14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

PHASE NOISE MITIGATION IN THE AUTOCORRELATION ESTIMATES WITH DATA WINDOWING: THE CASE OF TWO CLOSE SINUSOIDS

Mustafa A. Altınkaya[1], Emin Anarım[2] and Bülent Sankur[2]

[1] Department of Electrical & Electronics Engineering,
İzmir Institute of Technology, İzmir, Türkiye
phone: + (90) 232 750 6279, fax: + (90) 232 750 6196,
email: mustafaaltinkaya@iyte.edu.tr,
web: www.iyte.edu.tr/~mustafaaltinkaya

[2] Department of Electrical & Electronics Engineering,
Boğaziçi University, İstanbul, Türkiye
phone: + (90) 212 359 6414, fax: + (90) 212 2872465,
email: {anarim, bulent.sankur}@boun.edu.tr
web: busim.ee.boun.edu.tr/{academic/~anarim, ~sankur}

ABSTRACT

We address the phase noise and the superresolution problem in Toeplitz matrix-based spectral estimates. The Toeplitz autocorrelation (AC) matrix approach in spectral estimation brings in an order of magnitude computational advantage while the price paid is the phase noise that becomes effective at high signal-to-noise ratios (SNR). This noise can be mitigated with windowing the data though some concomitant loss in resolution occurs. The trade-offs between additive noise SNR, resolvability of sinusoids closer than the resolution limit, and behavior of the estimated AC lags and tone frequencies are investigated.

1. INTRODUCTION

Model-based frequency estimation methods are used extensively due to their computational advantages and high resolution property. Especially, for short data records model-based spectral estimation techniques show superior performance under adequate signal-to-noise ratio (SNR) conditions. Many of these techniques make use of autocorrelation (AC) matrix and its eigendecomposition. The AC matrix is usually generated using the covariance method of linear prediction [1]. Computing the eigenvalues and eigenvectors of such symmetric arbitrary real matrices, is of $O(M^3)$ complexity [3], where the notation $O(\cdot)$ denotes the order of required complex multiplications and M is the size of the AC matrices. On the other hand, if the AC matrix is generated using the AC method of linear prediction resulting in a Toeplitz structure, the complexity of the eigendecomposition process is of order $O(M^2)$ [2]. This complexity differential motivates us for subspace-based frequency estimators utilizing Toeplitz AC matrix estimates. However, linear prediction based on Toeplitz AC matrices suffer from the "phase noise" phenomenon. This phenomenon reveals itself at high SNR values as a "noise floor" effect. In other words, even though the influence of the additive noise has been reduced to a negligible level, the accuracy of the frequency estimate does not improve proportional to the ever increasing SNR. However, in high SNR conditions, for example, in time of arrival (TOA) and time difference of arrival (TDOA) estimation in order to determine the position of a mobile terminal, frequency estimators like non-Toeplitz AC matrix-based Multiple Signal Characterization (or Classification) (MUSIC) are successful [7]. In a previous work [4], we showed that in the frequency estimation of a single sinusoid in additive white Gaussian noise (AWGN) from short data records based on Toeplitz AC matrices, data windowing can mitigate the limitations caused by phase noise. In this work, we extend that study to the case of multiple sinusoids where we consider the interesting particular case of two close sinusoids.

In Section 2, the sinusoidal frequency estimation problem and the utilized frequency estimator, namely, MUSIC frequency estimator is briefly visited. In Section 3, for the case of two closely spaced sinusoids observed in AWGN the pdf of AC lags are found both by histograms of simulation experiments and through semi-analytic computations, and the similarities and differences with the single sinusoid case is discussed. Section 4 covers the experimental study of the resolution capability of the MUSIC frequency estimator for the windowed and non-windowed data. In Section 5, we investigate the joint distribution of two AC lags, in order to clarify the mechanism of success in phase noise mitigation for data windowing for the considered two close sinusoids problem. Conclusions are given in Section 6. Finally, the derivation of the pdf of AC lags for the windowed data for small sample sizes which also takes the phase noise into account is given in the appendix.

The first author is supported in part by European Commission FP6 "IYTE Wireless" Project, contract no: 017442.

