ESTIMATING THE CIRCULAR MEAN FROM CORRELATED OBSERVATIONS

Robby G. McKilliam[1] *and Barry G. Quinn*[2]

[1]Institute for Telecommunications Research, University of South Australia, Adelaide, AUSTRALIA

[2]Department of Statistics, Macquarie University, Sydney, AUSTRALIA

ABSTRACT

A common approach to estimating the mean direction from a sequence of observations on a circle is the *sample circular mean*. We describe the asymptotic properties of this estimator under the condition that the sequence of observations is stationary ergodic and satisfies a mixing condition. We describe the results of Monte-Carlo simulations. These simulations agree with our theoretical results.

Index Terms— Circular statistics, mean direction estimation, strong mixing processes.

1. INTRODUCTION

The field of circular statistics aims to describe the nature of data that is measured in angles or 2-dimensional unit vectors or complex numbers on the unit circle [1–6]. Such data occur frequently in science, particularly in astronomy, biology [7], and meteorology [8, 9], and also in engineering, particularly in communications and radar [10–16].

A fundamental and useful problem in circular statistics is that of estimating the *mean direction* of a circular random variable from a number, say N, of observations. For example, if you listen to the weather report you may be told (an estimate of) the direction of the wind. Obtaining an accurate estimate requires a method for accurately estimating the mean wind direction from a number of observations of the wind direction. The problem of estimating mean direction also appears in signal processing problems such as phase and frequency estimation [17–23], polynomial phase estimation [23–26], noncoherent detection [14, 23, 27] and delay and period estimation from sparse observations [10, 12, 13, 15, 23].

Various definitions of the mean direction exist, for example the *circular mean* [1, 5], the *circular median* [28, 29] and the *unwrapped mean* [23, 30]. The most common definition in the literature is the circular mean and this is the definition considered in this paper.

The circular mean is derived by considering circular random variables as unit vectors in 2-dimensions or equivalently complex numbers on the unit circle. Given N observations the appropriate estimator for the circular mean is the *sample circular mean* which can be computed efficiently by averaging N complex numbers. The asymptotic properties of the sample circular mean have been studied under various assumptions [1, 6, 17, 23]. To our knowledge all of the existing analysis make the assumption that the observations are independent and identically distributed (i.d.d.). In this paper we drop the i.i.d. assumption and instead assume that the sequence of observations is stationary ergodic and satisfies a mixing condition. The same mixing condition was assumed by Craig to derive asymptotic results about sequences of wrapped normal circular random variables [31, p. 19].

This paper is organised as follows. Section 2 and 3 introduces circular random variables, the circular mean, and relevant notation. This section is similar to that from [30], but we have included it again so that this paper is self contained. Section 4 introduces the mixing condition that we shall use and Section 5 describes the sample circular mean estimator and derives its asymptotic properties. Sections 6 and 7 describe the results of Monte-Carlo simulations.

1.1. Notations

We write random variables using capital letters, such as X and Y and circular random variables using the capital Greek letters Θ and Φ. When describing estimators we use a subscript zero, as in μ_0, to denote the *true* value of a parameter and a hat, as in $\hat{\mu}$, to denote an estimator of μ_0. We use $\langle x \rangle$ to denote x taken to its representative inside $[-\pi, \pi)$, i.e. taken modulo 2π.

2. CIRCULAR RANDOM VARIABLES

In this paper a circular random variable takes values on $[-\pi, \pi)$. It is common in the literature to define a special *circular* probability density function (pdf) f to be periodic with period 2π so that $f(\theta + 2\pi k) = f(\theta)$ for integers k and the integral $\int_T f(\theta) d\theta = 1$ where T is any interval of length on 2π. We will not use this definition. Here, a circular random variable is just a random variable with pdf that has support on $[-\pi, \pi)$. A circular random variable inherits all the properties of a *regular* random variable. For example, if Θ is a circular random variable with pdf f then the expected value of a function $g(\Theta)$ of Θ is given in the usual way by

$$E[g(\Theta)] = \int_{-\infty}^{\infty} g(\theta)f(\theta)d\theta = \int_{-\pi}^{\pi} g(\theta)f(\theta)d\theta.$$

This leads to the usual definitions of mean and variance for a circular random variable $E[\Theta]$ and $\operatorname{var}[\Theta] = E[\Theta^2] - E[\Theta]^2$. A little thought must be given here. The mean $E[\Theta]$ does not necessarily correspond to the *mean direction* of Θ in the sense one might expect. For example, consider a distribution with equal point masses at -0.49π and 0.49π. The expected value of this random variable is zero, but intuitively a more reasonable estimate for the mean direction is $-\pi$. This motivates the definition of the *circular mean*.

