MAXIMUM-LIKELIHOOD PERIOD ESTIMATION FROM SPARSE, NOISY TIMING DATA

Robby McKilliam and I. Vaughan L. Clarkson

School of Information Technology and Electrical Engineering
The University of Queensland, Qld., 4072, AUSTRALIA
{robertm,v.clarkson}@itee.uq.edu.au

ABSTRACT

The problem of estimating the period of a periodic point process is considered when the observations are sparse and noisy. There is a class of estimators that operate by maximizing an objective function over an interval of possible periods, notably the periodogram estimator of Fogel & Gavish [1] and the line-search algorithms of Sidiropoulos *et al.* [2] and Clarkson [3]. For numerical calculation, the interval is sampled. However, it is not known how fine the sampling must be in order to ensure statistically accurate results. In this paper, a new estimator is proposed which eliminates the need for sampling. For the proposed statistical model, it calculates a maximum-likelihood estimate. It is shown that the expected arithmetic complexity of the algorithm is $O(n^3 \log n)$ where n is the number of observations. Numerical simulations demonstrate the superior statistical performance of the new estimator.

***Index Terms*—** Maximum likelihood estimation, Synchronization, Frequency hop communication

1. INTRODUCTION

The problem of estimating the period of a signal from measurements that are both sparse and noisy is considered. Solutions to the problem have the following applications: bit-synchronization from zero-crossings for telecommunications devices [1]; estimation of the pulse repetition interval (PRI) in Electronic Support [4]; and the estimation of the hop rate of a frequency-hopped-spread-spectrum signal from observations of occupancy times in only some of the frequency bands [2].

A number of estimators have been proposed which involve maximization of an objective function over an interval of possible periods. An intuitively appealing and statistically accurate estimator is the *periodogram estimator* proposed by Fogel and Gavish [1]. The periodogram estimator considers the timing data as a train of Dirac delta functions at the observation times. The estimate is the period which maximizes the periodogram.

Sidiropoulos *et al.* [2] proposed a family of estimators called *separable least squares line search* (SLS2) estimators. The objective function in SLS2 has an advantage over the periodogram in that there is no need to evaluate trigonometric functions.

A recent estimator proposed by Clarkson is the *lattice line search* (LLS) estimator [3]. The estimator developed in this paper is derived from it. Clarkson showed how the *maximum likelihood (ML)* estimator for the sparse, noisy period estimation problem can be understood as a lattice problem [3, 5]. In particular, the problem is reduced to finding the nearest lattice point in the lattice A_{n-1}^* [5] along a line segment representing the interval of possible periods.

When implemented numerically, each of these estimators operates by sampling the objective function over the interval. If the chosen sampling interval is too large then the best estimate may be missed. Currently, no theory exists to determine the required sampling interval.

In this paper, we propose a statistical model and an estimator that is guaranteed to find the ML estimate. The estimator avoids the need for sampling. We call the new estimator the *integer lattice line search (ZnLLS)*. We will show that, under mild conditions, ZnLLS has an expected arithmetic complexity of $O(n^3 \log n)$. This represents a major improvement on the only other known ML estimation algorithm for this model, which requires enumeration of lattice points inside a sphere and has exponential arithmetic complexity [3]. Further, we will demonstrate through simulations that the ZnLLS algorithm has superior statistical accuracy.

This paper is organized as follows. In Section 2, we outline the required lattice theory. In Section 3, we introduce the statistical model used to describe the period estimation problem. In Section 4, we review the LLS estimator and, in Section 5, we propose the new ZnLLS algorithm, showing that it is an ML estimator for period. In Section 6, we derive the computational complexity of the algorithm. Simulated results concerning the statistical performance of the ZnLLS estimator and comparable estimators are presented in Section 7.

2. THEORETICAL BACKGROUND

We begin by recalling some elements of lattice theory that will be necessary in the sequel.

A *lattice*, L, is a set of points in $\mathbb{R}^n$ such that

$$L = \{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{x} = \boldsymbol{B}\boldsymbol{w}, \boldsymbol{w} \in \mathbb{Z}^n\}$$

where $\boldsymbol{B}$ is termed the *generator matrix*.

