

Adaptive Signal Source Estimation Using Eigenvalue Analysis and Optimized Dynamic Thresholding

Uyarlamalı Sinyal Kaynağı Tahmini: Özdeğer Analizi ve Optimize Edilmiş Dinamik Eşikleme Yaklaşımı

Gökhan Yüksek¹, Timur Lale²

 0000-0002-6832-8622

 0000-0002-6958-5057

^{1,2}Department of Electrical and Electronics Engineering, Batman University, Batman, Türkiye
gokhan.yuksekk@batman.edu.tr¹, timur.lale@batman.edu.tr²

Öz— Bu çalışma, radar, kablosuz haberleşme ve sensör ağlarında kullanılmak üzere uyarlamalı eşiklemeye dayalı bir sinyal kaynağı belirleme yöntemi sunmaktadır. Minimum Tanım Uzunluğu (MDL), Akaike Bilgi Kriteri (AIC) ve yeni bir özdeğer tabanlı uyarlamalı eşikleme yaklaşımı olmak üzere üç yöntem; gürültülü, kısa menzilli ve birbirine yakın kaynakların bulunduğu koşullar için tasarlanan ultra geniş bant (UWB) sensör dizileri üzerinde gerçekleştirilen simülasyonlarla karşılaştırılmıştır. MDL ve AIC, model karmaşıklığı ile veri uyumunu dengelemek amacıyla kovaryans matrisi özdeğerlerini kullanırken, önerilen uyarlamalı eşikleme yöntemi özdeğer dağılımına göre gürültü eşiğini dinamik olarak ayarlayarak yüksek gürültü koşullarında kaynak sayısı tahmin performansını artırmaktadır. Sonuçlar, tüm yöntemlerin düşük gürültü koşullarında iyi performans gösterdiğini, ancak önerilen yöntemin yüksek gürültü seviyelerinde veya kaynakların birbirine yakın olduğu durumlarda daha yüksek başarı sağladığını göstermektedir. Ayrıca eşik fonksiyonundaki duyarlılık parametreleri hafif hesaplama yüküne sahip bir arama yöntemiyle optimize edilerek farklı gürültü seviyeleri ve kaynak konfigürasyonları altında kararlı performans elde edilmektedir.

Anahtar Kelimeler— Uyarlamalı Eşikleme, Kaynak Tahmini, Özdeğer Analizi, Optimizasyon

Abstract— This study presents a method for identifying signal sources using adaptive thresholding, crucial in radar, wireless communication, and sensor networks. Three approaches, Minimum Description Length (MDL), Akaike Information Criterion (AIC), and a novel eigenvalue-based adaptive thresholding are compared through simulations on ultra-wideband (UWB) sensor arrays designed for noisy, low-range conditions with closely spaced sources. MDL and AIC employ covariance matrix eigenvalues to balance model complexity and data fit, while the proposed adaptive thresholding dynamically adjusts the noise floor according to the eigenvalue distribution, improving source-enumeration performance under high-noise conditions. Results show that all methods perform well under low-noise conditions, but the proposed method excels when noise is high or sources are closely located. Additionally, sensitivity parameters in the threshold function are

optimized via a lightweight search, providing stable performance across varying noise and source configurations.

Key Words— Adaptive Thresholding, Source Estimation, Eigenvalue Analysis, Optimization

I. Introduction

Estimating the number of active signal sources remains a major challenge in signal processing, which underpins applications in sensor networks, radar systems, wireless communication, and medical imaging. Accurate source estimation improves network efficiency and channel allocation in wireless systems, enables reliable multi-target tracking in radar applications, and supports fault detection and environmental monitoring in sensor arrays. [1], [2], [3], [4], [5], [6], [7].

Eigenvector-based techniques are widely applied to separate signal and noise components by analysing the eigenvalues of the covariance matrix, which represent energy distribution across dimensions. [8] The present work employs this property to estimate the number of sources in real time without explicitly constructing subspaces. Compared with conventional time-domain approaches, eigenvalue-based methods provide source-number information directly from the covariance eigenvalue spectrum without requiring prior knowledge of the exact source waveforms or explicit construction of an array manifold [9], [10].

Several notable studies have contributed to this field. Hyvärinen and Oja advanced independent component analysis algorithms [11]; Huang et al. introduced an MMSE-based MDL estimator [12]. Wiesel and Hero explored decomposable principal component analysis [13];

Ouedraogo et al. developed a non-negative blind source separation approach [14]; and Han and Nehorai enhanced source estimation using nested and linear arrays [15].

This study compares three representative techniques: Minimum Description Length (MDL), Akaike Information Criterion (AIC), and a proposed eigenvalue-driven adaptive thresholding scheme. MDL [16] and AIC [17], [18] are established model-selection criteria balancing model complexity and data fit through eigenvalues. While MDL favours simplicity to prevent overfitting, AIC emphasises prediction accuracy. In contrast, adaptive thresholding dynamically adjusts the noise floor from the eigenvalue distribution, maintaining effectiveness under varying noise and source configurations.

The paper makes several contributions. It offers a systematic comparison of MDL, AIC, and adaptive thresholding under different noise environments, demonstrating the latter's advantage when noise is high or sources are closely spaced. The study also integrates a lightweight optimization framework for threshold parameters (α, β) , reducing heuristic dependence and ensuring consistent performance. A brief computational analysis confirms that the proposed approach shares the same cost order as MDL and AIC but avoids extra overheads, making it suitable for practical ultra-wideband (UWB) sensor implementations.

