THRESHOLDING THE AMBIGUITY FUNCTION

Heidi Hindberg[a], Alfred Hanssen[a] and Sofia C. Olhede[b]

[a]Dept. of Physics and Technology, University of Tromsø , Tromsø, Norway,
email: heidih@phys.uit.no, alfred@phys.uit.no
[b]Dept. of Statistical Science, University College London, London,UK, email: s.olhede@ucl.ac.uk

ABSTRACT

In this paper we propose a new method for estimating the Ambiguity Function (AF) of a random process with limited spreading support. The observed process is modelled as the aggregation of a non-stationary signal of interest and noise. As the AF has limited spreading, thresholding is a suitable estimation procedure. Some key stochastic properties of the Empirical Ambiguity Function are derived to obtain a suitable threshold. Based on a median absolute deviation estimator for the variance, we derive a suitable threshold, which forms the basis for our proposed estimator. The estimator is tested on both artificial and real signals, and our results demonstrate a remarkably high resolution and reduced variance.

Index Terms— Ambiguity function, chirp signal, Doppler-lag function, underspread process, thresholding.

1. INTRODUCTION

The Ambiguity Function (AF) is a measure of dependence between a given signal, and its translates in the time-frequency plane. It has been studied for the characterisation of deterministic signals [1], and was initially introduced by Woodward [2] for radar applications. The AF is defined in terms of local variables, τ and ν, that do not represent global properties of the signal, only relative dependencies. The values (ν, τ) range over correspond to the ambiguity plane. If there is a non-negligible contribution of the AF at a given (ν, τ) then this indicates that at some global time t there is a dependence with time lag τ, and likewise at some global frequency f there is a dependence with frequency lag ν.

The AF has been established as a fundamental object in determining whether the generating mechanism of a zero-mean second order stochastic signal can be consistently characterised using different second order descriptions [3, 4]. Priestley coined the phrase 'semi-stationary processes', for processes whose AFs were limited in support near $\nu = 0$. Matz and Hlawatsch argued that only processes with underspread AFs (AFs concentrated near $(\nu, \tau) = (0, 0)$) were suited to time-frequency analysis. Later developments include those of Pfander and Walnut [5], who argue that non-stationary systems can be identified *without* assuming a concentration around $(\nu, \tau) = (0, 0)$, as long as the spread of the AF is limited in the ambiguity plane.

Thus estimation of the AF is in itself of considerable interest. Despite this fact previous estimation methods have been surprisingly simplistic. Standard methods calculate the AF directly from the observed sample of the signal of interest, as if the observed signal was noise free and exhibited no variation in sample path, see [6]. Without several realisations of a non-stationary process the Empirical AF (EAF) has large variance. To reduce variability the AF is sometimes smoothed, by either averaging the empirical auto-covariance or by performing a windowed Fourier Transform (FT) of the empirical auto-covariance. However both these actions cause a loss of resolution in the ambiguity domain.

In this paper we propose to estimate the AF for signals whose spreading in the ambiguity plane is limited. Such signals with adequate sampling in time are *AF compressible*. Thus we propose to adapt thresholding methods for the estimation of the AF [7]. We shall demonstrate that our proposed estimation procedure can greatly outperform existing (if somewhat naive) non-parametric estimation methods.

2. THE AMBIGUITY FUNCTION

Some confusion exists in the literature when both deterministic and stochastic signals are treated in the same framework. We define the AF of a discrete-time *analytic* signal $X[t]$, with finite and summable second order structure, by [6]

$$A_{XX^*}(\nu, \tau] = \sum_{t=-\infty}^{\infty} E\{X[t]X^*[t-\tau]\}\, e^{-j2\pi\nu t}. \qquad (1)$$

Here, () indicate a continuous valued argument, and [] indicate a discrete valued argument.