2. FREQUENCY ESTIMATION PROBLEM AND MUSIC FREQUENCY ESTIMATOR

The signal model under consideration consists of multiple real sinusoids observed in AWGN, i.e.,

$$x_k = s_k + n_k = \sum_{i=1}^{K} \sqrt{2A_i}\cos[\omega_i kT + \phi_i] + n_k \quad k = 1,2,...N \qquad (1)$$

where A_i, ϕ_i and ω_i are the non-random amplitude, the random phase angle uniformly distributed on $(-\pi,\pi)$ and the angular tone frequency of the ith real sinusoid, respectively, and T is the sampling period, $\{n_k\}$ is a real white Gaussian noise sample sequence with zero mean and power σ_n^2 and N is the number of data samples. We are interested only in the angular frequency parameter. The other parameters are considered as nuisance factors. We assume without loss of generality that the number of sinusoids is either known or can be estimated from the data. In this work we assess a widely used frequency estimator which uses high dimensional noise-subspace resulting in well resolved frequency estimates, namely, MUSIC, which we will define briefly. We will assume $T = 1$ throughout the paper in order to simplify the notation.

To obtain the MUSIC frequency estimator first the Toeplitz sample estimate of the $M \times M$ AC matrix is built as $\{\mathbf{R}_M(i,j) = r(|i-j|), i,j = 1,\ldots,M\}$ where $r(k)$ denotes the kth autocorrelation coefficient of the input samples and it is estimated as:

$$r(k) = \frac{1}{N-k}\sum_{i=1}^{N-k} x_i x_{i+k}. \qquad (2)$$

The MUSIC power spectrum is given as:

$$\text{MUSIC}(\omega) = \frac{1}{\sum_{k=2K+1}^{m} |\mathbf{p}^{\dagger}(\omega)\mathbf{v}_k|^2} \qquad (3)$$

where the set $\{\mathbf{v}_i, i = 2K+1,...,M\}$ contains the M-dimensional noise subspace eigenvectors corresponding to the $M-2K$ smallest eigenvalues of the AC matrix estimate and $\mathbf{p}^{\dagger}(\omega) = [1\ e^{-j\omega}\ \ldots\ e^{-j(M-1)\omega}]$. The estimated tone frequency (or frequencies) is (are) found by picking the K peak-pairs of this power spectrum which correspond to K real tones.

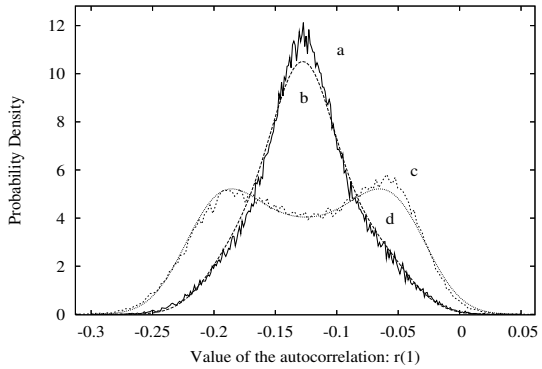
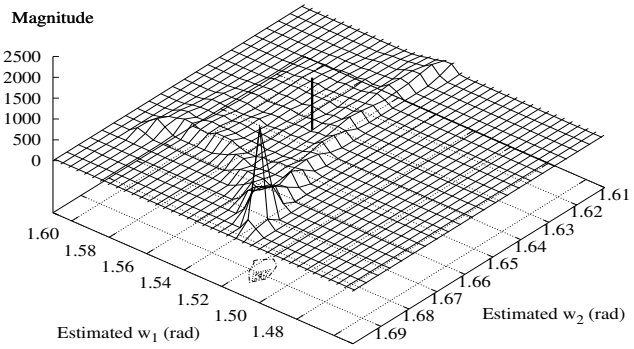


Figure 1: Pdf of $r_y(1)$, a: histogram (non-windowed data), b: pdf (semi-analytic compution, non-windowed data), c: histogram (windowed data), d: pdf (semi-analytic compution, windowed data) ($\omega_1 = \pi/2$ rad, $\omega_2 = \pi/2 + \pi/25$ rad, $\omega_2 - \omega_1 = \pi/N =$ DFT resolution limit, SNR $= 10$ dB, $N = 50$, $M = 20$, 100 000 noise realizations, raised cosine window [6])

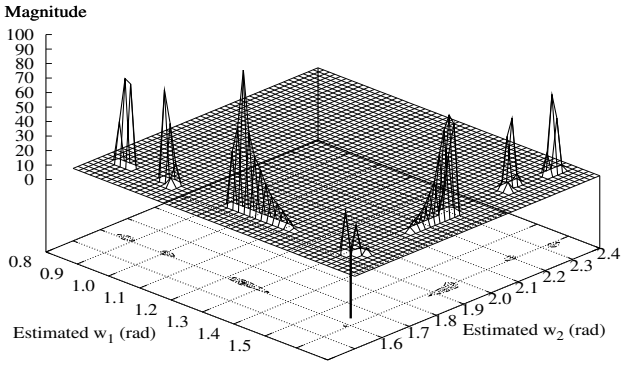
3. BEHAVIOR OF AC LAGS FOR WINDOWED DATA