The main contribution of this study is an optimized adaptive source enumeration method that combines spectral curvature information, eigenvalue-based noise-floor estimation, and data-driven threshold parameter selection within a single procedure. Unlike MDL and AIC, which determine the source number through penalized model-order selection, the proposed method directly evaluates the structure of the ordered covariance eigenvalues. First, the second discrete difference of the eigenvalue sequence is used to identify sharp spectral transitions between signal- and noise-dominated components. This information is incorporated into a weighted adaptive threshold rather than treating all eigenvalues equally. Second, an MP-inspired spread term is included to account for the finite-sample spread of noise eigenvalues as a function of the sensor-to-sample ratio. Third, the threshold sensitivity parameters are selected through a lightweight coarse-search and local-refinement procedure and are subsequently fixed for the reported simulations. Thus, the contribution lies in the combined use of discrete spectral-curvature weighting, finite-sample noise correction, and optimized threshold selection for source enumeration, rather than in eigenvalue

thresholding alone.

The remainder of the paper is organized as follows. Section II introduces MDL and AIC and presents the proposed adaptive thresholding method, including the weighting strategy, threshold formulation, parameter optimization, and computational complexity. Section III describes the implementation and simulation setup. Section IV presents the comparative results under different SNR levels and source separations and discusses the main limitations of the method. Finally, Section V summarizes the main findings and conclusions.

II. Methods

A. Minimum Description Length

The MDL [16], [18] criterion is a model selection approach based on information theory, aiming to minimize the total description length of both the model and the observed data. The number of active sources in sensor array systems can be estimated using MDL for signal processing. This is done by analysing the eigenvalues of the sample covariance matrix. The main principle behind MDL is to balance model accuracy with complexity, preventing overfitting by advancing simpler models. Given a covariance matrix R derived from received signals at an array of N sensors, its eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ represent the variance of the signal components along different dimensions. The MDL criterion for estimating the number of sources k is formulated as:

$$MDL(k) = -T \sum_{i=k+1}^N \log \left(\lambda_i^{\frac{1}{N-k}} \right) + k \frac{(2N - k + 1)}{2} \log(T), \quad (1)$$

where T is the number of time samples. The first term in this expression represents the likelihood of the data fitting the assumed signal model, while the second term penalizes complex models to prevent overestimation. The optimal number of sources k is determined by selecting the value that minimizes $MDL(k)$. This method is effective in distinguishing between the signal and noise subspaces, as it identifies the significant eigenvalues corresponding to the signal sources.

Compared with AIC, MDL applies a stronger penalty to model complexity, which generally reduces the tendency toward model-order overestimation. But this stronger penalty

does not imply better performance under high-noise conditions. At low SNR levels, weak signal eigenvalues may approach the noise eigenvalue cluster, increasing the tendency of MDL to underestimate the number of active sources. So, its performance depends on the separability of the signal- and noise-related eigenvalues as well as the applied complexity penalty [16], [17], [18], [19], [20], [21].

B. Akaike Information Criterion

The AIC [16], [17], [18] is a model selection method that seeks to strike a balance between model intricacy and data congruence. Statisticians generally use it in the context of statistical signal processing to estimate the number of active sources for sensor arrays. They do this by analysing the eigenvalues of the sample covariance matrix. AIC is different from the MDL criterion in that it is designed to minimise the relative information loss by selecting the model with the lowest penalised likelihood score. This is in contrast to the MDL criterion, which imposes a stronger penalty on model complexity. Given a covariance matrix R derived from the received signals at an array of N sensors, its eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ represent the variance of the signal components along different dimensions. The AIC criterion for estimating the number of sources k is as follows:

$$AIC(k) = -2T \sum_{i=k+1}^N \log\left(\lambda_i^{\frac{1}{N-k}}\right) + 2k(2N - k + 1), \quad (2)$$

where T is the number of time samples. The first term in this equation corresponds to the likelihood function, measuring how well the model fits the data, while the second term serves as a penalty to prevent overfitting by discouraging excessive model complexity. The optimal number of sources k is determined by selecting the value that minimizes $AIC(k)$. AIC uses a weaker penalty than MDL, so it often chooses more complex models. This makes it useful when underestimating the number of sources is a bigger problem than overestimating. However, in high-noise conditions, AIC may overestimate the number of sources due to its lower penalization of complexity. [16], [17], [18], [19], [20], [21].

C. Adaptive Thresholding

The adaptive thresholding method dynamically determines the number of active signal sources by analysing the eigenvalue spectrum. Unlike traditional model selection criteria such as MDL and AIC, which rely on predefined penalty functions, adaptive thresholding adjusts itself based

on the statistical distribution of eigenvalues, allowing it to perform well across different noise levels and signal conditions. However, conventional adaptive thresholding methods often consider all eigenvalues equally, potentially leading to incorrectness when sources have varying power levels. To address this, a weighted adaptive thresholding approach is introduced, enhancing sensitivity to dominant sources while suppressing noise components.

In an array signal model, a linear array that formed by N sensors receives signals from k active sources. The received signal matrix X can be expressed as;

$$X = AS + W, \quad (3)$$

where A is the mixing matrix including the steering vectors of the sources, S represents the transmitted source signals over T time samples, and W is an additive noise matrix, typically modeled as zero-mean Gaussian noise. From this signal model, the covariance matrix R is computed as;

$$R = \frac{1}{T} XX^H, \quad (4)$$

where X^H is the Hermitian transpose of X . The eigenvalues of R , denoted as $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$, provide insight into the energy distribution of the signal and noise components. The larger eigenvalues correspond to active sources, while the smaller eigenvalues represent noise. The signal model considered in this study assumes additive spatially white Gaussian noise and non-coherent source signals. Under these conditions, the noise covariance is proportional to the identity matrix, while statistically independent source components contribute distinct signal-related eigenvalues to the covariance spectrum. These assumptions provide the basis for separating signal- and noise-dominated eigenvalues. Therefore, the proposed method does not assume an arbitrary signal environment; rather, it avoids requiring prior knowledge of the exact source waveforms or an explicitly constructed array manifold.