If $X[t]$ is a zero-mean harmonizable random process then $M_{XX^*}[t, \tau] = E\{X[t]X^*[t-\tau]\}$, corresponds to the auto-covariance of the process, where t denotes a global time variable and τ is a local time shift. The AF is defined as the discrete time FT of $M_{XX^*}[t, \tau]$ with respect to t, represented in frequency variable ν.The AF can also be expressed as an inverse FT of a dual-frequency representation known as the Loève spectrum [8], The auto-covariance function measures the dependence between the process at time t and itself at time $t-\tau$. The dual-frequency spectrum measures the dependence between the process at frequency f and the frequency $f-\nu$. Thus the variable ν is a frequency shift or offset. The AF is a measure of dependencies in the local time – local frequency plane.

In this paper, we use the analytic signal in order to avoid interference between negative and positive frequencies, and to avoid aliasing. Note that unlike general perception an analytic non-stationary process may be non-circular [9], and we need to consider the Complementary AF (CAF) to fully characterise the second order structure of the process. We here focus on circular analytic processes, noting that for example real-valued stationary processes will have circular analytic signals. In general our methods can be extended to also estimate the CAF.

Since the AF is the expectation of a (potentially) random quantity, it will in practice need to be estimated. We base the estimation on a single finite length realisation (length N say). Standard approaches to this estimation problem are based on time averaging and/or frequency smoothing. We start of by defining the EAF as the FT of the method of moments estimator of the auto-covariance,

$$\widehat{A}_{XY^*}(\nu,\tau)=\sum_{t=\max(0,\tau)}^{N-1+\min(0,\tau)} X[t]Y^*[t-\tau]e^{-j2\pi\nu t}. \qquad (2)$$

This estimator has a large variance and should not be used directly without modification.

3. MODELLING

To propose estimation methods we now model our observed signal $Y[t]$ as follows,

$$Y[t]=X[t]+\sigma_\epsilon\epsilon[t]. \qquad (3)$$

We want to estimate the AF of the signal of interest, $X[t]$, that has been immersed in a zero-mean analytic white noise process, $\epsilon[t]$, with finite variance σ_ϵ^2. We also assume that the signal and the noise are independent. The EAF of $Y[t]$ is given by

$$\begin{aligned}\widehat{A}_{YY^*}(\nu,\tau)=&\widehat{A}_{XX^*}(\nu,\tau]+\widehat{A}_{X\epsilon^*}(\nu,\tau]+\widehat{A}_{\epsilon X^*}(\nu,\tau]\\&+\sigma_\epsilon^2\widehat{A}_{\epsilon\epsilon^*}(\nu,\tau].\end{aligned}$$

Thus, $\mathrm{E}\{\widehat{A}_{YY^*}(\nu,\tau]\}=A_{XX^*}(\nu,\tau]+\sigma_\epsilon^2 A_{\epsilon\epsilon^*}(\nu,\tau]$ and

$$\begin{aligned}\mathrm{var}\{\widehat{A}_{YY^*}(\nu,\tau]\}=&\mathrm{var}\{\widehat{A}_{XX^*}(\nu,\tau]\}+\sigma_\epsilon^4\mathrm{var}\{\widehat{A}_{\epsilon\epsilon^*}(\nu,\tau]\}\\&+\mathrm{var}\{\widehat{A}_{X\epsilon^*}(\nu,\tau]\}+\mathrm{var}\{\widehat{A}_{X\epsilon^*}(-\nu,-\tau]\}\end{aligned}$$

as all the cross covariance terms can be shown to be zero. We note that $A_{XX^*}(\nu,\tau]$ is assumed to be of negligible magnitude apart from at a few values of $(\nu,\tau]$, and that

$$\begin{aligned}\mathrm{E}\left\{\widehat{A}_{\epsilon\epsilon^*}[\nu,\tau]\right\}=&\frac{1}{2}D_{N-|\tau|}(\pi\nu)\mathrm{sinc}(\pi\tau/2)\\&\times e^{-j\pi(\nu(N+\tau-1)-\tau/2)},\end{aligned}$$

where $D_N(\pi f)=\sin(\pi fN)/\sin(\pi f)$ and $\mathrm{sinc}(x)=\sin(x)/x$. Likewise,

$$\mathrm{var}\left\{\widehat{A}_{\epsilon\epsilon^*}[\nu,\tau]\right\}=[N-|\tau|][1/2-|\nu|]. \qquad (4)$$

