Theoretical analysis of a new approach to order determination for a modified Prony method in swath mapping application

Piotr GRALL, Jacek MARSZAL

Gdańsk University of Technology
Faculty of Electronics, Telecommunications and Informatics
Department of Marine Electronic Systems
Narutowicza 11/12,80-233 Gdańsk, Poland
jacek.marszal@eti.pg.gda.pl, grallu@poczta.onet.pl

This article presents a new approach to determine the model order (number of principal components) in the modified Prony method applied to swath acoustic mapping. Determination of the number of principal components is a crucial step in the modified Prony method. In the proposed approach the model order is chosen based on the underlying physical model of the underwater acoustic environment, and utilised signal processing operations. This data-driven approach, attempts to make use of all available information to assess the number of signals arriving at the receiver using pipeline processing in lieu of iterative processing.

Keywords: swath bathymetry, interferometry, principal components, DOA, MBES.

1. Introduction

Application of the Prony method to Direction-of-Arrival (DOA) estimation is an interesting alternative for conventional beamforming and basic interferometry in underwater acoustics for depth determination [8, 10, 13, 20]. It is capable of simultaneously determining multiple angles of arrival (similarly to beamforming) while utilizing a reduced number of hydrophones in the receiving array (as in interferometry) [3]. Although other sub-space (model-based) methods exist, such as variants of ESPRIT and MUSIC, they are highly computationally intensive and their performance abruptly deteriorates below a certain SNR level (known as "threshold effect") [13–15, 18]. On the other hand, interferometry is capable of determining only one DOA at a given instant, which limits its performance and object detection capabilities in complex geometry environments [9].

The heart of the modified Prony (MP) method is the usage of a reduced rank approximation of the correlation matrix and solution of a set of linear equations. When choosing this low-rank approximation of the correlation matrix it is necessary to determine the model order to separate signal eigenvalues from noise eigenvalues [17]. While general methods exist for determination of the model order (also referred to as the number of principal components-PCs or principal eigenvectors-PEs) such as Aike Information Criterion and Minimum Description Length [1, 2, 19], they are not suitable for acoustic mapping applications due to the inherent non-stationariness of the processed signal (i.e. constant change of the spatial frequency) [8]. An alternative solution is to use an iterative model order determination algorithm to solve the linear set of equations in a Least Squares sense, which stops after a pre-set number of steps, or after reaching the convergence criterion [8]. A different, new approach is proposed in this article (see Section 3), which takes into account the underlying physical model of the signal propagation environment and signal processing operations applied in the processing chain. Based on the assumption that signal eigenvalues are greater than noise eigenvalues, the proposed method dynamically determines the threshold below which eigenvalues are removed from the sample covariance matrix to obtain its low-rank approximation.

2. Basic concepts of Prony method

The method proposed in this article is an extension to Computed Angle-of-Arrival Transient Imaging (CAATI) under the following assumptions [8]:

1. A linear N-element equispaced array is used to measure backscatter arrivals propagating in the same plane as the array.
2. At each instant in time exactly M independent, coplanar plane waves are incident on the receiving array.
3. The acoustic backscatter is narrowband.
4. The receiving array output signals are in steady state across the entire array (received amplitude is only affected by exponential damping between receive elements and noise).

Let there be a set of N equispaced ($d = \lambda/2$) identical receivers impinged by M signals perturbed by the noise (Fig. 1). The received signals, $s(n)$ at individual hydrophones can be written as:

$$s(n) = \sum_{i=1}^{M} a_i \, e^{(\alpha_i + j\, u_i)\, d\, (n-1)} + w(n) \tag{1}$$

$$a_i = A_i \, e^{j\Theta_i},\, u_i = k \sin\theta_i \,, n = 1, 2, 3, \cdots , N$$

where u_m is the acoustic wavenumber, $k = 2\pi/\lambda$, λ is the wavelength at the central frequency, α is the exponential damping factor and $w(n)$ is additive noise at each array element. Eq.(1) is equivalent to the statement that each hydrophone, at the same time, receives a phase delayed and exponentially attenuated copy of the signal received by the reference element. In eq. (1) time dependency is suppressed and a_i is complex amplitude of the backscatter signal at the reference receiver $(n = 1)$.

