FREQUENCY ESTIMATION, PHASE UNWRAPPING AND THE NEAREST LATTICE POINT PROBLEM

I. V. L. Clarkson

Department of Electrical and Electronic Engineering
The University of Melbourne
Parkville 3052
AUSTRALIA
v.clarkson@ee.mu.oz.au

ABSTRACT

In this paper, we examine the relationship between frequency estimation and phase unwrapping and a problem in algorithmic number theory known as the nearest lattice point problem. After briefly reviewing the theory of these three topics, we introduce an interpretation of the maximum likelihood frequency estimation problem as a nearest lattice point problem. We develop an algorithm based on this approach and present numerical results to compare its performance with other estimation techniques. We find that the algorithm has good powers of estimation.

1. INTRODUCTION

The purpose of this paper is to elucidate the connection between the problem of frequency estimation and its solution by maximisation of the periodogram and by methods of phase unwrapping and the number theoretic problem of determining the nearest point in a lattice to a given point with respect to a norm.

The frequency estimation problem, by which we mean the estimation of the frequency of a single tone in noisy, sampled data, is one of the core studies in signal processing. It has applications in radar, sonar, telecommunications and medicine. In their paper [1], Rife & Boorstyn proposed a model for the problem in which complex data is used, containing a cisoid corrupted by additive white complex Gaussian noise. They found that the maximum likelihood solution to the problem is obtained through maximisation of the periodogram of the sampled data.

Although the amount of computational effort which must be expended to maximise the periodogram is not too great, research has been conducted on algorithms which can estimate frequency nearly as well but with less effort. Tretter [2] proposed using the unwrapped phase of a signal to perform frequency estimation. He showed that linear regression can be performed on the unwrapped phases, yielding an accurate estimate of frequency when the signal-to-noise ratio (SNR) is sufficiently high. Simple phase unwrapping and linear regression can both be performed more quickly than the maximum likelihood technique for sufficiently large data sets.

A problem in algorithmic number theory is the determination of the nearest element of a point lattice to any other given point according to a prescribed norm. It has been conjectured [3] that this problem is computationally infeasible under some conditions. Therefore, there is some interest in discovering ways of finding a provably almost-nearest lattice point with only a modest amount of computation. One such algorithm has been formulated by Babai [4], drawing on the lattice reduction algorithm of Lenstra, Lenstra & Lovász [5].

Our aim in this paper is to demonstrate the similarity between the three problems and to apply Babai's algorithm to the frequency estimation problem. We will show that this algorithm can be interpreted as explicitly unwrapping the phase of the signal and performing a weighted linear regression on the result. We will demonstrate through numerical simulation that it can operate at reasonably low SNRs. However, our algorithm is not intended to be a practical replacement to the other algorithms, but rather serves to illustrate the connection between the two seemingly disparate practices of lattice reduction and frequency estimation.

2. SIGNAL MODEL

We will assume a record of complex data of length N containing a constant amplitude cisoid in additive white Gaussian noise. Each datum z_n, $n = 1, 2, \ldots, N$ is an instance of a complex random variable Z_n where

$$Z_n = b\exp[i(2\pi f_c n + \theta)] + \Xi_n, \tag{1}$$

b, f_c and θ are parameters representing the amplitude, frequency and phase of the signal. The Ξ_n are independent, identically-distributed complex normal random variables with variance σ^2.

For convenience, we write $A_n = |z_n|$ and $\phi_n = \angle z_n$ to represent the instantaneous amplitude and phase, respectively, of the signal.

3. THE PERIODOGRAM AND PHASE UNWRAPPING

Rife & Boorstyn [1] showed that, when b, f_c and θ are unknown parameters, maximum likelihood estimates can be found by maximising the periodogram of the data, it being defined as

$$\zeta(f) = \left| \sum_{n=1}^{N} z_n \exp(-i2\pi f n) \right|.$$

The maximum likelihood estimate for f is

$$\hat{f} = \arg \max_{-\frac{1}{2} < f \leqslant \frac{1}{2}} \zeta(f). \tag{2}$$