For small sample sizes the AC lags can be represented as a sum of two assumably independent random variables, a phase noise-dependent component and an AWGN-dependent component with a deterministic shift. Then, the pdf of AC lags becomes a convolution of the individual pdfs and its shape will be dominated by the component of larger support which is the Gaussian and phase noise-dependent component at low and high SNR values, respectively. In [4] a direct dependence between the supports of the pdfs of AC lag and frequency estimates was observed and it was shown that data windowing helps shrinking the support of the pdfs of phase noise-dependent AC lag component and frequency estimates, and in turn decreases the variance lower bound caused by phase noise. The cost of windowing will be a power loss of typically 2 dB in SNR due to the increased equivalent noise bandwidth [1, 6] in the case of a single sinusoid. This same effect is observed in the the case of well resolved sinusoids where the frequency difference of the sinusoids is much larger than the Fourier resolution limit of $2\pi/N$. If the distance of two peaks in the power spectrum gets smaller, the spectral broadening caused by data windowing can be expected to make the two peaks interact earlier due to the accompanying loss in resolution. However, as it will be shown in the subsequent sections data windowing helps the frequency estimator even at high resolution requiring frequency constellations.

In Figure 1 we plotted the pdfs of $r_y(1)$ for the windowed and non-windowed data together with the histograms obtained as a result of simulations with 100,000 independent runs of two sinusoids in AWGN where $\omega_1 = \pi/2$ rad, $\omega_2 = \pi/2 + \pi/25$ rad, $N = 50$ and the subscript "y" in $r_y(1)$ denotes the association with windowed data. The pdfs are computed semi-analytically as described in the appendix. The frequency difference of the sinusoids is equal to the Fourier resolution limit $2\pi/N$ and SNR = 10 dB. The windowing function is the raised cosine window which showed a good compromise of contradicting phase noise reduction and small power loss requirements, in the case of a single sinusoid in AWGN. The figure depicts that there is close matching of the histograms and the computed pdfs. This validates our assumptions related to the summands in (5). But, Figure 1 depicts that windowing does not shrink the support of the pdf of phase noise in the case of two close sinusoids. Then there should be another mechanism by which data windowing contributes to the resolution ability of the considered frequency estimators, if there is any. We will consider this after investigating the frequency estimation performance with and without data windowing.



(I) Resolved frequency estimates (57.0 %),



(II) Unresolved frequency estimates (43.0 %),

Figure 2: Joint histograms of the two MUSIC frequency estimates for **non-windowed** data ($\omega_1 = 1.5708$ rad, $\omega_2 = 1.6336$ rad, $\omega_2 - \omega_1 = \pi/N$ rad (half of the DFT resolution limit), SNR $= 40$ dB, $N = 50$, $M = 20$, 100 000 noise realizations)

4. EXPERIMENTAL STUDY OF THE RESOLVING CAPABILITY

When data windowing is applied in the case of two close sinusoids, one should consider the price paid for the phase noise reduction, that is, the spectral broadening and the concomitant loss of resolution. As a criterion to measure the resolution capability of the estimators we adopt the resolution probability where it is required that the inequality

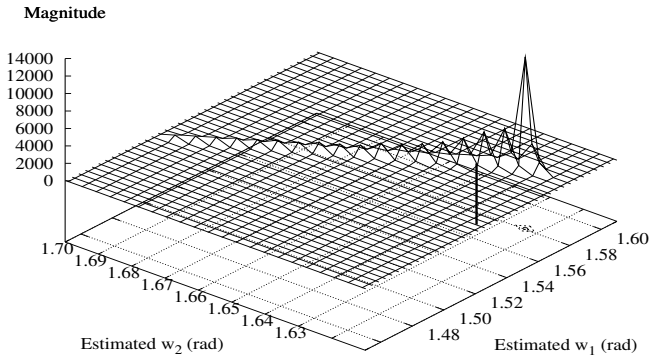
$$(\hat{\omega}_1 - \omega_1)^2 + (\hat{\omega}_2 - \omega_2)^2 \quad < \quad \min\Big((\hat{\omega}_1 - \omega_{\text{mid}})^2 + (\hat{\omega}_2 - \omega_{\text{mid}})^2\,, \; 2(\omega_2 - \omega_1)^2\Big) \tag{4}$$

is satisfied in order to consider two frequencies resolved, where $\{\hat{\omega}_i, i = 1, 2\}$ and ω_{mid} denote the estimates and the arithmetic average of the sinusoidal frequencies $\{\omega_i, i = 1, 2\}$, respectively, and $\min(\cdot, \cdot)$ denotes the minimum of it's arguments. The first argument of the $\min(\cdot, \cdot)$ function assures that the estimated peaks are not both located in the vicinity of ω_{mid} whereas the second argument is to reject the outliers in the histograms.