Noticing the analogies between this model and linear prediction-error filter ([16, 17]), we might find DOAs of the backscatter signals by finding complex roots of the polynomial:

$$H(z) = 1 + \sum_{l=1}^{L} g_l \, z^{-l} = 0 \tag{2}$$

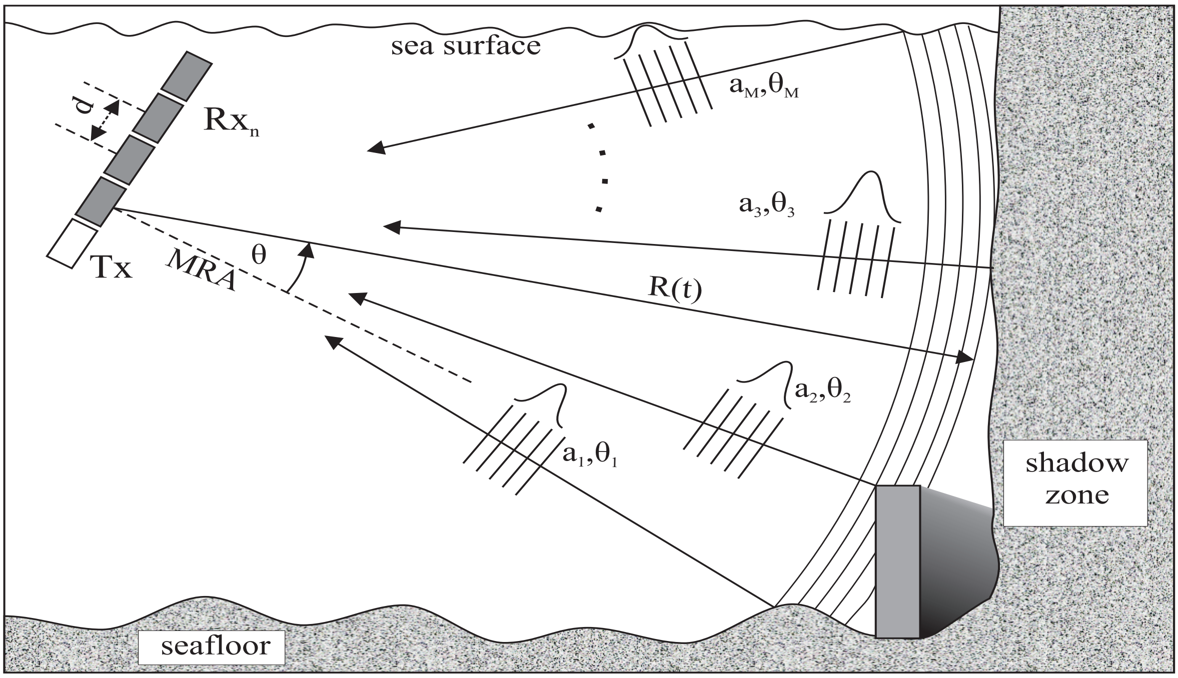

Fig. 1. Multiple signals arriving at the receiver array.

where L is the filter order and $z = e^{\sigma + j\omega}$. If L satisfies the inequality:

$$M \leqslant L \leqslant (N - M/2) \tag{3}$$

complex coefficients g_l may be obtained by solving a set of linear complex forward–backward equations [17]:

$$\begin{bmatrix} s(L) & s(L-1) & \cdots & s(1) \\ s(L+1) & s(L) & \cdots & s(2) \\ \vdots & \vdots & & \vdots \\ s(N-1) & s(N-2) & \cdots & s(N-L) \\ \hdashline s^*(2) & s^*(3) & \cdots & s^*(L+1) \\ s^*(3) & s^*(4) & \cdots & s^*(L+2) \\ \vdots & \vdots & & \vdots \\ s^*(N-L+1) & s^*(N-L+2) & \cdots & s^*(N) \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_L \end{bmatrix} = - \begin{bmatrix} s(L+1) \\ s(L+2) \\ \vdots \\ s(N) \\ \hdashline s^*(1) \\ s^*(2) \\ \vdots \\ s^*(N-L) \end{bmatrix} \tag{4a}$$

or shortly:

$$Ag = -h \tag{4b}$$

where * denotes a complex conjugate. In the noiseless case, out of L roots of the eq.(2) M closest to the unit circle on the complex plane are equal to $e^{(\alpha_i + j\,u_i)\,d}$ components in eq. (1). Additional $L - M$ roots lying further away from the unit circle are rejected. The unknowns α_i and u_i might be obtained from:

$$\alpha_i = \frac{\ln |z_i|}{d}\ (\mathrm{m}^{-1}),\ u_i = \frac{\arg(z_i)}{d}\ (\mathrm{m}^{-1}) \tag{5}$$

Finally, DOA relative to main response axis (MRA) of each plane wave is obtained from:

$$\theta_i = \arcsin(\frac{u_i}{k}) \text{ (rad)} \tag{6}$$

Complex amplitudes might be calculated by solving the second set of linear equations:

$$\begin{aligned} s &= Fa \text{ , where :} \\ s &= \begin{bmatrix} s(1) & s(2) & \cdots & s(N-1) \end{bmatrix}^T \\ a &= \begin{bmatrix} a_1 & a_2 & \cdots & a_M \end{bmatrix}^T \\ F &= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-ju_1} & e^{-ju_2} & & e^{-ju_M} \\ \vdots & & \ddots & \vdots \\ e^{-j(N-1)u_1} & e^{-j(N-1)u_2} & \cdots & e^{-j(N-1)u_M} \end{bmatrix}. \end{aligned} \tag{7}$$