From a computational point of view, it is not immediately clear how we should go about obtaining the maximum likelihood estimate. Rife & Boorstyn suggest a two-step approach. In the first step, we compute the Fast Fourier Transform (FFT) of the data, zero-padded to twice or four times its original length (which is assumed to be a power of two). This makes available to us values of $\zeta(f)$ at $f_n = n/(MN)$, $n = 0, 1, \ldots, MN-1$ where $M = 2$ or $M = 4$ is the factor by which the original record has been zero-padded. In the second step, we choose that index l such that the absolute value of the FFT of the data is maximised at f_l. We then refine the estimate by executing a gradient ascent method, such as the secant method, from f_l. We terminate the method when the difference, ϵ, between successive refinements in the estimate is appreciably less than the Cramér-Rao bound, which is to say that

$$\epsilon \ll \frac{3\sigma^2}{\pi^2 b^2 N(N^2-1)}.$$

Rife & Boorstyn do not prove that this procedure generates an estimate which satisfies (2). Indeed, they recommend $M = 2$ or $M = 4$ since a choice of $M = 1$ leads to poor performance for low SNRs. Thus, the procedure only approximates maximum likelihood. Nevertheless, for proper choice of M, numerical simulations show that the algorithm has very good statistical performance.

Tretter [2] noticed that if the phase of the data could be correctly unwrapped then linear regression could be used to efficiently estimate the frequency. To see this, consider the complex argument (instantaneous phase) of each sample. The instantaneous phases, ϕ_n, are instances of random variables Φ_n where

$$\Phi_n = \angle Z_n = 2\pi f_c n + \theta + \Delta_n \tag{3}$$

and the Δ_n are zero-mean, independent random variables expressing the phase error arising from the additive noise Ξ_n in (1) and bounded between $-\pi$ and π. The Δ_n are approximately normal with variance $2\sigma^2/b^2$ when $\sigma \ll b$. Therefore, given the linear nature of (3), we should be able to perform linear regression on the phases to obtain very accurate estimates of f. However, the observed phases are calculated using an arctangent operation which contains a branch cut, usually at $\pm\pi$. That is, they are calculated modulo 2π. In order to perform linear regression, we must perform *phase unwrapping*. To each $-\pi < \phi_n \leqslant \pi$, we associate an integer s_n to construct the unwrapped phase $\phi_n^* = \phi_n + 2\pi s_n$. Linear regression is then performed on the ϕ_n^*. Although not originally expressed in terms of phase unwrapping or linear regression, Kay [6] proposes a scheme in which we set $\phi_1^* = \phi_1$ and then construct the unwrapped phase by setting $\phi_n^* = \phi_{n-1}^* + \angle z_n z_{n-1}^\dagger$ for each $n = 2, 3, \ldots, N$ in turn, where $\dagger$ denotes the complex conjugate. Other methods, such as numerical integration of the phase derivative [7, 8] have been proposed, but they require more computation.

The weakness of the linear regression method of Tretter lies in the fact that phase unwrapping errors can cause relatively large errors in the frequency estimate, in addition to those caused by the phase noise. For low SNRs, the errors due to phase unwrapping become appreciable and the method becomes statistically inefficient.

4. LATTICE REDUCTION AND THE NEAREST LATTICE POINT PROBLEM

The problem of finding short vectors and nearest points in a lattice is important in algorithmic number theory. A *(point) lattice* in $\mathbb{R}^m$ is an additive subgroup of $\mathbb{R}^m$. A lattice is any set which can be expressed

$$\Omega = \{a_1\mathbf{b}_1 + a_2\mathbf{b}_2 + \cdots + a_d\mathbf{b}_d \mid a_1, a_2, \ldots, a_d \in \mathbb{Z}\}$$

where $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_d\}$ are d linearly independent vectors in $\mathbb{R}^m$, $m \geqslant d$. We say that Ω has *rank* d and that $\mathcal{B}$ is a *basis* of Ω. Any unimodular transformation of the basis is also a basis of the same lattice.

Of all the possible vectors in a lattice Ω, we may want to find the shortest non-zero vector, $\mathbf{u}$, so that $\|\mathbf{v}\| \geqslant \|\mathbf{u}\|$ for all $\mathbf{v} \in \Omega \setminus \{\mathbf{0}\}$ and for some choice of norm $\|\cdot\|$. The problem is interesting from the computational point of view because it has been shown that this problem is $\mathfrak{NP}$-complete when the sup-norm is used. The problem is also closely related to the process of finding a canonical basis from any given basis, a process which is called lattice *reduction*, especially when the canonical or reduced basis consists of short or shortest vectors in some sense.

