*Progress In Electromagnetics Research C, Vol. 107, 65–79, 2021*

# Improved Enumeration of Scatterers Using Multifrequency Data Fusion in MDL for Microwave Imaging Applications

**Roohallah Fazli[1, *], Hadi Owlia[1], and Majid Pourahmadi[2]**

**Abstract**—This paper presents a modified version of minimum description length (MDL) method, referred as multifrequency MDL (FMDL), for scatterers enumeration before using the multiple signal classification (MUSIC) algorithm in microwave imaging applications. The inclusion of data from multiple frequencies should make an attempt to reduce the error in number estimation due to noise and multiple scattering. Data fusion in multiple frequencies is performed based on two schemes called averaging and maximization rules. The solution for MDL criterion which is a minimum for one frequency is not likely to be the solution for other frequencies, so by averaging the MDL criterion over the total frequencies or by maximization of the solution for each frequency, we can achieve the correct source number. Furthermore, a whitening step before applying FMDL method is employed to compensate the aspect limitations of the measured data due to the limited number of antennas. The superiority of the proposed FMDL approach with respect to the other competing methods is confirmed by both the numerical examples and the Institut Fresnel experimental dataset. The results indicate that the FMDL yields more accurate estimate of the targets number specially for the cases of low SNR values and very closely spaced scatterers.

## 1. INTRODUCTION

Microwave imaging (MI) for biomedical applications, such as breast cancer detection, is a continually emerging field of research, and it is a promising method for a low-cost, compact, safe, and real-time medical imaging. The considerable contrast in the dielectric properties between healthy and malignant tissues at microwave frequencies and the non-ionizing radiation are two main interesting factors for the development of the microwave tomography systems over the last decades. Probably, the main challenge to make microwave tomography a competitive medical imaging modality is its lower resolution than magnetic resonance imaging (MRI) and x-ray computed tomography (CT) [1]. However, achieving high resolution fast imaging of the scatterers (targets) remains a challenge in the area of electromagnetic inverse scattering.

In microwave tomography, we attempt to solve an inverse scattering problem, i.e., finding the dielectric profile of an object from measurements of the field scattered by the object. In general, there are two main classes of methods for solving such inverse scattering problems. First, nonlinear methods such as Gauss-Newton (GN) approach, which can perform a quantitative reconstruction of the scatterers, but they suffer from intensive computations and inaccurate initial guess of the object profile. Linear methods, such as MUSIC algorithm, are the second group of algorithms which cannot obtain useful information about the dielectric properties of the scatterers, but they make it possible to acquire the location of the scatterers in a few seconds. So, it seems that the linear techniques are promising for real time inverse scattering as can be seen in references [2–6].

*Received 18 September 2020, Accepted 24 November 2020, Scheduled 30 November 2020*
* Corresponding author: Roohallah Fazli (rofazli@ardakan.ac.ir).
[1] Department of Electrical Engineering, Faculty of Engineering, Ardakan University, P. O. Box 184, Ardakan, Iran. [2] Department of Electrical Engineering, Yazd Branch Islamic Azad University, Yazd, Iran.

The MUSIC algorithm is a well-known technique that has been shown to be effective for the localization of electromagnetic small inclusions in microwave imaging applications [7]. The main idea of the MUSIC algorithm estimating the location of sources is to employ the orthogonality between signal and noise subspaces and to search the minimum projections of the array Green's functions onto a noise subspace [8]. The common way to find the noise subspace is to enumerate the signal and noise eignvalues via singular value decomposition (SVD) of the so-called multistatic response matrix (MSR). The significant eigenvalues map to the signal subspace (targets) and the noise subspace determined by the smallest eigenvalues. However, due to the measurement errors affecting the MSR matrix, there is often no enough gap in an eigenvalue map to recognize them, and one may be incorrectly estimate the number of targets. Hence, the performance of the MUSIC algorithm will be significantly deteriorated with inaccurate enumeration of targets. So, the method for estimating the number of scatterers has a vital role for the accurate MUSIC imaging.

The problem of scatterers enumeration before MUSIC imaging can be considered as source number estimation studied a lot in signal processing. Many algorithms have been presented and developed for source number estimation such as the hypothesis testing approach [9] and information theoretic criteria based methods. In [9], a hypothesis testing method has been proposed, in which an artificial threshold value must be assigned based on the confidence interval of the noise eigenvalue. However, the threshold needs to be tuned, and the method performance is drastically affected by a low number of samples. The information theoretic based methods such as the minimum description length (MDL) and Akaike information criterion (AIC) are two mostly used methods that apply selection criteria to detect differences between eigenvalues of the signal and noise subspaces. These methods suffer from low signal to noise ratio and small number of snapshots. The other two methods for source enumeration based on the eigenvalue gap measure are the second order statistic of eigenvalues (SORTE) [10] and the ratio of adjacent eigenvalues (RAE) [11]. These methods have the advantages of simple implementation and low computational cost compared to the MDL and AIC methods, but the performances are considerably degraded in the cases of using low sensors number and for low SNR cases.

Although the methods presented for source number estimation have been extensively used in the field of direction of arrival (DOA) estimation for various applications such as radar, sonar, communication, and speech processing, the utilization of these methods in microwave imaging is much less studied. In [12], Pourahmadi et al. demonstrated that the mathematical model behind the scattering from the small scatterers was well compatible with the MDL model, and they employed the MDL to estimate the number of scatterers for microwave imaging applications. We have also previously proposed [13] a method that utilized an analytical methodology to estimate the number and position of 2D small targets. We show that the method can extend the ability of the MUSIC algorithm to localize the small targets in the case of noisy data and when the targets are closely located.

In all works mentioned above, the scattering data matrices measured at a fixed frequency are used for scatterers location and number estimation. So, small variations in the measured data can lead to large errors in reconstruction and cause to wrong target enumeration. Thus, the motivation for the use of multifrequency data is to achieve increased stability for the problem at hand. On the other hand, the multifrequency data are an attempt to compensate for the aspect limitations of the measured data. Furthermore, it is clear that to improve the resolution, i.e., the detection of very closely located targets, we must use an incident field with a shorter wavelength or a higher frequency to illuminate the scatterer. However, we are limited to use very high frequencies because the inverse problem becomes very oscillatory and makes many more local minima, and the accuracy of source number estimation will be affected.

The present work introduces an improved version of MDL, frequency based MDL (FMDL), applying the received signals from independent transmitters at different frequencies. In the proposed scheme, the information at several frequencies is employed in MDL to reduce the error in the estimation. The use of MDL and FMDL in MUSIC algorithm for number detection shows that the FMDL yields more accurate determination of the target number, even in the presence of strong noise and multiple scatterings. It is illustrated that in the case of real data, the noise may behave like non-white Gaussian noise, and the FMDL produces non-consistent estimates as the SNR increases from low to high values. To remedy this problem, we propose a preprocessing method for whitening the eigenvalues of the noise. The resulting algorithm, called whitened FMDL, is consistent for the total range of SNR values.

The remaining of the paper is organized as follows. The statement and formulation of the scattering from small targets are introduced in Section 2. This section also reviews the theory of subspace signal processing and the development of the MUSIC algorithm for imaging. Section 3 introduces the MDL and FMDL and discusses their applications to the problem of estimating the number of scatterers. In this section, we present the modified version of MDL (FMDL) for target enumeration by combination of multifrequency data. Simulation results, which illustrate the performance of the proposed algorithm for both synthetic and real experimental data, are described in Section 4.

## 2. FORMULATION OF THE MUSIC ALGORITHM

The MUSIC imaging algorithm employs the multistatic response matrix to detect and localize the targets. First, we describe the mathematical model for the generation of this matrix and, then, explain the MUSIC algorithm.

