SAMP++: Robust Structure-Aware Matrix Pencil

Yehonatan-Itay Segman, Alon Amar, and Ronen Talmon

Viterbi Faculty of Electrical and Computer Engineering,
Technion – Israel Institute of Technology, Haifa 32000, Israel.
yehonatans@campus.technion.ac.il, aamar@technion.ac.il, ronen@ef.technion.ac.il

***Abstract*—We address the problem of detecting the number of complex exponentials and estimating their parameters from a noisy signal using the Matrix Pencil (MP) method. Leveraging the recently introduced MP modes and their spectral structure, we propose SAMP++, a new MP algorithm that accurately detects the model order by exploiting mode interdependencies. Empirical results show that SAMP++ significantly outperforms standard MP implementations, prevalent information-based criteria, and the recently proposed SAMP algorithm.**

***Index Terms*—Matrix Pencil Method, model-order detection, parameter estimation, clustered frequencies, super-resolution.**

I. Introduction

THE classical signal model consisting of a sum of complex exponentials and additive noise has many applications in communications, audio, radar, and biomedical signal analysis [1], [2] and is given by:

$$y(n) = \sum_{i=1}^{M} b_i e^{(-d_i + j\theta_i)n} + w(n), \quad 0 \le n \le N-1. \tag{1}$$

where $w(n)$ is the additive noise, M is the number of complex exponentials (or model order), $b_i = |b_i|e^{j\phi_i}$ are the complex amplitudes, $\phi_i \in \mathbb{R}$ are the initial phases, $0 \le d_i \in \mathbb{R}$ are the damping factors and $\theta_i \in \mathbb{R}$ are the normalized frequencies. The complex exponentials $z_i = e^{-d_i + j\theta_i}$ are assumed to be distinct and termed the signal poles.

Given N samples of $y(n)$, the goal is twofold: (i) detect the (discrete-valued) model order M, and (ii) estimate the (continuous-valued) parameters of interest $\{b_i\}_{i=1}^M$, $\{d_i\}_{i=1}^M$, and $\{\theta_i\}_{i=1}^M$. This is a fundamental detection-estimation problem in signal processing, and numerous methods have been proposed for this problem, including maximum likelihood-based techniques [3], subspace-based approaches such as MVDR [4], MUSIC [5], ESPRIT [6], MODE [7], SPICE [8], and more recently, neural network-based approaches [9]–[11].

We consider the MP method, introduced in [12] as a super-resolution technique for estimating the parameters in (1). Numerous MP method variants exist, e.g., [13], [14], and they typically assume a known model order or determine it via singular value truncation. Recently, a novel approach for model order detection was proposed in [15], leveraging the concept of *MP modes*, which was inspired by Dynamic Mode Decomposition [16]. There, the left and right MP modes are defined as the columns and rows of the pseudoinverse matrix of the generalized eigenvectors matrix obtained by the MP method. These modes are shown to be of two types: signal modes and noise modes. The model order is determined by identifying the signal modes, using their expected structure. Unlike traditional approaches, this method exploits spectral information embedded in the eigenvectors, offering a fundamentally different perspective than existing methods, which almost exclusively rely on singular values.

In this paper, we present a new approach to leveraging the spectral information within MP modes, resulting in a robust MP algorithm, termed the Structure-Aware Matrix Pencil++ (SAMP++). We show that SAMP++ is highly effective for clusters of closely-spaced frequencies and low signal-to-noise ratio (SNR) values. Specifically, we show that in the presence of noise, signal modes exhibit interdependencies that can be exploited through correlation analysis, providing a reliable basis for enhanced model-order detection method.

We show in simulations that SAMP++ outperforms the recent SAMP method [15], which disregards mode interdependencies, existing MP implementations, which rely on singular value truncation for model-order detection, and the prevalent information-based criteria, which depend on the often-unknown likelihood function. This superiority is demonstrated in challenging scenarios with clustered frequencies and low SNR values.