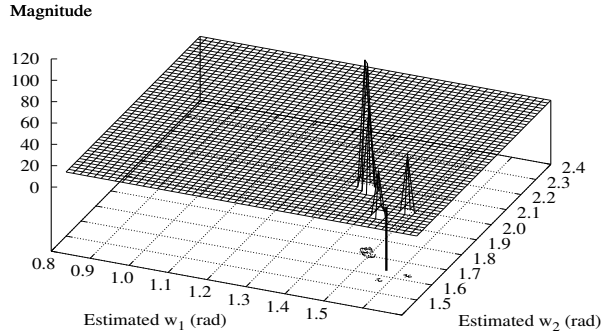
In the simulations, the data window is the raised cosine window which achieves moderate phase noise reduction at a small expense in the estimation variance.

4.1 Histograms of frequency estimates

In order to focus on the effect of phase noise in the resolution performance of the frequency estimators, firstly we investigate the histograms of the estimated frequencies in the case of two sinusoids with a frequency difference of π/N and where SNR = 40 dB.



(I) Resolved frequency estimates (84.9 %),



(II) Unresolved frequency estimates (15.1 %),

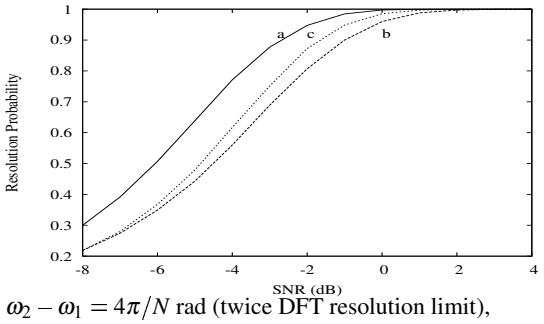
Figure 3: Joint histograms of the two MUSIC frequency estimates for **windowed** data ($\omega_1 = 1.5708$ rad, $\omega_2 = 1.6336$ rad, $\omega_2 - \omega_1 = \pi/N$ rad (half of the DFT resolution limit), SNR $= 40$ dB, $N = 50$, $M = 20$, 100 000 noise realizations)

In Figures 2 and 3 we plot the joint histograms of the two MUSIC frequency estimates for the non-windowed and windowed data, respectively. The location of the true sinusoidal frequencies, $\omega_1 = \pi/2 = 1.5708$ rad and $\omega_2 = \pi/2 + \pi/50 = 1.6336$ rad, is shown with a vertical line segment in the figures. In the non-windowed case sinusoids are resolved in 57 % of realizations and windowing increases this ratio to 84.8 % even at this frequency difference of the sinusoids which is equal to the half of the Fourier resolution limit.

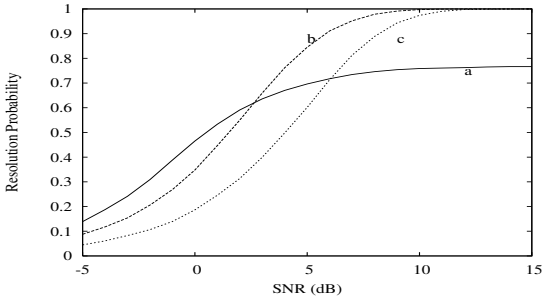
Figure 2-I depicts that without windowing the majority of the resolved frequency estimates are located at $\omega_1 \approx 1.51$ rad and $\omega_2 \approx 1.69$ rad which corresponds to a large bias in the frequency estimates whereas the majority of the resolved frequency estimates are in the 0.01 rad approximity of their true values in the windowed case as shown in Figure 3-I and all of the frequency estimates are symmetrically distributed around the average value of the true frequency parameters.

It is shown in Figure 2-II that for the non-windowed case the frequency estimates are labelled "unresolved" when one of them is in large error. Actually, only one sinusoid is detected in this case with the frequency around the average of the two sinusoids and for the frequency of the other sinusoid, MUSIC finds another frequency determined according to the severe phase noise in the AC lag estimates. On the other hand, for the windowed case the unresolved estimates are either on one side of the average of the true frequencies corresponding to the two small peaks in Figure 3-II or located symmetrically and somewhat far from the true frequencies.