### 2.1. Multistatic Response Matrix

Consider the geometry of the imaging problem given in Fig. 1, where the imaging domain, containing a set of $M$ small scatterers with the unknown locations $\mathbf{X}_m$ and the unknown scattering strengths $\tau_m$, $m = 1, 2, \ldots, M$, is successively irradiated by $N_t$ transmitters, and the scattered electric fields are measured by $N_r$ receivers. The antenna array is assumed to be circular and coincident, in which each sensor is commonly used for transmitter and receiver at single frequency simultaneously, similar to the prototype MI systems built for breast cancer detection. The homogeneous and non-magnetic background medium is considered, and the imaging is performed under the 2D TM incident. The aim of imaging is to get the number, location, and if possible the shape and dielectric properties of the scatterers, using the total measured scattered field at receivers. Under the assumptions of small transceivers and small targets and also by neglecting all multiple scattering between the targets (Born approximation), if the $j$th antenna is excited by an input voltage $e_j$, then the generated field at position $\mathbf{r}= (x, y)$ in the frequency domain equals

$$\psi_j^{inc}(\mathbf{r}, \omega) = G_0 \left(\mathbf{r}, \mathbf{R}_j^t, \omega\right) e_j(\omega) \tag{1}$$

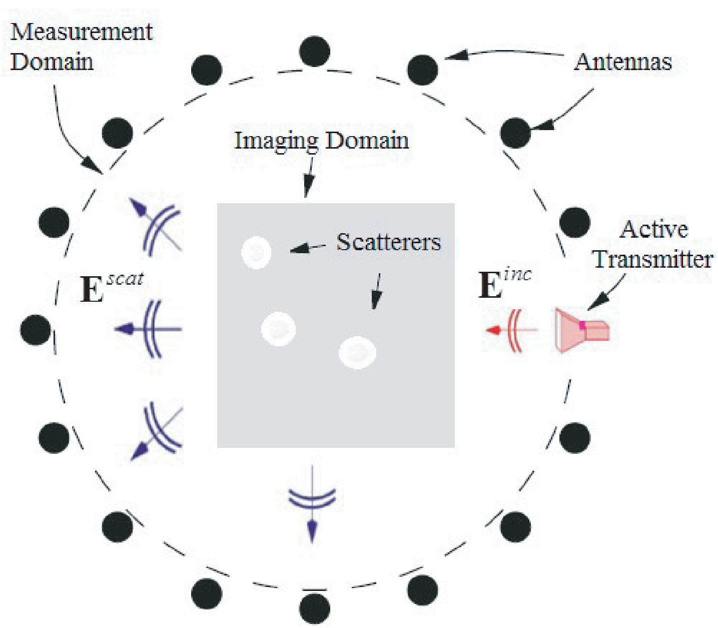



**Figure 1.** General setup for a 2D microwave imaging system.

where $\omega$ is the frequency; $G_0\ (\mathbf{r},\ \mathbf{R}_j^t,\ \omega)$ is the Green function corresponding to the background medium in which the targets are embedded; and $\mathbf{R}_j^t$ is the location of the $j$th transmitter antenna. The above Green's function for homogeneous medium is $G_0(\mathbf{r},\ \mathbf{R}_j^t) = i/4 * H_0(k_0|\mathbf{r} - \mathbf{R}_j^t|)$ in which $H_0$ is the zero order Hankel function of the first kind, and $k_0 = 2\pi/\lambda$ is the free space wavenumber. If the field $\psi_j^{inc}(\mathbf{r}, \omega)$ is incident on the $m$th scatterer, it generates the scattered field at point $\mathbf{r}$ as

$$\psi_j^{scat}(\mathbf{r}, \omega) = G_0(\mathbf{r}, \mathbf{X}_m, \omega)\tau_m(\omega)\psi_j^{inc}(\mathbf{X}_m, \omega) \tag{2}$$

where $\tau_m$ is the strength of the $m$th scatterer. Thus, the total scattered field due to the presence of $M$ scatterers, when $N_t$ transmitter antennas are simultaneously excited using the voltages $e_j(\omega)$, becomes

$$\psi^{scat}(\mathbf{r}, \omega) = \sum_{j=1}^{N_t}\sum_{m=1}^{M} G_0\left(\mathbf{r}, \mathbf{X}_m, \omega\right)\tau_m(\omega)G_0\left(\mathbf{X}_m, \mathbf{R}_j^t, \omega\right)e_j(\omega) \tag{3}$$

The output voltage $V_i\ (\omega)$ at the $i$th receiver antenna is assumed to be equal to the amplitude of the scattered field measured at the $i$th antenna as

$$V_i(\omega) = \sum_{j=1}^{N_t}\psi_j^{scat}\left(\mathbf{R}_i^r, \omega\right) = \sum_{j=1}^{N_t}\sum_{m=1}^{M} G_0\left(\mathbf{R}_i^r, \mathbf{X}_m, \omega\right)\tau_m(\omega)G_0\left(\mathbf{X}_m, \mathbf{R}_j^t, \omega\right)e_j(\omega) \tag{4}$$

in which $\mathbf{R}_i^r$ is the location of the $i$th receiver antenna. Using the matrix notation, Eq. (4) can be rewritten as

$$V_i(\omega) = \sum_{j=1}^{N_t} K_i(\omega)e_j(\omega) = \mathbf{K}(\omega)\mathbf{e}(\omega) \tag{5}$$

where $\mathbf{e}(\omega) = [e_1(\omega),\ e_2(\omega), \ldots, e_{N_t}(\omega)]^T$ is an $N_t \times 1$ column vector formed from the set of input voltages applied at the antenna terminals, and $\mathbf{K}(\omega)$ is a $N_r \times N_t$ matrix which is known as multistatic response matrix (MSR). So, we can obtain

$$\mathbf{K}(\omega) = \sum_{m=1}^{M} G_0\left(\mathbf{R}_i^r, \mathbf{X}_m, \omega\right)\tau_m(\omega)G_0\left(\mathbf{X}_m, \mathbf{R}_j^t, \omega\right) = \sum_{m=1}^{M}\tau_m(\omega)\mathbf{g}_{0,r}\left(\mathbf{X}_m, \omega\right)\mathbf{g}_{0,t}^T\left(\mathbf{X}_m, \omega\right) \tag{6}$$

where $T$ stands for the transposition, and $\mathbf{g}_{0,r}(\mathbf{X}_m, \omega)$ and $\mathbf{g}_{0,t}(\mathbf{X}_m, \omega)$, respectively, are the receiving and transmitting background Green's function vectors at the target location $\mathbf{X}_m$ that can be written as

$$\mathbf{g}_{0,r}(\mathbf{X}_m, \omega) \equiv \left[G_0\left(\mathbf{R}_1^r, \mathbf{X}_m, \omega\right), G_0\left(\mathbf{R}_2^r, \mathbf{X}_m, \omega\right), \ldots, G_0\left(\mathbf{R}_{N_r}^r, \mathbf{X}_m, \omega\right)\right]^T \tag{7}$$

$$\mathbf{g}_{0,t}\left(\mathbf{X}_m, \omega\right) \equiv \left[G_0\left(\mathbf{X}_m, \mathbf{R}_1^t, \omega\right), G_0\left(\mathbf{X}_m, \mathbf{R}_2^t, \omega\right), \ldots, G_0\left(\mathbf{X}_m, \mathbf{R}_{N_t}^t, \omega\right)\right]^T \tag{8}$$

Eq. (6) is valid when the targets are well separated. Therefore, for the cases of closely located targets, there is the strong multiple scattering between the targets, and we cannot apply the Born model anymore. So, we should use the Foly-Lax model [17] to get the MSR matrix as

$$\mathbf{K}(\omega) = \sum_{m=1}^{M}\sum_{m'=1}^{M}\mathbf{A}_{m,m'}(\omega)\mathbf{g}_{0,r}(\mathbf{X}_m, \omega)\mathbf{g}_{0,t}^T(\mathbf{X}_{m'}, \omega) \tag{9}$$

where the generalized multiple scattering amplitudes are $\mathbf{A}_{m,m'}(\omega) = \tau_m\mathbf{H}_{m,m'}^{-1}(\omega)$. The $M \times M$ matrix $\mathbf{H}$ is defined by

$$\mathbf{H}_{m,m'}(\omega) = \delta_{m,m'} - (1 - \delta_{m,m'})\tau_{m'}(\omega)G_0\left(\mathbf{X}_m, \mathbf{X}_{m'}, \omega\right) \tag{10}$$

in which $\delta_{m,m'}$ denotes the Kronecker delta function. It is possible to define the correlation matrix, called time reversal matrix as $\mathbf{T}(\omega) = \mathbf{K}(\omega)\mathbf{K}^{\dagger}(\omega)$ in which $\dagger$ denotes the Hermitian. So we can conclude that the eigenvalues of $\mathbf{T}(\omega)$ are the positive squared of the singular values of $\mathbf{K}(\omega)$. Thus, the eigenvalues of matrix $\mathbf{K}(\omega)$ and the singular values of matrix $\mathbf{T}(\omega)$ can be used interchangeably for scatterers enumeration. The MUSIC algorithm can now be used as follows to determine the location of scatterers.