If there are M distinct solutions, we may calculate:

$$a = F^{+} s \tag{8}$$

where $^+$ denotes Moore-Penrose pseudoinverse. Amplitudes and phases of the backscattered signals can be then calculated from:

$$A_1 = |a_i|, \quad \Theta_i = \arg(a_i) \tag{9}$$

Equation (4a) may be solved using various methods [14, 15]. In the modified Prony (MP) method proposed by Tufts and Kumaresan [17], we first calculate correlation matrix and the correlation vector:

$$R = A^* A \tag{10}$$

$$r = -A^* h \tag{11}$$

where * denotes conjugate transpose. We may then rewrite eq.(4a) as:

$$g = R^{+} r \tag{12}$$

where $^+$ denotes the Moore–Penrose pseudoinverse. Eigenvalue/eigenvector decomposition of R is performed as:

$$R = U \Lambda U^* \tag{13}$$

where $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_M$ and $\gamma_{M+1} \geq \gamma_{M+2} \geq \cdots \geq \gamma_L = 0$ are the eigenvalues of R sorted in descending order, situated on the main diagonal of Λ. The columns of U are orthonormal eigenvectors of R. Using this decomposition we may rewrite g as a linear combination of correlation vector, eigenvectors and eigenvalues of R:

$$g = \hat{g} = \sum_{i=1}^{M} \frac{u_i^* r \, u_i}{\gamma_i} \tag{14}$$

where u_i is the eigenvector of U associated with i^{-th} eigenvalue. In the presence of noise the smallest eigenvalues are not equal to each other and their values are different than zero. By selecting M greatest eigenvalues, and setting other eigenvalues to zero, we obtain the low-rank approximation $\hat{R}$ of the correlation matrix R (see section 3 for further details). We are free to choose any value of L in the interval (3) to construct matrix A and vector h. If we use a much larger number of receivers than the number of backscattered signals ($N \gg M$) we might significantly improve the performance of the Prony method. The optimal value of L, in terms of accuracy of DOAs, in case of two equal amplitude signals, is equal to approximately $\frac{3}{4}N$ [17]. Additionally, as suggested in [8], we may use several snapshots of data to obtain better approximations of correlation matrix and correlation vector:

$$A_{t,K} = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \cdots \\ A_K \end{bmatrix}, h_{t,K} = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ \cdots \\ h_K \end{bmatrix} \tag{15}$$

where K is the number of snapshots associated with a given instant t. The number of snapshots is limited, however, by spatial and temporal decorrelation of the signal, and the available number of degrees of freedom [8].

3. New method for determination of the model order

The key step of the MP method is estimation of the order of the model $\hat{M}$ i.e the number of signals impinging the array. There are three possible situations [17]:

$$\hat{M} = \begin{cases} = M, & \text{properly modelled,} \\ > M, & \text{overmodelled order,} \\ < M, & \text{undermodelled order.} \end{cases}$$

In the first case, the model order is resolved properly and solutions are only slightly perturbed. For the majority of SNR scenarios $\hat{R}$ is a better estimator of R than the least-squares solution which includes all L eigenvalues [17]. In the second case, the model order is overestimated and noise is resolved as a valid signal giving false detection, and perturbing $\hat{g}$ vector. In the last case, one or more signals are not resolved at all, resulting in a depleted solution (vector $\hat{g}$ does not contain all PCs). It is evident that proper determination of $\hat{M}$ is crucial for the performance of this method. The solution proposed in this paper is aimed at adjusting the model order determination scheme to take advantage of knowledge of the wave propagation environment, and applied signal processing methods.

Noise present in a digital signal processing (DSP) unit might be divided into two categories, depending of its origin:

1. Additive noise - noise independent of signal (present when the signal is absent).
2. Multiplicative noise - signal–dependent noise (present only when signal is present).

Noise independent of signal consists mainly of thermal noise present in the water environment and the noise of the electrical circuit affecting the signal between receiver elements