II. Background: The Matrix Pencil Method

The MP method begins by constructing an $(N-L)\times(L+1)$ Hankel matrix $\mathbf{Y}$ from the noisy measurements $y(n)$:

$$\mathbf{Y} = \begin{bmatrix} y(0) & y(1) & \cdots & y(L) \\ y(1) & y(2) & \cdots & y(L+1) \\ \vdots & \vdots & \ddots & \vdots \\ y(N-L-1) & y(N-L) & \cdots & y(N-1) \end{bmatrix}, \tag{2}$$

where L is termed the *pencil parameter*. It is assumed that $M \le L \le N-M$ for the MP theorem to hold (see [12, Theorem 2.1] for more details), and that $\frac{N}{3} \le L \le \frac{N}{2}$ for minimal variance in the estimator [17], [18].

Next, two sub-matrices $\mathbf{Y}_0$ and $\mathbf{Y}_1$ are derived from $\mathbf{Y}$ by removing its last and first columns, respectively. Without noise, the ranks of $\mathbf{Y}_0$ and $\mathbf{Y}_1$ are M, making the detection problem trivial. However, in the presence of noise, $\mathbf{Y}_0$ and $\mathbf{Y}_1$ attain full rank L. Consequently, $L-M$ extraneous,

This work was supported by the European Union's Horizon 2020 research and innovation programme (grant No. 802735-ERC-DIFFOP).

noise-related components are introduced, and the model order needs to be determined in addition to the estimation of the parameters.

Standard implementations of the MP method are based on the Singular Value Decomposition (SVD) of $\mathbf{Y}_0$:

$$\mathbf{Y}_0 = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H, \tag{3}$$

where $\mathbf{U} \in \mathbb{C}^{(N-L)\times L}$ and $\mathbf{V} \in \mathbb{C}^{L\times L}$ are unitary matrices, $\mathbf{\Sigma} = \text{diag}(\sigma_1, \ldots, \sigma_L)$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_L$ are the singular values, and $(\cdot)^H$ denotes the complex conjugate transpose. The model order is usually detected by truncating the singular values in $\mathbf{\Sigma}$. Commonly used techniques are the Significant Decimal Digit (SDD) and the spectral gap [13], [19], [20]. Another widely used approach is the information-theoretic based estimators (ITE) [2], [21], which require prior knowledge of the likelihood function, and are computationally inefficient, as the likelihood must be computed repeatedly for each hypothesized model order.

After detecting the model order $\widehat{M}$, (3) is rewritten using the largest $\widehat{M}$ singular values and vectors. This yields the MP (a generalized eigenvalue problem):

$$\mathbf{Y}_1 - \lambda\mathbf{Y}_0 = \mathbf{U}\mathbf{\Sigma}(\mathbf{A} - \lambda\mathbf{I})\mathbf{V}^H, \tag{4}$$

where the square, non-symmetric matrix $\mathbf{A} \in \mathbb{C}^{\widehat{M}\times\widehat{M}}$ is:

$$\mathbf{A} = \mathbf{\Sigma}^{-1}\mathbf{U}^H\mathbf{Y}_1\mathbf{V}, \tag{5}$$

and λ denotes the MP eigenvalue. The MP method is based on finding the eigenvalues of $\mathbf{Y}_1 - \lambda\mathbf{Y}_0$ as an estimate for the signal poles z_i. Instead of directly solving (4), the estimates of the signal poles are obtained by the eigenvalues of $\mathbf{A}$ [12], which is favorable because $\widehat{M} \leq L = \text{rank}(\mathbf{Y}_0)$. Accordingly, an eigenvalue decomposition (EVD) of $\mathbf{A}$ is performed:

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}, \tag{6}$$

where $\mathbf{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_{\widehat{M}})$ and $\mathbf{Q} \in \mathbb{C}^{\widehat{M}\times\widehat{M}}$ contains the eigenvectors as columns. The estimates of the damping factors and normalized frequencies are given by:

$$\hat{d}_i = \log|\lambda_i| \ , \ \hat{\theta}_i = \arg(\lambda_i) \ , i = 1, 2, \ldots, \widehat{M}, \tag{7}$$

and the amplitudes $\{b_i\}_{i=1}^{\widehat{M}}$, are estimated by solving the linear least squares problem [13]:

$$\hat{b}_i = \arg\min_{b_i} \sum_{n=0}^{N-1} |y(n) - \lambda_i^n b_i|^2, \ \ 1 \leq i \leq \widehat{M}. \tag{8}$$

III. MP Modes and Their Interdependence

A new approach for detecting the model order, utilizing the temporal information encoded in the spectral components, was proposed in [15] with the introduction of the MP *modes*.