A key challenge in adaptive thresholding is the dynamic determination of a threshold τ that effectively distinguishes signal-dominant and noise-dominant eigenvalues. Instead of relying on a fixed threshold, this method estimates the noise floor using statistical properties of eigenvalues. A common approach is to define τ as a function of the mean or median of the eigenvalues;

$$\tau = \alpha \cdot \text{median}(\lambda_i) \text{ or } \tau = \alpha \cdot \text{mean}(\lambda_i), \quad (5)$$

where α is a scaling factor that adjusts the threshold sensitivity. While effective in moderate noise environments, this formulation does not account for variations in eigenvalue importance. To increase the truthfulness of estimation, a new weighted thresholding formulation incorporating eigenvalue characteristics is proposed;

$$\tau = \alpha \cdot \sum_{i=1}^N \omega_i \lambda_i + \beta \cdot \sigma_n^2 (1 + \sqrt{C}). \quad (6)$$

Where ω_i denotes the weight assigned to each eigenvalue according to the local spectral variation described below, while α controls the sensitivity of the eigenvalue-based threshold component. The proposed weighting approach assigns greater importance to regions where the ordered eigenvalue spectrum exhibits sharp local variations. Since the eigenvalue index is discrete, this variation is characterized by the second discrete difference;

$$\Delta^2 \lambda_i = \lambda_{\{i+1\}} - 2\lambda_i + \lambda_{\{i-1\}} \quad \left. \vphantom{\Delta^2 \lambda_i} \right\} \quad i = 2, \dots, M-1. \quad (7)$$

A larger magnitude of $(\Delta^2 \lambda_i)$ represents a stronger local change in the spectral slope and therefore indicates a potential transition between signal- and noise-dominated eigenvalues.

The weighting strategy is based on the local structure of the ordered eigenvalue spectrum. A pronounced change in the second discrete difference indicates a stronger variation in spectral curvature and therefore provides information about a possible transition between signal- and noise-dominated components. This transition information is incorporated into the adaptive threshold so that eigenvalues are not treated equally. In this formulation, regions with stronger local curvature receive greater emphasis, allowing the threshold to respond to the spectral separation observed in the covariance eigenvalues. Since the calculation uses only the ordered eigenvalues already obtained from the covariance matrix, no additional eigenvector-based processing is required.

The subspace separation in the proposed adaptive thresholding method is performed by analysing the ordered eigenvalue spectrum rather than explicitly using the corresponding eigenvectors. The second discrete difference provides local spectral-curvature information, and eigenvalues around stronger transition regions are assigned

greater importance in the adaptive threshold. Thus, the weighting strategy incorporates the local structure of the eigenvalue spectrum instead of treating all eigenvalues equally. The procedure operates on the eigenvalues already obtained from the covariance matrix and does not require additional eigenvector-based processing.

The threshold formulation includes an MP-inspired correction term [10], [22], to account for the finite-sample spread of noise-related eigenvalues. In random matrix theory, this spread depends on the ratio between the number of sensors and the number of temporal observations. Accordingly, the correction term incorporates the ratio $C = N/T$, together with the estimated noise variance σ_n^2 . In the proposed formulation, β controls the contribution of this noise-dependent correction to the adaptive threshold.

$$C = N/T, \quad (8)$$

the ratio of the number of sensors (N) to the number of time samples (T). The estimated number of sources is then determined by;

$$n = \sum_{i=1}^N \mathbb{I}(\lambda_i > \tau). \quad (9)$$

Here, $\mathbb{I}(\cdot)$ is an indicator function that assigns 1 if the condition is met and 0 in other cases. This function forms the output that only eigenvalues significantly above the noise floor are classified as signal components.

Adaptive thresholding is an effective solution for source estimation in array signal processing. This is achieved by leveraging dynamic noise floor estimation, weighted eigenvalue distributions and statistical noise modelling. Incorporating noise-adaptive scaling factors and weighted thresholding enhances its capacity to function in real-world applications such as radar, ultra-wideband (UWB) communication and multi-sensor networks, where signal characteristics fluctuate dynamically. Furthermore, its utility is demonstrated in civil defence scenarios, for example in identifying the number of heartbeats beneath a collapsed structure.

D. Optimization of Threshold Parameters α and β

The sensitivity parameters in α and β in (10) and (11) directly control the threshold's position relative to the eigenvalue spectrum and play a decisive role in

distinguishing signal and noise components. Instead of assigning fixed values, these parameters are optimized as an integral part of the proposed method.

To evaluate each candidate pair (α, β) , a success rate metric is used. For a given scenario d , the success rate $\eta(\alpha, \beta | d)$ is defined as

$$\eta(\alpha, \beta | d) = \frac{\text{Number of correct source est.}}{\text{Total number of trials}}. \quad (10)$$

The success rate $\eta(\alpha, \beta | d)$ corresponds to the fraction of trials in which the estimated number of sources matches the ground truth. Maximizing its average across the validation set $\mathcal{D}$ defines the objective function of the optimization.

The optimization objective is then expressed as

$$J(\alpha, \beta) = \frac{1}{|\mathcal{D}|} \sum_{d \in \mathcal{D}} \eta(\alpha, \beta | d), \quad (11)$$

where $\mathcal{D}$ denotes a validation set including representative conditions such as SNR levels between -10 and $+5$ dB, sensor array sizes of 7 and 9 elements, and minimum source spacings between 0.05 m and 0.20 m. The optimal parameters are determined by

$$(\alpha^*, \beta^*) = \arg \max_{\alpha, \beta} J(\alpha, \beta). \quad (12)$$

In practice, the optimization follows a two-step procedure:

- Coarse grid search [23]: A discrete grid of candidate values is generated for $\alpha \in [0.8, 2.0]$ and $\beta \in [0, 0.3]$. Each pair is evaluated by computing $J(\alpha, \beta)$ over the validation set $\mathcal{D}$.
- Local refinement: The best-performing candidates are further refined using a Nelder–Mead [24] simplex search, which efficiently locates the optimum without requiring derivative information.

This lightweight procedure converges quickly and introduces negligible additional cost, since the dominant computation remains the eigenvalue decomposition of the covariance matrix. The optimized pair (α^*, β^*) is subsequently fixed and applied consistently in all reported simulations.