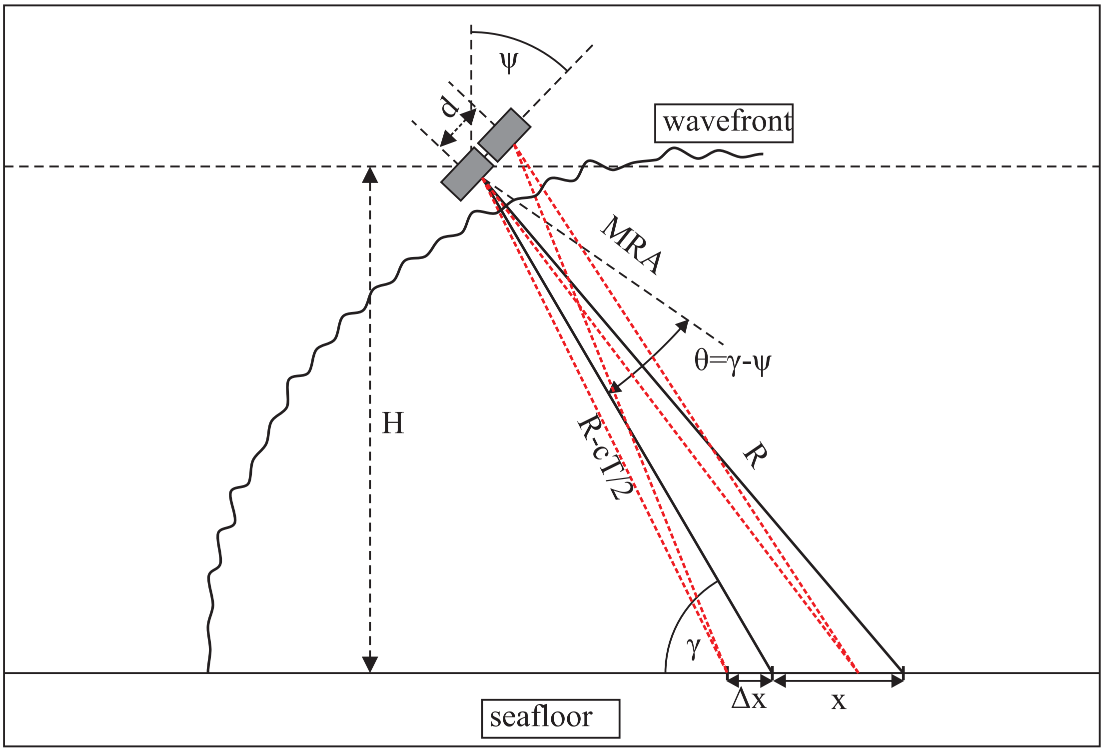

Fig. 2. Basic geometric relations for calculation of baseline decorrelation and shifting footprint equivalent noise calculations– flat seafloor assumption.

and DSP unit (for simplicity electrical noise is not included in further calculations). For high–frequencies (above 200 kHz) thermal noise might be modelled as isotropic Additive White Gaussian Noise. We might determine model order by calculating the numerical rank of $A_{n,K}$ [6]:

$$\hat{M}_1 = rank(A_{t,K}, tol_1) \tag{16}$$

$$tol_1 = \sigma^2(t)\sqrt{\frac{2 \cdot (N - L) \cdot L \cdot K}{(L - M)}} \tag{17}$$

where the term in the nominator is equal to the number of terms in $A_{n,K}$ and $\sigma^2(t)$ is the standard deviation of noise due to thermal. The term $(L - M)$ in the denominator results from the fact that noise energy is spread between $L-M$ eigenvalues. The time dependence in eq. (17) indicates that thermal noise level at the DSP unit is dependant on the time–varied gain (TVG) applied to the acquired signal. This method of determining model order would give accurate predictions if the signal was stationary (constant noise variance, constant DOAs and complex signal envelope).

The other category of noise is proportional to the signal level, and is a result of coherence loss between signals received by adjacent array elements. This might be attributed primarily to the baseline decorrelation noise and the shifting (sliding) footprint effect ([7, 12]). In essence, it is the part of backscattered signal that is not common to all receivers and acts as random noise (Fig. 2). This noise is related to input (external) SNR at the receiver and to the receiver, bottom and signal parameters. For flat sea-floor baseline decorrelation equivalent SNR might be calculated from the following equations [7, 12]:

$$d_{bd} = (\frac{\nu}{1-\nu}) \tag{18}$$
$$\nu = \frac{\sin\eta}{\eta}$$
$$\eta = \frac{kd_{L+1}}{H}\frac{c\tau}{4}\sin\gamma\tan\gamma\cos(\gamma-\psi)$$

where T is the acoustic pulse length, ψ is the depression angle, γ is the grazing angle, and d_{L+1} is the mean distance between $(L+1)$ elements.