3. THE CIRCULAR MEAN AND VARIANCE

Given a circular random variable Θ with pdf f the most common analogue of 'mean' and 'variance' in the literature is the *circular mean* given by

$$\mu_{\text{circ}} = \angle c$$

and the *circular variance* given by

$$\nu = 1 - |c|$$

where

$$c = E\left[e^{j\Theta}\right] = \int_{-\pi}^{\pi} e^{j\theta} f(\theta)d\theta, \tag{1}$$

and $\angle c$ and $|c|$ are the complex argument and the magnitude of c [1, p. 29][5]. Consider the case when $c = 0$. The circular variance is 1 but the circular mean is undefined. We say that the distribution *has no circular mean* when $c = 0$.

For example, consider the *circular uniform distribution* with pdf displayed in Figure 1. From the symmetry of this distribution it can be seen that the value of c given by the integral (1) will be zero and we therefore conclude that this distribution has no circular mean. This result conforms well with our intuition. The circular uniform distribution is not the only case where the circular mean is undefined. For example, consider the bimodal pdf displayed in Figure 2. Its symmetry clearly ensures that $c = 0$.

4. STRONG MIXING PROCESSES

In Section 5 we consider the sample circular mean estimator of the circular mean given a sequence of circular data. This section describes the types of sequences we consider. Let $\{X_n, n \in \mathbb{Z}\}$ be a sequence of random variables, let $(\Omega, \mathcal{A}, P)$ be the probability space over which each of the X_n is defined, and let $\mathcal{M}_{-\infty}^{k}$ be the σ-algebra of events generated by $\{X_n, -\infty < n \leq k\}$ and $\mathcal{M}_{k}^{\infty}$ be the σ-algebra of events

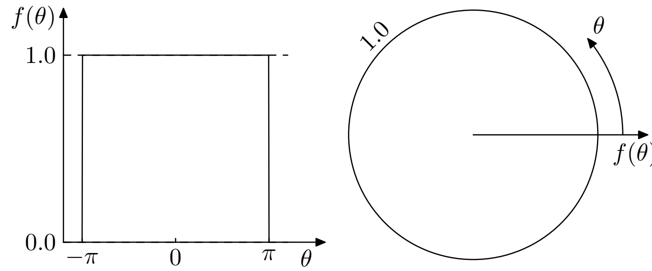

Fig. 1. The circular uniform distribution which has no circular mean.

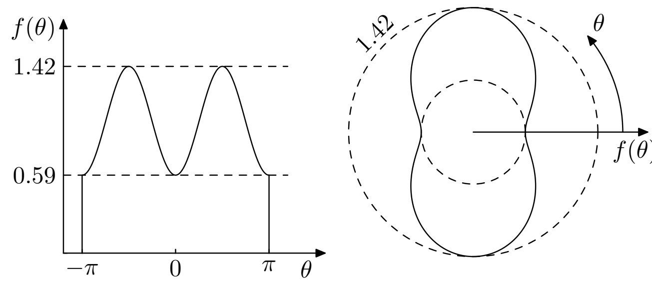

Fig. 2. A bimodal distribution with no circular mean.

generated by $\{X_n, k \leq n < \infty\}$. The sequence is called *strong mixing* if

$$\sup_{A \in \mathcal{M}_{-\infty}^{k}, B \in \mathcal{M}_{k+s}^{\infty}} |\operatorname{P}(A \cap B) - \operatorname{P}(A)\operatorname{P}(B)| \leq \alpha(s)$$

where $\alpha(s)$, called the *strong mixing coefficient*, goes to zero as s goes to ∞. Intuitively this implies that elements far apart in the sequence are almost independent. The strong mixing condition is placed on the underlying probability space, so for any scalar function g the sequence $\{g(X_n)\}$ is also strong mixing with coefficient $\alpha(s)$.