The *Voronoi region* or *nearest-neighbor region* $V(\boldsymbol{x})$ of a lattice point $\boldsymbol{x}$ is the subset of $\mathbb{R}^n$ such that, with respect to a given norm, all points in $V(\boldsymbol{x})$ are nearer to $\boldsymbol{x}$ than to any other point in the lattice. The Voronoi regions are n dimensional polytopes [5].

We say that the Voronoi regions $V(\boldsymbol{x})$ and $V(\boldsymbol{x'})$ of lattice points $\boldsymbol{x}$ and $\boldsymbol{x}'$, and by extension the lattice points themselves, are *adjacent* if they share a face.

The cubic lattice $\mathbb{Z}^n$ is the set of n dimensional vectors with integer elements. The Voronoi regions of $\mathbb{Z}^n$ are hypercubes of side length 1.

The *zero-mean plane* or *zero-sum plane* is the subspace of dimension $n-1$ in $\mathbb{R}^n$ that is orthogonal to $\mathbf{1}$, where $\mathbf{1}$ denotes the vector $(1\dots1)^T$. A vector $\boldsymbol{b} \in \mathbb{R}^n$ is orthogonally projected into the zero-mean plane by multiplication by the matrix

$$\boldsymbol{Q} = \left(\boldsymbol{I} - \frac{\mathbf{1}\mathbf{1}^T}{n}\right) \tag{1}$$

where $\boldsymbol{I}$ is the identity matrix.

The lattice A_{n-1} can be defined as the intersection of the cubic lattice $\mathbb{Z}^n$ with the zero-mean plane, *i.e.*,

$$A_{n-1} = \left\{\boldsymbol{x} \in \mathbb{Z}^n \middle| \sum_{i=1}^{n} x_i = 0\right\}. \tag{2}$$

One way to define A_{n-1}^* is as the projection of the cubic lattice $\mathbb{Z}^n$ onto the zero-mean plane. That is,

$$A_{n-1}^* = \{\boldsymbol{Q}\boldsymbol{x} \mid \boldsymbol{x} \in \mathbb{Z}^n\} \tag{3}$$

where $\boldsymbol{Q}$ is the generator matrix for A_{n-1}^* defined in (1). A_{n-1}^* can also be generated by the union of n translates of the lattice A_{n-1} [5]. That is,

$$A_{n-1}^* = \bigcup_{i=0}^{n-1} ([i] + A_{n-1}) \tag{4}$$

where the $[i]$ are known as *glue vectors* and are defined in this case [5] as

$$[i] = \frac{1}{n}\big(\underbrace{i,\dots,i}_{i \text{ times}}, \underbrace{-j,\dots,-j}_{j \text{ times}}\big) \tag{5}$$

for $i = 1,\dots,n-1$ with $i + j = n$.

In relation to the line $\boldsymbol{y}\mathbb{R}$, the lattice point $\boldsymbol{x}$ is a *Bresenham point* if, at some position on $\boldsymbol{y}\mathbb{R}$, $\boldsymbol{x}$ is the nearest lattice point. The term 'Bresenham point' is taken from the well-known algorithm in computer graphics for approximating a line by points on a square grid [6]. The set of all Bresenham points for the line $\boldsymbol{y}\mathbb{R}$ in the lattice L is denoted $\vartheta(\boldsymbol{y}\mathbb{R}, L)$. $\vartheta(\boldsymbol{y}\mathbb{R}, L)$ is called the *Bresenham set*. A consequence of $\boldsymbol{x} \in \vartheta(\boldsymbol{y}\mathbb{R}, L)$ is that the line $\boldsymbol{y}\mathbb{R}$ must pass through the Voronoi region $V(\boldsymbol{x})$.

The Bresenham set is an ordered set. The ordering is defined as follows.

Definition 1. *Let $\boldsymbol{x}, \boldsymbol{x}' \in \vartheta(\boldsymbol{y}\mathbb{R}, L)$. Then*

$$\boldsymbol{x} < \boldsymbol{x}' \Leftrightarrow f < f'$$

where $f, f' \in \mathbb{R}$, $f\boldsymbol{y} \in V(\boldsymbol{x})$ and $f'\boldsymbol{y} \in V(\boldsymbol{x}')$.