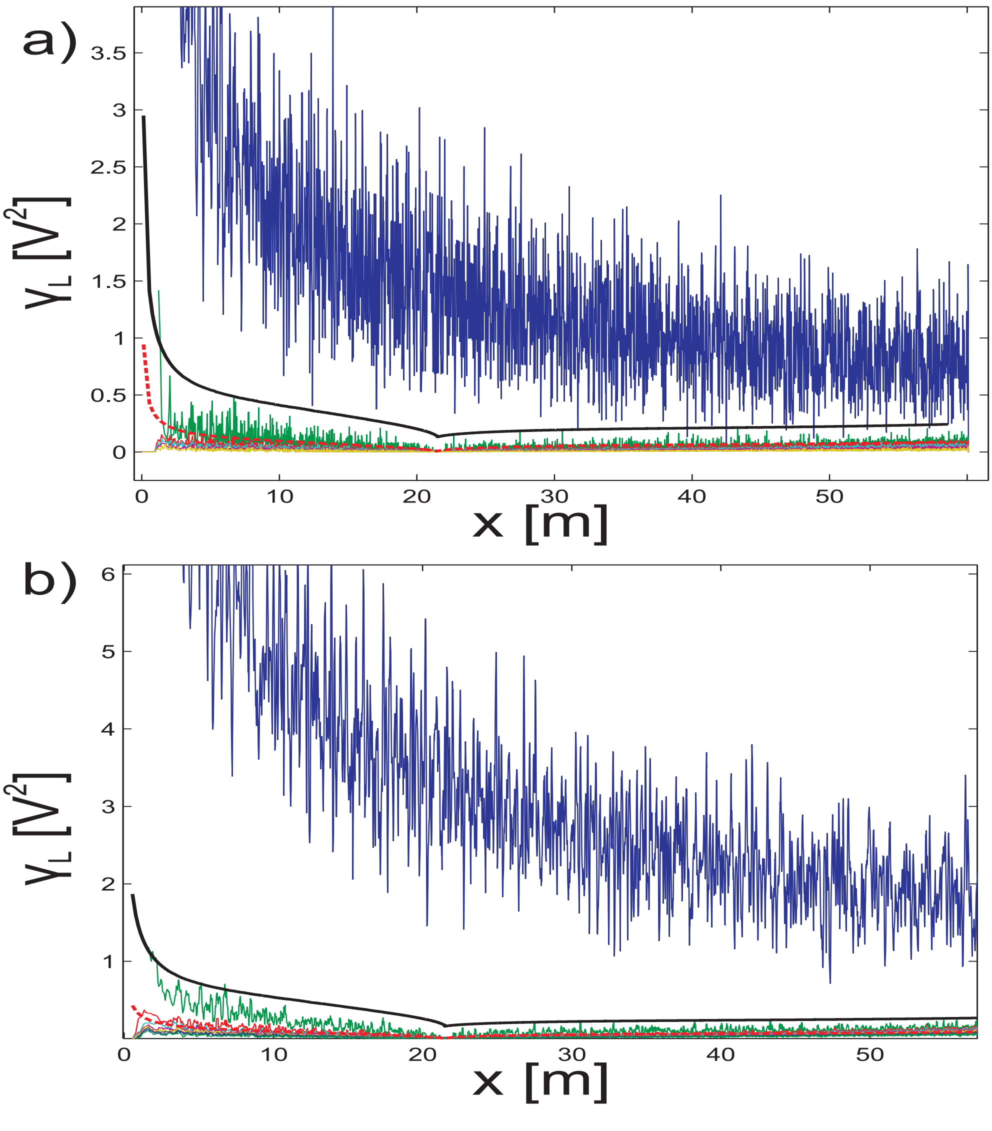

Fig. 3. Eigenvalues and calculated thresholds.
a) $K=1, H=10, \psi=25°$, b) $K=5, H=10, \psi=25°$
dashed line (red) – $(a=0, b=0)$; solid line (black) – (a, b) adjusted for signal variability.