Although the problem of finding the shortest vector is apparently infeasible, the algorithm of Lenstra, Lenstra & Lovász [5], known as the LLL algorithm, is able to find almost shortest vectors in a reasonable amount of time. The algorithm calculates a *Lovász-reduced* basis from a given basis of a lattice Ω. An important property of the basis is that, for all non-zero $\mathbf{v} \in \Omega$, $\|\mathbf{b}_1\|_2^2 \leqslant 2^{d-1}\|\mathbf{v}\|_2^2$, where $\mathbf{b}_1$ is the shortest vector in the reduced basis and $\|\cdot\|_2$ denotes the Euclidean norm. The LLL algorithm is able to construct a Lovász-reduced basis in $\mathrm{O}\big(md^2(d + \log\rho)\big)$ arithmetic operations, where ρ is the ratio of the longest vector in the input basis to the shortest vector in Ω.

A similar question to the shortest vector question is how to find the closest lattice point to another given point. That is, given a point $\mathbf{x} \in \mathbb{R}^m$, how do we go about finding a lattice point $\mathbf{u} \in \Omega$ such that $\|\mathbf{x} - \mathbf{v}\| \geqslant \|\mathbf{x} - \mathbf{u}\|$ for all $\mathbf{v} \in \Omega$? Again, the problem appears likely to be computationally infeasible. However, building on the LLL algorithm, Babai [4] proposed an algorithm which is able to find an almost closest lattice point in reasonable time, in that it finds a point $\mathbf{b} \in \Omega$ such that, for all non-zero $\mathbf{v} \in \Omega$,

$$\|\mathbf{b} - \mathbf{x}\|_2^2 \leqslant 2^d \|\mathbf{v} - \mathbf{x}\|_2^2. \tag{4}$$

The number of arithmetic operations required for Babai's extension to the LLL algorithm does not increase its order.

5. FREQUENCY ESTIMATION AS A NEAREST LATTICE POINT PROBLEM

We now consider how the problem of maximum likelihood frequency estimation can be interpreted as a nearest lattice point problem. Consider the function

$$\psi(x_1, x_2, \ldots, x_{N-1}) = \left| A_N + \sum_{n=1}^{N-1} A_n e^{i2\pi x_n} \right|^2.$$

Rewriting $\zeta(f)$ so that

$$\zeta(f) = \left| A_N + \sum_{n=1}^{N-1} A_n e^{i[2\pi f(N-n) + (\phi_n - \phi_N)]} \right|$$

and abusing notation slightly by expressing ψ as a function of a column vector, we can quickly confirm that

$$\zeta(f)^2 = \psi(f\mathbf{t} + \boldsymbol{\delta}), \tag{5}$$

where

$$\mathbf{t} = (N-1, N-2, \ldots, 1)^T,$$
$$\boldsymbol{\delta} = \frac{1}{2\pi}(\phi_1 - \phi_N, \phi_2 - \phi_N, \ldots, \phi_{N-1} - \phi_N)^T.$$

Now, ψ is periodic in each of its arguments with period 1. Assume hereafter that $A_n > 0$, $n = 1, 2, \ldots, N$ (this will be true with probability 1 when $b > 0$). We can show that ψ attains its global maxima on each element of $\mathbb{Z}^{N-1}$ and nowhere else, nor does it have any local maxima (see [9] for a proof in the case $N = 3$).

Therefore, (5) implies that the periodogram can be interpreted as the value of ψ along the line $f\mathbf{t}+\boldsymbol{\delta}$. As we shall see, near points in the integer lattice $\mathbb{Z}^{N-1}$, ψ behaves very much like a distance to that lattice point. Thus, in the case of high SNR, the value of ψ along the line is maximised at its closest approach, in a sense we will make clear, to a lattice point.

