PHASE-ONLY INFORMATION LOSS

Barry Quinn

Department of Statistics, Faculty of Science
Macquarie University, Sydney, NSW, 2109, Australia
Email: Barry.Quinn@mq.edu.au

ABSTRACT

In many areas of signal processing, the phases of complex-valued random variables are used to estimate system parameters, the magnitudes being discarded. In this paper, we consider the implications of doing this: the loss of statistical information and subsequent increase in asymptotic variance. Two particular cases, those of estimating the phase of the mean of a complex distribution, and estimating the frequency of a complex sinusoid in white noise, are considered. The estimators are motivated by estimation under von Mises distributional assumptions. The asymptotic distributional properties are obtained under general assumptions, and are tested using a small number of simulations.

Index Terms— Phase estimation, frequency estimation.

1. INTRODUCTION

[1] considered the problem of estimating the phase ϕ of the mean μ of independent and identically distributed (i.i.d.) random variables X_t, when the moduli of the $|X_t|$ are discarded. The problem was motivated by the more complicated one of estimating the frequency ω in a noisy complex sinusoid, using only the phases, in other words from $\{\angle X_t; 1 \leq t \leq T\}$ when X_t satisfies

$$X_t = \rho e^{i(\omega t+\phi)} + \varepsilon_t, \tag{1}$$

and the ε_t are i.i.d. with zero mean. It was suggested that ϕ be estimated by minimizing the sum of squares of the "distances" between the phases and ϕ, where the distances are computed modulo 2π, i.e. as angular distances on the unit circle, with the angle $-\pi$ identified with $+\pi$.

Recently [2] this approach was applied to the problem of estimating ϕ and ω in (1) . In both [1] and [2] the estimators were proven to be strongly consistent and to satisfy a central limit theorem.

In this paper, we consider another approach, which is based on fitting a von Mises distribution to the angular noise, even when it does not have a von Mises distribution. The estimators in both cases are shown to have excellent asymptotic properties, which are compared theoretically, and via simulations, with the properties of the estimators proposed in [1] and [2].

2. ESTIMATING THE PHASE OF THE MEAN

We consider the problem of estimating ϕ in

$$X_t = \mu + \varepsilon_t, \tag{2}$$

where $\mu = \rho e^{i\phi}, \rho > 0, \phi \in [-\pi, \pi)$, the ε_t are i.i.d. with mean 0, and $\theta_t = \angle\varepsilon_t$ is uniformly distributed on $[-\pi, \pi)$. This will be the case, for example, when ε_t has the standard complex normal distribution. It is imposed so that the distribution of $e^{i\alpha}\varepsilon_t$ is the same as the distribution of ε_t, for any $\alpha \in \mathbb{R}$. Now, under Gaussian assumptions on $\{\varepsilon_t\}$, the maximum likelihood estimator (MLE) of μ given $\{X_1, \ldots, X_T\}$ is

$$\overline{X}_T = T^{-1} \sum_{t=1}^{T} X_t = \overline{\rho}_T e^{i\overline{\phi}_T}, \tag{3}$$

say, where $\overline{\rho}_T > 0, \overline{\phi}_T \in [-\pi, \pi)$, and the MLE of ϕ is then $\overline{\phi}_T$. We shall not, however, assume Gaussianity in what follows.

Suppose that it is the $X_t / |X_t| = e^{iY_t}$ that are observed, where $Y_t \in [-\pi, \pi)$, and not the X_t. Let

$$\{x\} = x - \lfloor x \rceil ,$$

the difference between x and its nearest integer $\lfloor x \rceil$. Then

$$e^{iY_t} = e^{i\phi} \frac{1 + \rho^{-1} R_t e^{i(\theta_t - \phi)}}{\left|1 + \rho^{-1} R_t e^{i(\theta_t - \phi)}\right|},$$

where $R_t = |\varepsilon_t|$, and so

$$Y_t = 2\pi \left\{ \frac{\phi + Z_t}{2\pi} \right\}, \tag{4}$$

where

$$e^{iZ_t} = \frac{1 + \rho^{-1} R_t e^{i(\theta_t - \phi)}}{\left|1 + \rho^{-1} R_t e^{i(\theta_t - \phi)}\right|}. \tag{5}$$