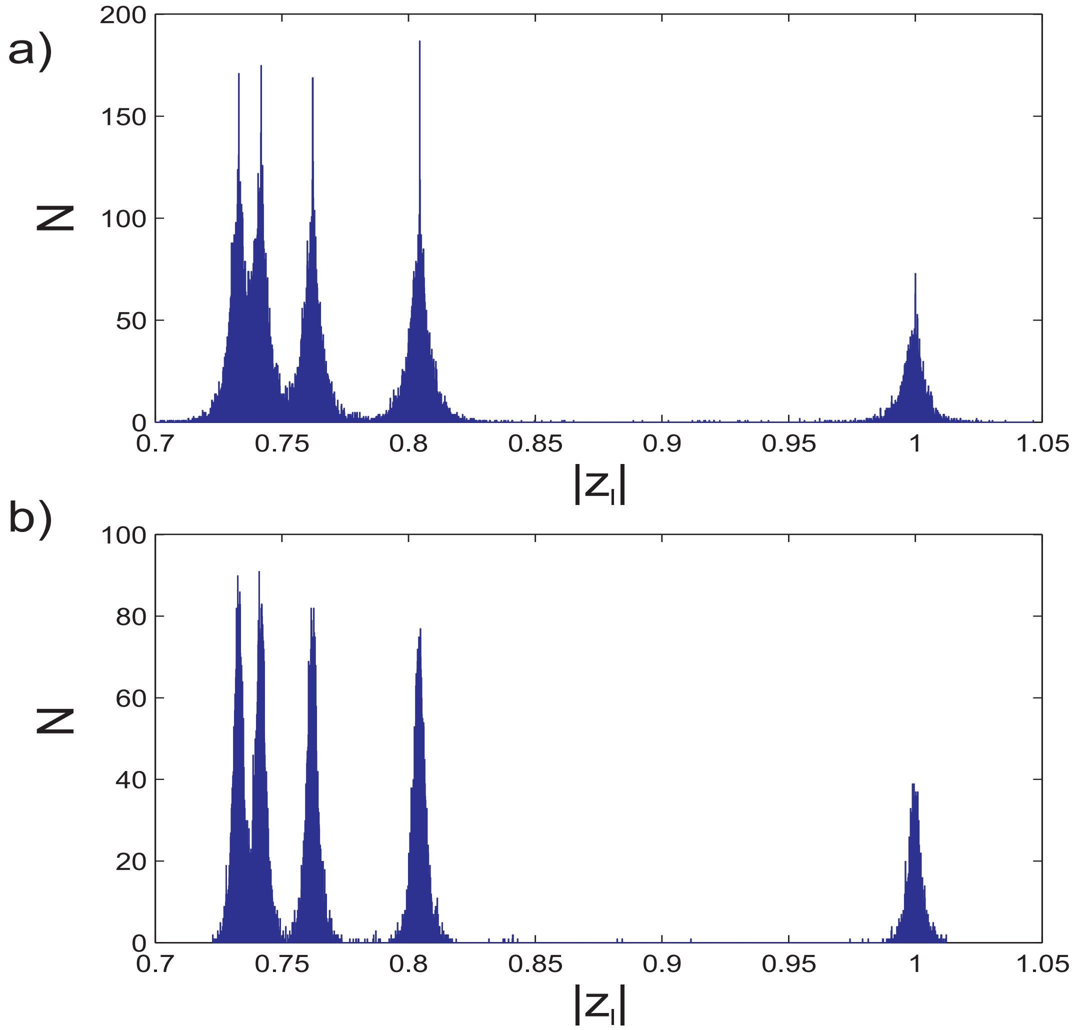

Fig. 4. Histogram of the amplitude of complex roots.
a) $K = 5, H = 5, \psi = 25°$, b) $K = 5, H = 20, \psi = 25°$

Shifting footprint equivalent signal to noise ratio, for the same bottom and receiver configuration, might be calculated as [12]:

$$d_{sf} \approx \left(\frac{c\tau}{2d_{L+1} \left| \sin(\gamma - \psi)) \right|} - 1 \right) \tag{19}$$

Equivalent SNR from the combined effect of the two above factors is calculated as:

$$d_d = \frac{1}{\frac{1}{d_{bd}} + \frac{1}{d_{sf}}} \tag{20}$$

Total equivalent SNR is equal to:

$$d_{eq} = \frac{1}{\frac{1}{d_{dd}} + \frac{1}{d_{ex}}} = \frac{s_{eq}^2}{\sigma_{eq}^2} \tag{21}$$

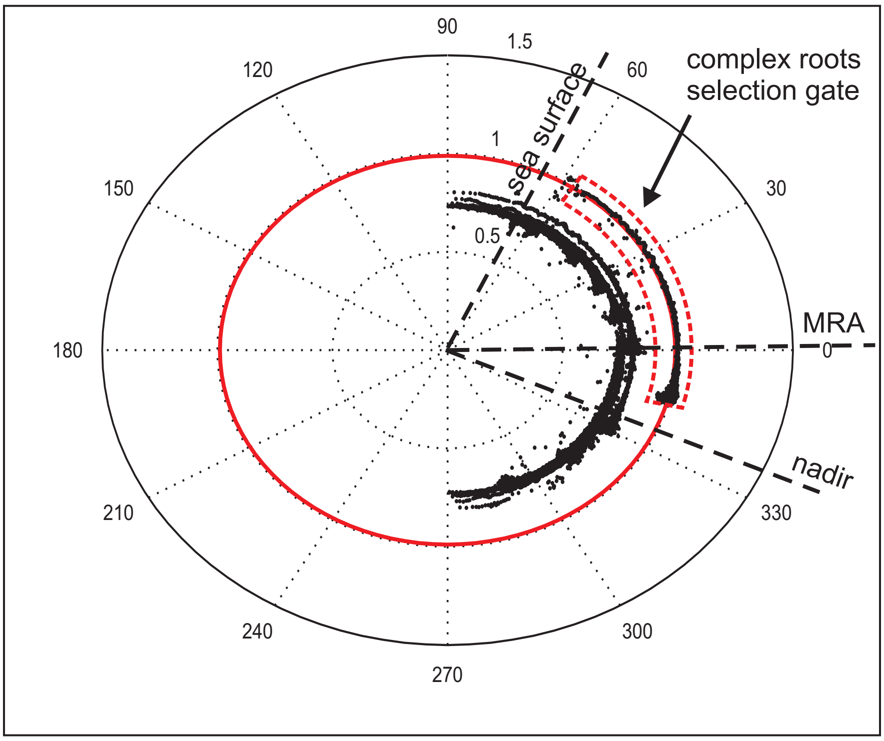

Fig. 5. Polar plot of resolved DOAs, complex roots amplitudes and selection gate.

