gradual, mejorando la fiabilidad a largo plazo de la red.

Por tanto, se concluye que la selección de sensores puede ser una herramienta útil no tanto para maximizar la probabilidad de detección, sino para optimizar el uso de recursos energéticos, reducir la complejidad del sistema y prolongar la vida útil de la red, sin comprometer de manera significativa el rendimiento en la tarea de detección. Esto refuerza la idea de que, en entornos con restricciones energéticas y de procesamiento, como las redes de sensores inalámbricos, una estrategia de selección adecuada puede ofrecer un compromiso eficiente entre rendimiento y consumo energético.

## ANEXO

### 1. Función Matlab para la inicialización de los parámetros:

```
function [f, d_reales, e, X] = inicializacion_parametros(K, N, u)
%Inicializa los parámetros necesarios para el algoritmo EM
%
% Entradas:
% K - Número de sensores
% N - Número de observaciones
% u - Probabilidad de que una observación verdadera sea 1
%
% Salidas:
% f - Vector de probabilidades de falsa alarma por sensor
% d_reales - Vector de probabilidades reales de detección por sensor
% e - Vector de probabilidades de error del canal por sensor
% X - Matriz de observaciones binarias (K sensores x N muestras)

%----- Inicialización de parámetros -----%

% Inicializar la probabilidad de falsa alarma:  $f_k \sim U(0, 0.1)$ 
f = rand(K, 1) * 0.1;
% Inicializar la probabilidad real de detección:  $Pd_k - Pf_k \sim U(0, 0.9)$ 
d_reales = f + rand(K, 1) * 0.9;

%----- Probabilidad de error del canal -----%

sigma_h = 1; % Varianza del canal de Rayleigh (valor fijado)
sigma_w = 0.2595; % Ruido del canal =>  $E[e] \approx 0.05$ 

% Muestras de la amplitud del canal (valor absoluto de h)
muestras_abs_h = raylrnd(sigma_h * ones(K, 1)); % Vector Kx1

% Cálculo de la probabilidad de error del canal BSC para cada sensor
e = qfunc(muestras_abs_h / (2 * sigma_w));

%----- Generación de observaciones -----%

% Generar estados verdaderos z (1 con probabilidad u)
z = rand(1, N) < u;

% Inicializar matriz de observaciones
X = zeros(K, N);

% Generar observaciones X según el estado verdadero y parámetros de los sensores
for n = 1:N
    for k = 1:K
        if z(n) == 1
            % Si el estado verdadero es 1, usar la probabilidad de detección
            X(k, n) = rand() < d_reales(k);
        else
            % Si el estado verdadero es 0, usar la probabilidad de falsa alarma
            X(k, n) = rand() < f(k);
        end
    end
end
end
```

## 2. Función Matlab para estimar las probabilidades de detección usando el algoritmo EM:

```

function d = EM_algorithm(X, f, e, u)
%
% Entradas:
% X -> Matriz de observaciones binarias (K x N)
% f -> Vector de probabilidades de falsa alarma por sensor (K x 1)
% e -> Vector de probabilidades de error de canal por sensor (K x 1)
% u -> Probabilidad a priori de H1
%
% Salida:
% d -> Vector estimado de probabilidades de detección por sensor (K x 1)

%----- Inicialización -----%

[K, N] = size(X);           % K sensores, N observaciones
epsilon_t = 1e-3;          % Pequeño valor para evitar extremos en t
t = zeros(1, N);           % Vector de probabilidades posteriores

% Inicialización heurística de t según la mayoría de sensores
sum_X = sum(X, 1);
t(sum_X < K / 2) = epsilon_t;
t(sum_X == K / 2) = 0.5;
t(sum_X > K / 2) = 1 - epsilon_t;

% Inicialización de d a partir de t inicial
N_1_1 = X * t';
N_0_1 = (1 - X) * t';
N_1 = N_1_1 + N_0_1;

b = N_1_1 ./ N_1;
d = max(f, min(b, 1 - e)); % Aplicar restricciones:  $f \leq d \leq 1 - e$ 

% Inicializar d_old para el criterio de parada
d_old = zeros(K, 1);

%----- Bucle EM -----%
while max(abs(d - d_old)) > 1e-4 % Umbral de convergencia elegido

    d_old = d;
    % ----- E-step: calcular t -----
    p_x_H0 = prod((f .^ X) .* ((1 - f) .^ (1 - X)), 1); % P(X|H0)
    p_x_H1_d = prod((d .^ X) .* ((1 - d) .^ (1 - X)), 1); % P(X|H1, d)

    t = (u * p_x_H1_d) ./ ((1 - u) * p_x_H0 + u * p_x_H1_d); % P(H1|X)
    % ----- M-step: actualizar d -----
    N_1_1 = X * t';
    N_0_1 = (1 - X) * t';
    N_1 = N_1_1 + N_0_1;

    b = N_1_1 ./ N_1;

    % Aplicar restricciones  $f \leq d \leq 1 - e$ 
    for k = 1:K
        if b(k) > f(k) && b(k) < 1 - e(k)
            d(k) = b(k);
        elseif b(k) <= f(k)
            d(k) = f(k);
        else
            d(k) = 1 - e(k);
        end
    end
end
end
end