In other words,

$$Z_t = \angle \left(1 + \rho^{-1} R_t e^{i(\theta_t - \phi)}\right).$$

Note that the distribution of Z_t is the same for all ϕ. Since $e^{i(\theta_t - \phi)}$ is uniformly distributed on the unit circle, Z_t is symmetrically distributed about 0. It is obvious that $\overline{Y}_T = T^{-1} \sum_{t=1}^{T} Y_t$ is not a good estimator of ϕ, especially if ϕ is close to $\pm\pi$. In fact, since the Y_t are i.i.d., $\overline{Y}_T$ converges almost surely to

$$\begin{aligned} E(Y_t) &= E\left(2\pi \left\{ \frac{\phi + Z_t}{2\pi} \right\}\right) \\ &= E(\phi + Z_t - 2\pi k_t) \\ &= \phi - 2\pi E(k_t), \end{aligned}$$

where

$$k_t = \begin{cases} 1 & ; \quad \phi + Z_t > \pi \\ -1 & ; \quad \phi + Z_t < -\pi \\ 0 & ; \quad \text{otherwise.} \end{cases}$$

Thus $\overline{Y}_T$ converges almost surely to the wrong value, unless $E(k_t) = 0$. It is easily shown that

$$E(k_t) = \Pr(Z_t > \pi - |\phi|).$$

Hence $\overline{Y}_T$ is only strongly consistent when the support of the distribution of Z_t is not all of $[-\pi, \pi)$, in other words, when there is no possibility of 'wrapping'.

In [1], it was proposed that ϕ instead be estimated by

$$\widetilde{\phi}_T = \arg\min_{\phi} \sum_{t=1}^{T} \{(Y_t - \phi) / (2\pi)\}^2.$$

It was shown there that, if the distribution of Z_t is also unimodal, with mode at 0, then $\widetilde{\phi}_T$ converges almost surely to ϕ, and that the distribution of

$$T^{1/2}\left(\widetilde{\phi}_T - \phi\right)$$

converges to the normal with mean 0 and variance

$$\frac{\sigma_Z^2}{(1 - 2\pi f_Z(\pi))^2}, \tag{6}$$

where $\sigma_Z^2 = \operatorname{var} Z_t$ and $f_Z(z)$ is the common probability density function of the Z_t. We now describe another estimator, which is easier to compute, and whose asymptotic behavior is much simpler to deduce. In order to motivate the estimator, we consider the model in (4), with Z_t having a von Mises distribution and probability density function

$$f_Z(z) = \begin{cases} \dfrac{1}{2\pi I_0(\kappa)} e^{\kappa \cos z} & ; \quad z \in [-\pi, \pi) \\ 0 & ; \quad \text{otherwise}, \end{cases} \tag{7}$$

where I_0 is the modified Bessel function of order 0. The log-likelihood given $Y_1, \ldots, Y_T$ is then

$$l(\phi, \kappa; Y) = -T \log(2\pi I_0(\kappa)) + \kappa \sum_{t=1}^{T} \cos(Y_t - \phi),$$

which is linear in $\sum_{t=1}^{T} \cos Y_t$ and $\sum_{t=1}^{T} \sin Y_t$. All MLE's are functions of these statistics, and asymptotically attain the relevant Cramér-Rao lower bounds in the sense that the asymptotic distributions have these variances. It thus makes sense to consider functions of these statistics when estimating parameters. Now

$$\begin{aligned} \kappa^{-1} \frac{\partial l(\phi, \kappa; Y)}{\partial \phi} &= \sum_{t=1}^{T} \sin(Y_t - \phi) \\ &= \cos\phi \sum_{t=1}^{T} \sin Y_t - \sin\phi \sum_{t=1}^{T} \cos Y_t, \end{aligned}$$

and so $\widehat{\phi}_T$, the MLE of ϕ, satisfies $e^{i\widehat{\phi}_T} = Q \sum_{t=1}^{T} e^{iY_t}$, where $Q = \frac{\sin \widehat{\phi}_T}{\sum_{t=1}^{T} \sin Y_t} \in \mathbb{R}$. Thus

$$e^{i\left(\widehat{\phi}_T - \phi\right)} = Q \sum_{t=1}^{T} e^{i(Y_t - \phi)} = Q \sum_{t=1}^{T} e^{iZ_t}. \tag{8}$$