Definition 1 (MP modes)**.** *The right MP modes are defined as the rows of:*

$$\mathbf{\Phi}_R := \mathbf{Q}^{-1}\mathbf{V}^H \in \mathbb{C}^{L\times L}. \tag{9}$$

Similarly, the left MP modes are defined as the columns of:

$$\mathbf{\Phi}_L := \mathbf{U}\mathbf{\Sigma}\mathbf{Q} \in \mathbb{C}^{(N-L)\times L}. \tag{10}$$

In [15] it was shown that the L left modes can be divided into two types: M signal modes and $L - M$ noise modes, where each signal mode is associated with a signal pole. Without loss of generality, it is assumed that the M signal modes are the leftmost columns of $\mathbf{\Phi}_L$. It was further shown that under mild assumptions, any left signal mode $\mathbf{\Phi}_L^i$ can be recast as:

$$\mathbf{\Phi}_L^i \cong \begin{bmatrix} 1 \\ z_i \\ \vdots \\ z_i^{N-L-1} \end{bmatrix} + \sum_{\substack{m=1 \\ m\neq i}}^{M} \gamma_{i,m} \begin{bmatrix} 1 \\ z_m \\ \vdots \\ z_m^{N-L-1} \end{bmatrix} + \boldsymbol{\xi}^i \tag{11}$$

where $\mathbf{\Phi}_L^i$ is the i-th column of $\mathbf{\Phi}_L$. The coefficients $\gamma_{i,m}$ and noise-related vector $\boldsymbol{\xi}^i$ are controlled by the $\text{SNR}_i := \frac{|b_i|^2}{\sigma_w^2}$ and the i-th pole separation: $\min_{m\neq k}|z_i - z_m|$, assuming that the additive noise in (1) follows $w(n) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_w^2)$.

Equation (11) implies that the i-th signal mode $\mathbf{\Phi}_L^i$ carries information not only about its corresponding signal pole z_i but also about the remaining $M-1$ signal poles $\{z_m\}_{m\neq i}$. First, in the following proposition, we show that this interdependence is well characterized and controlled.

Proposition 1. *For any $\varepsilon \in (0, 1)$, $\gamma_{i,m}$ is bounded by:*

$$\sqrt{log\left(\frac{2}{2-\varepsilon}\right)}\frac{B_i}{3\Delta_i} \leq |\gamma_{i,m}| \leq \sqrt{log\left(\frac{2}{\varepsilon}\right)}\frac{B_i}{|z_i - z_m|}, \tag{12}$$

w.p of at least $1 - \varepsilon$, provided that $2|\delta z_i| \leq |z_i - z_m| \leq \Delta_i$, for some constant Δ_i, where: $B_i = 2\sqrt{\frac{2}{SNR_i}}\left\|\mathbf{u}_m * \mathbf{q}_i\right\|_2$.

Proof. See Appendix. □

The vector $\mathbf{u}_m$ is defined by: $\mathbf{u}_m^T := \left[0, \ \mathbf{p}_m^H\right] - z_i\left[\mathbf{p}_m^H, \ 0\right]$, where $\mathbf{p}_m$ and $\mathbf{q}_i$ are the m-th and i-th left and right eigenvectors of the noiseless MP, respectively (see [15, Section II.C] for more details).

Second, we assert that signal mode interdependence can be leveraged by computing the correlations between them, quantified through the sample covariance between the i-th and k-th signal modes[1]:

$$[\mathbf{C}]_{i,k} := \left(\mathbf{\Phi}_L^i\right)^H \mathbf{\Phi}_L^k. \tag{13}$$

Using (11), equation (13) can be written as[2]:

$$[\mathbf{C}]_{i,k} = \mathcal{D}(\theta_i - \theta_k) + \alpha_{i,k} + \beta_{i,k}, \tag{14}$$

where $\mathcal{D}(\theta) = \frac{\sin(0.5\theta(N-L))}{\sin(0.5\theta)}e^{-j0.5\theta(N-L-1)}$ is the Dirichlet kernel of order $N - L$, and θ_i and θ_k are the corresponding

[1]Assume the modes are centered, by subtracting their respective means.