### 2.2. Music Imaging

The MUSIC exploits the matrix $\mathbf{T}(\omega)$ in order to image the small targets using the SVD of it. The SVD of $\mathbf{T}(\omega)$ can be written by

$$\mathbf{T}(\omega)\mathbf{v}_p(\omega) = \lambda_p(\omega)\mathbf{u}_p(\omega), \quad \mathbf{T}^\dagger(\omega)\mathbf{u}_p(\omega) = \lambda_p(\omega)\mathbf{v}_p(\omega) \tag{11}$$

where $\lambda_p(\omega) \geq 0$ are the singular values, and $\mathbf{u}_p(\omega)$ and $\mathbf{v}_p(\omega)$ are the left and right singular vectors, respectively. Using the orthogonality of signal and noise subspaces, considering the noise subspace spanned by the eigenvectors of $\mathbf{T}(\omega)$ having zero eigenvalues and spanning the signal subspace by the Green's vectors at target locations, it is possible to form a pseudospectrum through the following equation

$$P_{r,t}(\mathbf{X},\omega) = \frac{1}{\sum\limits_{p=M+1}^{\min(N_t,N_r)} \left|\mathbf{u}_p^\dagger(\omega)\mathbf{g}_{0,r}(\mathbf{X},\omega)\right|^2 + \sum\limits_{p=M+1}^{\min(N_t,N_r)} \left|\mathbf{v}_p^\dagger(\omega)\mathbf{g}_{0,t}^*(\mathbf{X},\omega)\right|^2} \tag{12}$$

where $*$ denotes the conjugation. The pseudospectrum in Eq. (12) will create a theoretically infinite peak (without noise) at each target location $\mathbf{X}_m$, $m = 1, 2, \ldots, M$ in a deterministic way when the number of targets ($M$) is *a priori* known. In low noise conditions, we can detect the number of targets from the number of significant eigenvalues of matrix $\mathbf{T}(\omega)$. However, the presence of noise and multiple scattering between the targets can make it difficult to determine the correct number of targets from the magnitudes of the eigenvalues of matrix $\mathbf{T}(\omega)$, and hence, the performance of the MUSIC algorithm for imaging is strongly degraded. So, several methods have been proposed to resolve this problem.

## 3. WELL KNOWN SCATTERERS ENUMERATION METHODS

In MI applications, by assuming the scatterers as sources and different illuminations of transmitters taken to be the multiple snapshots, the methods utilized for source number estimation in signal processing can be employed for the problem of scatterers enumeration. So, using Eqs. (6) and (9), we can consider the following equation for each column of matrix $\mathbf{K}$ as

$$\mathbf{K}_i(\omega) = \mathbf{A}\mathbf{S}_i(\omega) + \mathbf{V}_i(\omega) \tag{13}$$

where $\mathbf{S}_i(\omega) = [\tau_1 G_0(\mathbf{X}_1,\, \mathbf{R}_i^t\omega), \ldots, \tau_M G_0(\mathbf{X}_M,\, \mathbf{R}_i^t\omega)]^T$ considered as sources and the matrix $\mathbf{A}(\omega) = [\mathbf{a}_1(\omega), \ldots, \mathbf{a}_M(\omega)]$ in which $\mathbf{a}_i(\omega) = [G_0(\mathbf{X}_i,\, \mathbf{R}_1^r\omega), \ldots, G_0(\mathbf{X}_i,\, \mathbf{R}_N^r\omega)]^T$ are the steering vector. Also, $\mathbf{v}_i(\omega)$ is a column vector which models the measurement noise added to the scattered field measured at sensors outputs. Therefore, for the $i$th transmitter radiation, we assume $M$ sources, which are defined in vector $\mathbf{s}_i(\omega)$ present at the location of $M$ scatterers. Then, we must estimate the location and number of these sources (scatterers). However, it must be kept in mind that the algorithms should be chosen for enumeration to be effective for the cases where the number of snapshots equals the number of sensors. Here, some of the effective methods for scatterers enumeration are introduced.

### 3.1. RAE Method

Using SVD of matrix $\mathbf{T}(\omega)$ and sorting the resulted eigenvalues in descending order, we can obtain the eigenvalues $\lambda_1(\omega) \geq \ldots \geq \lambda_M(\omega) \geq \lambda_{M+1}(\omega) \geq \ldots \geq \lambda_{N_t}(\omega) \geq 0$, which contain $M$ larger eigenvalues $\lambda_1(\omega) \geq \ldots \geq \lambda_M(\omega)$ associated with the scatterers, and the remaining $N_t - M$ eigenvalues are, theoretically, equal, i.e., $\lambda_{M+1}(\omega) = \ldots = \lambda_{N_t}(\omega) = \sigma^2$, where $\sigma^2$ is the variance of the noise. Then, we calculate the ratio between the adjacent singular values as

$$RAE(m) = \frac{\lambda_m(\omega)}{\lambda_{m+1}(\omega)} \tag{14}$$

where $m = 1, \ldots, N_t - 1$. The method chooses $m$ as targets number for which the criterion in Eq. (14) is maximized.

### 3.2. SORTE Method

The method can identify the number of targets by searching the gap between $\lambda_M(\omega)$ and $\lambda_{M+1}(\omega)$ using a gap measure called SORTE that is defined as follows

$$SORTE(m) = \begin{cases} \dfrac{var\left[\{\nabla\lambda_i(\omega)\}_{i=m+1}^{N_t-1}\right]}{var\left[\{\nabla\lambda_i(\omega)\}_{i=m}^{N_t-1}\right]}, & \text{if} \quad var\left[\{\nabla\lambda_i(\omega)\}_{i=m}^{N_t-1}\right] \neq 0 \\ +\infty & \text{if} \quad var\left[\{\nabla\lambda_i(\omega)\}_{i=m}^{N_t-1}\right] = 0 \end{cases} \tag{15}$$

where $m = 1, \dots, N_t - 1$ and the sample variance of the sequence $\{\nabla\lambda_i(\omega)\}_{i=m}^{N_t-1}$ is

$$var\left[\{\nabla\lambda_i(\omega)\}_{i=m}^{N_t-1}\right] = \frac{1}{N_t - m}\sum_{i=m}^{N_t-1}\left(\nabla\lambda_i(\omega) - \frac{1}{N_t - m}\sum_{i=m}^{N_t-1}\nabla\lambda_i(\omega)\right)^2 \tag{16}$$

in which $(\omega)(\omega)(\omega)\nabla\lambda_i = \lambda_i - \lambda_{i+1}$, $i = 1, \dots, N_t - 1$. Then, we can determine the number of scatterers using the following criterion: SORTE $(m)$ is minimal.

### 3.3. MDL Method

The MDL method is an information-based approach consisting of minimizing a criterion over the number of scatterers that are detectable. By assuming that the number of snapshots equals the number of transmitters (antennas), the criterion for MDL is given by [14, 15]

$$\begin{aligned} MDL(m) \;=\; & N_t \ln\left(\left(\frac{1}{N_t - m}\sum_{i=m+1}^{N_t}\lambda_i(\omega)\right)^{N_t-m} \Big/ \prod_{i=m+1}^{N_t}\lambda_i(\omega)\right) \\ & + \frac{1}{2}m(2N_t - m)\ln(N_t), \quad m = 0, \dots, N_t - 1 \end{aligned} \tag{17}$$

The first term in Eq. (17) shows the ML (maximum likelihood) criterion, while the second one is a penalty function based on the number of free parameters in the model [15].

In all of the methods described above, the single frequency data have been used for number estimation. The single frequency source enumeration approach suffers from two main problems. First, when multiple scatterers are present, the detection performance quickly degrades. Secondly, when the number of transmitters (snapshots) is decreased, artifacts may rise and prevent the scatterer detection. We now introduce a multifrequency MDL method for number of scatterers estimation in order to enhance the detection performance.

## 4. MULTIFREQUENCY MDL METHOD

Consider that the matrix $\mathbf{T}(\omega)$, defined in Eq. (4), is given in $Q$ frequencies, i.e.,

$$\mathbf{T}(\omega_j) = \mathbf{K}(\omega_j)\mathbf{K}^\dagger(\omega_j), \quad j = 1, 2, \dots, Q \tag{18}$$

By performing SVD on the matrix $\mathbf{T}(\omega_j)$, we can obtain the vector $\lambda(\omega_j) = [\lambda_1(\omega_j), \lambda_2(\omega_j), \dots, \lambda_{Nt}(\omega_j)]$ containing the singular values of matrix $\mathbf{T}(\omega_j)$ at frequency $\omega_j$. Then, by embedding $\lambda(\omega_j)$ into Eq. (17), the MDL criteria is

$$\begin{aligned} MDL(m, \omega_j) \;=\; & N_t \ln\left(\left(\frac{1}{N_t - m}\sum_{i=m+1}^{N_t}\lambda_i(\omega_j)\right)^{N_t-m} \Big/ \prod_{i=m+1}^{N_t}\lambda_i(\omega_j)\right) \\ & + \frac{1}{2}m(2N_t - m)\ln(N_t), \quad m = 0, \dots, N_t - 1 \end{aligned} \tag{19}$$

### 4.1. Averaging Method

One method for the utilization of multifrequency data in MDL is the averaging of the single frequency MDL criterion over the total range of frequencies. So, assuming that the coefficients $\lambda(\omega_j)$ corresponding to different frequencies are statistically independent and employing the data from multiple frequencies, say, $\omega_1$, $\omega_2, \ldots, \omega_Q$, the MDL criterion for detecting the number of scatterers can be given by the sum of Eq. (19) over the frequency range of interest as

$$\sum_{j=1}^{Q} MDL(m,\omega_j) = \sum_{j=1}^{Q} \left( N_t \ln \left( \left( \frac{1}{N_t - m} \sum_{i=m+1}^{N_t} \lambda_i(\omega_j) \right)^{N_t - m} \Big/ \prod_{i=m+1}^{N_t} \lambda_i(\omega_j) \right) \right) + \frac{1}{2} Qm(2N_t - m) \ln(N_t), \quad m = 0, \ldots, N_t - 1 \tag{20}$$

The scatterers number is estimated as the value of $m$ that minimizes the Eq. (20). If we apply only a single frequency MDL for target enumeration, the accuracy of high frequency MDL and low frequency MDL can be changed relative to each other according to the various target locations and their strengths. So, for some target configuration, the high frequency MDL outperforms the low frequency MDL, and this condition may be reversed for another configuration. However, employing multiple frequencies in MDL criterion guarantees better results than applying a single frequency (low or high), and the simulation results verify its effectiveness.