Incorporating this parameter optimisation makes the adaptive thresholding method statistically adaptive through eigenvalue distribution analysis, as well as parametrically tuned. This ensures stable performance across different SNR and geometry scenarios.

E. Computational Complexity

The dominant cost in all methods is the eigenvalue decomposition of the covariance matrix, which scales as $O(N^3)$ for an array of N sensors. MDL and AIC additionally require scoring candidate orders, resulting in $O(N^3 + KN)$, where K is the maximum number of sources considered. The proposed adaptive thresholding performs only a threshold comparison after the decomposition, maintaining $O(N^3)$ complexity. The lightweight optimization of α and β (Sec. 2.D) involves a coarse grid and a small number of refinement steps; since each step reuses the same decomposition, the additional cost remains negligible compared to the baseline operations.

III. Implementation

A. Pseudo-code

The steps of the proposed method are outlined in the pseudocode below;

Pseudocode:

i. Input: Sensor data matrix X , number of sensors M , number of temporal samples T , and noise variance σ_n^2 .

ii. Compute: Sample covariance matrix:

$$R = (1/T) XX^H$$

iii. Compute: Eigenvalues of R and sort them in descending order:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$$

iv. Compute: Second discrete differences of the ordered eigenvalues:

$$\Delta^2 \lambda_i = \lambda_{i+1} - 2\lambda_i + \lambda_{i-1}, \quad i = 2, \dots, M-1$$

v. Identify: Dominant local variations in the second discrete-difference sequence and determine the corresponding eigenvalue weights w_i .

vi. Compute: Adaptive threshold τ using the weighted eigenvalue information, optimized sensitivity parameters α and β , and the MP-inspired correction term.

vii. Count: Eigenvalues above the adaptive threshold:

$$\lambda_i > \tau$$

viii. Output: Estimated number of active signal sources:
 $\hat{K} = \sum_{i=1}^M I(\lambda_i > \tau)$

B. Simulation

The present simulation example is concerned with the estimation of the number of active signal sources in a linear sensor array configuration by means of a technique known as eigen-value distribution analysis. The example setup is shown in Fig. 1. The setup consists of an array that has been formed by 9 sensors which have been arranged linearly along the x-axis, with a spacing of 0.12 metres to optimise resolution.

Each sensor in the linear array receives signals from 7 sources located at varying distances (ranging from 1.7 m to 4 m) and angles (ranging from -15° to 15°). The signals received at array experience time delays depending on the distances they travel, so interpolation is used to align the signals at each sensor. The sensor outputs are then used to compute the sample covariance matrix, from which eigenvalue decomposition is performed.

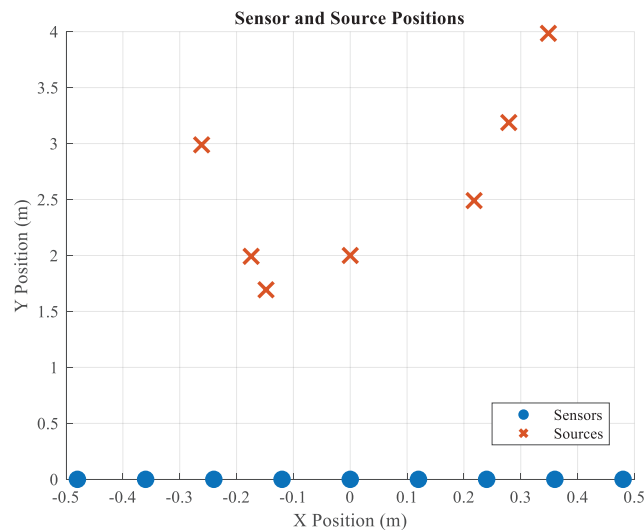


Fig. 1. Distribution of the linear sensor array and signal sources.

The system works at a centre frequency of 2.4 GHz, with a sampling rate of 24 GHz to make sure that time is measured very accurately. The SNR calculation is used to evaluate performance. This is done by simulating realistic signal conditions using Gaussian noise with a proper variance. Broadband characteristics suitable for UWB applications are ensured by the transmitted signals being high-order Morlet

pulses.

The simulations are conducted under a temporally oversampled condition, where the number of temporal observations (T) is greater than the number of array sensors (M), i.e., ($T > M$). Accordingly, the received data matrix contains sufficient temporal observations for estimating the ($M \times M$) sample covariance matrix. This condition reduces the risk of rank deficiency associated with an insufficient number of observations and provides a more suitable eigenvalue spectrum for signal-noise separation. It is also consistent with the ratio ($c = M/T < 1$) used in the MP-inspired correction term. Therefore, the reported performance corresponds to the temporally oversampled operating condition considered in this study.

Fig. 2 presents the Morlet wavelet signal and its corresponding Fourier spectrum at a centre frequency of approximately 2.5 GHz. The top plot illustrates the Morlet wavelet in the time domain, locally showing a sinusoidal oscillation. The bottom plot displays its Fourier spectrum, which exhibits a sharp peak around 2.4 GHz, confirming the dominant frequency of the wavelet. The frequency spectrum also shows minor side lobes, which result from the finite time window of the signal. The presence of low-level noise in both the time and frequency domains suggests realistic signal conditions, which may impact signal resolution but remain within acceptable limits for the UWB applications.

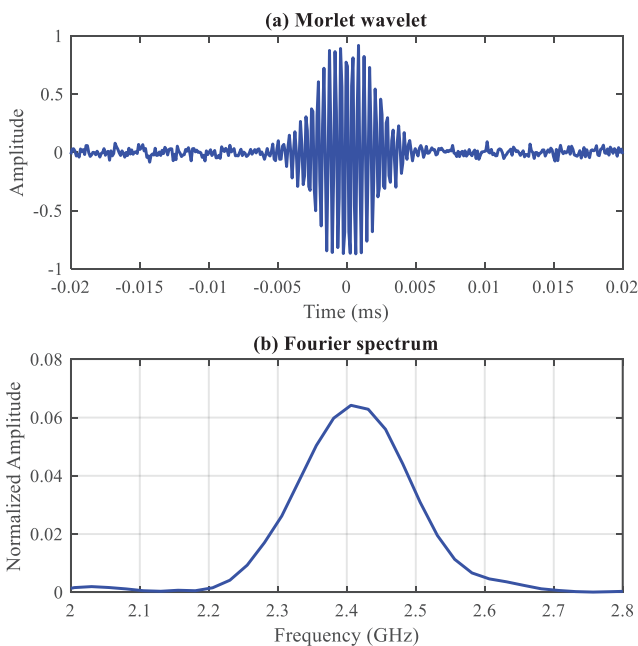


Fig. 2. Ultra-Wideband signal used in implementation.