[2]For simplicity, we assume that the signal is undamped; however, as shown in Sec. V, SAMP++ performs well for damped signals.

frequencies of z_i and z_k, respectively. The term $\alpha_{i,k}$ consists of the cross-terms that include only the coefficients $\gamma_{n,m}$:

$$\alpha_{i,k} = \sum_{\substack{m=1 \\ m\neq k}}^{M} \gamma_{k,m}\mathcal{D}(\theta_m - \theta_i) + \sum_{\substack{m=1 \\ m\neq i}}^{M} \overline{\gamma}_{i,m}\mathcal{D}(\theta_k - \theta_m) + \sum_{\substack{m=1 \\ m\neq i}}^{M}\sum_{\substack{n=1 \\ n\neq k}}^{M} \overline{\gamma}_{i,m}\gamma_{k,n}\mathcal{D}(\theta_m - \theta_n), \tag{15}$$

and $\beta_{i,k}$ represents the remaining cross-terms, which include the noise vectors $\boldsymbol{\xi^i}$ and $\boldsymbol{\xi^k}$ controlled by SNR_i and SNR_k, respectively [15, Section III, Proposition 3].

We argue that for clustered frequencies, $\alpha_{i,k}$ contributes non-negligibly to the sample covariance in (14). We call θ_i and θ_m clustered if $|\theta_i - \theta_m| \ll \frac{2\pi}{N-L}$. Applying proposition 1 for $\Delta_i \ll \frac{2\pi}{N-L}$, we have:

$$|\gamma_{i,m}| \gg B_i\sqrt{\log\left(\frac{2}{2-\varepsilon}\right)}\frac{N-L}{6\pi}, \quad m \neq i. \tag{16}$$

Observing (15), this suggests that for clustered frequencies $|\alpha_{i,k}|$ contributes non-negligibly to $|[\mathbf{C}]_{i,k}|$ in (14). This stems from (i) non-negligible contributions of the coefficients $\gamma_{i,m}$ and $\gamma_{k,m}$ based on (16), and (ii) the corresponding terms $\mathcal{D}(\theta_k - \theta_m)$ and $\mathcal{D}(\theta_m - \theta_i)$ that attain large values, as $|\mathcal{D}(\theta)| \approx N - L$ for $|\theta| \ll \frac{2\pi}{N-L}$.

The analysis above suggests that when frequency clusters are present, the i-th row of $\mathbf{C}$ contains several high-value entries corresponding to signal modes associated with frequencies in the same cluster as the i-th frequency. In the sequel, these values serve as explicit cues for developing a robust model-order detection method. Conversely, since noise modes lack a specific structure, their corresponding rows are expected to have uniformly low values, reflecting weak correlations with other modes (see Fig. 1 for illustration).

IV. Proposed Algorithm

Based on Sec. III, we propose a new model-order detection method that identifies signal modes by considering their interdependence. We define the following feature to determine whether the i-th mode is associated with a signal or noise:

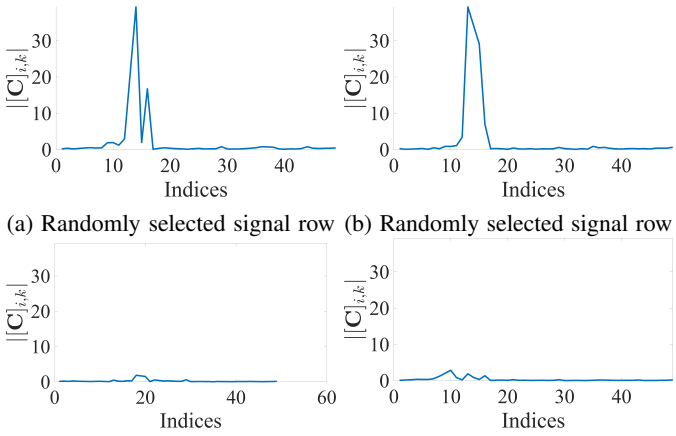
$$\epsilon_i := a\rho_i + (1-a)\sigma_i, \quad 1 \leq i \leq L, \quad 0 \leq a \leq 1. \tag{17}$$