### 4.2. Maximization Method

Throughout various numerical tests, we observed that by formation of the MDL criterion for each frequency bin $\omega_j$ (Eq. (19)) and by detection of the maximum value of the criterion for the total number of frequencies, i.e., $j = 1, \ldots, Q$, one can achieve a method for number estimation that outperforms the averaging method. Hence, the number of scatterers using the above method is estimated as follows

$$\widehat{M} = \max\left( \min_m \left[ MDL(m,\omega_1) \right], \min_m \left[ MDL(m,\omega_2) \right], \ldots, \min_m \left[ MDL(m,\omega_Q) \right] \right) \tag{21}$$

To show the correction of the above statement, we perform various simulation tests, and this relation is verified by employing various scatterers and transceivers geometries. From simulation results, it can be seen that the noise affects the MDL criterion at each frequency leading to several minima at various frequencies. The correct source number can be estimated as the maximum of created minimums. Whatever the noise is decreased (SNR is increased), the MDL criterion at each frequency converges to a same minimum or correct source number. In Section 7, we present one sample experiment where calculates the MDL criterion in terms of variable $m$ for total frequencies $\omega_1$, $\omega_2, \ldots, \omega_Q$, and the results validate that the correct source number is equal to the maximum $m$ for minimization of the MDL criterion at total frequencies. Also, the simulation results show that the introduced method for number estimation has advantage over the single frequency MDL method and also superior relative to the averaging method for both low and high frequency conditions.

## 5. PREPROCESSING APPROACH

The noise in data is considered, theoretically, as white Gaussian noise, and hence, the noise eigenvalues become all equal to the variance of the noise $\sigma^2$. However, it is observed that the eigenvalues of the noise in the real data present significant gradient whereas the gradient of the eigenvalues for the synthetic data is much less. This is due to using the limited number of snapshots (transmitters) that makes a non-white noise in situations with real data. Even with the addition of white Gaussian noise in the simulations, the results indicate that the synthetic data also have considerable gradient for the noise eigenvalues in the case of restricted number of snapshots. The gradient in the noise eigenvalue makes the MDL overestimate the number of targets even in high SNRs as discussed in [16]. Our research shows that this is dedicated not only to the MDL, but also to other modified versions of the MDL and AIC methods [17]. The change of penalty function in the MDL [18] is a method to solve this problem. However, the method in [18] works well for high SNRs, but its performance is weaker than the classical

MDL in low SNRs. Hence, we propose a preprocessing method that will produce consistent estimate in both low and high SNRs.

The preprocessing method is based on whitening the noise eigenvalues, so that their gradient will decrease without any destructive effect on the signal eigenvalues. We propose a diagonal loading technique that modifies the eigenvalues $\lambda(\omega_j)$ to reduce the impact of the noise in the estimation performance for real applications. For each frequency $\omega_j$, we can add the diagonal loading value $\lambda_{DL}$ to the singular values $\lambda_i$ as

$$\mu_i(\omega_j) = \lambda_i(\omega_j) + \lambda_{DL}(\omega_j), \quad i = 1, \ldots, N_t, \quad j = 1, \ldots, Q \tag{22}$$

where $\mu_i(\omega_j)$ are the final singular values at frequency $\omega_j$. The diagonal loading value makes little effect on the larger eigenvalues corresponding to the source signals, and the smaller eigenvalues associated with the noise will converge near the loading value $\lambda_{DL}$. Hence, by diagonal loading method, the noise eigenvalues are approximately equal, and the noise is smoothed. Choosing the value of $\lambda_{DL}$ at each frequency makes a great effect on the method. A feasible selection of $\lambda_{DL}$ can be considered as follows

$$\lambda_{DL}(\omega_j) = P_r * \left( \sum_{i=1}^{N_t} \lambda_i(\omega_j) \right), \quad j = 1, \ldots, Q \tag{23}$$

where $P_r$ is the percentage of the eigenvalues summation that must be included in a diagonal loading method. The simulation results for different numbers of targets show that choosing low amount of $P_r$ leads to the overestimation of the targets, whereas the higher one yields underestimation. The concluded amount is 1% that gives suitable estimations for both real and simulated data for practical array configurations.

## 6. PROPOSED METHOD FOR ENUMERATION OF SCATTERERS

The total stages of the proposed method for scatterers enumeration can be summarised by the following steps:

Step 1: Calculate the time reversal matrix $\mathbf{T}(\omega_j)$ at each frequency $\omega_j$ as $\mathbf{T}(\omega_j) = \mathbf{K}(\omega_j)\mathbf{K}^{\dagger}(\omega_j)$.

Step 2: Perform SVD on the matrix $\mathbf{T}(\omega_j)$ to get the eigenvalues $\lambda_i(\omega_j)$, $i = 1, \ldots, N_t$ at each frequency $\omega_j$.

Step 3: Apply the diagonal loading method to whiten the noise using Eq. (22). The whitening is done by calculating the value $\lambda_{DL}(\omega_j)$ at each frequency $\omega_j$ using Eq. (23) and adding to the eigenvalues $\lambda_i(\omega_j)$ to obtain the final eigenvalues $\mu_i(\omega_j)$, $i = 1, \ldots, N_t$.

Step 4: Employ the final eigenvalues $\mu_i(\omega_j)$ instead of $\lambda_i(\omega_j)$ into Eq. (20) to obtain the averaged multifrequency MDL criterion and then estimate the number of targets by minimization of this criterion (FMDL by averaging).

Step 5: Compute the MDL criterion for each frequency $\omega_j$ from Eq. (19) and estimate the number of targets by finding the maximum values as denoted in Eq. (21) (FMDL by maximization).

## 7. SIMULATION RESULTS

In this section, several simulations are carried out to validate the superiority of the proposed algorithm over the classical methods for scatterers enumeration for both simulated and experimental data. The synthetic data have been generated by both the Foldy-Lax model (Eq. (9)) and the method of moments (MOM) by imposing the geometry of the experimental MI systems [5]. The complex additive white Gaussian noise (AWGN) is added to all synthetic data sets as $\mathbf{K}_n(\omega) = \mathbf{K}(\omega) + \mathbf{N}(\omega)$ where $\mathbf{N}(\omega)$ is an uncorrelated zero mean noise, and the signal-to-noise ratio in dB is $SNR = 20 \log_{10}(\|\mathbf{K}(\omega)\|/\|\mathbf{N}(\omega)\|)$.

The numerical experiments taken over 100 times Monte-Carlo simulations, unless otherwise stated. The performance metric is the probability of detecting the correct number of scatterers defined by $P = N_K/N_t$ in which $N_t$ means the number of Monte-Carlo simulations, and $N_K$ denotes the number of the correct detections. In all simulations, the whitening preprocessing method is applied before using the FMDL method. Moreover, the proposed FMDL method also consists of two approaches, a) the FMDL by averaging rule and b) the FMDL by maximization rule. These FMDL methods are also examined using the sample experimental data provided by the Institut Fresnel, Marseille, France [19].

