Frequency Estimation of Real-Valued Single-Tone in Colored Noise Using Multiple Autocorrelation Lags

1Rim Elasmi-Ksibi, 1Hichem Besbes, 2Roberto López-Valcarce and 1Sofiane Cherif

Abstract

The problem of frequency estimation of a real-valued sinusoid in colored noise is addressed. A new estimator using higher-order lags of the sample autocorrelation is proposed, and a statistical analysis is provided. By choosing the number of lags to include in the estimation process, a tradeoff between performance and complexity can be easily achieved.

I. INTRODUCTION

Detection of sinusoidal components and estimation of their frequencies in the presence of broadband noise are common problems in signal processing with a broad range of areas of application; numerous techniques have been developed for the white noise case [1]. The Maximum Likelihood (ML) method is statistically efficient, achieving the Cramer-Rao Lower Bound (CRLB) asymptotically, but it is computationally demanding [2]. Simpler, suboptimal frequency estimators can be obtained using the Linear Prediction properties of sinusoidal signals, such as the Pisarenko Harmonic Decomposer (PHD) [3], Reformed PHD [4] and Modified Covariance (MC) [5] methods; a number of statistical analyses have shown their inefficiency [6], [7].

For the single sinusoid case, these simple estimators make use of just two coefficients of the sample autocorrelation. Using multiple autocorrelation lags has been considered in [8], [9] for the complex-valued (cisoid) case, as well as in [10] (the so-called $P$-estimator) for the real-valued case. Although these approaches result in performance improvement, they require phase unwrapping to resolve the

(1) Unité de Recherche TECHTRA, Sup’Com, Tunisia

(2) Dept. Signal Theory and Communications, University of Vigo, Spain

e-mail: asmi.rim@planet.tn, {hichem.besbes, sofiane.cherif}@supcom.rnu.tn, valcarce@gts.tsc.uvigo.es

This work was supported by the Spanish Agency for International Cooperation (AECI) and the Tunisian Government under project A/5405/06, and by the Spanish Ministry of Science and Innovation under projects SPRACTIVE (ref. TEC2007-68094-C02-01/TCM) and COMONSENS (CONSOLIDER-INGENIO 2010 CSD2008-00010).
frequency ambiguities that then appear. The phase unwrapping stage adds extra complexity to the procedure, which is clearly undesirable in applications requiring rapid frequency estimation.

We focus on the real-valued case, and present a novel estimator that exploits higher lags of the sample autocorrelation. It enjoys low computational complexity since no phase unwrapping is required. The new method compares favorably to its forerunners, and it is applicable to environments with colored noise of finite memory (e.g. Moving Average (MA) noise models).

Section II states the problem and presents the new estimator. A statistical analysis is given in Section III. Section IV shows some numerical examples, and conclusions are drawn in Section V.

II. NOVEL FREQUENCY ESTIMATOR

Consider the problem of estimating the unknown frequency \( \omega_0 \in [0, \pi] \) of a real-valued sine wave \( s_n \) immersed in colored noise \( u_n \). The \( N \) available samples of the observed signal, \( y_n \), are given by

\[
y_n = s_n + u_n
\]

\[
y_n = \alpha \sin(\omega_0 n + \varphi) + u_n, \quad 1 \leq n \leq N,
\]

where \( \alpha \) is the sinusoid amplitude and \( \varphi \) is a random phase uniformly distributed in \([-\pi, \pi]\). The noise process \( \{u_n\} \) is zero-mean wide-sense stationary and Gaussian, and independent of \( \{s_n\} \). It is modeled as an MA(\( M \)) process, generated by passing a white Gaussian process with variance \( \sigma^2 \) through an order-\( M \) FIR filter. The noise autocorrelation is denoted \( E\{u_n u_{n-k}\} = \sigma^2 \rho_k \); thus, \( \rho_k = 0 \) for \( k > M \). The power spectral density of \( \{u_n\} \) is \( S_u(e^{j\omega}) = \sigma^2 \sum_{k=-M}^{M} \rho_k e^{-jk\omega} \). It will also be useful to define \( \beta_l = \sum_{k=-M}^{M} \rho_k \rho_{k-l} \). The Signal to Noise Ratio is defined as \( \text{SNR} = \alpha^2 / (2\sigma^2) \).