This definition is consistent by virtue of the convexity of Voronoi regions. Adjacency in this ordering is equivalent to adjacency of the Voronoi regions.

It will be seen in Section 5 that our estimation problem can be framed in terms of finding all Bresenham points in A_{n-1}^* for a particular line segment.

3. STATISTICAL MODEL

In our statistical model for period estimation, we make sparse, noisy observations y_i of a periodic point process. Collecting the observations as a vector $\boldsymbol{y}$ of length n, we assume that [1–3]

$$\boldsymbol{y} = T\boldsymbol{s} + \mathbf{1}\theta + \boldsymbol{w} \tag{6}$$

where $\boldsymbol{s}$ is a vector of integers or *indices* indicating which of the events were observed, T is the period of the process, θ is its phase and $\boldsymbol{w}$ is a vector of Gaussian white noise, each element having variance σ^2.

The aim is to estimate the period, T. Clarkson [3] derives the ML estimator as

$$\hat{f} = \arg\min_{f \in [f_{\min}, f_{\max}]} \min_{\boldsymbol{s} \in \mathbb{Z}^n} \frac{1}{f^2} \|f\boldsymbol{Q}\boldsymbol{y} - \boldsymbol{Q}\boldsymbol{s}\|_2^2 \tag{7}$$

where $\hat{f} = 1/\hat{T}$, $\boldsymbol{Q}$ is the projection matrix defined in (1), $\|\cdot\|_p$ is the vector p-norm and $[f_{\min}, f_{\max}]$ is the known and finite interval containing f.

Following [3], we will see how (7) can be described in terms of lattice theory. The matrix $\boldsymbol{Q}$ from (1) is the generator matrix for the lattice A_{n-1}^* [3,5,7]. Let $\boldsymbol{z} = \boldsymbol{Q}\boldsymbol{y}$ and $\boldsymbol{v} = \boldsymbol{Q}\boldsymbol{s}$. We rewrite (7) as $\hat{f} = M_\angle(A_{n-1}^*)$ where, for set S,

$$M_\angle(S) = \arg\min_{f \in [f_{\min}, f_{\max}]} \min_{\boldsymbol{x} \in S} \frac{1}{f^2} \|f\boldsymbol{z} - \boldsymbol{x}\|_2^2.$$

Given $\boldsymbol{s}$, the ML estimate of f is

$$\hat{f} = \frac{\boldsymbol{z}^T\boldsymbol{v}}{\boldsymbol{z}^T\boldsymbol{z}}. \tag{8}$$

Given f, the ML estimate $\hat{\boldsymbol{s}}$ is that for which the corresponding $\hat{\boldsymbol{v}}$ is a closest lattice point in A_{n-1}^* to $f\boldsymbol{z}$.

4. THE LLS ALGORITHM

We see then that the Bresenham set in A_{n-1}^* of the line segment $G = \{f\boldsymbol{z} \mid f \in [f_{\min}, f_{\max}]\}$, *i.e.*, $\vartheta(G, A_{n-1}^*)$, contains elements $\boldsymbol{v}$ that correspond to all candidate ML values of $\boldsymbol{s}$. By enumerating these elements, calculating the corresponding values of $\hat{f}$ as in (8) and choosing the best of these according to (7), we obtain the ML period estimate.

The LLS algorithm samples the line segment G and calculates the closest lattice point at each sample point [3, 7]. This yields a subset of the Bresenham set. The subset becomes more complete as we reduce the sampling interval but it is not known how small the interval must be to ensure set equality with the Bresenham set. Without a guarantee of set equality, the LLS algorithm cannot therefore be guaranteed to produce ML period estimates.

5. THE ZNLLS ALGORITHM

For the ZnLLS estimator, we will generate the complete Bresenham set rather than a subset. For this reason, the method is guaranteed to find $\hat{\boldsymbol{v}}$ and thus the ML estimate of period. The proof of this result is sketched below for Theorem 1.