### 7.1. Simulation with Foldy-Lax Data

The first simulation is performed using the array of 36 antennas which act as transmitters and receivers equally spaced on a circle with radius 720 mm. The Foldy-Lax Equation (9) is employed to generate the synthetic data. The simulation is conducted using four point targets located at positions (0, 50), (50, 0), (−50, 0), and (0, −50) in the $x$-$y$ plane with a distance unit of millimeter. The simulation is carried out over the frequency range of 1 to 5 GHz with steps of 500 MHz and for two cases of choosing the targets scattering strengths. Moreover, the diagonal loading value is selected as 1% of total eigenvalues summation. The simulation results for various SNR levels are shown in Fig. 2. It can be seen from Figs. 2(a) and 2(b) that the multifrequency method gives higher accuracy for number detection than the MDL method at single frequency, and also, the method by maximization strategy has better results than the other methods. By the comparison between Figs. 2(a) and 2(b), it can be observed that the FMDL approach gives better accuracy in lower SNR values than the single frequency MDL method in various configuration of targets.

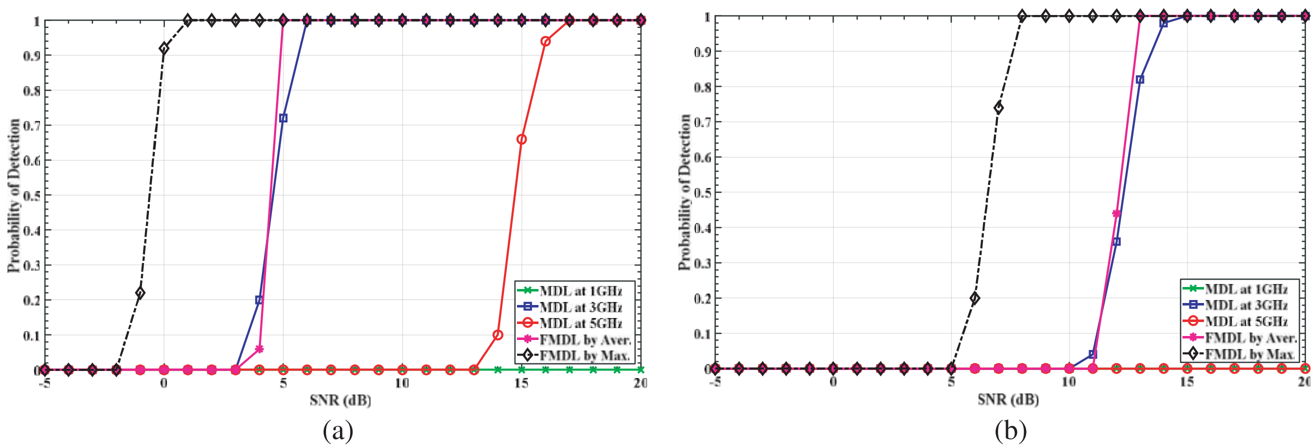
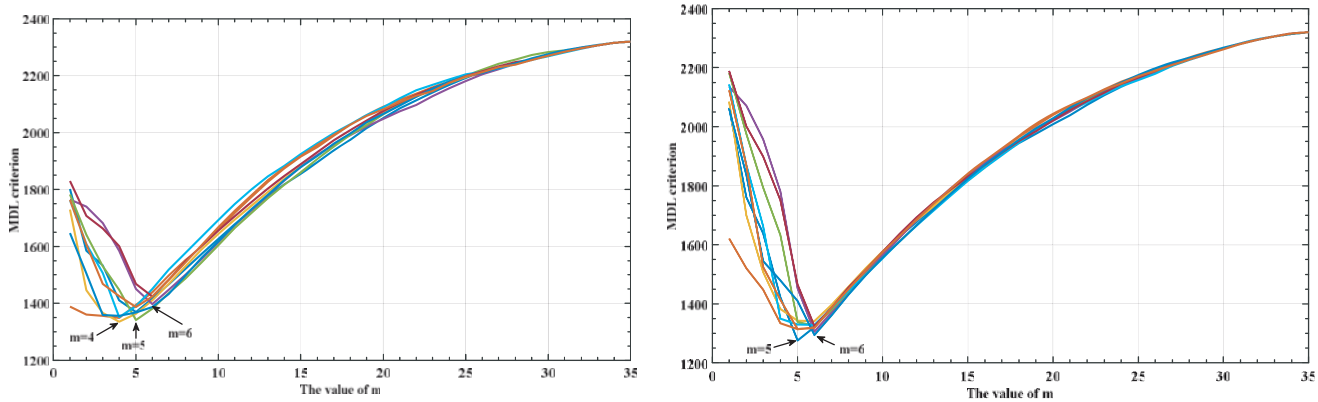



**Figure 2.** Probability of correct decision in terms of SNR for four closely spaced targets with diagonal loading value $P_r = 0.01$. (a) The scatterers with $\tau_m$ =1, $m = 1, \ldots, 4$. (b) The scatterers with $\tau_m = m$, $m = 1, \ldots, 4$.

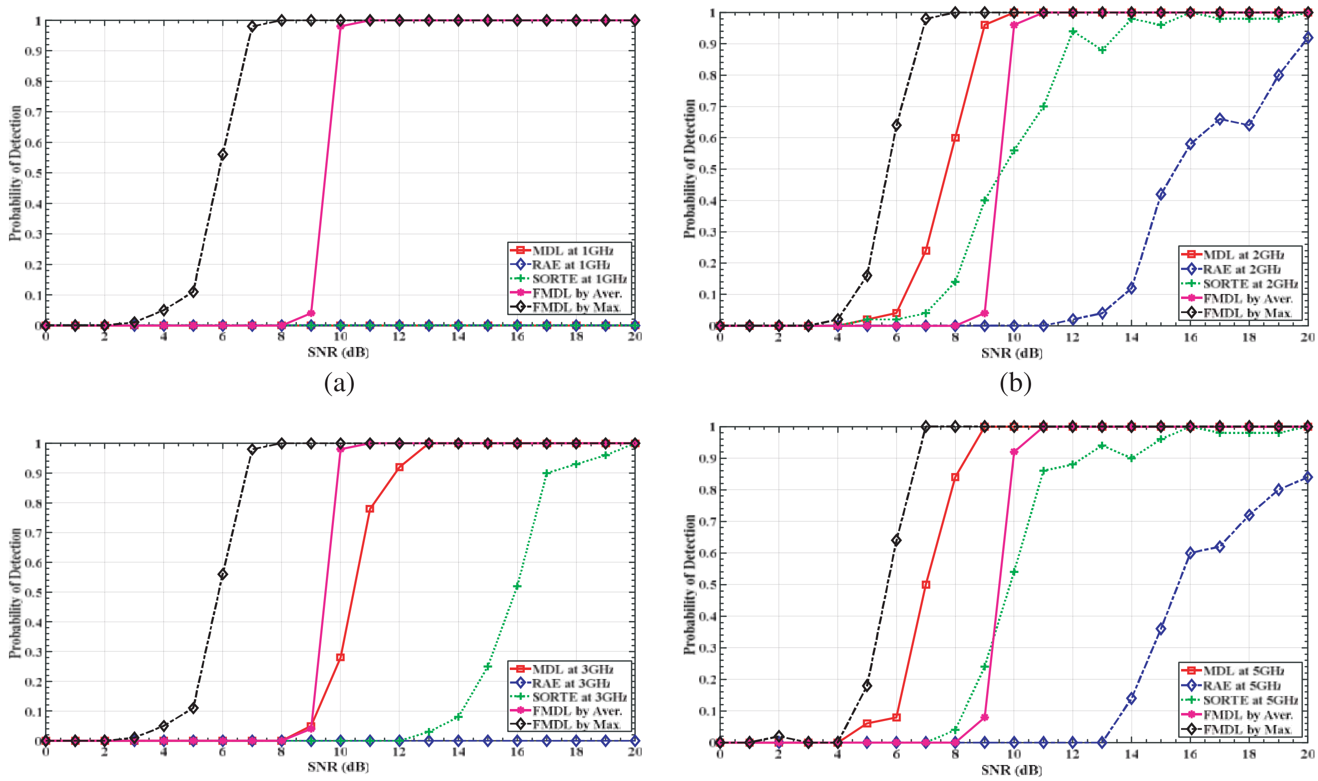
Another simulation is conducted using the previous transceiver geometry and by assuming 6 point targets located at positions (0, 200), (200, 0), (−200, 0), (0, −200), (100, 0), and (0, 100), all in unit of millimeter. The scattering strengths of all targets are assumed to be one. The MDL criterion computed for total frequencies $\omega_j$, $j = 1, \ldots, Q$, i.e., all terms of Eq. (21), are shown in Figs. 3(a) and 3(b) for two SNR values. As can be seen, the MDL criterion in terms of variable $m$ makes multiple minimums at several frequencies, and the correct source number is equal to the maximum $m$ which creates the minimum of criterion. Moreover, it is obvious by increasing the SNR value that the number of minimums is decreased, and the minimums gradually converge to the correct source number.