The autocorrelation sequence of the observations is

\[
r_k = E\{y_n y_{n-k}\} = \frac{\alpha}{2} \cos k\omega_0 + \sigma^2 \rho_k.
\]

Thus, using trigonometric relations one can show that

\[
r_{k-1} + r_{k+1} = 2r_k \cos \omega_0, \quad \text{for all } k > M + 1.
\]

Hence, we observe that for any integers \( q > p > M + 1 \),

\[
\sum_{k=p}^{q} r_k (r_{k-1} + r_{k+1}) = 2 \cos \omega_0 \sum_{k=p}^{q} r_k^2.
\]

Eq. (4) suggests a means to obtain an estimate \( \hat{a}_0 = \cos \hat{\omega}_0 \) as

\[
\hat{a}_0 = \frac{\sum_{k=p}^{q} \hat{r}_k (\hat{r}_{k-1} + \hat{r}_{k+1})}{2 \sum_{k=p}^{q} \hat{r}_k^2},
\]

where \( \hat{r}_k \) is the unbiased autocorrelation estimate

\[
\hat{r}_k = \frac{1}{N-k} \sum_{n=k+1}^{N} y_n y_{n-k}.
\]
III. PERFORMANCE ANALYSIS

Let \( \hat{\mathbf{r}} \equiv [ \hat{r}_{p-1} \; \hat{r}_p \; \cdots \; \hat{r}_{q+1} ]^T \). We can express the novel estimator as \( \hat{a}_0 = f(\hat{r}) \), where \( f \) is defined by (5). A first-order Taylor expansion of \( f(\hat{r}) \) around \( \mathbf{r} = [ r_{p-1} \; r_p \; \cdots \; r_{q+1} ]^T \) yields

\[
\hat{a}_0 \approx a_0 + v^T (\hat{r} - \mathbf{r}),
\]

with \( v = \nabla f(\mathbf{r})_{|\mathbf{r}=\mathbf{r}}. \) Thus, \( \text{E}\{ (\hat{a}_0 - a_0)^2 \} \approx v^T \mathbf{C} v, \) with \( \mathbf{C} \) the covariance matrix of \( \hat{r} \). One has

\[
v = \frac{1}{\alpha^2 \Delta(\omega_0)} \left[ \begin{array}{cccc} \cos p\omega_0 & -\cos(1-1)\omega_0 & \cdots & 0 \\ 0 & \cdots & 0 & -\cos(q+1)\omega_0 \end{array} \right],
\]

with \( \Delta(\omega) \equiv \sum_{k=p}^{q} \cos^2 k\omega. \) It is shown in the Appendix that

\[
\mathbf{C} \approx \frac{1}{N} \left[ \sigma^4 \mathbf{B} + 2\alpha^2 \Sigma_{\omega}(e^{i\omega}) \mathbf{w} \mathbf{w}^T \right],
\]

with \( \mathbf{B} \) a Toeplitz matrix with elements \( [\mathbf{B}]_{i,j} = \beta_{i-j} \), and \( \mathbf{w} \) a vector with elements \( [\mathbf{w}]_i = \cos(p+i-1)\omega_0 \). It turns out that \( v^T \mathbf{w} = 0, \) and thus, using (8)-(9) and \( \text{E}\{ (\hat{\omega}_0 - \omega_0)^2 \} \approx \text{E}\{ (\hat{a}_0 - a_0)^2 \} / \sin^2 \omega_0, \) one obtains

\[
\text{E}\{ (\hat{\omega}_0 - \omega_0)^2 \} \approx \frac{\Gamma(\omega_0)}{4\text{NSR}^2 \Delta^2(\omega_0) \sin^2 \omega_0},
\]

where \( \Gamma(\omega) = \sum_{i \in \mathcal{I}} \beta_i \Gamma_i(\omega), \) with \( \mathcal{I} = \{ 0, 1, q - p, q - p + 1, q - p + 2 \} \) and

\[
\Gamma_0(\omega) \equiv \cos^2(p-1)\omega + \cos^2 p\omega + \cos^2 q\omega + \cos^2(q+1)\omega \quad (11)
\]

\[
\Gamma_1(\omega) \equiv -2[\cos(p-1)\omega \cos p\omega + \cos q\omega \cos(q+1)\omega] \quad (12)
\]

\[
\Gamma_{q-p}(\omega) \equiv 2 \cos(p-1)\omega \cos(q+1)\omega \quad (13)
\]

\[
\Gamma_{q-p+1}(\omega) \equiv -2[\cos p\omega \cos(q+1)\omega + \cos(p-1)\omega \cos q\omega] \quad (14)
\]

\[
\Gamma_{q-p+2}(\omega) \equiv 2 \cos p\omega \cos q\omega. \quad (15)
\]