We can rewrite ψ so that

$$\psi(x_1, x_2, \ldots, x_{N-1}) = A_N^2 + \sum_{j=1}^{N-1} A_j A_N \cos 2\pi x_j + \sum_{j=1}^{N-1}\sum_{k=1}^{N-1} A_j A_k \cos(2\pi(x_j - x_k)).$$

Taking a Taylor series expansion to second order about $x_n = 0$, $n = 1, 2, \ldots, N-1$, we then have

$$\begin{aligned}\psi(x_1, x_2, \ldots, x_{N-1}) &\approx p(x_1, x_2, \ldots, x_{N-1}) \\ &= A_N^2 + \sum_{j=1}^{N-1} A_j A_N \left(1 - 2\pi^2 x_j^2\right) \\ &\quad + \sum_{j=1}^{N-1}\sum_{k=1}^{N-1} A_j A_k \left[1 - 2\pi^2 (x_j - x_k)^2\right].\end{aligned} \tag{6}$$

We can express p in matrix notation as $p(\mathbf{x}) = \mathbf{A}^T\mathbf{A} - 2\pi^2\mathbf{x}^T\mathbf{M}\mathbf{x}$, where we have expressed the A_n and x_n as column vectors, $\mathbf{A}$ and $\mathbf{x}$, in $\mathbb{R}^N$ and $\mathbb{R}^{N-1}$, respectively,

$$\mathbf{M} = \begin{pmatrix} A_1 S - A_1^2 & -A_1 A_2 & \cdots & -A_1 A_{N-1} \\ -A_1 A_2 & A_2 S - A_2^2 & \cdots & -A_2 A_{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ -A_1 A_{N-1} & -A_2 A_{N-1} & \cdots & A_{N-1} S - A_{N-1}^2 \end{pmatrix}$$

and $S = A_1 + A_2 + \cdots + A_N$. Noticing that $\mathbf{M}$ is a positive definite matrix, we have by Cholesky decomposition that $\mathbf{M} = \mathbf{R}^T\mathbf{R}$ where $\mathbf{R}$ consists of the first $N-1$ columns of

$$\mathbf{R}^* = \left(\mathbf{I} - \frac{\mathbf{D}\mathbf{1}\mathbf{1}^T\mathbf{D}^T}{\mathbf{1}\mathbf{D}\mathbf{D}^T\mathbf{1}^T}\right)\mathbf{D}$$

where $\mathbf{D} = \operatorname{diag}(A_1^{\frac{1}{2}}, A_2^{\frac{1}{2}}, \ldots, A_N^{\frac{1}{2}})$, $\mathbf{I}$ is the $N \times N$ identity matrix and $\mathbf{1}$ is the column vector of dimension N consisting of unit elements.

Now, consider maximising $p(f\mathbf{t} - \mathbf{y})$ over f for some $\mathbf{y} \in \mathbb{R}^{N-1}$. The maximum is attained when

$$f = \frac{\mathbf{t}^T\mathbf{M}\mathbf{y}}{\mathbf{t}^T\mathbf{M}\mathbf{t}}. \tag{7}$$

We then find that

$$\max_{f\in\mathbb{R}} p(f\mathbf{t} - \mathbf{y}) = \mathbf{y}^T\mathbf{R}^T\mathbf{T}^T\mathbf{T}\mathbf{R}\mathbf{y} = \|\mathbf{T}\mathbf{R}\mathbf{y}\|_2^2$$

where $\mathbf{T}$ is the projection matrix

$$\mathbf{T} = \mathbf{I} - \frac{\mathbf{R}\mathbf{t}\mathbf{t}^T\mathbf{R}^T}{\mathbf{t}^T\mathbf{R}^T\mathbf{R}\mathbf{t}}.$$