To evaluate $\vartheta(G, A_{n-1}^*)$, we will make use of the algorithm of Ryan *et al.* [8] to efficiently evaluate $\vartheta(\boldsymbol{b}\mathbb{R}, \mathbb{Z}^n)$ for any line $\boldsymbol{b}\mathbb{R}$. Given the relationships between the lattices A_{n-1}^*, A_{n-1} and $\mathbb{Z}^n$ described in (3) and (4), it is possible to enumerate $\vartheta(G, A_{n-1}^*)$ by enumerating $\vartheta(G, \Lambda)$ where Λ is the lattice

$$\Lambda = \bigcup_{i=0}^{n-1} ([i] + \mathbb{Z}^n) \tag{9}$$

and $[i]$ are the glue vectors described in (5). $\vartheta(G, \Lambda)$ can be obtained by evaluating $\vartheta(G, [i] + \mathbb{Z}^n)$ for each $[i]$ individually using the algorithm described in [8]. For brevity, we only sketch the proof. The following theorem encapsulates the important result.

Theorem 1. *The ML estimate of frequency can be expressed as*

$$\hat{f} = M_\angle(\vartheta(G, \Lambda)).$$

Proof. The proof follows by showing that

$$\hat{f} = M_\angle(\vartheta(G, A_{n-1}^*)),$$
$$\boldsymbol{x} \in \vartheta(G, A_{n-1}^*) \Rightarrow \boldsymbol{x} \in \vartheta(G, \Lambda),$$
$$\boldsymbol{x} \in \Lambda \Rightarrow \boldsymbol{Q}\boldsymbol{x} \in A_{n-1}^* \Rightarrow \boldsymbol{Q}\boldsymbol{x} \in \Lambda.$$

To complete the proof, we show that, for any $\boldsymbol{x} \in \Lambda$, $\|f\boldsymbol{z} - \boldsymbol{x}\|_2 \geq \|f\boldsymbol{z} - \boldsymbol{Q}\boldsymbol{x}\|_2$. □

The ZnLLS algorithm is set out as Algorithm 1. The function BBP$(\cdot)$, *best Bresenham point*, is the algorithm proposed in [8], modified to take account of the translation by $[i]$ of $\mathbf{Z}^n$. BBP$(\boldsymbol{z}, f_{\min}, f_{\max}, [i])$ returns $M_\angle(\vartheta(G, [i] + \mathbb{Z}^n))$. Also returned is L, the log-likelihood of $M_\angle(\vartheta(G, [i] + \mathbb{Z}^n))$.

The for loop at Line 2 runs BBP$(\cdot)$ for the lattice $\mathbb{Z}^n + [i]$ for each glue vector, $[i]$. The ML estimate of frequency is returned at Line 7.

The running time for the BBP$(\cdot)$ function is dependent on the number of Bresenham points it must test. Each Bresenham point requires $O(\log n)$ calculations to test.

Algorithm 1: Finds a nearest point to G in A_{n-1}^*

```
1 z = y − ȳ
2 for i = 0 to n − 1 do
3     [f, L] = BBP(z, f_min, f_max, [i])
4     if L > L_best then
5         L_best = L
6         f_best = f
7 return f_best
```

6. COMPUTATIONAL COMPLEXITY

We will now show that, with an additional assumption on the statistical model, the *expected* complexity of the algorithm is $O(n^3 \log n)$. The additional assumption is that the *index differences* $s_{i+1} - s_i$ are random variables from a geometric distribution with parameter p, they are mutually independent and independent of the noise and $s_1 = 0$.

First, consider the number of lattice points in $\vartheta(G, \Lambda)$. This is the number of points the ZnLLS algorithm must examine. For any translation vector, $\boldsymbol{t}$,

$$|\vartheta(G, \mathbb{Z}^n + \boldsymbol{t})| \leq |\vartheta(G, \mathbb{Z}^n)| + n$$

where $|\cdot|$ denotes cardinality. Recalling (9), this leads to

$$|\vartheta(G, \Lambda)| \leq n(|\vartheta(G, \mathbb{Z}^n)| + n) \tag{10}$$

Recall that the Bresenham set is an ordered set (Definition 1) and that every pair of adjacent Bresenham points in $\vartheta(G, \mathbb{Z}^n)$ is also adjacent in $\mathbb{Z}^n$. Hence, the $(j+1)^{\text{th}}$ Bresenham point can be expressed as