Note that if \( q, p \) are chosen such that \( q - p > 2M \), then \( \beta_{q-p} = \beta_{q-p+1} = \beta_{q-p+2} = 0. \) For the case of white Gaussian noise, \( \beta_i = \delta_i \) and thus the Mean Square Error (MSE) of the estimator is

\[
\text{E}\{ (\hat{\omega}_0 - \omega_0)^2 \} \approx \frac{\cos^2(p-1)\omega_0 + \cos^2 p\omega_0 + \cos^2 q\omega_0 + \cos^2(q+1)\omega_0}{4\text{NSR}^2 \left( \sum_{k=p}^{q} \cos^2 k\omega_0 \right)^2 \sin^2 \omega_0}. \quad (16)
\]

The influence of the design parameters \( p, q \) in the MSE is via the factor \( \Gamma(\omega_0) / \Delta^2(\omega_0). \) Although this term is highly oscillatory, the amplitude of the oscillations decreases as \( q \) increases. It is instructive to consider (16) for \( \omega_0 = \frac{\pi}{2}, \) for which \( \Gamma(\omega_0) = 2 \) and \( \Delta(\omega_0) \approx (q - p + 1) / 2; \) thus, the MSE behaves as \( 1 / (q - p + 1)^2. \) Roughly speaking, by doubling the number of lags included in the estimation process, the MSE is decreased by 6 dB. Of course, this reduction of the MSE cannot be achieved indefinitely by increasing \( q, \) since the finite sample size limits the number of autocorrelation estimates available; at the same time, the approximations used when deriving (10) become less accurate with larger \( q. \)
IV. SIMULATION RESULTS

Fig. 1 shows the MSE of several estimators in a white noise setting as a function of $\omega_0$, for $N = 100$ and SNR = 10 dB. The RPHD [4] and MC [5] schemes use two autocorrelation coefficients only, performing worse than methods exploiting higher-order lags. Among these, the $P$-estimator [10] suffers from the so-called 'edge frequency' problem: performance degrades for $\omega_0$ close to $k\pi/P$, with $0 \leq k \leq P - 1$ ($P = 20$ in this example). For the new estimator (5), $p = 2$ was taken, and two different values of $q$ (7 and 20) were considered. As expected, performance improves with increasing $q$, since more autocorrelation lags are then included in the estimation process. Good agreement with the approximate MSE (16) is observed. With $q = 20$, the novel estimator is very close to the CRLB [11] for a wide range of frequencies.

Next we consider a colored noise case in which the noise coloring filter has transfer function $B(z) = 1 + 0.1z^{-1} + 0.7z^{-2} + 0.05z^{-3} + 0.3z^{-4}$. Since $M = 4$, now $p = 6 > M + 1$ is adopted for the novel estimator. The RPHD and MC estimates are biased in the presence of noise coloring, resulting in the poor MSE behavior observed in Fig. 2. The novel estimator still offers a means to trade off performance and complexity by increasing $q$. Fig. 3 shows the MSE of the estimators as a function of the SNR, fixing $\omega_0 = 0.4\pi$. Using $(p, q) = (6, 30)$, the novel estimate asymptotically achieves the CRLB.
V. CONCLUSION

A frequency estimator for real-valued sinusoids in noise exploiting higher-order lags of the sample autocorrelation has been presented. For MA noise, if an upper bound is available for the order of the noise coloring filter, the ’corrupted’ low-order lags can be left out, and thus bias is avoided. Knowledge of the MA noise model parameters is not needed. The proposed method compares favorably to previous...
approaches and allows a tradeoff between performance and complexity by choosing the number of
lags to be included in the estimation process.