Let us again consider the approximation of ψ by p in (6), especially with regard to approximating the periodogram using the identity (5). When there is no noise, the line $f\mathbf{t} + \boldsymbol{\delta}$ runs through an integer point, attaining its maximum at that point. This is because, in the noiseless case, the vector $\boldsymbol{\delta}$ has the form $\boldsymbol{\delta} = \mathbf{t}f_c + \mathbf{k}$ where the k_n (of $\mathbf{k}$), $n = 1, 2, \ldots, N-1$ are integers induced by the branch cut in the arctangent. The maximum is attained not only at $f = f_c$ but at all points $f = f_c + \mathbb{Z}$, as we expect. When noise is added but the SNR is high, we expect that the line $f\mathbf{t} + \boldsymbol{\delta}$ would not run through an integer point but would pass very close by and would attain its maximum in that vicinity. Therefore, it is reasonable to use the approximation (6) at that point. A good estimate of f_c should be obtained using

$$\hat{f} = \arg \max_{\mathbf{k}\in\mathbb{Z}^{N-1}} \max_{f\in\mathbb{R}} p(f\mathbf{t} + \boldsymbol{\delta} - \mathbf{k}).$$

The innermost maximisation is effected for any given value of $\mathbf{k}$ in the way discussed above. The maximisation problem is so reduced to finding the integer point $\mathbf{k}^*$ such that

$$\mathbf{k}^* = \arg \min_{\mathbf{k}\in\mathbb{Z}^{N-1}} \|\mathbf{T}\mathbf{R}(\mathbf{k} - \boldsymbol{\delta})\|_2^2. \tag{8}$$

Rather than minimising over the integer lattice, we could equivalently consider minimising over the lattice

$$\Lambda = \mathbf{T}\mathbf{R}\mathbb{Z}^{N-1}, \tag{9}$$

the lattice constructed from the columns of $\mathbf{TR}$. We have

$$\min_{\mathbf{k}\in\mathbb{Z}^{N-1}} \|\mathbf{T}\mathbf{R}(\mathbf{k} - \boldsymbol{\delta})\|_2^2 = \min_{\mathbf{u}\in\Lambda} \|\mathbf{u} - \mathbf{v}\|_2^2,$$

where

$$\mathbf{v} = \mathbf{T}\mathbf{R}\boldsymbol{\delta}. \tag{10}$$

In this way, we have approximated the maximisation of the periodogram by converting the problem into a nearest lattice point problem.

We make two further observations. Firstly, the rank of the lattice Λ is $N-2$ since $\mathbf{T}\mathbf{R}\mathbf{t} = \mathbf{0}$. Therefore, a basis of Λ can be constructed from the first $N-2$ columns of $\mathbf{TR}$. Secondly, the process of finding the $\mathbf{k}^*$ corresponding to the nearest integer point can be interpreted as the unwrapping of the phase of the data. The unwrapped phase ϕ_n^* can be obtained by calculating $\phi_n^* = \phi_n + 2\pi k_n^*$, $n = 1, 2, \ldots, N-1$, with $\phi_N^* = \phi_N$. Equivalently, we have approximated the frequency estimation problem as a problem of optimising over all possible phase unwrappings.

6. NUMERICAL RESULTS

In this section, we examine the performance of an algorithm based on the nearest lattice point construction developed in the previous section. The algorithm, which we call the *nearby lattice point frequency estimator* (NLPFE), calculates an frequency estimate from the input data in a way we now describe. Firstly, we calculate the basis of Λ according to (9) and $\mathbf{v}$ from (10). We find a nearby lattice point $\mathbf{u} \in \Lambda$ to $\mathbf{v}$ using Babai's algorithm. Although we cannot be guaranteed that the lattice point $\mathbf{u}$ is the closest, the bound (4) suggests that the closest point will be found when the SNR is high. Using an integer point $\mathbf{k}^*$ relating $\mathbf{u}$ and $\mathbf{TR}$, we calculate the estimate $\hat{f}$ using (7).

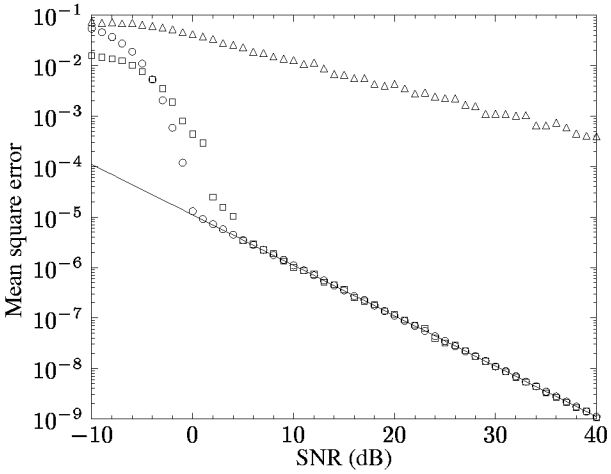

Figure 1: Comparison of experimental mean square error for the AMLFE, NLPFE & LRFE against the Cramér-Rao bound. (Key: — CRLB, ◯ AMFLE, □ NLPFE, △ LRFE.)