It is easy to show that we need $Q > 0$ in order that the likelihood be *maximized* at $\widehat{\phi}_T$. We now discard the von Mises assumption, assuming only that the Z_t are i.i.d, symmetrically about 0. Then

$$T^{-1} \sum_{t=1}^{T} e^{iZ_t} \to \mu_C, \tag{9}$$

almost surely as $T \to \infty$, and

$$T^{-1} \sum_{t=1}^{T} e^{iZ_t} = \mu_C + T^{-1/2}(U_T + iV_T),$$

where

$$U_T = T^{1/2}\left(\sum_{t=1}^{T} \cos Z_t - \mu_C\right), V_T = T^{1/2} \sum_{t=1}^{T} \sin Z_t$$

and $\mu_C = E \cos Z_t$. Now $E \sin Z_t = 0$ and

$$E(\cos Z_t \sin Z_t) = \frac{1}{2} E \sin(2Z_t) = 0.$$

Thus U_T and V_T are asymptotically normally distributed with means 0, and variances $\sigma_C^2 = \operatorname{var} \cos Z_t$ and$\sigma_S^2 = E \sin^2 Z_t$, respectively. Assuming that $\mu_C > 0$, it follows from (9) that $\widehat{\phi}_T$ converges almost surely to ϕ, that

$$\widehat{\phi}_T - \phi = \arg\left(1 + T^{-1/2} \mu_C^{-1}(U_T + iV_T)\right)$$

and so

$$T^{1/2}\left(\widehat{\phi}_T - \phi\right) = \mu_C^{-1} V_T + o_P(1).$$

Hence, as long as $\mu_C > 0$, we have

Theorem 1 *$\widehat{\phi}_T$ converges almost surely to ϕ and $T^{1/2}\left(\widehat{\phi}_T - \phi\right)$ is asymptotically normally distributed with mean 0 and variance $\mu_C^{-2} \sigma_S^2$.*

If $\mu_C < 0$, the estimator will be asymptotically π out of phase. As well, if $\mu_C = 0$, which occurs, for example, when Z_t is uniformly distributed, $e^{i(Z_t + \phi)}$ is uniformly distributed on the unit circle, and so the parameter ϕ is not identified.

It is of interest to compare this asymptotic variance with that given in [1]. We shall do this for the special cases where Z_t has the von Mises distribution given in (7) and where ε_t in (2) is Gaussian.

Lemma 2 *When Z_t has the distribution given in (7), $T^{1/2}\left(\widehat{\phi}_T - \phi\right)$ is asymptotically normally distributed with mean 0 and variance*

$$\frac{I_0(\kappa)}{\kappa I_1(\kappa)}.$$

Proof. Since

$$\int_0^{\pi} \cos(jx) e^{\kappa \cos x} dx = \pi I_j(\kappa),$$

where I_j is the modified Bessel function of order j, it follows that $\mu_C = \{I_0(\kappa)\}^{-1} I_1(\kappa)$,

$$\begin{aligned} \sigma_S^2 &= \frac{1}{2\pi I_0(\kappa)} \int_0^{\pi} (1 - \cos(2x)) e^{\kappa \cos x} dx \\ &= \frac{1}{2}\left(1 - \frac{I_2(\kappa)}{I_0(\kappa)}\right), \end{aligned}$$

and the asymptotic variance is thus

$$\frac{(I_0(\kappa) - I_2(\kappa)) I_0(\kappa)}{2(I_1(\kappa))^2} = \frac{I_0(\kappa)}{\kappa I_1(\kappa)}. \tag{10}$$

■

The results of representative simulations are shown in figure 1. The Best-Fisher algorithm [3] was used to simulate von Mises random variables. Each point represents the sample mean square error computed from 1000 replications, for the case $(T,\phi) = (50, \pi - 0.1)$. The 'wrapped' mean square error was computed, i.e. $2\pi\left\{\left(\widehat{\phi}_T - \phi\right)/(2\pi)\right\}$ was used as the difference between the estimator and the true value, for obvious reasons. Also shown are the 'asymptotic' variances calculate from (10) ('CRLB' here) and (6)('least squares'). The ratio of the two asymptotic variances, i.e. the asymptotic relative efficiency of the 'least squares' estimator, is easily shown to decrease from $\pi^2/6$, (the limit as $\kappa \to 0$) to 1, (the limit as $\kappa \to \infty$), and is about 1.016 when $\kappa = 5$.