Strong mixing was first suggested by Rosenblatt [32, 33] as a condition under which a central limit theorem could be obtained for (strictly) stationary processes. The strong mixing condition has since been studied extensively and various different central limit theorems exist [34–36]. The next theorem, due to Ibragimov [34], is the most useful for our purposes.

Theorem 1. (Ibragimov) *Let $\{X_n\}$ be a stationary ergodic sequence of zero mean random variables that is strong mixing with coefficient $\alpha(s)$ such that $\sum_{s=1}^{\infty} \alpha(s)$ converges and with the $|X_n| \leq K$ bounded above by some constant K. Let*

$$S_N = \sum_{n=1}^{N} X_n.$$

Then the distribution of $S_N/\sqrt{N}$ converges to the zero mean normal with variance

$$E[X_1^2] + 2\sum_{n=2}^{\infty} E[X_1 X_n].$$

Proof. The proof is given by Ibragimov [34, Theorem 1.6]. The theorem is also a trivial specialisation of that given by Hall and Heyde [35, Corollary 5.1 and Theorem 5.1]. A weaker, but more immediately accessible result is given by Durret [37, page 422]. □

The following corollary describes properties of strong mixing sequences of circular random variables. These properties will be useful for describing the asymptotic behaviour of the sample circular mean estimator in the Section 5.

Corollary 1. *Let $\{\Phi_n, n \in \mathbb{Z}\}$ be a stationary ergodic sequence of circular random variables that is strong mixing with $\sum_{s=1}^{\infty} \alpha(s) < \infty$. If the Φ_n have zero circular mean and circular variance ν then*

$$\mathcal{I}_N = \sum_{n=1}^{N} \sin(\Phi_n) \qquad \textit{and}$$
$$\mathcal{R}_N = \sum_{n=1}^{N} (\cos(\Phi_n) + \nu - 1)$$

are such that $\mathcal{I}_N/\sqrt{N}$ converges to the zero mean normal with variance

$$E[\sin^2(\Phi_1)] + 2\sum_{n=2}^{\infty} E[\sin(\Phi_1)\sin(\Phi_n)]$$

and $\mathcal{R}_N/\sqrt{N}$ converges to the zero mean normal with finite variance

Proof. Because the Φ_n have zero circular mean then

$$E[\sin(\Phi_n)] = 0 \;\text{ and }\; E[\cos(\Phi_n) + \nu - 1] = 0$$

follows from (1). We immediately have that $\mathcal{I}_N$ and $\mathcal{R}_N$ are zero mean. Also

$$|\sin(\Phi_n)| \le 1 \;\text{ and }\; |\cos(\Phi_n) + \nu - 1| \le 2$$

are both bounded and therefore satisfy the requirements of Theorem 1. The proof follows. □

5. THE SAMPLE CIRCULAR MEAN ESTIMATOR

Let $\Theta_1, \dots, \Theta_N$ be circular random variables. The sample circular mean is given by

$$\hat{\mu} = \angle \bar{C} \tag{2}$$

where

$$\bar{C} = \frac{1}{N}\sum_{n=1}^{N} e^{j\Theta_n}. \tag{3}$$

The next theorem describes the asymptotic behaviour of this estimator under some assumptions about the distribution of $\Theta_1, \dots, \Theta_N$. The proof is based on that given in [17].

Theorem 2. *Let $\Theta_1, \cdots, \Theta_N$ be N observations of the form*

$$\Theta_n = \langle \Phi_n + \mu_0 \rangle \tag{4}$$

where $\{\Phi_n, n \in \mathbb{Z}\}$ is a stationary ergodic sequence of circular random variables with zero circular mean, circular variance ν and marginal pdf f. Let $\hat{\mu}$ denote the sample circular mean estimator of the Θ_n, then:

1. *(Strong consistency) The difference $\langle \hat{\mu} - \mu_0 \rangle$ converges almost surely to zero as $N \to \infty$.*
2. *(Asymptotic normality) If the sequence $\{\Phi_n\}$ is strong mixing with coefficients $\alpha(s)$ such that $\sum_{s=0}^{\infty} \alpha(s)$ converges then the distribution of $\sqrt{N}\langle \hat{\mu} - \mu_0 \rangle$ approaches the normal with zero mean and variance*

$$\frac{h}{(1-\nu)^2} \tag{5}$$

where

$$h = E[\sin^2(\Phi_1)] + 2\sum_{k=2}^{\infty} E[\sin(\Phi_1)\sin(\Phi_k)].$$

Before we give the proof note that the theorem places conditions on the the difference modulo 2π, i.e. $\langle \hat{\mu} - \mu_0 \rangle$, between the *true* circular mean μ_0 and the estimated circular mean $\hat{\mu}$ rather than directly on the difference $\hat{\mu} - \mu_0$. This makes intuitive sense because the angles μ_0 and $\mu_0 + 2\pi k$ are equivalent for any integer k. So, for example, we intuitively expect the angles 0.49π and -0.49π to be *close* together, the difference between them being $|\langle -0.49\pi - 0.49\pi \rangle| = 0.02\pi$, and *not* $|-0.49\pi - 0.49\pi| = 0.98\pi$. In this way, the theorem asserts that the estimator converges to the circular mean whenever it is defined.

The term $E[\sin(\Phi_1)\sin(\Phi_k)]$ is called the *circular covariance* between Φ_1 and Φ_k in the literature [31, 38]. Methods for estimating the circular covariance have been studied by Fisher and Lee [38, 39] and also Craig [31]. These methods can be used to compute confidence intervals for the sample circular mean estimator.

Proof. From (2), (3) and (4),

$$\hat{\mu} = \angle \bar{C} = \angle\left(\frac{1}{N}\sum_{n=1}^{N} e^{j\Theta_n}\right) = \angle\left(\frac{1}{N}\sum_{n=1}^{N} e^{j\langle \mu_0 + \Phi_n \rangle}\right).$$

Subtracting μ_0 from both sides and taking the result modulo 2π,

$$\langle \hat{\mu} - \mu_0 \rangle = \angle\left(\frac{1}{N}\sum_{n=1}^{N} e^{j\Phi_n}\right). \tag{6}$$

Because the Φ_n have zero circular mean the expectation $E[e^{j\Phi_n}] = 1 - \nu$ is a positive real and as $\{\Phi_n\}$ is ergodic

$$N^{-1}\sum_{n=1}^{N} e^{j\Phi_n} \to E[e^{j\Phi_n}] = 1 - \nu$$

almost surely as N goes to infinity. As the complex argument $\angle(1-\nu)=0$ then $\langle\hat{\mu}-\mu_0\rangle \to 0$ almost surely as $N\to\infty$. This completes the proof of strong consistency.

To prove the central limit theorem let

$$\mathcal{I}_N = \sum_{n=1}^{N} \sin(\Phi_n) \qquad \text{and}$$

$$\mathcal{R}_N + N(1-\nu) = \sum_{n=1}^{N} \cos(\Phi_n)$$

denote the real and imaginary parts of $\sum_{n=1}^{N} e^{j\Phi_n}$. From Corollary 1 we have that $\mathcal{R}_N/N$ and $\mathcal{I}_N/N$ are $O_p(N^{-1/2})$ and $\mathcal{I}_N/\sqrt{N}$ converges to the zero mean normal with variance h. So,

$$\begin{aligned}\sqrt{N}\,\langle\hat{\mu}-\mu_0\rangle &= \frac{\sqrt{N}}{2\pi}\angle\left(1-\nu+\mathcal{R}_N/N + j\mathcal{I}_N/N\right)\\ &= \frac{\sqrt{N}}{2\pi}\left(\frac{\mathcal{I}_N/N}{1-\nu+\mathcal{R}_N/N} + O_p(N^{-1})\right)\\ &= \frac{\mathcal{I}_N/\sqrt{N}}{2\pi(1-\nu)} + O_p(N^{-1/2}),\end{aligned}$$

by a Taylor expansion of the $\angle\cdot$ function. So $\sqrt{N}\,\langle\hat{\mu}-\mu_0\rangle$ converges in distribution to the normal with zero mean and variance $\frac{h}{4\pi^2(1-\nu)^2}$ by Corollary 1. $\square$