Figure 1 shows the experimental mean square error obtained for $N = 24$ from the approximate maximum likelihood frequency estimator (AMLFE) of Rife & Boorstyn which attempts to maximise the periodogram, the nearby lattice point frequency estimator (NLPFE) and the linear regression frequency estimator (LRFE) proposed by Tretter using Kay's phase unwrapping technique. All are plotted against the Cramér-Rao lower bound. The experimental mean square errors were each calculated from at least 1 000 trials for each value of SNR. For each trial, the frequency f_c was chosen in a pseudo-random fashion from a uniform distribution between -0.5 and 0.5. Similarly, the initial phase θ was chosen between $-\pi$ and π. The square error for each trial was calculated modulo 1 as $(\hat{f} - f_c - \lfloor f - f_c \rceil)^2$, where $\lfloor \cdot \rceil$ represents the nearest integer function.

The results of Figure 1 show that the approximate maximum likelihood technique of Rife & Boorstyn performs best, closely approaching the lower bound for SNRs as low as 0 dB. However, we see that the nearby lattice point procedure produces quite good results also, approaching the lower bound at SNRs as low as 5 dB. On the other hand, the linear regression technique using Kay's phase unwrapping fails to approach the lower bound at any SNR. The problem is caused by phase unwrapping errors when the frequency is close to 0.5 or -0.5. In these cases, small errors in the instantaneous phase can then lead to large variations in the value of $\angle z_n z_n^\dagger$. We could alleviate the problem by bounding the frequencies away from these troublesome values.

Of course, the NLPFE is of very little practical use. In the worst case, Babai's algorithm requires $O(N^4)$ arithmetic operations to compute the nearby lattice point which we use to unwrap the phase (the number of operations required on average is unknown to the author).

7. CONCLUDING REMARKS

In this paper, we presented an interpretation of the maximum likelihood estimation of frequency as a nearest lattice point problem. We showed that the nearest lattice point can be interpreted as a best phase unwrapping of the input data, in a certain sense. Furthermore, we presented an algorithm, based on Babai's nearby lattice point algorithm, for estimating frequency. We showed that the algorithm was able to estimate the frequency efficiently for moderately low SNR (> 5 dB) for data sets of 24 points.

However, while the algorithm presented proved to be computationally inefficient for the signal model examined here, the underlying approach may yield algorithms which are better matched to related problems. For example, the nearest lattice point interpretation can be extended to frequency estimation in which the samples are not uniformly spaced, to frequency estimation in multidimensional data and perhaps to the estimation of parameters in polynomial-phase signals.

8. REFERENCES

[1] David C. Rife and Robert R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," *IEEE Trans. Inform. Theory*, vol. IT–20, no. 5, pp. 591–598, Sept. 1974.

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

[3] P. van Emde Boas, "Another NP-complete partition problem and the complexity of computing short vectors in a lattice," Tech. Rep. 81–04, Mathematisch Instituut, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands, Apr. 1981.

[4] L. Babai, "On Lovász' lattice reduction and the nearest lattice point problem," *Combinatorica*, vol. 6, no. 1, pp. 1–13, 1986.

[5] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász, "Factoring polynomials with rational coefficients," *Math. Ann.*, vol. 261, pp. 515–534, 1982.

[6] Steven M. Kay, "A fast and accurate single frequency estimator," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, no. 12, pp. 1987–1990, Dec. 1989.

[7] Alan V. Oppenheim and Ronald W. Schafer, *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, New Jersey, 1975.

[8] José M. Tribolet, "A new phase unwrapping algorithm," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP–25, no. 2, pp. 170–177, Apr. 1977.

[9] I. Vaughan L. Clarkson, *Approximation of Linear Forms by Lattice Points with Applications to Signal Processing*, Ph.D. thesis, The Australian National University, 1997.