For regions of the ambiguity plane where $|A_{XX^*}(\nu,\tau]|$ is small, the expected value of the EAF of $Y[t]$ will be approximately zero and

$$\begin{aligned}&\mathrm{E}\{\widehat{A}_{YY^*}(\nu,\tau]|\,\{X[t]\}_t\}\approx 0\\&\mathrm{var}\{\widehat{A}_{YY^*}(\nu,\tau]|\,\{X[t]\}_t\}\approx\sigma_\epsilon^4[N-\tau][1/2-\nu]\\&+\mathrm{var}\{\widehat{A}_{X\epsilon^*}(\nu,\tau]\}+\mathrm{var}\{\widehat{A}_{X\epsilon^*}(-\nu,-\tau]\}. \qquad (5)\end{aligned}$$

$\widehat{A}_{YY^*}(\nu,\tau]$ is conditional on $\{X[t]\}$ asymptotically Gaussian; this can be verified empirically and follows from the distributional assumptions on $\epsilon[t]$. We define the quantity

$$\widehat{A}^{(N)}_{YY^*}(\nu,\tau]=\frac{\widehat{A}_{YY^*}(\nu,\tau]}{\sqrt{[N-\tau][1/2-\nu]}}. \qquad (6)$$

If there is no signal present then this has variance σ_ϵ^4. Thus, under the assumption that only noise is present, an estimator of σ_ϵ^4 can be obtained using the Median Absolute Deviation (MAD) estimator on the sequence $\left\{\left|\widehat{A}^{(N)}_{YY^*}(\nu,\tau]\right|^2\right\}$. MAD has been used for estimating the scale of correlated data before, see [10]. However the estimator will need an adjustment factor that is different for $\frac{1}{2}\chi_2^2$ random variables compared to χ_1^2 random variables. We note that the median value of a $\frac{1}{2}\chi_2^2$ variable is $\ln(2)$. We thus replace our estimator by

$$\widehat{\sigma}_\varepsilon^4=\frac{\mathrm{median}\left\{\left|\widehat{A}^{(N)}_{YY^*}(\nu,\tau]\right|^2\right\}_{(\nu,\tau)}}{\ln(2)}. \qquad (7)$$

The imprecision of this procedure will depend on the lack of compression of $X[t]$ in the ambiguity domain. We note that MAD has a breakdown point of 50 %, and so with quite severe contamination the estimator will still be useful, if somewhat inefficient.

4. PROPOSED AF ESTIMATOR

We have proposed an estimation procedure for the AF. If no signal is present, then the distribution of the EAF is wholly determined by the distribution of $\widehat{A}_{\epsilon\epsilon^*}(\nu,\tau]$. In this case from the results of the previous section the distribution of $\widehat{A}_{YY^*}(\nu,\tau]/\sigma_\varepsilon^2$ can be approximated as complex Gaussian with mean $\mathrm{E}\left\{\widehat{A}_{\epsilon\epsilon^*}[\nu,\tau]\right\}$ and variance $[N-|\tau|][1/2-|\nu|]$. For many values of $(\nu,\tau]$ we can note the approximate distribution of $\widehat{A}^{(N)}_{YY^*}(\nu,\tau]/\widehat{\sigma}_\varepsilon^2$ as complex Gaussian with zero mean and a unit variance. $\widehat{A}^{(N)}_{YY^*}(\nu,\tau]/\widehat{\sigma}_\varepsilon^2$ is not in general zero. Olhede [7] derived the conservative threshold for correlated variables with this distribution, and independent of

$\widehat{A}^{(N)}_{YY^*}(\nu,\tau]/\widehat{\sigma}^2_\varepsilon$, the threshold value is $\lambda^2_N(C) = 2\log(N[\log(N)]^C)$, where $C > -1$ is chosen. We propose the threshold AF estimator of

$$\widehat{A}^{(\mathrm{th})}_{YY^*}(\nu,\tau] = \begin{cases} \widehat{A}_{YY^*}(\nu,\tau] & \text{if} \quad \left|\frac{\widehat{A}^{(N)}_{YY^*}(\nu,\tau]}{\widehat{\sigma}^2_\varepsilon}\right|^2 > \lambda^2_{2N}(1) \\ 0 & \text{if} \quad \left|\frac{\widehat{A}^{(N)}_{YY^*}(\nu,\tau]}{\widehat{\sigma}^2_\varepsilon}\right|^2 \leq \lambda^2_{2N}(1) \end{cases} \tag{8}$$