In the next simulation, we compare the efficiency of the proposed FMDL with three well-known source number estimation methods including MDL, RAE, and SORTE methods. The array configuration of this experiment is based on the Manitoba MI setup [20] in which 24 antennas are equally situated at a circle with radius 220 mm. This simulation is conducted using three point targets closely located at positions (0, 0), (45, 0), (0, 45) in the $x$-$y$ plane with a distance unit of millimeter. The corresponding scattering strengths of the targets are assumed to be $\tau_m = m$, $m = 1, 2, 3$, and the Foldy-Lax Equation (9) is applied to produce the synthetic data. The data are generated for a range of frequencies 1 to 5 GHz with steps of 500 MHz, and the value of $P_r$ for whitening process is set to 0.02. The simulation results for various SNR levels and for several frequency cases are shown in Fig. 4. The results for each SNR are the average of values obtained for different noise realizations according to the Monte Carlo method. Fig. 4 reveals that the FMDL method has better probability of source number detection than the other competing methods working at single frequency. The results for four frequencies are given in Figs. 4(a) to 4(d) that verify the superiority of the method over the classical




(a) (b)

**Figure 3.** MDL criterion in terms of variable $m$ for various frequencies. (a) $SNR = 4\,\text{dB}$. (b) $SNR = 7\,\text{dB}$.




(a) (b)

(c) (d)

**Figure 4.** Probability of correct decision in terms of SNR for classical MDL, RAE, SORTE and FMDL methods by assuming three closely spaced targets with scattering strengths $\tau_m = m$, $m = 1, 2, 3$.

MDL, RAE, and SORTE methods. It can also be seen that the FMDL with maximization rule resolves the targets better than the FMDL by averaging rule.

## 7.2. Simulation with MOM Data

Further simulation evaluates the detection accuracy of the proposed method with respect to the other methods using the array configuration of Manitoba setup and by assuming three small dielectric cylinders

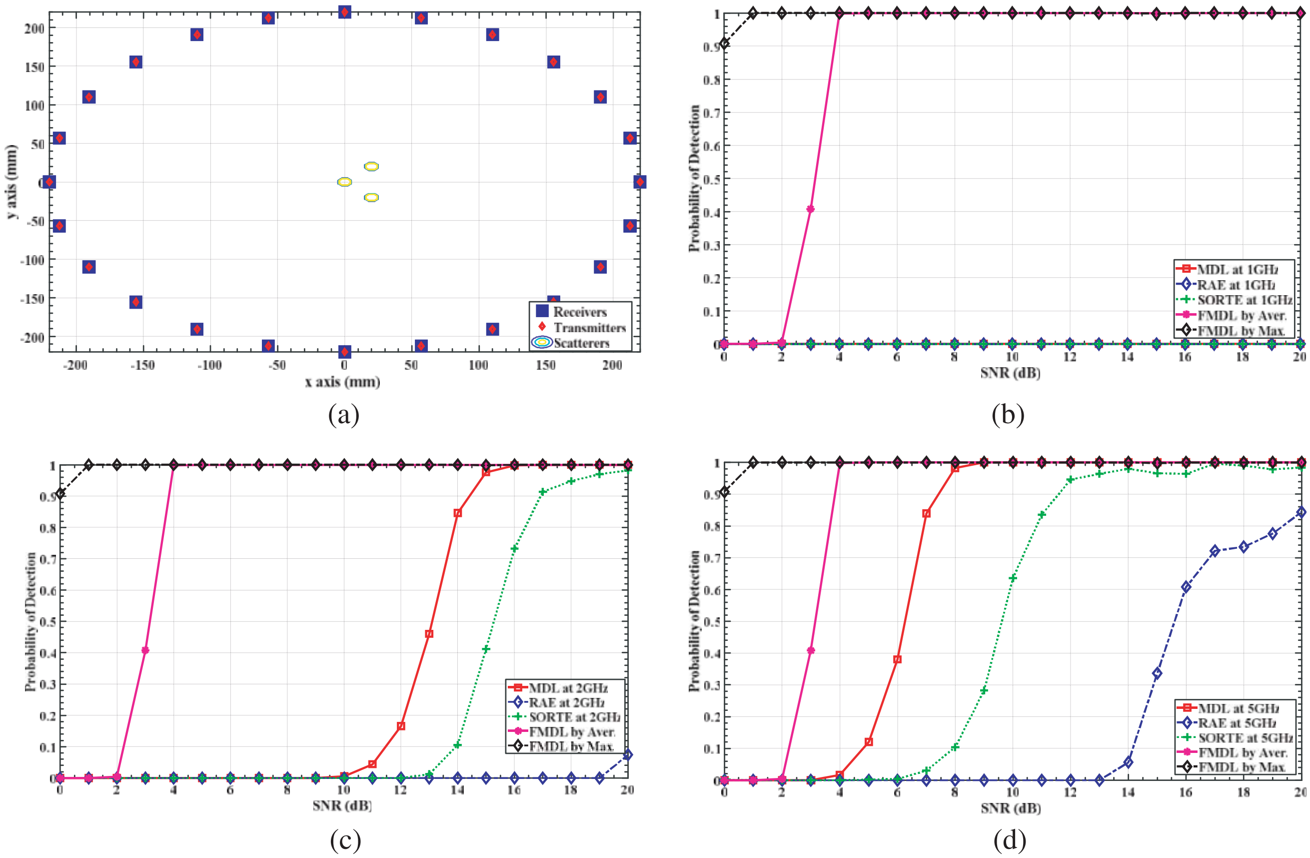



**Figure 5.** Comparing the probability of correct decision in terms of SNR for MDL, RAE, SORTE and FMDL methods using three same small dielectric cylinders with diameter 10 mm having the relative permittivity $\varepsilon_r = 3$ and conductivity $\sigma = 0\,\mathrm{S/m}$ as scatterers.

having diameter 10 mm, relative permittivity $\varepsilon_r = 3$, and conductivity $\sigma = 0\,\mathrm{S/m}$, shown in Fig. 5(a). The relative permittivity and conductivity of the background medium are $\varepsilon_b = 1$ and $\sigma_b = 0$. In this simulation, the numerical data for various frequencies ranging from 1 GHz to 5 GHz with steps of 500 MHz are generated by MOM method for 2-D TM (transverse magnetic) incident electromagnetic wave to compute the $\mathbf{T}$ matrix, and we ignored the variation of targets relative permittivities with frequency. The cylinders are located very closely at positions (0, 0) mm, (20, 20) mm, (20, $-20$) mm, and the parameter $P_r$ for whitening process of FMDL method is set to 0.02. Figs. 5(b) to 5(d) depict the results of this simulation at various SNR levels. It can be found that the proposed FMDL methods give superior accuracy over the other single frequency methods. As can be seen at frequency 1 GHz, all of the methods failed to detect the targets number at total ranges of SNR levels whereas two FMDL methods could detect the targets number correctly at SNR level up to 4 dB.

The next simulation is performed using the same geometry as the previous example to analyze the effect of the diagonal loading value $P_r$ on the detection accuracy of FMDL method for the total range of SNR values (low to high). It can be realized from the results, shown in Fig. 6, that without any whitening process ($P_r = 0$), the classical MDL compared with the FMDL has lower accuracy in low SNR values while for high SNR values, the single frequency classical MDL method has better performance than the FMDL method. However, when the prewhitening method is employed, the FMDL method outperforms the classical MDL for both low and high SNR values. Furthermore, the FMDLs by averaging rule and the FMDL by maximization rule are compared for different $P_r$ values, and the results are shown in Fig. 7. The results verify that by choosing a appropriate $P_r$ value ($P_r = 0.02$), the FMDL by maximization rule performs better than the FMDL by averaging rule in both low and high SNR values.