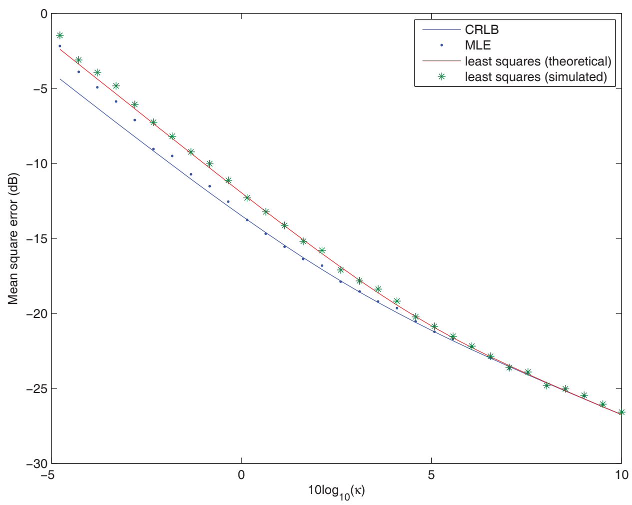

Fig. 1. Von Mises simulated mses, $T = 50, \phi = \pi - 0.1$

When $\{Z_t\}$ is derived from (2), and the distribution of the ε_t is specified, the asymptotic computations are more complicated. We consider as an example the Gaussian case.

Lemma 3 *Suppose the ε_t are complex normal, with means 0, and real and imaginary parts independent with the same variance σ^2. Then $T^{1/2}\left(\widehat{\phi}_T - \phi\right)$ is asymptotically normally distributed with mean 0 and variance*

$$\frac{8\nu^4\left(e^{\frac{1}{2\nu^2}} - 1\right)}{\pi\left(I_0\left(\frac{1}{4\nu^2}\right) + I_1\left(\frac{1}{4\nu^2}\right)\right)^2}$$

Proof. Although it is tempting to consider computing σ_S^2 and μ_C directly from the distribution of Z_t, it proved simpler not to. From (5),

$$e^{iZ_t} = \frac{1+U+iV}{\sqrt{(1+U)^2+V^2}},$$

where U and V are normal and independent with means 0 and variances $\nu^2 = \sigma^2/\rho^2$. Thus

$$E\sin^2 Z_t = E\frac{V^2}{(1+U)^2+V^2}$$

and

$$E\cos Z_t = E\frac{1+U}{\sqrt{(1+U)^2+V^2}}.$$

Making the substitution $v = |1+u|\,x$, we obtain

$$\begin{aligned} &\sigma_S^2 \\ &= \frac{1}{2\pi\nu^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{v^2}{(1+u)^2+v^2}e^{-\frac{1}{2\nu^2}\left(u^2+v^2\right)}dvdu \\ &= 1 - e^{-\frac{1}{2\nu^2}}\int_0^{\infty} u\,\mathrm{erfc}\,u\left(e^{\frac{u\sqrt{2}}{\nu}} + e^{-\frac{u\sqrt{2}}{\nu}}\right)du \\ &= \nu^2\left(1 - e^{-\frac{1}{2\nu^2}}\right), \end{aligned}$$

and

$$\begin{aligned} &\mu_C \\ &= \frac{1}{2\pi\nu^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1+u}{\sqrt{(1+u)^2+v^2}}e^{-\frac{1}{2\nu^2}\left(u^2+v^2\right)}dvdu \\ &= \frac{1}{2\pi\nu^2}e^{-\frac{1}{2\nu^2}}\int_0^{\infty} ue^{-\frac{u^2}{4\nu^2}}\left(e^{\frac{u}{\nu^2}} - e^{-\frac{u}{\nu^2}}\right)K_0\left(\frac{u^2}{4\nu^2}\right)du \\ &= \sqrt{\frac{\pi}{8\nu^2}}e^{-\frac{1}{4\nu^2}}\left(I_0\left(\frac{1}{4\nu^2}\right) + I_1\left(\frac{1}{4\nu^2}\right)\right), \end{aligned}$$

where K_0 is the modified Bessel function of the second kind, of order 0. Note that $\mu_C > 0\ \forall\nu$, since the modified Bessel functions are non-negative. The asymptotic variance is thus

$$\frac{8\nu^4\left(e^{\frac{1}{2\nu^2}} - 1\right)}{\pi\left(I_0\left(\frac{1}{4\nu^2}\right) + I_1\left(\frac{1}{4\nu^2}\right)\right)^2}. \tag{11}$$

■

This variance should be compared with two asymptotic variances: the asymptotic variance of the 'least squares' estimator, given by (6), and the asymptotic variance of the Gaussian maximum likelihood estimator constructed from the X_t, i.e. using amplitude and phase. The latter asymptotic variance is given in the next lemma.