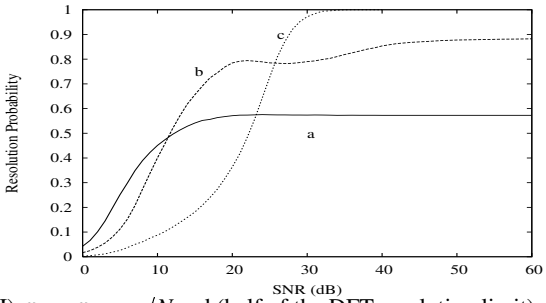
Figures 2 and 3 show that the adopted criterion successfully classifies the frequency estimates into resolved and unresolved sets as shown by the histograms and data windowing is beneficial even in the case of very closely spaced sinusoids.



(I) $\omega_2 - \omega_1 = 4\pi/N$ rad (twice DFT resolution limit),



(II) $\omega_2 - \omega_1 = 2\pi/N$ rad (DFT resolution limit),



(III) $\omega_2 - \omega_1 = \pi/N$ rad (half of the DFT resolution limit)

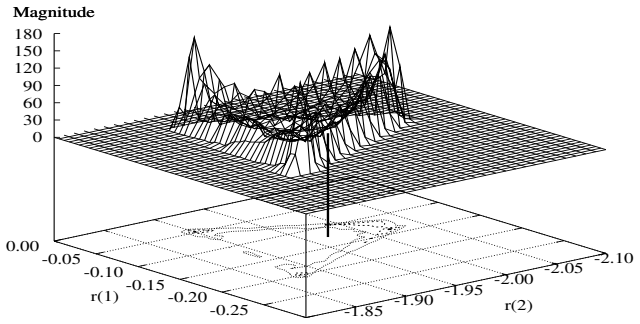
Figure 4: Resolution probability of the MUSIC frequency estimators versus SNR, a: rectangular window (Toeplitz AC matrix), b: raised cosine window (Toeplitz AC matrix), c: rectangular window (non-Toeplitz AC matrix) ($\omega_1 = 1.57$ rad, $N = 50$, $M = 20$)

4.2 Resolution Probabilities

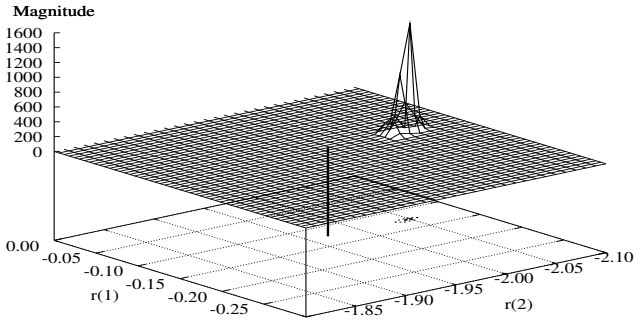
In Figures 4-I, II and III, we plot the resolution probabilities of MUSIC frequency estimator with and without data windowing against SNR, for two sinusoids where their frequency difference is $4\pi/N$ rad, $2\pi/N$ rad and π/N rad, respectively, going from a well resolved case towards a case with a frequency difference of half of the Fourier resolution limit. In these figures we plot the resolution probabilities of the non-windowed, windowed, and non-Toeplitz AC matrix-based variants of the MUSIC frequency estimator.

4.2.1 Twice DFT Resolution Limit Separation:

Since the sinusoids are well resolved, the phase noise variance is ineffective on the resolution of the frequency estimator and the use of a data window only decreases the resolution probability of the frequency estimator. At a resolution probability of 0.9, MUSIC frequency estimator based on the non-Toeplitz AC matrix estimate and the one utilizing the raised cosine window require approximately 1.6 dB and 2.4 dB more signal power as compared to no data windowing.



(I) Corresponding to **resolved** frequency estimates (77.1 %),



(II) Corresponding to **unresolved** frequency estimates (22.9 %),

Figure 5: Joint histograms of the $r_y(1)$ and $r_y(2)$ estimates for **non-windowed** data ($\omega_1 = 1.5708$ rad, $\omega_2 = 1.6965$ rad, $\omega_2 - \omega_1 = 2\pi/N$ rad (DFT resolution limit), SNR = 40 dB, $N = 50$, $M = 20$, 100 000 noise realizations)

4.2.2 Separation Equal to DFT Resolution Limit:

When the frequency difference equals $2\pi/N$, the phase noise rejection becomes more effective than the spectral broadening caused by the data window, at SNRs higher than 5 dB and the windowed MUSIC frequency estimator attains unity resolution probability at an SNR of 10 dB whereas the non-windowed estimator can only achieve a resolution probability of 0.78 asymptotically. Despite the spectral broadening caused by the data window the resolution performance of the non-Toeplitz AC matrix-based estimator is about 2.5 dB inferior to that of the windowed estimator for resolution probabilities greater that 0.5. This is due to the fact that the inferiority of the non-Toeplitz AC matrix-based frequency estimators with respect to their Toeplitz AC matrix based counterparts at low SNRs increases as the frequency difference of the sinusoids decreases and this effect is higher when compared to the resolution loss due to data windowing.