6. GENERATING CIRCULAR PROCESSES

To test the results from Theorem 2 we require a method for generating correlated sequences of circular random variables. We use the following proceedure. Let f and F be the marginal pdf and marginal cdf of the circular random variables we wish to generate. Construct a Gaussian process $\{X_n\}$ with known autocorrelation and marginal cdf G. A sequence of correlated circular random variables is

$$\{F^{-1}G(X_n)\}$$

where F^{-1} is the inverse of the cdf F. To compute the asymptotic variance of the sample circular mean estimator (5) we require to evaluate

$$E[\sin(F^{-1}G(X_1))\sin(F^{-1}G(X_n))]$$

for integers $n\geq 1$. For the circular distributions that we will consider there do not appear to be closed form expressions for these expectations so we have resorted to numerical computation.

The above describes an approach for generating correlated random variables with arbitrary marginal distribution. Fisher and Lee [39] have previously described specific methods for generating sequences of correlated circular random variables with projected normal and wrapped normal marginal distributions.

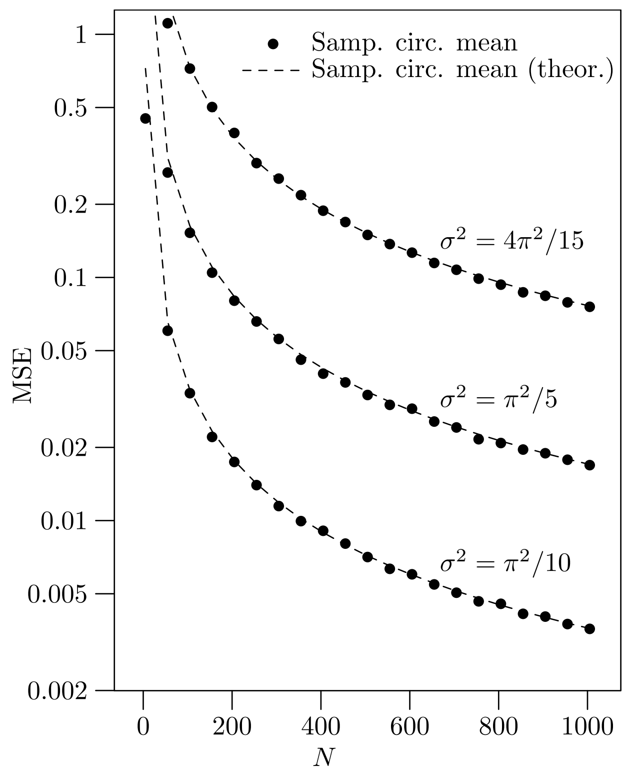

Fig. 3. MSE versus N with the wrapped uniform distribution.

7. SIMULATIONS

Figures 3 and 4 display the results of Monte Carlo simulations. In all simulations the sequence of circular random variables is generated according to Section 6 with the Gaussian process $\{X_n\}$ defined by

$$X_n = Z_n + \tfrac{1}{2}Z_{n-1} + \tfrac{1}{4}Z_{n-2}$$

where $\{Z_n\}$ is an i.i.d. sequence of Gaussian random variables with zero mean and variance 1. The figures display the mean square error (MSE) of the estimator with $N = 5+50k$ for $k=0,1,\dots,20$ where 10^5 trials are run for each value of N. In Figure 3 the marginal circular distribution is the zero mean wrapped uniform with variance $\frac{4\pi^2}{15}$, $\frac{\pi^2}{5}$ and $\frac{\pi^2}{10}$. In Figure 4 the marginal circular distribution is the zero mean wrapped normal with variance $0.4\pi^2$, $0.14\pi^2$ and $0.06\pi^2$. Both figures display the asymptotic variance predicted by Theorem 2. The theorem accurately models the behaviour of the sample circular mean estimator.