Lemma 4 *$\overline{\phi}_T$, given in (3), converges almost surely to ϕ, and the distribution of $T^{1/2}\left(\overline{\phi}_T - \phi\right)$ converges to the normal with mean 0 and variance $\nu^2 = \sigma^2/\rho^2$.*

Proof.

$$\overline{\rho}_T e^{i\overline{\phi}_T} = \rho e^{i\phi}\left(1 + T^{-1}\sum_{t=1}^{T}\rho^{-1}e^{-i\phi}\varepsilon_t\right)$$

Thus

$$\frac{\overline{\rho}_T}{\rho}e^{i\left(\overline{\phi}_T - \phi\right)} = 1 + T^{-1/2}\left(U_T + iV_T\right),$$

where $U_T = T^{-1/2}\sum_{t=1}^{T}\rho^{-1}\,\mathrm{Re}\left(e^{-i\phi}\varepsilon_t\right)$ and $V_T = T^{-1/2}\sum_{t=1}^{T}\rho^{-1}\,\mathrm{Im}\left(e^{-i\phi}\varepsilon_t\right)$ are asymptotically normal and independent with zero means and variances $\nu^2 = \sigma^2/\rho^2$. In fact, they are exactly normally distributed if the ε_t are Gaussian. Thus

$$T^{1/2}\left(\overline{\phi}_T - \phi\right) = V_T + o_P(1)$$

and consequently is asymptotically normal with mean 0 and variance ν^2. Note that the result does not depend at all on $\{X_t\}$ being Gaussian. ■

Depicted in figure 2 are the results of simulations for the Gaussian case. Each point represents the sample mean square error computed from 1000 replications, for the case $(T, \phi) = (50, \pi - 0.1)$, and for signal to noise ratios (SNRs) between -10 and 20 dB, with SNR defined to be $10 \log_{10} \left(\rho^2 / \left(2\sigma^2 \right) \right)$. The 'wrapped' mean square errors were again computed. Also shown are the 'asymptotic' mean square errors calculated from (11) ('CRLB') and (6) ('least squares').

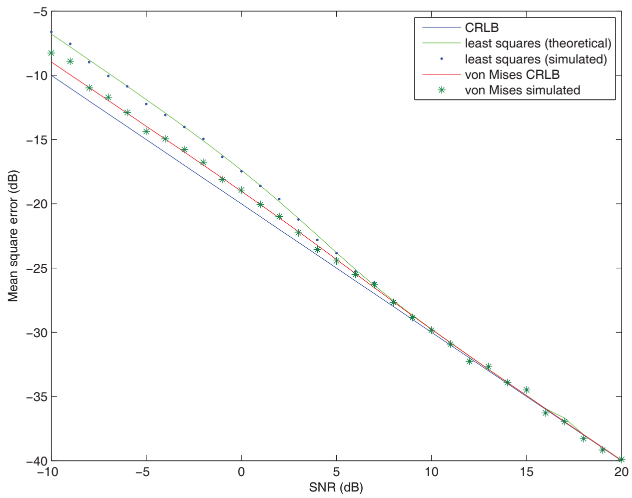

Fig. 2. Normal simulated mses, $T = 50, \phi = \pi - 0.1$

The asymptotic efficiencies, relative to the Cramér-Rao lower bound, are at worst $4/\pi$ and $2\pi/3$ respectively for the phase-only (von Mises) and least squares estimators, the worst cases occurring in the limit as $\nu^2 \to \infty$, i.e. as the SNR diverges to $-\infty$. The worst losses are thus approximately 1.05 dB and 3.21 dB respectively.