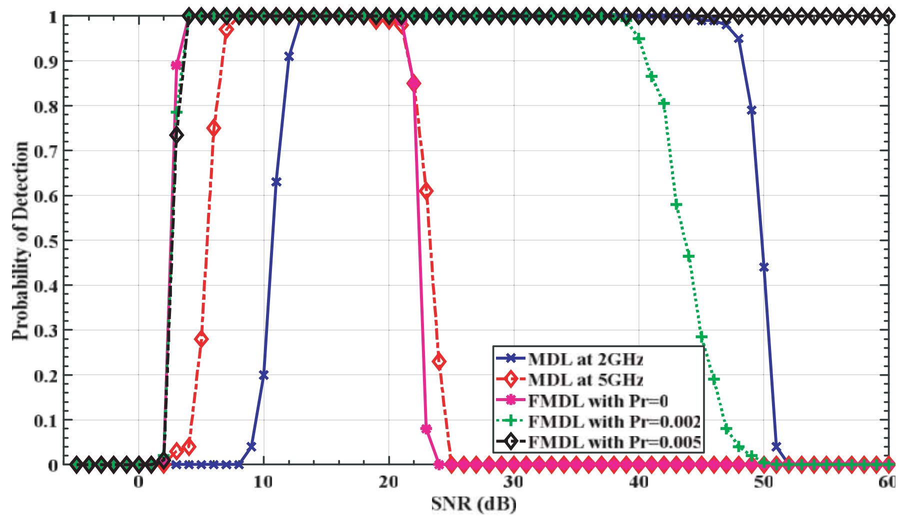



**Figure 6.** Comparing MDL and averaging FMDL using the whitening stage for various $P_r$ values.

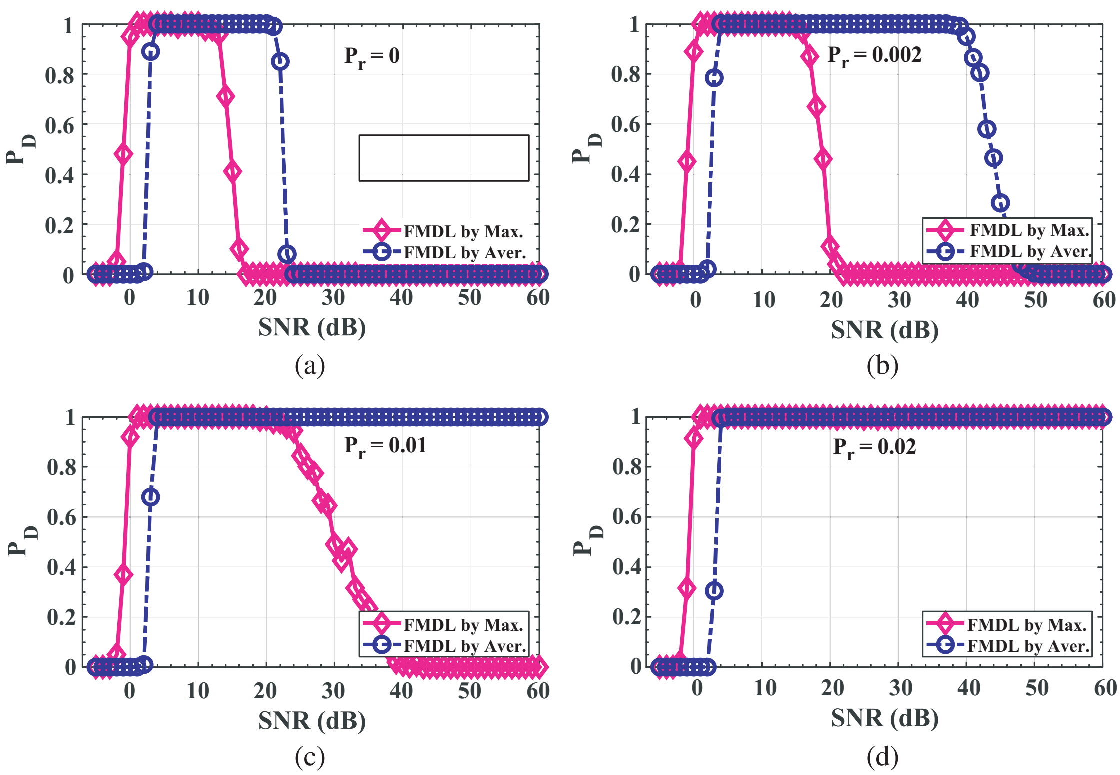



**Figure 7.** Comparing the FMDL methods by averaging rule and maximization scheme for various SNR values.

### 7.3. Simulation with Fresnel Experimental Data

The final simulation is carried out using the experimental data sets provided by the Institut Fresnel, Marseille, France [19] to verify the proposed algorithm. This real data set was employed in various works [21–24] for evaluating the inverse scattering methods against the experimental data. In this real data set, 36 antennas are used for transmitters and 49 antennas employed for receivers. The activated receivers are changed for any active transmitter, so we must use the selected transceivers geometry

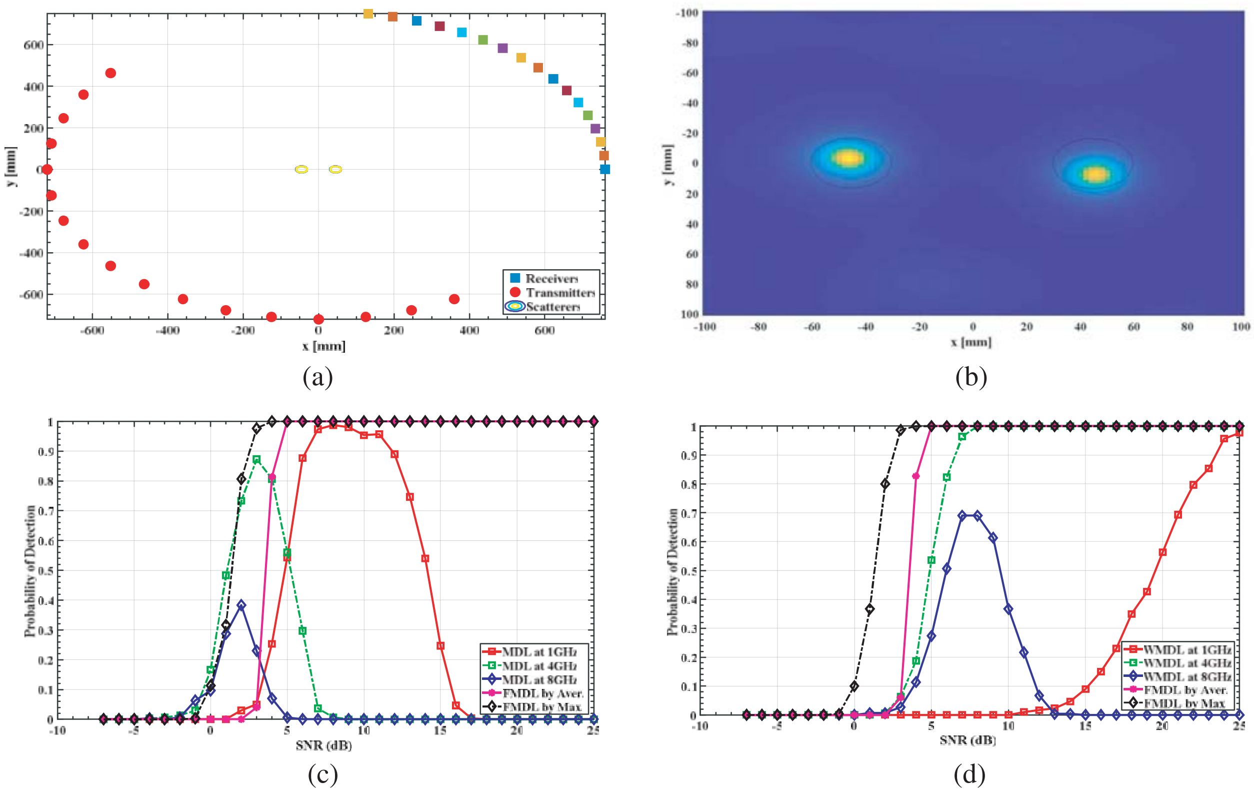



**Figure 8.** Comparing the accuracy of classical MDL, whitened MDL (WMDL) and FMDL using real experimental dataset, (a) selected transceivers, (b) MUSIC reconstructed image at $SNR = 10\,\text{dB}$, (c) comparison between MDL and FMDL and (d) comparison between whitened MDL and FMDL.

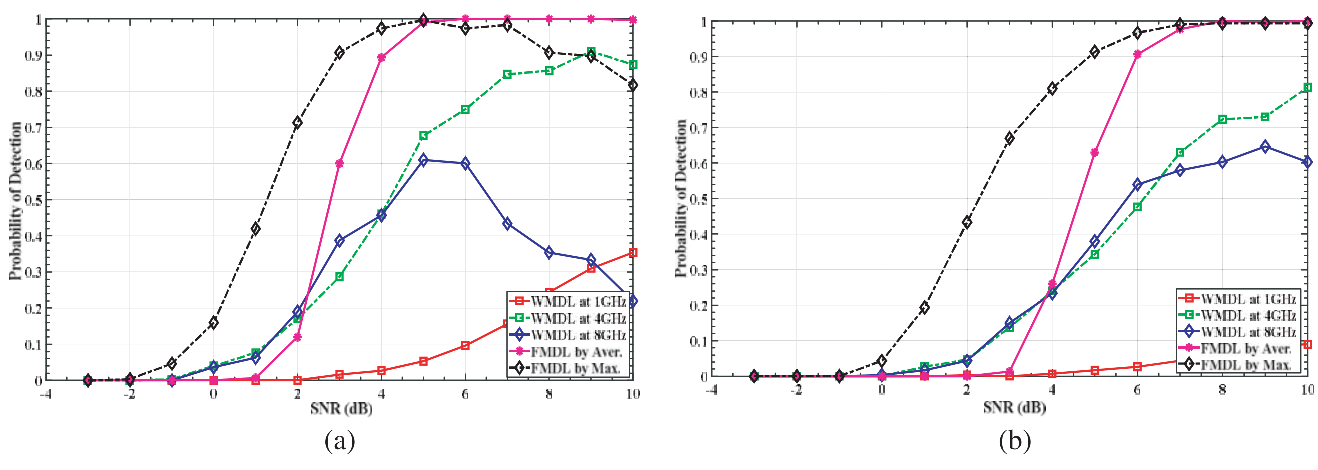



**Figure 9.** Comparing the proposed method and WMDL in the presence of random multiplicative noise, (a) $P_r = 0.2$ and (b) $P_r = 0.3$.