8. CONCLUSION

We considered the sample circular mean estimator of the circular mean of a circular random variable. We derived the asymptotic properties of the sample circular mean under the

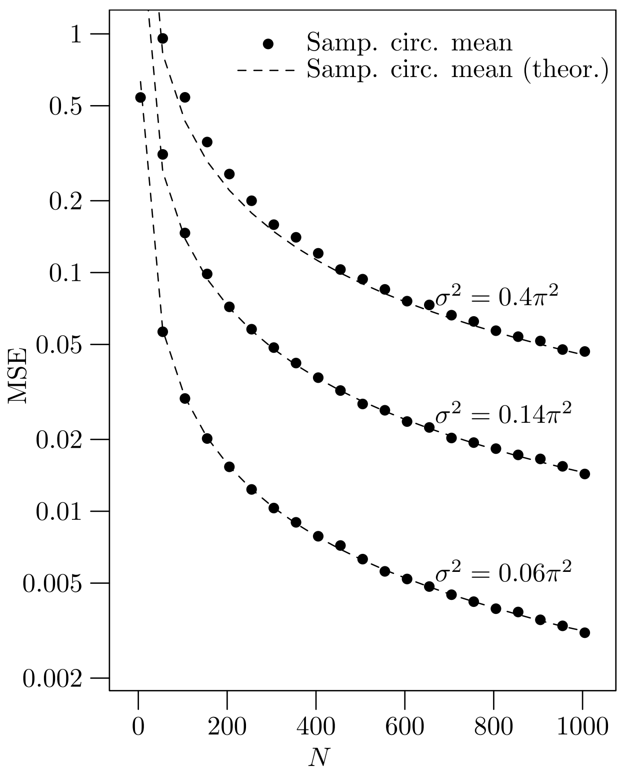

Fig. 4. MSE versus N with the wrapped normal distribution.

assumption that the sequence of observations is stationary ergodic and strongly mixing. The results of Monte-Carlo simulations were presented. These simulation support our theoretical anaylsis.

References

[1] K. V. Mardia and P. Jupp, *Directional Statistics*, John Wiley & Sons, 2nd edition, 2000.

[2] P. E. Jupp and K. V. Mardia, "A unified view of the theory of directional statistics, 1975-1988," *International Statistical Review*, vol. 57, no. 3, pp. 261–294, 1989.

[3] K. V. Mardia, "Statistics of directional data," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 37, no. 3, pp. 349–393, 1975.

[4] K. V. Mardia, *Statistics of directional data*, Academic Press, London, New York, 1972.

[5] N. I. Fisher, *Statistical analysis of circular data*, Cambridge University Press, 1993.

[6] A. Sengupta S. R. Jammalamadaka, *Topics in circular statistics*, World Scientific, 2001.

[7] E. Batschelet, *Circular statistics in biology*, Academic Press, New York, 1981.

[8] J. A. Bowers, I. D. Morton, and G. I. Mould, "Directional statistics of the wind and waves," *Applied Ocean Research*, vol. 22, no. 1, pp. 13–30, 2000.

[9] K. V. Mardia, "Directional statistics in geosciences," *Communications in Statistics - Theory and Methods*, vol. 10, no. 15, pp. 1523–1543, 1981.

[10] E. Fogel and M. Gavish, "Parameter estimation of quasi-periodic sequences," *Proc. Internat. Conf. Acoust. Spe. Sig. Process.*, vol. 4, pp. 2348–2351, 1988.

[11] S. D. Elton and D. A. Gray, "The application of circular statistics to specific radar pulse train detection," *Proceedings EUSIPCO*, vol. 1, pp. 93–96, 1994.

[12] S. D. Elton and B. J. Slocumb, "Robust parameter estimation for periodic point process signals using circular statistics," *Proceedings EUSIPCO*, 1996.

[13] I. V. L. Clarkson, "Approximate maximum-likelihood period estimation from sparse, noisy timing data," *IEEE Trans. Sig. Process.*, vol. 56, no. 5, pp. 1779–1787, May 2008.

[14] R. G. McKilliam, I. V. L. Clarkson, D. J. Ryan, and I. B. Collings, "Linear-time block noncoherent detection of PSK," *Proc. Internat. Conf. Acoust. Spe. Sig. Process.*, pp. 2465–2468, Apr. 2009.

[15] R. G. McKilliam and I. V. L. Clarkson, "Maximum-likelihood period estimation from sparse, noisy timing data," *Proc. Internat. Conf. Acoust. Spe. Sig. Process.*, pp. 3697–3700, Mar. 2008.