This signal type and system geometry is well-organized for electromagnetic near-field conditions [25], [26], because in this region the source distance is comparable to or smaller than the wavelength, causing the wavefront to take a spherical form rather than planar propagation. Conventional methods, which assume planar wavefronts, may introduce errors in source estimation under such circumstances. In contrast, the proposed adaptive thresholding method does not depend on explicit array manifold models or steering vectors; instead, it relies solely on the statistical separation of signal- and noise-dominated eigenvalues. This separation continues to hold in near-field propagation, as the covariance matrix still reflects the number of independent sources regardless of the wavefront shape. Moreover, the integration of parameter optimization and MP-based correction stabilizes the threshold even when eigenvalue spreads are affected by near-field curvature or broadband effects. As a result, the method remains effective in indoor radar and short-range communication systems, where high-resolution source detection is required despite the complex wave interactions inherent to near-field geometries.

For each time step, the sorted eigenvalues of the covariance matrix are analysed. MDL and AIC compute likelihood-based penalties, while the adaptive threshold dynamically identifies significant eigenvalues above the noise floor. The estimated number of sources is derived by selecting the minimizing index for MDL and AIC, while adaptive thresholding counts eigenvalues exceeding the computed noise threshold [27], [28].

For each scenario in the validation set $\mathcal{D}$, the optimized parameter pair (α^*, β^*) obtained in Sec. 2.D was applied, ensuring consistent parameter settings across different SNR levels and array geometries.

IV. Results and Discussion

A. Source Enumeration under Noise Variation

The performance of MDL, AIC, and proposed adaptive thresholding is evaluated under different SNR levels and source configurations. Each method is tested multiple times in a loop to determine its ability to accurately estimate the number of sources in different randomized noise environments. The results show how noise affects their performance and highlight the advantages of adaptive thresholding in challenging conditions. Fig. 3 shows the results for noise environments by their SNR in the following.

Formulation of the success rate is determined in (10). In

Fig. 3 one can see the performance metrics of all three algorithms. Measurements are taken by 1000 independent runs for each SNR level between -15dB and 15dB. The result differences on each run are caused by the randomness of the noise figure. Algorithm's truthfulness is one of the key expectations of the evaluation.

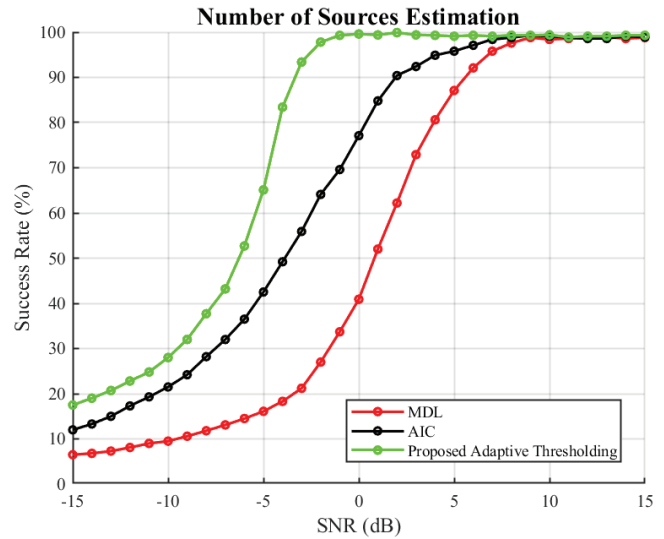


Fig. 3. performance evaluation of the algorithms (for 1000 runs)

The performance evaluation of the number of source estimation methods whose results are given in Fig. 3 which demonstrates the advantages of adaptive thresholding versus the traditional approaches. In low-SNR conditions, MDL and AIC fail to distinguish signal components from noise that results with low success rates.

The proposed adaptive thresholding method outperforms both, maintaining relatively higher accuracy by dynamically adjusting the noise floor. However, this improvement does not necessarily imply acceptable overall performance. As the SNR increases, AIC shows improvement, but adaptive thresholding reaches a better success rate faster. This indicates that the adaptive method is more effective in handling varying noise levels and closely spaced sources.

Table I. presents the summary of the achievements and comments for categorized SNR levels. At high SNR values, all methods eventually achieve nearly 100% accuracy, but the adaptive approach converges more quickly. MDL remains the slowest, often underestimating the number of sources in moderate-SNR conditions. AIC performs better than MDL but still lags behind the adaptive method in complex scenarios.

TABLE I
AVERAGE SUCCESS EVALUATION FOR CATEGORIZED SNR LEVELS

	SNR ranges					
	(-15, -10) dB	(-10, -5) dB	(-5, 0) dB	(0, 5) dB	(5, 10) dB	(10, 15) dB
MDL Avg (%)	7.5	11.9	23.2	61.7	94.2	98.6
AIC Avg (%)	15.3	28.4	56.2	87.9	97.8	98.8
Proposed Avg (%)	20.9	38.7	87.8	99.1	99.2	99.5
Observation	Success rates are low. MDL is the lowest, the proposed method performs better.	Success rates are increasing. MDL remains low, while AIC and the proposed method perform better.	Success rates rise rapidly. The proposed method shows the best performance.	High success levels. The proposed method is nearly perfect	All methods achieve very high success rates.	All methods show near-optimal performance in this range.