We might also calculate the external noise gain factor:

$$d_{NG} = \frac{d_{ex} + d_d + 1}{(d_d + 1)} = \frac{\sigma_{eq}^2}{\sigma^2(t)} \tag{22}$$

This value might be used to obtain the final version of eq.(17) by replacing $\sigma_{tn}^2(t)$ by σ_{eq} :

$$tol_2 = a \cdot \sigma(t) \sqrt{\frac{2 \cdot d_{NG} \cdot (N - L) \cdot L \cdot K}{L - M}} + b \tag{23}$$

Factors a and b might be used to account for modelling errors and signal fluctuations (Fig. 3). Finally, model order is determined as:

$$\hat{M}_2 = rank(A_{t,K}, tol_2) \tag{24}$$

Any eigenvalue lower than the tolerance threshold in eq.(23) is rejected. In case no signal is present (before the backscattered signal reaches array or in the regions of acoustic shadow), $\hat{M}_2$ is equal to zero and invalid solutions are stopped in the processing chain. This also implies that in the case of multiple signals arriving at the receiver, if any signal eigenvalue is lower than this threshold, it will be rejected and the model order will be under-modelled $\left(\hat{M}_2 < M\right)$ introducing errors in the solution. This situation might occur when there is a secondary, weak target or two comparable energy targets with low angular separation. As a result, this algorithm properly estimates PCs only if all signal eigenvalues exceed this threshold. In other words, when the received signal eigenvalues fall below this threshold, noise eigenvalues might intertwine with signal eigenvalues so that PCs cannot be easily extracted [5, 11].

Fig. 3 depicts plots of the calculated eigenvalues of simulated signals for two different number of snapshots. The values calculated directly by eq. (23) capture the mean behaviour of eigenvalues (red, dashed line). In order to exclude all noise eigenvalues from low-rank matrix approximation values a, b are introduced and adjusted to obtain final limits (black, solid line). It can be noticed that an increase in the number of snapshots provides better separation between the signal and noise eigenvalues.

Tab. 1. Simulation parameters

Parameter	Value(s)
Altitude (H)	5, 10, 20 m
Depression Angle (ψ)	25°
Frequency	200 kHz
Pulse Length τ	100 µs
Horizontal Beamwidth (θ_{3dB})	1.5°
Receive Elements (N)	12
Filter Order (L)	9
Number of snapshots (K)	1, 5
Baseband Complex Envelope Sample Rate	25 kHz
SL+OCV+TS (Equivalent of 74,2 dB initial SNR for H=10 m)	0 dB
Noise Floor (NL+OCV)	1 µVrms
Scatterers Density (SD) - uniformly distributed	200 m^{-1}
Sound Speed (c)	1486 m/s
Sound Absorption (α)	0.05 dB/m

4. Selection of valid solutions

Having estimated the model order $\hat{M}_2$, L roots of eq.(2) can be calculated and valid M solutions ought to be extracted. In different papers [8,20], the limit for the amplitude of complex roots is suggested to be $1 \pm (0.1 - 0.15)$. The values obtained during simulations (Tab. 1) for the MP method conform to those reported in earlier publications. It should be noted, however, this limit is SNR and range dependent (Fig. 4).

The last step is to limit the resulting DOAs by taking into consideration transmit and receive elements beampatterns or directions of interest. In principle solutions, due to noise, propagation and modelling errors, might lie anywhere within the $\pm 90°$ interval in relation to MRA. The limits for amplitude of complex roots together with acceptable DOAs form a selection gate (Fig. 5).

5. Conclusions

This article presented a new approach for determination of the model order (number of PCs/PEs) in the modified Prony method applied to the underwater DOA estimation. Low-rank approximation of correlation matrix is possible owing to utilization of a much greater number of receivers than the number of expected backscattered signals. The proposed method resolves model order by dynamically determining the noise eigenvalue threshold. Although, there are number of simplifying assumptions which might limit the validity of this method, the preliminary simulation results show good performance for simple, generic scenarios. Additional sources of coherence loss, such as volume reverberation and multipath, might also be included in calculation of equivalent SNR. Moreover, parameters required as input for calculations might be determined adaptively on-line or during post-processing. The proposed method of model order determination is closely tied to the underlying physical problem, and utilises as much as possible the information about the signal propagation environment. Single step model order determination might perform better than the iterative approach, and might be easier to implement due to lower algorithm complexity. This method will be further tested for its accuracy, resolution, performance and detection capabilities in more complicated acoustical environment scenarios; and compared with other methods utilized for the same problem formulation (partial results have already been published this year [4]). The next step of validation of this new approach will be carried out on signals recorded during sea surveys, and compared against results of different approaches applied to the same input signal.