The proposed feature is a convex sum of two terms; the first term ρ_i is adopted from [15]:

$$\rho_i = \frac{1}{d_{\theta_i}} \max_{1\leq k\leq N-L} |\mathcal{F}(\mathbf{\Phi}_L^i)[k]|, \tag{18}$$

where $\mathcal{F}(\mathbf{\Phi}_L^i)[k]$ is the DFT of $\mathbf{\Phi}_L^i$ at the k-th frequency bin, $d_{\theta_i} = \sum_{m=1}^{L} |\hat{\theta}_i - \hat{\theta}_m|^2$ is a measure of the frequencies spread around the spectrum peak, and $\{\hat{\theta}_m\}_{m=1}^L$ are the estimated frequencies extracted from the corresponding poles, according to (7). The second term σ_i is the standard deviation of the magnitudes of the off-diagonal elements in the i-th row of $\mathbf{C}$:

$$\sigma_i := \text{std}(\{|[\mathbf{C}]_{i,k}|\}_{k\neq i}). \tag{19}$$



(a) Randomly selected signal row (b) Randomly selected signal row

(c) Randomly selected noise row (d) Randomly selected noise row

Fig. 1: Signal-related rows of $\mathbf{C}$ (excluding the main diagonal) in (a) and (b), and noise-related rows of $\mathbf{C}$ (excluding the main diagonal) in (c) and (d). We simulate the signal (1) with $M = 4$ undamped poles grouped into two frequency clusters where $N = 151$, $L = 50$ and $\text{SNR} = 20$ [dB].

The scalar a balances between mode energy and mode correlations; in our experiments, we set $a = 0.5$, assuming that ρ_i and σ_i are normalized to the interval $[0, 1]$.

Intuitively, large values of ϵ_i indicate that $\mathbf{\Phi}_L^i$ is likely a signal mode due to high energy or strong correlations to other signal modes associated with frequencies within the same cluster of frequencies. Conversely, as noise modes lack a particular structure, small values of ϵ_i indicate that $\mathbf{\Phi}_L^i$ is likely a noise mode. The set of features $\{\epsilon_i\}_{i=1}^L$ in (17) is used to determine the model order by dividing them into two distinct subsets: signal-related and noise-related. The cardinality of the signal-related subset then determines the model order. We propose to use the k-means algorithm with $k = 2$ for implementing this division.

The proposed algorithm, termed SAMP++ , is summarized in Algorithm 1. SAMP++ reduces to the SAMP algorithm from [15] by setting $a = 1$ in (17). In Sec. V, we will show that SAMP++ demonstrates enhanced model order detection capabilities for clusters of frequencies with $M > 2$. Similarly to SAMP, SAMP++ is robust to noise and distribution-free. As in SAMP [15, Section VI], we apply a weak-truncation method to the SVD of $\mathbf{Y}_0$ in step 1 of Algorithm 1.

V. Numerical results

We present simulation results considering two clusters of frequencies and low SNR values. We focus on the detection capabilities of the proposed SAMP++ (Algorithm 1), as the estimation step is performed in the same manner in all the competing methods using the eigenvalues as described in (7). We compare the proposed SAMP++ with five other methods, including SAMP.

We consider two ITE methods: the MDL [2] and EVT [21] criteria. These methods require prior knowledge of the noise distribution and involve computing the likelihood function for

Algorithm 1 SAMP++ proposed algorithm

Input: Noisy measurements $\{y(n)\}_{n=1}^{N}$, pencil parameter L.
Output: Estimates of the signal frequencies, damping factors, and amplitudes.

1: Modes and Eigenvalues Computation:
- Construct $\mathbf{Y}_0, \mathbf{Y}_1$ from the measurements $\{y(n)\}_{n=0}^{N-1}$ by (2), and compute the SVD of $\mathbf{Y}_0 = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$.
- Compute the EVD of $\mathbf{A} = \mathbf{\Sigma}^{-1}\mathbf{U}^H\mathbf{Y}_1\mathbf{V}$, by (6).
- Calculate the MP modes $\mathbf{\Phi}_R$ and $\mathbf{\Phi}_L$ by (9) and (10).