$$\boldsymbol{v}_{j+1} = \boldsymbol{v}_j + \operatorname{sgn}(z_i)\boldsymbol{e}_i \tag{11}$$

for some $i \in \{1, \ldots, n\}$ and where $\operatorname{sgn}(\cdot)$ is the signum function. From (11), it follows that the number of Bresenham points is

$$\begin{aligned} |\vartheta(G, \mathbb{Z}^n)| &= \sum_{i=1}^{n} |\lceil f_{\max}\boldsymbol{z} \rfloor_i - \lceil f_{\min}\boldsymbol{z} \rfloor_i| \\ &\leq (f_{\max} - f_{\min})(||\boldsymbol{z}||_1 + n) \end{aligned} \tag{12}$$

where $\lceil \cdot \rfloor$ denotes the nearest integer to its argument. Using (10), an upper bound for $|\vartheta(G, \Lambda)|$ is then

$$|\vartheta(G, \Lambda)| \leq n(f_{\max} - f_{\min})(||\boldsymbol{z}||_1 + n) + n^2. \tag{13}$$

We now derive an expression for the expected 1-norm of $\boldsymbol{z}$ in terms of n. A property of vector norms is that $\|\boldsymbol{z}\|_1 \leq \sqrt{n}\|\boldsymbol{z}\|_2$ and, since $\boldsymbol{z} = \boldsymbol{Q}\boldsymbol{y}$ and $\boldsymbol{Q}$ is an orthogonal projection matrix, we also have that $\|\boldsymbol{z}\|_2 \leq \|\boldsymbol{y}\|_2$. From Jensen's inequality, we have that $E[\|\boldsymbol{y}\|_2] \leq E[\|\boldsymbol{y}\|_2^2]^{1/2}$. Further,

$$\begin{aligned} E[\|\boldsymbol{y}\|_2^2] &= \sum_{i=1}^{n} \left\{ E[s_i^2 T^2] + 2E[s_i T \theta] + E[\theta^2] + E[w_i^2] \right\} \\ &= \frac{T}{6p^2} n(n-1)(2Tn - 3Tp + 6\theta p - 2T) + n(\theta^2 + \sigma^2) \end{aligned}$$

Thus, $E[\|\boldsymbol{z}\|_1] = O(n^2)$.

In the ZnLLS algorithm, $O(\log n)$ calculations are required to identify and test each lattice point in $\vartheta(G, \Lambda)$, as explained in Section 5. It follows that the expected computational complexity of the algorithm is $O(n^3 \log n)$.

7. SIMULATION RESULTS

The ZnLLS estimator is compared with the periodogram estimator, the LLS estimator and the SLS2-ALL estimator. We refer to these other estimators as *sampling estimators* because of the need to choose a sampling interval.

A number of Monte Carlo simulations were conducted to assess the statistical performance of the estimators. In the simulations, the parameters $f_{\min} = 0.7$ and $f_{\max} = 1.3$ were used. The value of f was selected pseudo-randomly and uniformly on this interval, excluding the top and bottom 10% of the interval. The number of samples was set at 40 for each of the sampling estimators. Two thousand independent trials were run for each plotted data point.

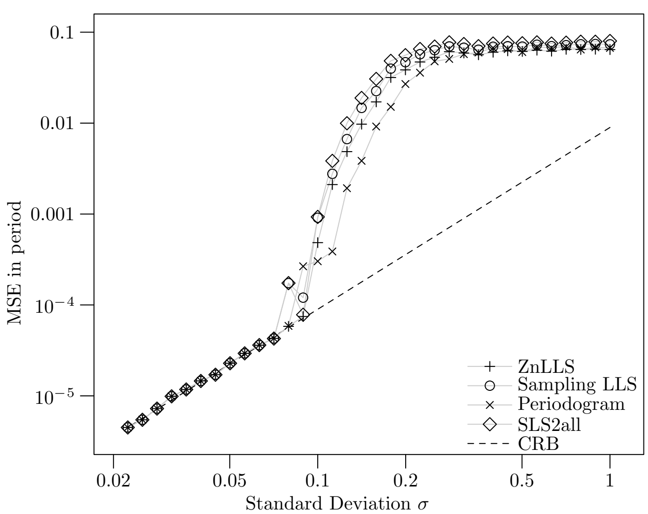

Fig. 1. MSE in period when $n = 10$ as σ varies.