```

## REFERENCIAS

- [1] I.F. Akyildiz et al., "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer Networks*, vol. 50, no. 13, pp. 2127–2159, 2006.
- [2] Ian F. Akyildiz, Brandon F. Lo, Ravikumar Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Physical Communication*, vol. 4, no. 1, pp. 40–62, 2011.
- [3] Zeng, Y., Liang, YC., Hoang, A.T. *et al.* A Review on Spectrum Sensing for Cognitive Radio: Challenges and Solutions. *EURASIP J. Adv. Signal Process.* **2010**, 381465 (2010)
- [4] T. Renk, C. Kloeck and F. K. Jondral, "A Cognitive Approach to the Detection of Spectrum Holes in Wireless Networks," *2007 4th IEEE Consumer Communications and Networking Conference*, Las Vegas, NV, USA, 2007, pp. 1118-1122, doi: 10.1109/CCNC.2007.225.
- [5] Tevfik Yücek y Hüseyin Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, pp. 116-130, 2009.
- [6] C. Sun, W. Zhang and K. Ben Letaief, "Cooperative Spectrum Sensing for Cognitive Radios under Bandwidth Constraints," *2007 IEEE Wireless Communications and Networking Conference*, Hong Kong, China, 2007, pp. 1-5,
- [7] E. Visotsky, S. Kuffner and R. Peterson, "On collaborative detection of TV transmissions in support of dynamic spectrum sharing," *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005.*, Baltimore, MD, USA, 2005, pp. 338-345
- [8] E. Peh and Y.-C. Liang, "Optimization for cooperative sensing in cognitive radio networks," in *Proc. IEEE Wireless Commun. and Networking Conf.*, Hong Kong, Mar. 2007, pp. 27–32.
- [9] D. Ciuonzo, G. Romano and P. S. Rossi, "Channel-Aware Decision Fusion in Distributed MIMO Wireless Sensor Networks: Decode-and-Fuse vs. Decode-then-Fuse," in *IEEE Transactions on Wireless Communications*, vol. 11, no. 8, pp. 2976-2985, August 2012
- [10] J. -Y. Wu, C. -W. Wu, T. -Y. Wang and T. -S. Lee, "Channel-Aware Decision Fusion With Unknown Local Sensor Detection Probability," in *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1457-1463, March 2010
- [11] D. Ciuonzo and P. Salvo Rossi, "Decision Fusion With Unknown Sensor Detection Probability," in *IEEE Signal Processing Letters*, vol. 21, no. 2, pp. 208-212, Feb. 2014
- [12] P. K. Varshney, *Distributed Detection and Data Fusion*, Springer New York, 1996

- [13] J. Pérez, I. Santamaría, A. Pagès-Zamora, Blind learning of the optimal fusion rule in wireless sensor networks, submitted to Signal Processing, Elsevier, 2025
- [14] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice Hall, Signal Processing Series, 1998.
- [15] H. V. Poor, *An Introduction to Signal Detection and Estimation*. Springer-Verlag, 1994.
- [16] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [17] G. J. McLachlan y T. Krishnan, *The EM Algorithm and Extensions*. John Wiley & Sons, 2008.
- [18] F. R. Mohammad, D. Ciuonzo, and Z. A. K. Mohammed, "Mean-based blind hard decision fusion rules," *IEEE Signal Processing Letters*, vol. 25, no. 5, pp. 630–634, 2018
- [19] M. Najimi, A. Ebrahimzadeh, S. M. H. Andargoli and A. Fallahi, "Lifetime Maximization in Cognitive Sensor Networks Based on the Node Selection," in *IEEE Sensors Journal*, vol. 14, no. 7, pp. 2376-2383, July 2014
- [20] M. Najimi, A. Ebrahimzadeh, S. M. H. Andargoli and A. Fallahi, "A Novel Sensing Nodes and Decision Node Selection Method for Energy Efficiency of Cooperative Spectrum Sensing in Cognitive Sensor Networks," in *IEEE Sensors Journal*, vol. 13, no. 5, pp. 1610-1621, May 2013
- [21] S. Maleki, S. P. Chepuri, and G. Leus, "Optimization of hard fusion based spectrum sensing for energy-constrained cognitive radio networks," *Physical Communication*, vol. 9, pp. 193–198, 2013