[16] I. V. L. Clarkson, S. D. Howard, and I. M. Y. Mareels, "Estimating the period of a pulse train from a set of sparse, noisy measurements," *Proc. Internat. Sympos. Signal Process. Appl.*, vol. 2, pp. 885–888, August 1996.

[17] B. G. Quinn, "Phase-only information loss," *Proc. Internat. Conf. Acoust. Spe. Sig. Process.*, pp. 3982–3985, Mar. 2010.

[18] B. C. Lovell, P. J. Kootsookos, and R. C. Williamson, "The circular nature of discrete-time frequency estimates," *Proc. Internat. Conf. Acoust. Spe. Sig. Process.*, vol. 5, pp. 3369–3372, April 1991.

[19] R. G. McKilliam, B. G. Quinn, I. V. L. Clarkson, and B. Moran, "Frequency estimation by phase unwrapping," *IEEE Trans. Sig. Process.*, vol. 58, no. 6, pp. 2953–2963, June 2010.

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

[21] I. V. L. Clarkson, "Frequency estimation, phase unwrapping and the nearest lattice point problem," *Proc. Internat. Conf. Acoust. Spe. Sig. Process.*, vol. 3, pp. 1609–1612, Mar. 1999.

[22] S. A. Tretter, "Estimating the frequency of a noisy sinusoid by linear regression," *IEEE Trans. Inform. Theory*, vol. 31, no. 6, pp. 832–835, November 1985.

[23] R. G. McKilliam, *Lattice theory, circular statistics and polynomial phase signals*, Ph.D. thesis, University of Queensland, Australia, December 2010.

[24] R. G. McKilliam, I. V. L. Clarkson, B. G. Quinn, and B. Moran, "Polynomial-phase estimation, phase unwrapping and the nearest lattice point problem," *Asilomar Conference on Signals, Systems, and Computers*, pp. 493–495, Nov. 2009.

[25] J. Kitchen, "A method for estimating the coefficients of a polynomial phase signal," *Signal Processing*, vol. 37, no. 1, pp. 463–470, Feb 1994.

[26] M.R. Morelande, "Parameter estimation of phase-modulated signals using bayesian unwrapping," *IEEE Trans. Sig. Process.*, vol. 57, no. 11, pp. 4209–4219, Nov. 2009.

[27] K.M. Mackenthun, "A fast algorithm for multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 42, pp. 1471–1474, Feb/Mar/Apr 1994.

[28] Sumitra Purkayastha, "An almost sure representation of sample circular median," *Journal of Statistical Planning and Inference*, vol. 46, no. 1, pp. 77 – 91, 1995.

[29] N. I. Fisher, "Spherical medians," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 47, no. 2, pp. 342–348, 1985.

[30] R. G. McKilliam, B. G. Quinn, and I. V. L. Clarkson, "Direction estimation by minimum squared arc length," submitted to IEEE Trans. Sig. Proc., Mar. 2011.

[31] P. S. Craig, *Time Series Analysis for Directional Data*, Ph.D. thesis, Trinity College Dublin, Sep. 1988.

[32] M. Rosenblatt, "A central limit theorem and a strong mixing condition," *Proc. Natl. Acad. Sci.*, vol. 42, pp. 43–47, 1956.

[33] R. Blum and M. Rosenblatt, "A class of stationary processes and a central limit theorem," *Proc. Natl. Acad. Sci.*, vol. 42, pp. 412–413, 1956.

[34] I. A. Ibragimov, "Some limit theorems for stationary processes," *Theory of probability and its applications*, vol. 7, no. 4, pp. 349–382, April 1962.

[35] P. Hall and C. C. Heyde, *Martingale limit theory and its application*, Academic press New York, 1980.

[36] R. Yokoyama, "Moment bounds for stationary mixing sequences," *Probability Theory and Related Fields*, vol. 52, pp. 45–57, 1980.

[37] R. Durrett, *Probability: theory and examples*, Thomson, 2005.

[38] N. I. Fisher and A. J. Lee, "A correlation coefficient for circular data," *Biometrika*, vol. 70, no. 2, pp. 327–332, Aug. 1983.

[39] N. I. Fisher and A. J. Lee, "Time series analysis of circular data," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 56, no. 2, pp. 327–339, 1994.