One advantage of the proposed method is that it does not directly depend on the geometric distribution of sources. Traditional subspace-based methods often require well-separated sources for accurate estimation, but adaptive thresholding can detect sources even in complex or irregular geometries. This makes it useful in environments where sources are randomly placed or closely spaced. Another strength is its ability to resolve neighbouring sources more effectively. Since the method adapts dynamically to the eigenvalue distribution, it can separate weak and strong sources without assuming a predefined structure.

B. Close-Source Discrimination Performance

In Fig. 4, an additional simulation experiment is conducted to comparatively evaluate the ability of the algorithms to resolve closely spaced active sources. Two separate sources are added to the same experimental scene, positioned at an average distance of 2 meters from the array centre. The distance between these two closely spaced sources varies between 0.05 and 0.25 meters. Each distance configuration is simulated 1000 times, and the following results are obtained regarding the success rates.

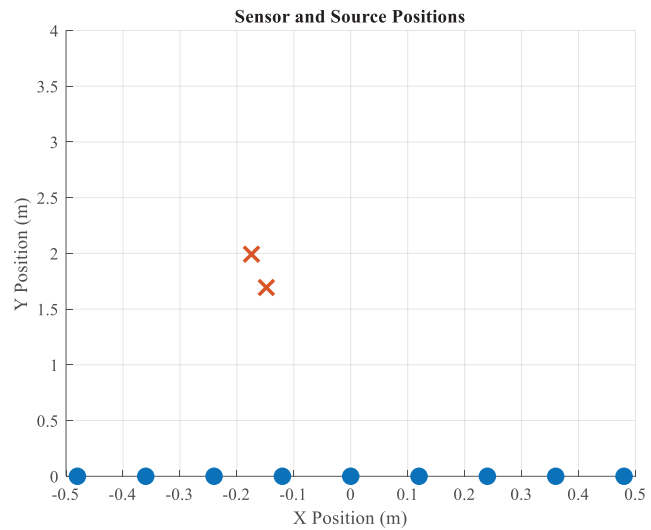


Fig. 4. Simulation scenario for detection of close sources.

Table II presents the success rates of the three methods in detecting closely spaced sources. The adaptive thresholding method significantly outperforms MDL and AIC, especially when the distance between sources is small. At 0.05 meters separation, the adaptive method achieves nearly 60% success, while MDL and AIC struggle below 30%. As the source distance increases, all methods improve, but the adaptive thresholding still provides the highest success rate. At 0.25 meters, all methods approach good accuracy, but the adaptive method converges to this level much faster.

When the distance between sources was very small (0.05–0.10 m), the optimized parameters (α^*, β^*) provided a clearer separation between signal and noise eigenvalues. These results include the optimized parameter setting described in Sec. 2.D, which ensures stable performance even in close-source scenarios.

The close-source results show that the performance differences between the methods are most evident at small source separations. At 0.05 m, the proposed method achieves a success rate of 58.9%, compared with 28.4% for AIC and 19.5% for MDL. At 0.10 m, the corresponding success rates increase to 75.3%, 43.2%, and 32.1%, respectively. These results indicate that the eigenvalue-based adaptive threshold provides a clearer distinction between signal- and noise-related components under the tested close-source conditions. The optimized parameter setting described in Section II-D is used throughout these experiments and contributes to the observed performance across the evaluated source spacings.

TABLE II
SUCCESS RATES FOR CLOSE SOURCE DISCRIMINATION

Distance Between Sources (m)	MDL Success Rate (%)	AIC Success Rate (%)	Adaptive Thresholding Success Rate (%)
0,05	19,5	28,4	58,9
0,10	32,1	43,2	75,3
0,15	50,8	71,7	84,2
0,20	86,4	88,9	97,6
0,25	96,3	98,2	99,3

As the source spacing increases, the performance difference between the methods gradually decreases. At 0.20 m, the proposed method reaches 97.6% success, while MDL and AIC reach 86.4% and 88.9%, respectively; at 0.25 m, all three methods exceed 96%. The proposed method determines the threshold from the ordered eigenvalue spectrum and uses the second discrete difference to characterize local spectral transitions. Since these operations are performed directly on the eigenvalues obtained from the covariance matrix, the procedure does not require additional eigenvector-based processing. This keeps the source-enumeration procedure computationally simple while preserving the performance advantage observed under the more challenging source-spacing conditions.

C. Statistical Confidence Analysis

To quantify the statistical uncertainty of the reported success rates, a confidence analysis was performed using the 1000 independent trials conducted for each source-spacing condition. Since each trial results in either correct or incorrect source-number estimation, the success rate is treated as a binomial proportion. For a given method and test condition, the estimated success probability is,

$$\hat{p} = \frac{N_{\text{correct}}}{N_{\text{runs}}}, \quad (13)$$

where (N_{correct}) denotes the number of correct source-number estimates and ($N_{\text{runs}} = 1000$). The corresponding standard error is

$$SE(\hat{p}) = \sqrt{\frac{\hat{p}(1 - \hat{p})}{N_{\text{runs}}}}. \quad (14)$$

Based on this formulation, the approximate 95% confidence interval is calculated as,

$$CI_{95\%} = \hat{p} \pm 1.96SE(\hat{p}). \quad (15)$$

Table III presents the resulting confidence intervals for the close-source experiment. The statistical analysis supports the differences observed in Table II. At the most challenging spacing of 0.05 m, the proposed method achieves a success rate of 58.9% with a 95% confidence interval of 55.9–61.9%, whereas AIC and MDL yield intervals of 25.6–31.2% and 17.0–22.0%, respectively. The confidence intervals of the proposed method and both conventional methods remain non-overlapping between 0.05 m and 0.20 m. This indicates that the observed performance differences within this range are substantially larger than the sampling uncertainty associated with the 1000 independent trials. At 0.25 m, the intervals become closer as all methods reach high source-enumeration success rates.