2: Parameter Estimation:
- Estimate the frequencies and damping factors by (7):

$$\hat{\theta}_i = \arg(\lambda_i), \quad \hat{d}_i = \log|\lambda_i|, \quad 1 \le i \le L.$$

- Estimate the complex amplitudes by [15, Eq. (43)]:

$$\hat{\mathbf{b}} = (\mathbf{e}_1^T\mathbf{\Phi}_L)^T \odot (\mathbf{\Phi}_R\mathbf{e}_1).$$

3: Model Order Detection:
- Compute the L features by (17):

$$\epsilon_i = a\rho_i + (1-a)\sigma_i, \quad 1 \le i \le L.$$

- Partition the set $\{\epsilon_i\}_{i=1}^{L}$ into two distinct subsets and select the subset with the highest average value.
- Set the model order to: $\widehat{M} = |\mathcal{S}|$, where $\mathcal{S}$ is the index set of the subset with the highest average value.

4: Parameter Selection:
- Select the signal components by: $\hat{\theta}_s, \hat{d}_s, \hat{b}_s, \quad s \in \mathcal{S}$.

each hypothesized model order, making them computationally inefficient. We note that we also evaluated the AIC criterion [2] which achieved much lower results and was omitted. The remaining methods rely solely on analyzing the singular values in (3), the standard approach in the MP literature. Specifically, we evaluate the SDD and GAP methods, which are widely used for SVD truncation [13], [19], [20].

We simulate a sum of four complex exponentials grouped into two clusters, each containing two closely-spaced frequencies, following (1) with $M = 4$, $\{b_i = 1\}_{i=1}^{4}$, and $\{\phi_i = 0\}_{i=1}^{4}$. Additionally, we set $L = \text{round}(N/3)$. We conduct two experiments. In the first experiment, we consider four *undamped* exponentials by setting $\{d_i = 0\}_{i=1}^{4}$ in (1), while in the second experiment, we consider four *damped* exponentials by setting $d_1 = d_3 = 0.03$, and $d_2 = d_4 = 0.05$ in (1). Given a finite sample of the noisy signal $\{y(n)\}_{n=0}^{N-1}$, as in (1), our goal is to detect the model order (the number of complex exponentials), which is 4 in our simulations. The code to reproduce these results is available in this GitHub link.

We determine the probability of correctly identifying the model order, $p_d = \mathbb{P}(\widehat{M} = M) \cong \frac{1}{N_{\text{exp}}}\sum_{i=1}^{N_{\text{exp}}} \mathbb{1}_{\widehat{M}=M}$, where $N_{\text{exp}} = 1000$ is the number of independent Monte-Carlo trials, and $\mathbb{1}_{\widehat{M}=M}$ is an indicator function that equals 1 if the estimated model order $\widehat{M}$ matches the true model order M, and 0 otherwise. We also present the Area Under the Curve (AUC) for each method to simplify the comparisons.