The usage of $2N$ is due to calculating the AF at both negative and positive lags. The risk of $\widehat{A}^{(\mathrm{th})}_{YY^*}(\nu,\tau]$ is stated in [7].

5. EXAMPLES

We test our estimator on two data sets, also estimating the AF by calculating a tapered FFT in (2).

5.1. Simulated linear chirp in noise

We generated samples of the linear chirp signal

$$x[t] = \exp\left[j\pi(2\alpha t + \beta t^2)\right] + \eta[t],\ t = 0,\ldots,N-1, \tag{9}$$

with $N = 256$, starting frequency $\alpha = 0.1$ and chirp rate $\beta = 7.8 \times 10^{-4}$ immersed in additive zero-mean analytic white noise with variance $\sigma^2 = 0.5$. We calculate the EAF, which is shown in Fig. 1(a) (all plots show the absolute value of the function on a dB scale). The variance of the EAF is estimated from Eq. (7), and the AF is estimated using Eq. (8). The result of the thresholding is shown in Fig. 1(b). For comparison we also include a smoothed estimator in Fig. 1(c). The AF of a linear chirp should be nonzero along the line $\nu = \beta\tau$. We see that the thresholding has removed most of the points outside this line, and that the line remains intact, whilst the EAF is extremely noisy, and the smoothed estimator has spread the line, without removing all of the noise contributions.

5.2. Bat signal

The second example is a recorded digitized signal of the echolocation of large brown bats. The sample is 400 points long, with sampling period $\Delta t = 7\mu$s. It contains multiple components, which leads to interference terms in the AF. See Fig 2(a) for a plot of the EAF, Fig 2(b) for the thresholded estimate and Fig 2(c) for the smoothed estimate. We are not seeking to remove the interference, but only the effects of the noise. The structure of the chirps, and the interaction between the components, is much more clearly made out in Fig 2(b). $A_{XX^*}(\nu,\tau] \neq 0$ can be estimated as zero; this happens when the noise is large compared to the signal contribution at (ν,τ).

6. CONCLUSION

The AF is a fundamental quantity for characterising the second order structure of a non-stationary signal. For many signals of interest, the AF is highly compressed and only substantially distinct from non-zero in special regions of the ambiguity plane. Using this compression we can estimate the AF with higher resolution than smoothing approaches, by applying thresholding methods. Improved estimation performance is shown on both synthetic and real data.

Acknowledgements

The authors thank the Research Council of Norway for support under project 162831/V00 and acknowledge C. Condon, K. White, and A. Feng (Beckman Center, University of Illinois) for use of the bat data-set.

7. REFERENCES

[1] R. E. Blahut, "Theory of remote surveillance algorithms," in *Radar and Sonar, Part I*, R. E. Blahut, W. Miller, and C. H. Wilcox, Eds. Springer Verlag, New York, USA, 1991.

[2] P. M. Woodward, *Probability and Information Theory with Applications to Radar*, New York: Pergamon, 1953.

[3] G. Matz and F. Hlawatsch, "Nonstationary spectral analysis based on time-frequency operator symbols and underspread approximations," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1067–1086, 2006.

[4] M. B. Priestley, "Evolutionary spectra and nonstationary processes," *Roy. Stat. Soc.*, vol. B 47, no. 2, pp. 204–237, 1965.

[5] G. E. Pfander and D. F. Walnut, "Measurement of time-variant linear channels," *IEEE Trans. on Info Theory*, vol. 52, pp. 4808–4820, 2006.