TABLE III
95% CONFIDENCE INTERVALS FOR CLOSE-SOURCE ENUMERATION SUCCESS RATES

Source Spacing (m)	MDL 95% CI (%)	AIC 95% CI (%)	Proposed 95% CI (%)
0.05	17.0–22.0	25.6–31.2	55.9–61.9
0.10	29.2–35.0	40.1–46.3	72.6–78.0
0.15	47.7–53.9	68.9–74.5	81.9–86.5
0.20	84.3–88.5	87.0–90.8	96.7–98.5
0.25	95.1–97.5	97.4–99.0	98.8–99.8

To quantify the statistical uncertainty of the reported success rates, 95% confidence intervals were calculated using the binomial proportion model, with ($N=1000$) independent trials for each source-spacing condition. The resulting intervals are presented in Table III. The separation between the methods is especially clear at short source distances. At 0.05 m, the proposed method achieves a 95% confidence interval of 55.9–61.9%, compared with 25.6–31.2% for AIC and 17.0–22.0% for MDL. The confidence intervals remain non-overlapping up to 0.20 m between the proposed method and both conventional methods, indicating that the observed advantage is substantially larger than the Monte Carlo sampling uncertainty. At 0.25 m, the intervals become closer as all three methods approach their high-SNR/high-separation performance region.

D. Practical Considerations and Limitations

The ability to resolve closely spaced sources is influenced not only by the estimation method but also by various physical and signal processing parameters. Line-of-Sight (LOS) conditions play a crucial role, as multipath interference and non-line-of-sight (NLOS) effects can alter the eigenvalue structure, affecting source separation. Operating frequency and bandwidth also impact resolution, where higher frequencies and wider bandwidths generally improve spatial resolution but may introduce additional noise sensitivity.

The present evaluation assumes spatially white Gaussian noise and non-coherent sources. Under this assumption, the noise covariance satisfies $R_n = \sigma_n^2 I$, which supports the formation of a compact noise-eigenvalue region used by the proposed threshold. For colored noise, where $R_n \neq \sigma_n^2 I$, the noise eigenvalues may exhibit a wider and non-uniform distribution, and the present MP-inspired correction may therefore require modification. Similarly, strong source correlation or coherence can reduce the effective rank of the signal covariance and decrease the separation between signal- and noise-related eigenvalues. These cases are outside the present simulation scope and constitute important directions for further evaluation, including covariance prewhitening or spatial-smoothing strategies.

Despite the many advantages of adaptive thresholding, there are challenges when implementing it in real-world scenarios. The approach depends on assumptions about the behaviour of the eigenvalues, which may require modifications for different contexts. While adaptive

thresholding is generally effective, sudden changes in noise can affect its accuracy. Subsequent studies could improve upon this methodology by refining the threshold calculation to accommodate a broader range of scenarios.

The proposed method's ability to function in both near-field and wideband scenarios is another of its key strengths. The eigenvalue-based thresholding strategy is different from traditional space-based approaches, which rely on planar wavefront assumptions. It does not require explicit manifold modelling. Even in situations involving spherical wavefronts, such as in short-range UWB conditions, the method maintains its capacity to effectively differentiate between signal and noise eigenvalues. The broadening of the applicability of the approach to indoor radar and short-range communication systems, where near-field effects are dominant, is a feature of this.

V. Conclusion

This study presents an optimized adaptive thresholding method for estimating the number of active signal sources and compares its performance with MDL and AIC. The results show that the proposed method achieves higher source-enumeration success rates under low-SNR conditions by dynamically adjusting the noise threshold. It also reaches high success rates at lower SNR levels and performs consistently across the tested noise conditions. The method requires only eigenvalue-based operations after covariance estimation, which keeps the computational structure simple. In addition, the threshold parameters are optimized before testing and then fixed for the reported simulations, allowing the method to adapt to different SNR and source-spacing conditions without adding substantial computational cost. Overall, the proposed method provides an effective alternative to MDL and AIC for source enumeration in radar, wireless communication, and sensor-array applications.

References

- [1] D. G. . Manolakis, V. K. . Ingle, and S. M. . Kogon, "Statistical and adaptive signal processing: spectral estimation, signal modeling, adaptive filtering, and array processing," p. 796, 2005.
- [2] D. H. D. H. Johnson and D. E. Dudgeon, "Array signal processing: concepts and techniques," *Book*, p. 533, 1993, Accessed: Mar. 16, 2025. [Online]. Available: <https://search.worldcat.org/title/25546138>
- [3] J. Bin, L. An'Nan, and X. Jie, "An improved signal number estimation method based on information theoretic criteria in array processing," *2019 IEEE 11th International*