References

[1] J. Akaeěl , Fitting Models in Time Series Analysis, Journal Series Statistics, vol. 13, no. 1, pp 121-148, 1982.

[2] B. Douglas, D. H. Johnson, Using the Sphericity Test for Source Detection with Narrow-Band Passive Arrays, IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 38, no. 11, pp 2008-2015, 1990.

[3] P. Grall, J. Marszal, Investigation into Interferometric Sonar System Accuracy, Hydroacoustics, vol.18, pp 69-76, 2015.

[4] P. Grall, J. Marszal, Zastosowanie zmodyfikowanej metody Prony'ego do detekcji dna morskiego i obiektów podwodnych, Advances in Acoustics 2017, pp 563-574, 2017.

[5] R.R. Hocking, R.L. Leslie, Selection of the Best Subset in Regression Analysis, Technometrix, vol. 9, no. 4, pp 531-540, 1967.

[6] J. N. Holt, R. J. Antill, Determining the Number of Terms in a Prony Algorithm Exponential Fit, Mathematical Biosciences , vol. 36, pp 319-323, 1977.

[7] G. Jin, D. Tang, Uncertainties of Differential Phase Estimation Associated with Interferometric Sonars, IEEE Journal of Oceanic Engineering, vol. 21, no. 1, pp 53-64, 1996.

[8] P. H. Kraeutner, J. S. Bird, Principal Components Array Processing for Swath Acoustic Mapping, Proceedings of the IEEE Oceans'97 Conference, pp 1246-1255, 1997.

[9] P. H. Kraeutner, J. S. Bird, Beyond Interferometry, Resolving Multiple Angles-of-Arrival in Swath Bathymetric Imaging, Proceedings of the IEEE Oceans'99 Conference, pp 37-46, 1999.

[10] P. H. Kraeutner, J. S. Bird, B. Charbonneau, D. Bishop, F. Hegg, Multi-Angle Swath Bathymetry Sidescan Performance Analysis, Proceedings of the IEEE Oceans'02 Conference, pp 2253-2264, 2002.

[11] R. Kumaresan, D.W. Tufts, L.L. Scharf, A Prony method for noisy data: Choosing the signal components and selecting the order in exponential signal models, Proceedings of the IEEE, vol. 72, no. 2, pp 230-233, 1984.

[12] X. Lurton, Swath Bathymetry Using Phase Difference: Theoretical Analysis of Acoustical Measurement Precision, IEEE Journal of Oceanic Engineering, vol. 25, no. 3, pp 351-364, 2000.

[13] Z. Li, H. Li, T. Zhou, Y. Yuan, Multiple sub-array beamspace CAATI algoritm for multibeam bathymetry system, Journal of Marine Science and Application, vol. 6, no. 1, pp 47-53, 2007.

[14] B. D. Rao, K. S. Arun, Model Based Processing of Signals: A State Space Approach, Proceedings of the IEEE, vol. 80 no. 2, pp 283-310, 1992.

[15] A. Rahman, K-B. Yu, Total Least Squares Approach for Frequency Estimation Using Linear Prediction, IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 35, no. 10, pp 1440-1455, 1987.

[16] M. P. Ribeiro, D. J. Ewins, D. A. Robb, Non-Stationary Analysis and Noise Filtering Using a Technique Extended from the Original Prony Method, Mechanical Systems and Signal Processing, vol. 17, no. 3, pp 533-550, 2003.

[17] D. W. Tufts, R. Kumaresan, Estimation of Frequencies of Multiple Sinusoids: Making Linear Prediction Perform Like Maximum Likelihood, Proceedings of the IEEE, vol. 70, no. 9, pp 975-990, 1982.

[18] A-J. van der VEEN, E. F. Deprettere, A. L. Swindelehurst, Subspace-Based Signal Analysis Using Singular Value Decomposition, Proceedings of the IEEE, vol. 81, no. 9, pp 1277-1309, 1993.

[19] M. Wax, T. Kaliath, Detection of Signals by Information Theoretic Criteria, IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 33, no. 2, pp 387-393, 1985.

[20] T. Zhou, H. Li, Z. Zhu, Y. Yuan, Application of modified multiple subarrays detection method to multibeam bathymetry system, Journal of Marine Science and Application, vol. 4, no. 2, pp 39-44, 2005.