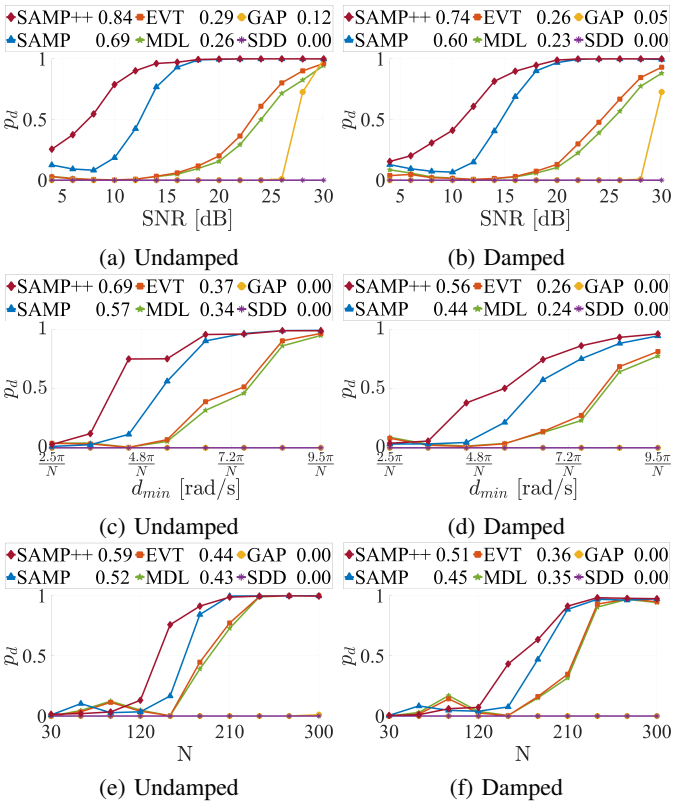


Fig. 2: Probability of correct model order detection versus the SNR_{dB} in (a) and (b), the distance between clusters of frequencies d_{min} in (c) and (d), and the number of samples N in (e) and (f). SNR $= 10$ [dB] in (c)-(f).

Fig. 2a-2b display p_d versus the SNR in dB, for $N = 151$ samples, $\theta_1 = 2$ rad/sample, $\theta_3 = 4$ rad/sample and $\theta_2 = \theta_1 + \frac{1.5\pi}{N}$, $\theta_4 = \theta_2 + \frac{1.5\pi}{N}$. The values of θ_2 and θ_4 are chosen such that the spacing $|\theta_2 - \theta_1| = |\theta_4 - \theta_3| = \frac{1.5\pi}{N}$ is smaller than the Rayleigh limit [22] and also smaller than $\frac{2\pi}{N-L}$, forming two clusters as defined in Sec. III. The SNR is defined as $\text{SNR} = \sum_{i=1}^{M}\frac{|b_i|^2}{\sigma_w^2}$. Fig. 2c-2d display p_d versus the minimal distance between the two clusters of frequencies, denoted as d_{min}, for SNR $= 10$ [dB], $N = 151$ samples, and the frequencies distance within each cluster is also $\frac{1.5\pi}{N}$. Fig. 2e-2f display p_d versus the number of samples N, for SNR $= 10$ [dB], and $\{\theta_i\}_{i=1}^{4}$ are the same as in Fig. 2a-2b.

We observe that SAMP++ significantly outperforms the SDD and GAP methods across all tested scenarios. Empirically, SDD tends to overestimate the model order, while GAP typically underestimates it, leading to poor results. Additionally, SAMP++ outperforms the MP-based MDL and EVT criteria, even though it does not assume any prior knowledge about the noise distribution. Finally, we see that SAMP++ achieves better results compared to the SAMP method in all the presented scenarios, demonstrating the contribution of considering modes correlations as described in Sec. IV.

VI. Conclusion

We suggested a modified MP method termed SAMP++ for the complex exponentials detection-estimation problem that leverages the concept of MP modes. Using the MP modes and their interdependencies, we developed a robust model-order detection method that integrates mode correlations and energy. Our simulations show that the proposed method outperforms the recently introduced SAMP method, as well as standard singular value-based methods, and common information-based approaches, which assume prior knowledge of the likelihood function and are computationally inefficient.

Appendix A
Upper and lower bound of $|\gamma_{i,m}|$

The coefficients $\gamma_{i,m}$ are defined by [15, Sec. III]: $\gamma_{i,m} = \frac{\mathbf{u}_m^T \mathbf{Q}_i \mathbf{w}}{(\widetilde{z}_i - z_m) b_i}$, where $\widetilde{z}_i$ is a perturbation of z_i. Using the proof of Proposition 3 in [15], it holds that $\mathbb{E}[\bar{\gamma}_{i,m}] = 0$ and $\mathrm{Var}(\bar{\gamma}_{i,m}) = \frac{\|\mathbf{u}_m * \mathbf{q}_i\|_2^2}{\mathrm{SNR}_i}$, for $\bar{\gamma}_{i,m} = \frac{\mathbf{u}_m^T \mathbf{Q}_i \mathbf{w}}{b_i}$, and that $|\bar{\gamma}_{i,m}|$ follows a Rayleigh distribution with scale parameter $\mathrm{Var}(\bar{\gamma}_{i,m})$. Assuming that $2|\delta z_i| \le |z_i - z_m| \le \Delta_i$, where $\delta z_i = \widetilde{z}_i - z_i$, we derive the following bound:

$$\frac{|z_i - z_m|}{2} \le |\widetilde{z}_i - z_m| \le \frac{3}{2}|z_i - z_m| \le \frac{3}{2}\Delta_i, \tag{20}$$