APPENDIX

Let \( k, m > M \). One has
\[
E\{(\hat{r}_k - r_k)(\hat{r}_m - r_m)\} = E\{\hat{r}_k\hat{r}_m\} - r_mr_k,
\]
and
\[
E\{\hat{r}_k\hat{r}_m\} = \frac{1}{N^2} \sum_{i=k+1}^{N} \sum_{j=m+1}^{N} E\{y_iy_jy_{i-k}y_{j-m}\}
\]
\[
= \frac{\alpha^4}{4} \cos k\omega_0 \cos m\omega_0 + T_0 + T_1,
\]
where \( T_0 \) and \( T_1 \) are defined as
\[
T_0 = \frac{1}{N^2} \sum_{i=k+1}^{N} \sum_{j=m+1}^{N} \left[ E\{s_{i-k}s_{j-m}\}E\{u_iu_j\} + E\{s_{i-k}s_j\}E\{u_i-ku_{j-m}\} \right.
\]
\[
+ E\{s_is_{j-m}\}E\{u_{i-k}u_j\} + E\{s_is_{j-k}\}E\{u_{i}u_{j-m}\} \right.
\]
\[
\left. + E\{s_is_{j-k}\}E\{u_{i}u_{j-m}\} + E\{s_is_{j-m}\}E\{u_{i-k}u_j\} \right] = 0 \text{ for } m > M
\]
\[
T_1 = \frac{1}{N^2} \sum_{i=k+1}^{N} \sum_{j=m+1}^{N} E\{u_iu_ju_{i-k}u_{j-m}\}.
\]
Assume w.l.o.g. that \( k \geq m \), and define a trapezoidal window \( w_l \) as
\[
w_l = \begin{cases} 
N - k - |l|, & -N + k + 1 \leq l \leq 0, \\
N - k, & 0 < l \leq k - m, \\
N - m - |l|, & k - m < l \leq N - m - 1.
\end{cases}
\]
Then, for large \( N \), the following approximation holds:
\[
\frac{1}{N^2} \sum_{i=k+1}^{N} \sum_{j=m+1}^{N} E\{s_{i-k}s_{j-m}\}E\{u_iu_j\} = \frac{1}{N^2} \sum_{l=k+1-N}^{N-m-1} w_l E\{s_{n-k}s_{n+l-m}\}E\{u_nu_{n+l}\}
\]
\[
\approx \frac{1}{N} \sum_{l=-\infty}^{\infty} E\{s_{n-k}s_{n+l-m}\}E\{u_nu_{n+l}\}
\]
\[
= \frac{1}{N} \sum_{l=-\infty}^{\infty} \frac{\alpha^2}{2} \cos(l - m + k)\omega_0 \cdot \sigma^2 p_l
\]
\[
= \frac{1}{N} \alpha^2 \cos(k - m)\omega_0 S(u)\left(e^{j\omega_0}\right).
\]
The remaining terms in (18) are obtained analogously, yielding
\[
T_0 \approx \frac{2\alpha^2}{N} S(u)\left(e^{j\omega_0}\right) \cos k\omega_0 \cos m\omega_0.
\]
Now, since \{\{u_n\} is Gaussian,
\[
E\{u_iu_ju_{i-k}u_{j-m}\} = E\{u_iu_j\}E\{u_{i-k}u_{j-m}\} + E\{u_iu_{j-m}\}E\{u_{i-k}u_j\} + E\{u_iu_{i-k}\}E\{u_ju_{j-m}\}
\]
\[
= 0 \text{ for } k > M = 0 \text{ for } m > M
\]
(22)
and therefore

\[
T_1 = \frac{1}{N^2} \sum_{i=k+1}^{N} \sum_{j=m+1}^{N} E\{u_i u_j\}E\{u_{i-k} u_{j-m}\} + E\{u_i u_{j-m}\}E\{u_{i-k} u_j\}
\]
\[
= \frac{\sigma^4}{N^2} \sum_{l=k+1}^{N-m-1} w_l (\rho_l \rho_{l-k+m} + \rho_{l+m} \rho_{l-k})
\]
\[
\approx \frac{\sigma^4}{N} \sum_{l=-M}^{M} \rho_l (\rho_{l-k+m} + \rho_{l-k-m}) = \frac{\sigma^4}{N} \sum_{l=-M}^{M} \rho_l \rho_{l-k+m},
\]

(23)

where the last step follows from the fact that \( l - k - m < -M \).

REFERENCES