4.2.3 Separation of half of the DFT Resolution Limit:

Figure 4-III depicts that when the frequency difference of the sinuosoids is half of the Fourier resolution limit, the windowed MUSIC frequency estimator can no longer attain unity resolution probability and saturates at a level of 0.87 whereas the non-Toeplitz AC matrix based frequency estimator achieves this at SNR = 35 dB. The asymptotic resolution probability for the non-windowed Toeplitz AC matrix based frequency estimator is as low as 0.58, in this case.

5. EXPERIMENTAL STUDY OF THE JOINT DISTRIBUTION OF AC LAGS

In order to understand the mechanism by which data windowing helps to improve the resolution performance of the considered frequency estimators we investigate the joint distribution of the AC lags. In Figures 5 and 6 we plot the joint histograms of $r_y(1)$ and $r_y(2)$ for the two sinusoids case when $\omega_1 = \pi/2 = 1.5708$ rad and $\omega_2 = \pi/2 + \pi/25 = 1.6965$ rad which corresponds to a frequency difference equal to Fourier resolution limit and when SNR = 40 dB. The location of the true AC lags, $r_y(1) = -0.12538$ and $r_y(2) = -1.9686$, is shown with a vertical line segment in the figures.

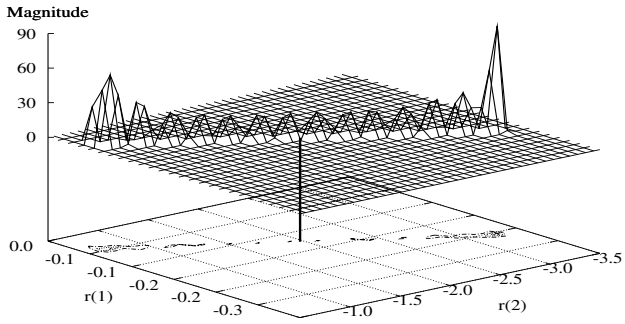


Figure 6: Joint histograms of the $r_y(1)$ and $r_y(2)$ estimates for **windowed** data (resolved frequencies = 100 %) ($\omega_1 = 1.5708$ rad, $\omega_2 = 1.6965$ rad, $\omega_2 - \omega_1 = 2\pi/N$ rad (DFT resolution limit), SNR = 40 dB, $N = 50$, $M = 20$, 100 000 noise realizations)

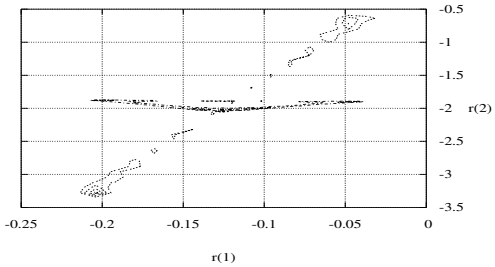


Figure 7: Joint histograms of the $r_y(1)$ and $r_y(2)$ estimates, a: non-windowed data (resolved frequencies), b: non-windowed data (unresolved frequencies), c: windowed data (resolved frequencies) ($\omega_1 = 1.5708$ rad, $\omega_2 = 1.6965$ rad, $\omega_2 - \omega_1 = 2\pi/N$ rad (DFT resolution limit), SNR = 40 dB, $N = 50$, $M = 20$, 100 000 noise realizations)

Figure 5-I shows that in the non-windowed case the support of the joint histogram corresponding to resolved frequency estimates constitutes the sides of a triangle and the one corresponding to unresolved frequency estimates constitutes one vertice of this triangle as shown in Figure 5-II. Data windowing changes this shape of the support drastically. Figure 6 depicts that the joint histogram of $r_y(1)$ and $r_y(2)$ has a linear support which means that the marginal pdfs of $r_y(1)$ and $r_y(2)$ have a similar shape which is a double hunched form as the curve "d" in Figure 1.

In Figure 7, we plot the supports of the joint histograms shown in Figures 5 and 6 to show their relations. In this figure the "triangular" shape belongs to the joint histogram of $r_y(1)$ and $r_y(2)$ corresponding to the resolved frequency estimates in the non-windowed case and the vertice of that triangle on the symmetry axis of the triangle corresponds to the unresolved frequencies in that case. The approximately linear support in the figure belongs to the joint histogram corresponding to the resolved frequency estimates in the windowed case. Actually, in that case all of the frequency estimates are resolved for the considered frequency difference of $2\pi/N$ in these simulations.

We also studied the joint histograms of other AC lag combinations and observed similar behavior.