Figure 1 shows the performance of the estimators for $n = 10$ as σ varies in the range $[10^{-1.7}, 1]$. We observe typical behavior for period estimators. There is a threshold value of σ, in this case $\sigma \approx 0.1$, below which the variances of the estimates are in close agreement with the so-called *clairvoyant* Cramer-Rao lower bound (CRLB) [2]. Above the threshold, performance deteriorates rapidly. The MSEs of all estimators are similar and all conform with the CRLB below the threshold.

However, in Figure 2, we see the effect of increasing n. To generate Figure 2, the simulation was run with $\sigma = 0.05$ and n was varied between 10 and 70 in steps of 5.

Figure 2 shows that the ZnLLS estimator performs very well for all values of n. The periodogram and SLS2-ALL estimators perform well for $n < 35$, the LLS estimator performs well for $n < 45$. However, as n increases, all but ZnLLS eventually diverge from the CRLB. This can be ascribed to the constant sampling interval used by the sampling estimators. As n increases, the sampling interval becomes too large and the sampling estimators tend to 'step over' the ML estimate. However, no theory currently exists for calculation of a suitable sampling interval. It is therefore a key advantage of the ZnLLS algorithm that it avoids sampling.

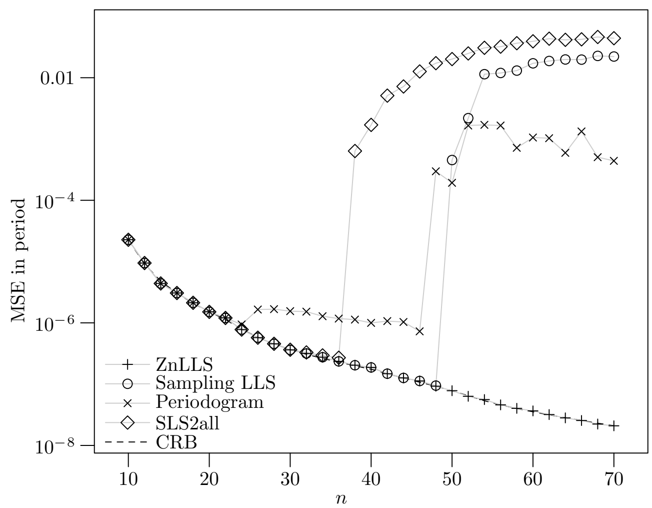

Fig. 2. Mean square error in period as n varies.

8. REFERENCES

[1] E. Fogel and M. Gavish, "Parameter estimation of quasi-periodic sequences," *Proc. Internat. Conf. Acoust. Speech Signal Proces.*, vol. 4, pp. 2348–2351, 1988.

[2] N. D. Sidiropoulos, A. Swami, and B. M. Sadler, "Quasi-ML period estimation from incomplete timing data," *IEEE Trans. Signal Process*, vol. 53, pp. 733–739, 2005.

[3] I. V. L. Clarkson, "Approximate maximum-likelihood period estimation from sparse, noisy timing data," *Accepted to appear at IEEE Trans. Signal Proces.*, July 2007.

[4] R. G. Wiley, *Electronic Intelligence: The Analysis of Radar Signals*, Artech House, Norwood, Massachusetts, 1982.

[5] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*, Springer, 3rd edition, 1998.

[6] J. E. Bresenham, "Algorithm for computer control of a digital plotter," *IBM Systems Journal*, vol. 4, no. 1, pp. 25–30, 1965.

[7] I. V. L. Clarkson, "An algorithm to compute a nearest point in the lattice A_n^*," in *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, volume 1719 of Lecture Notes in Computer Science*. 1999, pp. 104–120, Springer.

[8] D. J. Ryan, I. B. Collings, and I. V. L. Clarkson, "GLRT-optimal noncoherent lattice decoding," *IEEE Trans. Sig. Proces.*, vol. 55, pp. 3773–3786, 2007.