3. PHASE-ONLY FREQUENCY ESTIMATION

We now turn to the problem which motivated the above work, that of estimating frequency using phases only. Let $\{X_t\}$ satisfy

$$X_t = \rho e^{i(\phi + \omega t)} + \varepsilon_t,$$

where $\{\varepsilon_t\}$ again satisfies the conditions below (2). Now

$$\begin{aligned} e^{iY_t} &= \frac{X_t}{|X_t|} \\ &= e^{i(\phi+\omega t)} \frac{1 + \rho^{-1} R_t e^{i(\theta_t - \phi - \omega t)}}{\left|1 + \rho^{-1} R_t e^{i(\theta_t - \phi - \omega t)}\right|} \\ &= e^{i(\phi+\omega t)} e^{iZ_t}, \end{aligned}$$

and

$$Y_t = 2\pi \left\{ \frac{\phi + \omega t + Z_t}{2\pi} \right\}$$

where

$$Z_t = \angle \left(1 + \rho^{-1} R_t e^{i(\theta_t - \phi - \omega t)} \right).$$

Note that the Z_t are again i.i.d., with distributions not depending on ϕ, ω or t, since the $e^{i(\theta_t - \phi - \omega t)}$ are uniformly distributed on the unit circle. There is a substantial literature concerned with the estimation of ϕ and ω using only the Y_t, which appear to depend linearly on both ϕ and ω. However, 'wrapping' creates problems, and results in many techniques having poor asymptotic properties. See for example chapter 7 of [4]. Recently in [2] a 'least squares' technique has been proposed, the analogue of that proposed in [1]. It is shown that $T^{3/2} \left(\widetilde{\omega}_T - \omega \right)$ is asymptotically normal with mean 0 and variance

$$\frac{12 \operatorname{var} Z_t}{\left(1 - 2\pi f_Z(\pi)\right)^2},$$

where

$$\left(\widetilde{\phi}_T, \widetilde{\omega}_T \right) = \arg\min_{\phi, \omega} \sum_{t=1}^{T} \left\{ \frac{Y_t - \phi - \omega t}{2\pi} \right\}^2.$$

Suppose now that Z_t had a von Mises distribution, with probability density function given by (7). The log-likelihood would then be

$$l(\phi, \omega, \kappa; Y) = -T \log \left(2\pi I_0(\kappa) \right) + \kappa \sum_{t=1}^{T} \cos \left(Y_t - \phi - \omega t \right),$$

which, using the same method as before, is easily seen to have maximum value with respect to ϕ

$$-T \log \left(2\pi I_0(\kappa) \right) + \kappa \left| \sum_{t=1}^{T} e^{i(Y_t - \omega t)} \right|.$$

The MLE of ω would then be the maximizer $\widehat{\omega}_T$ of the periodogram of the $e^{iY_t} = X_t / |X_t|$. The asymptotic properties of $\widehat{\omega}_T$, which we propose as estimator in the general case, i.e. *without* von Mises assumptions, are given below.

Theorem 5 *$\widehat{\omega}_T$ converges almost surely to ω and $T^{3/2} \left(\widehat{\omega}_T - \omega \right)$ is asymptotically normally distributed with mean 0 and variance $12 \mu_C^{-2} \sigma_S^2$.*

The proof will not be given here, but is along lines similar to those in [6], which contains more general results for the case of a real sinusoid, with additive stationary (colored) noise.

Remark: Of interest is the fact that the asymptotic relative efficiencies of the 'least squares' and von Mises estimators are the same as in the previous section, since the 'asymptotic variance' of $T^{3/2} \left(\widehat{\omega}_P - \omega \right)$, where $\widehat{\omega}_P$ is the maximizer of the periodogram of the X_t, is $12\nu^2$. The losses in statistical efficiency are thus exactly the same as in the previous section. In particular, the asymptotic variance of the new estimator is at worst $4/\pi$ times the asymptotic variance of the ordinary periodogram estimator.

4. REFERENCES

[1] B.G. Quinn, "Estimating the mode of a phase distribution," in *Proceedings of the 2007 Asilomar conference on Signals, Systems and Computers*, IEEE, 2007, pp. 587–591.

[2] R.G. McKilliam, B.G. Quinn, I.V.L. Clarkson and B. Moran, "Frequency estimation by phase unwrapping", submitted.

[3] D.J. Best and N.I. Fisher, "Efficient Simulation of the von Mises Distribution," *Applied Statistics*, vol. 28, pp. 152–157, 1979.

[4] B.G. Quinn and E.J. Hannan, *The Estimation and Tracking of Frequency*, Cambridge University Press, 2001.

[5] B.G. Quinn and J.M. Fernandes, "A fast efficient technique for the estimation of frequency," *Biometrika*, vol. 78, pp. 489–98, 1991.

[6] E.J. Hannan, "The estimation of frequency," *Journal of Applied Probability*, vol. 10, pp. 510–519, 1973.