using standard triangle inequalities. Utilizing (20), we derive the following bounds, with a probability of at least $1 - \varepsilon$:

$$|\gamma_{i,m}| \ge \frac{2}{3\Delta_i}\sqrt{\frac{2\log(\frac{2}{2-\varepsilon})}{\mathrm{SNR}_i}}\,\|\mathbf{u}_m * \mathbf{q}_i\|_2\,, \tag{21}$$

$$|\gamma_{i,m}| \le 2\sqrt{\frac{2\log(\frac{2}{\varepsilon})}{\mathrm{SNR}_i}}\,\frac{\|\mathbf{u}_m * \mathbf{q}_i\|_2}{|z_i - z_m|} \tag{22}$$

References

[1] D. B. Percival and A. T. Walden, *Spectral analysis for physical applications*. cambridge university press, 1993.

[2] P. Stoica, R. L. Moses, *et al.*, *Spectral analysis of signals*. Pearson Prentice Hall Upper Saddle River, NJ, 2005, vol. 452.

[3] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 34, no. 5, pp. 1081–1089, 1986.

[4] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969.

[5] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, 1986.

[6] R. Roy and T. Kailath, "Esprit-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, no. 7, pp. 984–995, 1989.

[7] P. Stoica and K. C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 38, no. 7, pp. 1132–1143, 1990.

[8] P. Stoica, P. Babu, and J. Li, "New method of sparse parameter estimation in separable models and its use for spectral analysis of irregularly sampled data," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 35–47, 2010.

[9] Y. Xie, M. B. Wakin, and G. Tang, "Data-driven parameter estimation of contaminated damped exponentials," in *2021 55th Asilomar Conf. on Signals, Syst., and Computers*, IEEE, 2021, pp. 800–804.

[10] P. Pan, Y. Zhang, Z. Deng, and W. Qi, "Deep learning-based 2-d frequency estimation of multiple sinusoidals," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 33, no. 10, pp. 5429–5440, 2021.

[11] J. W. Smith and M. Torlak, "Frequency estimation using complex-valued shifted window transformer," *IEEE Geosci. Remote Sens. Lett.*, 2024.

[12] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 38, no. 5, pp. 814–824, 1990.

[13] T. K. Sarkar and O. Pereira, "Using the matrix pencil method to estimate the parameters of a sum of complex exponentials," *IEEE Antennas Propag. Mag.*, vol. 37, no. 1, pp. 48–55, 1995.

[14] F. Hu, T. K. Sarkar, and Y. Hua, "Utilization of bandpass filtering for the matrix pencil method," *IEEE Trans. Signal Process.*, vol. 41, no. 1, pp. 442–442, 1993.

[15] Y.-I. Segman, A. Amar, and R. Talmon, *Structure-aware matrix pencil method*, 2025. arXiv: 2502.17047.

[16] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *J. Fluid Mech.*, vol. 656, pp. 5–28, 2010.

[17] D. Rao, "Perturbation analysis of a svd based method for the harmonic retrieval problem," in *Int. Conf. on Acoust. Speech Signal Process. (ICASSP)*, IEEE, vol. 10, 1985, pp. 624–627.

[18] E.-H. Djermoune and M. Tomczak, "Perturbation analysis of subspace-based methods in estimating a damped complex exponential," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4558–4563, 2009.

[19] M. Bhuiyan, E. V. Malyarenko, M. A. Pantea, F. M. Seviaryn, and R. G. Maev, "Advantages and limitations of using matrix pencil method for the modal analysis of medical percussion signals," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 2, pp. 417–426, 2012.

[20] H. Yin, Z. Zhu, and F. Ding, "Model order determination using the hankel matrix of impulse responses," *Appl. Math. Lett.*, vol. 24, no. 5, pp. 797–802, 2011.

[21] B. Nadler and A. Kontorovich, "Model selection for sinusoids in noise: Statistical analysis and a new penalty term," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1333–1345, 2011.

[22] S. Dharanipragada and K. Arun, "Resolution limits in signal recovery," *IEEE Trans. Signal Process.*, vol. 44, no. 3, pp. 546–561, 1996.