- Conference on Communication Software and Networks, ICCSN 2019, pp. 193–197, Jun. 2019, doi: 10.1109/ICCSN.2019.8905347.
- [4] Jacob. Benesty, Israel. Cohen, and J. . Chen, “Fundamentals of signal enhancement and array signal processing,” 2019.
- [5] J. M. Kahn and J. R. Barry, “Wireless infrared communications,” *Proceedings of the IEEE*, vol. 85, no. 2, pp. 265–298, 1997, doi: 10.1109/5.554222.
- [6] R. F. Cinar, A. Demirkol, and Z. Demir, “An ambiguity function aided target density function for radar imaging,” *Optoelectronics and Advanced Materials, Rapid Communications*, vol. 9, no. 5–6, pp. 850–854, 2015.
- [7] R. F. Cinar and A. Demirkol, “An active sensor imaging for dense target environments,” *Optoelectronics and Advanced Materials, Rapid Communications*, vol. 14, no. 11–12, pp. 501–508, Nov. 2020.
- [8] J. Rissanen, “Modeling by shortest data description,” *Automatica*, vol. 14, no. 5, pp. 465–471, Sep. 1978, doi: 10.1016/0005-1098(78)90005-5.
- [9] W. Cheng, Z. Zhang, H. Cao, Z. He, and G. Zhu, “A Comparative Study of Information-Based Source Number Estimation Methods and Experimental Validations on Mechanical Systems,” *Sensors 2014, Vol. 14, Pages 7625–7646*, vol. 14, no. 5, pp. 7625–7646, Apr. 2014, doi: 10.3390/S140507625.
- [10] Y. Lee *et al.*, “Source Enumeration Approaches Using Eigenvalue Gaps and Machine Learning Based Threshold for Direction-of-Arrival Estimation,” *Applied Sciences 2021, Vol. 11, Page 1942*, vol. 11, no. 4, p. 1942, Feb. 2021, doi: 10.3390/APP11041942.
- [11] A. Hyvärinen and E. Oja, “Independent component analysis: algorithms and applications,” *Neural Networks*, vol. 13, no. 4–5, pp. 411–430, Jun. 2000, doi: 10.1016/S0893-6080(00)00026-5.
- [12] L. Huang, T. Long, E. Mao, and H. C. So, “MMSE-based MDL method for robust estimation of number of sources without eigendecomposition,” *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 4135–4142, 2009, doi: 10.1109/TSP.2009.2024043.
- [13] A. Wiesel and A. O. Hero, “Decomposable principal component analysis,” *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4369–4377, 2009, doi: 10.1109/TSP.2009.2025806.
- [14] W. S. B. Ouedraogo, A. Souloumiac, M. Jaïdane, and C. Jutten, “Non-negative blind source separation algorithm based on minimum aperture simplicial cone,” *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 376–389, Jan. 2014, doi: 10.1109/TSP.2013.2287683.
- [15] K. Han and A. Nehorai, “Improved source number detection and direction estimation with nested arrays and ULAs using jackknifing,” *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 6118–6128, 2013, doi: 10.1109/TSP.2013.2283462.
- [16] T. Salman, A. Badawy, T. M. Elfouly, A. Mohamed, and T. Khattab, “Estimating the number of sources: An efficient maximization approach,” in *2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*, IEEE, Aug. 2015, pp. 199–204. doi: 10.1109/IWCMC.2015.7289082.
- [17] H. Akaike, “Information Theory and an Extension of the Maximum Likelihood Principle,” in *Biogeochemistry*, vol. 1998, Springer Science and Business Media Deutschland GmbH, 1998, pp. 199–213. doi: 10.1007/978-1-4612-1694-0_15.
- [18] H. Akaike, “A New Look at the Statistical Model Identification,” *IEEE Trans. Automat. Contr.*, vol. 19, no. 6, pp. 716–723, 1974, doi: 10.1109/TAC.1974.1100705.
- [19] M. Wax and T. Kailath, “Detection of signals by information theoretic criteria,” *IEEE Trans. Acoust.*, vol. 33, no. 2, pp. 387–392, 1985, doi: 10.1109/TASSP.1985.1164557.
- [20] S. Kritchman and B. Nadler, “Determining the number of components in a factor model from limited noisy data,” *Chemometrics and Intelligent Laboratory Systems*, vol. 94, no. 1, pp. 19–32, Nov. 2008, doi: 10.1016/J.CHEMOLAB.2008.06.002.
- [21] K. P. . Burnham and D. Raymond. Anderson, “Model selection and multi-model inference: a practical information-theoretic approach,” 2011.
- [22] Z. Lu, “Source Enumeration in Sensor Array Processing: A Model Order Selection Problem,” Nov. 2012.
- [23] P. Liashchynskiy and P. Liashchynskiy, “Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS,” Dec. 2019, Accessed: Sep. 09, 2025. [Online]. Available: <https://arxiv.org/pdf/1912.06059>
- [24] S. Singer, J. N.- Scholarpedia, and undefined 2009, “Nelder-mead algorithm,” *var.scholarpedia.org*, Accessed: Sep. 09, 2025. [Online]. Available: http://var.scholarpedia.org/article/Nelder-Mead_algorithm
- [25] J. C. Chen, R. E. Hudson, and Kung Yao, “Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field,” *IEEE Transactions on Signal Processing*, vol. 50, no. 8, pp. 1843–1854, Aug. 2002, doi: 10.1109/TSP.2002.800420.
- [26] R. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986, doi: 10.1109/TAP.1986.1143830.
- [27] M. Shakeri, K. R. Pattipati, and D. L. Kleinman, “Optimal measurement scheduling for state estimation,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 2, pp. 716–729, Apr. 1995, doi: 10.1109/7.381919.
- [28] Q. Liang, W. Wang, X. Liu, Z. Na, M. Jia, and B. Zhang, Eds., *Communications, Signal Processing, and Systems*, vol. 571. Singapore: Springer Singapore, 2020. doi: 10.1007/978-981-13-9409-6.

Özgeçmişler



Gökhan Yüksek is currently working as an Associate Professor in the Department of Electrical and Electronics Engineering at Batman University, Turkey. He received his B.Sc. in Electrical and Electronics Engineering from Nigde University, Nigde, and his M.Sc. and Ph.D. degrees in Electrical and Electronics Engineering from Mersin University, Mersin, Turkey. His research interests include metaheuristic optimization algorithms, intelligent and adaptive control systems, renewable energy, battery management, and hybrid energy storage technologies. He has authored and co-authored several research papers in international journals and conferences indexed in SCI/SCI-Expanded. He also serves as a reviewer for various leading journals in the fields of control engineering, optimization, and energy systems.



Timur LALE received his B.Sc. degree in Electrical and Electronics Engineering from Niğde University in 2012 and his M.Sc. and Ph.D. degrees in Electrical and Electronics Engineering from Dicle University in 2018 and 2024, respectively. His doctoral research focused on developing methods for the detection of demagnetization and stator inter-turn short-circuit faults in permanent magnet synchronous motors. He is currently an Assistant Professor in the Department of Electrical and Electronics Engineering, Division of Electrical Machines, at Batman University, Türkiye. His research interests include fault diagnosis and condition monitoring of electrical machines, artificial intelligence and deep learning methods, metaheuristic optimization, forecasting applications in electrical energy systems, electric vehicles, and battery management systems. He has published several research articles in national and international peer-reviewed journals in these fields.