6. CONCLUSIONS

In this study we considered the phase noise problem specific to the sinusoidal frequency estimation using Toeplitz AC matrix estimates. In particular we investigated the phase noise mitigation ability of data windowing in the case of two close sinusoids. Our main conclusions can be listed as follows:

- The Toeplitz AC matrix estimates can be modeled as a sum of two independent variables, a phase-dominated term and a noise-dominated term, shifted by a deterministic constant. The pdf of the lag coefficient, computed as the convolution of the pdfs of the two random variables, matches very well the experimentally determined histogram as in the case of a single sinusoid [4].
- In the multiple sinusoids case data windowing is beneficial in resolving closely spaced sinusoids. For a spacing equal to DFT resolution limit data windowing makes the estimator attain unity resolution probability whereas in the non-windowed case an asymptote at 0.78 exists. It is worth noting that for such a frequency spacing a 2 dB SNR advantage is obtained for a specific resolution probability when compared to non-windowed and non-Toeplitz AC matrix-based estimation.
- Investigation of the joint histograms of the AC lags reveals that the success of data windowing in contributing to resolve close sinusoids is accompanied by the linear dependence of the AC lags obtained by data windowing.

We should note that explaining the cause of the windowing in linearizing the mutual dependence of AC lags requires further study which we will consider in a future work.

REFERENCES

[1] S.M. Kay, *Modern spectral estimation: theory and application.* Prentice Hall, New Jersey, 1988.

[2] William F. Trench, "Numerical solution of the eigenvalue problem for efficiently structured Hermitian matrices", *Linear Algebra and Its Applications*, Vol 154-156, pp. 415-432, 1991.

[3] William H. Press, Brian P. Flannery, Saul A. Teukolsky, William T. Vetterling, *Numerical Recipes in C : The Art of Scientific Computing.* Cambridge University Press, 1992.

[4] Mustafa A. Altınkaya, Emin Anarım and Bülent Sankur, "Removal Of The Phase Noise In The Autocorrelation Estimates With Data Windowing", in *Proc. of 13th European Signal Proc. Conf., EUSIPCO 2005, Turkey*, pp. 1037-1040, 4-8 Sept. 2005.

[5] P. Stoica, T. Söderström and F.-N. Ti, "Overdetermined Yule-Walker Estimation of the Frequencies of Multiple Sinusoids: Accuracy Aspects", *Signal Proc.*, Vol. 16, pp. 155-174, 1989.

[6] F. J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform", *Proc. IEEE*, Vol. 66, No. 1, pp. 53-83, January 1978.

[7] L. Krasny and H. Koorapaty, "Performance of successive cancellation techniques for time of arrival estimation", in *Proc. of VTC 2002-Fall*, Vol. 4, pp. 2278 - 2282, 24-28 Sept. 2002.

A. DERIVATION OF THE PDF OF AC LAGS FOR WINDOWED DATA

Let $y_k = w_k x_k$ where $\{w_k, k = 1,2,\ldots,N\}$ represents a data window [6]. Consider the sample AC coefficients of the windowed data, obtained via the AC method as in (2). The AC lag estimates can be written as

$$r_y(l) = r_{y,m}(l) + r_{y,h}(l) + r_{y,g}(l) \qquad \text{for } l = 0,\ldots,N-1. \tag{5}$$

For two sinusoids case the summands are determined as

$$\begin{aligned}
r_{y,m}(l) &= \frac{1}{N-l}\sum_{k=1}^{N-l} w_k w_{k+l}[A_1\cos(\omega_1 l) + A_2\cos(\omega_2 l) + \sigma_n^2\delta_{l,0}],\\
r_{y,h}(l) &= \frac{A_1}{N-l}[\cos(2\phi_1)C_1(N,l) - \sin(2\phi_1)S_1(N,l)]\\
&\quad + \frac{A_2}{N-l}[\cos(2\phi_2)C_2(N,l) - \sin(2\phi_2)S_2(N,l)]\\
&\quad + \frac{\sqrt{A_1A_2}}{N-l}[(\cos(\omega_2 l + \phi_2 - \phi_1) + \cos(-\omega_1 l + \phi_2 - \phi_1))\, C_3(N,l)\\
&\quad - (\sin(\omega_2 l + \phi_2 - \phi_1) + \sin(-\omega_1 l + \phi_2 - \phi_1))\, S_3(N,l)]\\
&\quad + \frac{\sqrt{A_1A_2}}{N-l}[(\cos(\omega_2 l + \phi_2 + \phi_1) + \cos(-\omega_1 l + \phi_2 + \phi_1))\, C_4(N,l)\\
&\quad - (\sin(\omega_2 l + \phi_2 + \phi_1) + \sin(-\omega_1 l + \phi_2 + \phi_1))\, S_4(N,l)],
\end{aligned}$$