shown in Fig. 8(a) for computing the MSR matrix $\mathbf{K}$ and also for calculating the matrix $\mathbf{T}$ because it is necessary that the receiver array remains unchanged for all columns of $\mathbf{K}$ matrix while the transmitter is switched from one column to another column. Two identical dielectric cylinder targets are used in this simulation having $\varepsilon_r = 3 \pm 0.3$ and radius 15 mm. The experimental data are provided at frequencies 1 to 8 GHz with steps of 1 GHz. The Fresnel measurement data set has a high SNR, so for checking the accuracy of the proposed method against noise level, we add some noise to these data. The results of scatterers number detection for various SNR levels are shown in Figs. 8(c) and 8(d). In these figures, the classical MDL method without whitening stage and with whitening stage is compared with the two FMDL methods, respectively. As can be seen, the FMDL methods give more accuracy for scatterers enumeration for various SNR values. Also, the image reconstruction using MUSIC algorithm is shown in Fig. 8(b) which verifies the ability of MUSIC method for microwave imaging in the case of real experimental data set by assuming that the number of scatterers is priori known.

The final simulation is conducted to further validate the performance of the proposed method against the multiplicative random noise. Thus, we multiply the random noise matrix $\mathbf{W}$ to the MSR matrix $\mathbf{K}$ to obtain a noisy MSR matrix as $\mathbf{K}_n = \mathbf{W} * \mathbf{K} + \mathbf{N}$ and repeat the previous experiment for the second time. The results for two values of $P_r$ are shown in Figs. 9(a) and 9(b). It can be seen by choosing appropriate $P_r$ value that the proposed method outperforms the WMDL method even in the case of multiplicative noise.

## 8. CONCLUSION

We have firstly reviewed the theory of the MUSIC algorithm for microwave imaging of small scatterers from scattered field measured by the antenna array. Afterwards, the methods for scatterers enumeration including MDL, RAE, AIC, and SORTE required in MUSIC algorithm for imaging have been investigated. These methods working at single frequency have been implemented and compared through some simulation tests, and then the new multifrequency MDL (FMDL) method combined with two averaging and maximization rules has been proposed. The FMDL fuses information at several frequencies using the proposed rules to reduce the error in the number estimation for various SNR levels. Also, because the limited number of antennas leading to the noise may behave like non-white Gaussian noise, a preprocessing step for whitening of noise has been introduced before employing the FMDL method. Many simulations using both the synthetic and experimental data have illustrated the merits of the proposed FMDL method with respect to the other competing methods for scatterers enumeration.

## REFERENCES

1. Benny, R., T. A. Anjit, and P. Mythili, "An overview of microwave imaging for breast tumor detection," *Progress In Electromagnetics Research B*, Vol. 87, 61–91, 2020.
2. Park, W.-K., K.-J. Lee, H.-P. Kim, and S.-H. Son, "Application of MUSIC to microwave imaging for detection of dielectric anomalies," *Progress In Electromagnetics Research Symposium — Spring (PIERS)*, 2908–2912, St. Petersburg, Russia, May 22–25, 2017.
3. Lee, K. J., S. H. Son, and W. K. Park, "A real-time microwave imaging of unknown anomaly with and without diagonal elements of scattering matrix," *Results in Physics*, Vol. 17, 103104, 2020.
4. Agarwal, K. and X. Chen, "Applicability of MUSIC-type imaging in two-dimensional electromagnetic inverse problems," *IEEE Trans. Antennas Propag.*, Vol. 56, No. 10, 3217–3223, 2008.
5. Ammari, H., E. Lakovleva, and D. Lesselier, "A MUSIC algorithm for locating small inclusions in a half-space from scattering amplitude at a fixed frequency," *SIAM Multiscale Model. Simul.*, Vol. 3, 597–628, 2005.
6. Fazli, R., M. Nakhkash, and A. A. Heidari, "Alleviating the practical restrictions for MUSIC algorithm in actual microwave imaging systems: Experimental assessment," *IEEE Trans. Antennas Propag.*, Vol. 62, No. 6, 3108–3118, 2014.
7. Solimene, R., A. DellAversano, and G. Leone, "Interferometric time reversal MUSIC for small scatterer localization," *Progress In Electromagnetics Research*, Vol. 131, 243–258, 2012.
8. Wax, W. and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. on Acoustic, Speech, and Signal Processing*, Vol. 33, 387–392, 1985.
9. Xiao, M., P. Wei, and H. M. Tai, "Estimation of the number of sources based on hypothesis testing," *Journal of Communications and Networks*, Vol. 14, No. 5, 481–486, 2012.

10. He, Z., A. Cichocki, S. Xie, and K. Choi, "Detection the number of clusters in n-way probabilistic clustering," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 32, No. 11, 2006–2021, 2010.
11. Liavas, A. P. and P. A. Regalia, "On the behavior of information theoretic criteria for model order selection," *IEEE Trans. on Signal Processing*, Vol. 49, No. 8, 1689–1695, 2001.
12. Pouramadi, M., M. Nakhkash, and A. A. Tadion, "Application of MDL criterion for microwave imaging by MUSIC algorithm," *Progress In Electromagnetics Research B*, Vol. 40, 261–278, 2012.
13. Fazli, R. and M. Nakhkash, "An analytical approach to estimate the number of small scatterers in 2D inverse scattering problems," *Inverse Prob.*, Vol. 28, No. 7, 75012–75033, 2012.
14. Wax, M. and I. Ziskind, "Detection of the number of coherent signals by the MDL principle," *IEEE Trans. on Acoustic, Speech, and Signal Processing*, Vol. 37, No. 8, 1190–1196, 1989.
15. Wax, W. and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. on Acoustic, Speech, and Signal Processing*, Vol. 33, 387–392, 1985.
16. Ding, Q. and S. Kay, "Inconsistency of the MDL: On the performance of model order selection criteria with increasing signal-to-noise ratio," *IEEE Trans. on Signal Processing*, Vol. 59, 1959–1969, 2011.
17. Ridder, F., R. Pintelon, J. Schoukens, and D. P. Gillikin, "Modified AIC and MDL model selection criteria for short data records," *IEEE Trans. on Instrumentation and Measurement*, Vol. 54, 144–150, 2005.
18. Kundu, D., "Estimating the number of signals in the presence of white noise," *Journal of Statistical Planning and Inference Elsevier*, Vol. 90, No. 4, 57–68, 2000.
19. Belkebir, K. and M. Saillard, "Testing inversion algorithms against experimental data," *Inverse Probl.*, Vol. 17, 1565–1571, 2001.
20. Gilmore, C., P. Mojabi, A. Zakaria, M. Ostadrahimi, C. Kaye, S. Noghanian, L. Shafai, S. Pistorius, and J. Lovetri, "A wideband microwave tomography system with a novel frequency selection procedure," *IEEE Trans. Biomedical Eng.*, Vol. 57, No. 4, 894–904, 2010.
21. Rocca, P., M. Donelli, G. L. Gragnani, and A. Massa, "Iterative multi-resolution retrieval of non-measurable equivalent currents for the imaging of dielectric objects," *Inverse Probl.*, Vol. 25, 055004–055018, 2009.
22. Caorsi, S., M. Donelli, A. Lommi, and A. Massa, "Location and imaging of two-dimensional scatterers by using a particle swarm algorithm," *Journal of Electromagnetic Waves and Applications*, Vol. 18, No. 4, 481–494, 2004.
23. Zheng, H., M. Z. Wang, Z. Q. Zhao, and L. L. Li, "A novel linear EM reconstruction algorithm with phaseless data," *Progress In Electromagnetics Research Letters*, Vol. 14, 133–146, 2010.
24. Caorsi, S., M. Donelli, and A. Massa, "Detection, location, and imaging of multiple scatterers by means of the iterative multiscaling method," *IEEE Trans. on Microwave Theory and Tech.*, Vol. 52, 1217–28, 2004.