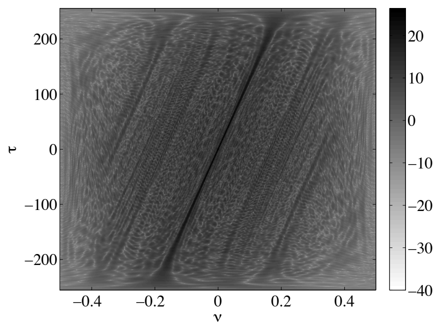
[6] R. Tolimieri and S. Winograd, "Computing the ambiguity surface," *IEEE Transactions on Acoust., Speech and Signal Processing*, vol. 33, pp. 1239–1245, 1985.

[7] S. Olhede, "Hyperanalytic denoising," *IEEE Trans. Image Proc.*, vol. 16, pp. 1522–1537, 2007.

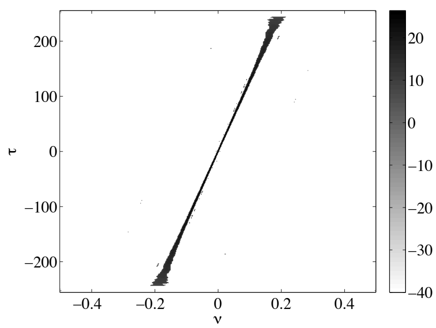
[8] M. Loève, *Probability Theory*, New York, NY: Springer-Verlag, fourth edition, 1978.

[9] P. J. Schreier and L. L. Scharf, "Stochastic time-frequency analysis using the analytic signal: why the complementary distribution matters," *IEEE Transactions on Signal Processing*, vol. 51, pp. 3071–3079, 2003.

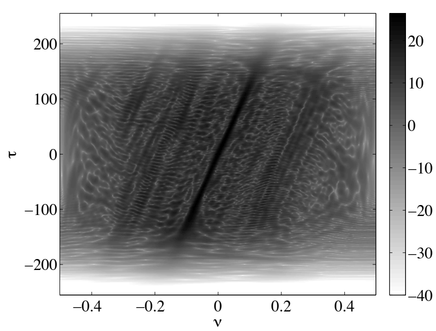
[10] I. M. Johnstone and B. W. Silverman, "Wavelet threshold estimators for data with correlated noise," *J. Roy. Statist. Society, Ser. B. Statistical Methodology*, vol. 59, pp. 319–351, 1997.



(a) The EAF of the chirp signal.

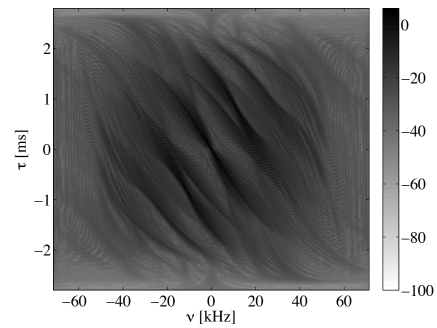


(b) The thresholded EAF of the chirp signal.

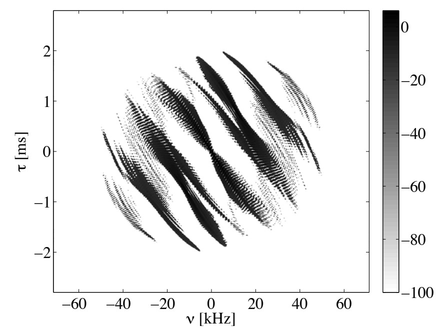


(c) The smoothed AF of the chirp signal.

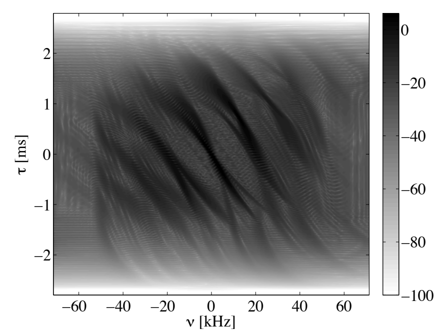
Fig. 1. Estimating the AF of the chirp signal.



(a) The EAF of the bat signal.



(b) The thresholded EAF of the bat signal.



(c) The smoothed AF of the bat signal.

Fig. 2. Estimating the AF of the bat signal.