and

$$\begin{aligned}
r_{y,g}(l) &= \frac{\sqrt{A_1}}{N-l}\sum_{k=1}^{N-l} w_k w_{k+l}[\cos(\omega_1 k + \phi_1)n_{k+l} + \cos(\omega_1(k+l) + \phi_1)n_k]\\
&\quad + \frac{\sqrt{A_2}}{N-l}\sum_{k=1}^{N-l} w_k w_{k+l}[\cos(\omega_2 k + \phi_2)n_{k+l} + \cos(\omega_2(k+l) + \phi_2)n_k]\\
&\quad + \frac{1}{N-l}\sum_{k=1}^{N-l} w_k w_{k+l} n_k n_{k+l}\Bigg\}
\end{aligned}$$

which denote the deterministic component, the phase noise and the Gaussian component, respectively, with the definitions:

$$\begin{aligned}
C_1(N,l) &= \sum_{k=1}^{N-l} w_k w_{k+l}\cos(\omega_1(2k+l)),\\
C_2(N,l) &= \sum_{k=1}^{N-l} w_k w_{k+l}\cos(\omega_2(2k+l)),\\
C_3(N,l) &= \sum_{k=1}^{N-l} w_k w_{k+l}\cos((\omega_2 - \omega_1)k),\\
C_4(N,l) &= \sum_{k=1}^{N-l} w_k w_{k+l}\cos((\omega_2 + \omega_1)k),
\end{aligned} \tag{6}$$

where $\delta_{l,0}$ denotes Kronecker delta and $\{S_i(N,l), i = 1,2,3,4\}$ are obtained replacing the $\cos(\cdot)$ function of $\{C_i(N,l), i = 1,2,3,4\}$ in (6) with $\sin(\cdot)$. The large sample statistics of $r_x(l) = r(l)$ was derived in [5] where it was shown to be Gaussian distributed. With a direct generalization to windowed data the large sample statistics of $r_y(l)$ will be also Gaussian with mean $r_{y,m}(l)$ and variance

$$\begin{aligned}
\sigma_y^2(l) &= \frac{\sigma_n^2}{(N-l)^2}\Bigg[\left(2(A_1+A_2) + \sigma_n^2(1+\delta_{l,0})\right)\sum_{k=1}^{N-l} w_k^2 w_{k+1}^2\\
&\quad + 2\sum_{k=1}^{N-2l} w_k w_{k+1}^2 w_{k+2l}\left(A_1\cos(2\omega_1 l) + A_2\cos(2\omega_2 l)\right)\Bigg],
\end{aligned} \tag{7}$$

for $l = 0,\ldots,N-1$.

Among the three summands of $r_y(l)$ given in (5), the first one is deterministic and the others are stochastic components. In the study [4] with a single sinusoid we neglected that the phase noise-dependent component $r_{y,h}(l)$ and the AWGN-dependent component $r_{y,g}(l)$ depend on eachother and furthermore we assumed that in $r_{y,g}(l)$ the additive noise components dominate this term and that the effects of phase are secondary. These assumptions were shown to be successful. In this study, we make the same assumptions. So, the pdf of $r_y(l)$ as a sum of independent random variables will be a convolution of the individual pdfs of $r_{y,h}(l)$ and $r_{y,g}(l)$ shifted by the deterministic quantity $r_{y,m}(l)$. Here, we make an additional assumption that $r_{y,g}(l)$ will also have approximately the Gaussian distribution valid for the large sample case with mean $r_{y,m}(l)$ and variance $\sigma_y^2(l)$ in the small sample size case.

For the considered two sinusoids case the computation of the pdf of phase noise, $f^{(l)}_{R_{y,h}}(r_{y,h})$, is not analytically tractable. That's why we adopt a semi-analytic procedure where we obtain the histogram of $r_{y,h}(l)$ through Monte Carlo experiments with independent realizations of ϕ_1 and ϕ_2. Then we fit a parametric function to this histogram using Levenberg-Marquardt method [3] to obtain a closed form expression for the pdf. Finally, the pdf of $r_y(l)$, $f^{(l)}_{R_y}(r_y)$, is obtained as the convolution $f^{(l)}_{R_y}(r_y) = f^{(l)}_{R_{y,h}}(r_{y,h} - r_{y,m}) * f^{(l)}_{R_{y,g}}(r_{y,g})$. The windowing causes a bias in the AC lag estimates. This bias will be removed by replacing $r_y(l)$ with $\frac{(N-l)r_y(l)}{\sum_{k=1}^{N-l} w_k w_{k+l}}$.