M. Stanciu, S. Azou, A. Serbanescu, “On the Blind Estimation of Chip Time of Time-Hopping Signals Through Minimization of a Multimodal Cost Function”, IEEE Trans. on Signal Processing, Vol. 59, No. 2, pp. 842-847, Feb 2011.

On the Blind Estimation of Chip Time of Time-Hopping Signals through Minimization of a Multimodal Cost Function

Mihai I. Stanciu[1,2,3], Stéphane Azou[1,2,*], and Alexandru Şerbănescu[3], *Senior Member, IEEE*

[1]Université Européenne de Bretagne, France.
[2]Université de Brest ; CNRS, UMR 3192 LabSTICC ; Brest, France
[3]Military Technical Academy, Bucharest, Romania

[*]Corresponding author (E-mail : azou@univ-brest.fr)

***Abstract*—In this paper, we focus on blind estimation of the chip duration of time-hopping signals by introducing a cost function based on Time Of Arrival (ToA) folding over multiple observation sets. An optimization algorithm that takes advantage of the highly oscillatory behavior in the nearby of the global minimum is proposed and a performance bound for the chip time estimate is derived. The pertinence of our approach is shown through numerical results, considering alternative methods like periodogram and separable least squares line search. The proposed technique enables a good tradeoff between statistical accuracy and computational complexity.**

***Index Terms*—Pulse Train Analysis, Time-Hopping, Impulse Radio, Blind estimation, Time-of-arrival, Period estimation, Chip duration, Multimodal optimization.**

I. Introduction

The issue of period estimation from a series of incomplete and noisy discrete events, arising from a periodical process, finds many applications in various domains including radar [1], communications [2], astrophysics [3] and neurology [4]. In the field of digital communications, this problem can arise when performing bitstream synchronization through zero-crossing analysis from the received Pulse Amplitude Modulated (PAM) signal [2]. Another example concerns blind estimation of the hop rate for Frequency-Hopping Spread Spectrum (FH-SS) signals [5], where the frequency band used by the transmitter is unknown.

Blind estimation is of central importance for passive listening, which is a main stage in Electronic Warfare and Signals Intelligence (SIGINT) [6], [7]. In this context, some technical characteristics should be estimated so as to be able to recognize the type of transmitter and locate it or extract the underlying message, with few or no prior knowledge. In this paper, we focus on the particular case where the transmitted signal is a non-periodic pulse train in which the time lapse between the beginning instants of two consecutive pulses is controlled by a pseudonoise code generator (time-hopping). In particular, the problem of blind estimation of the chip interval will be considered, which appears to be a key parameter of the system as it directly influences multiple-access performances, spectral properties or probability of intercept.

Over the past two decades, many results have been reported about the general problem of period extraction from sparse, timing measurements. Various contributions yield the period estimate from the arrival times by pursuing one of the following approaches: histogramming [11], Kalman filtering [12], [13], Euclidean algorithm [14], periodogram [15], [16] or function optimization [3], [5], [17], [18]. The present work is mainly inspired by a recent paper of Sidiropoulos *et al.* [5] in which the period estimation is achieved owing to an objective function based on the *round* operator. An interesting result pointed out by the authors is the performance increase of the line search procedure that can result from a pertinent choice of time differences. The basic approach (SLS2-ADJ) uses adjacent differences whereas the more sophisticated one (SLS2-ALL) operates over all pairwise differences. The question of which time differences to exploit with the SLS2 approach has been further investigated in [16].

The chip time estimator that will be developed in this paper shares two similarities with the SLS2 approach. First, our objective function also rely on the *round* operator, which has proved to be very effective in previous studies, and second, various arrival time differences are associated for function evaluation. Our work clearly distinguishes from the contribution of Sidiropoulos though, as the cost function takes a very different analytical form by using a “puncturing” principle to compute the total sum of observed data. As a result, we get a multimodal characteristic that brings useful information about the underlying received pulse train and enables an original optimization algorithm: the global minimizer is found through alternating a diversification strategy based on deterministic hops and Golden Section Search (GSS) for local exploration. Also, we propose a combining process of the local minima discovered for different record lengths to enhance the performance. Despite the increased computational cost, the complexity is still attractive for practical implementations as no trigonometric function is required. A few numerical results will be proposed to show the good statistical performances, considering SLS2-ADJ, SLS2-ALL and periodogram as alternate methods.

The paper is organized as follows. In section II, the statistical signal model is defined and a novel *round* operator-based objective function is introduced. A minimization process is then developed in section III and an approximate performance bound is stated. Finally, before some concluding remarks, the statistical performances of the proposed approach will be examined through Monte-Carlo simulations.

II. Problem Formulation

A. Signal Model

As pointed out by Clarkson in a recent paper [16], the model which is mostly taken into consideration for a series of sparse and noisy discrete events, arising from a periodical process, relies on a set of random variables $\{y_i\}_{i \in \{1,2,\ldots,n'\}}$, being expressed as $y_i = k_i T + \varphi + \eta_i$, where $T > 0$ is the unknown period, φ stands for the initial phase reflecting the transmitter-receiver time offset, indices $k_i \in \mathbb{N}^*$ specify the events that have been observed, and the elements η_i characterize the measurement noise, here being considered as identically distributed, zero-mean Gaussian random variables with standard deviation σ_η. The present paper focuses on the particular case where the received signal at the input of the detection stage has a time-hopping spread spectrum format [8]. Period T then corresponds to the so-called chip time T_c and the coefficients k_i are described through the generic model $k_i = iN_c + c_i$, where c_i are the pseudo-random code elements taking integer values in $[0, N_c - 1]$, N_c being the number of time delay bins per frame. A long pseudonoise code is considered so that no code repetition occurs over the recorded signal. Our aim is to recover T_c once pulses ToA have been collected. In a blind approach, the arrival times can be estimated by energy measurements and thresholding [19], which allows sub-Nyquist sampling. As a result, a set of ToAs $\mathcal{Y} = \{y_i\}_{i=1,\ldots,n'}$ which conform to the above model can be considered to develop our chip duration estimator. We must note that this model does not completely reflect the uncertainties resulting from ToA detection process : some false alarms can occur due to unknown transmitter parameters and propagation conditions (background noise and multipath). So, outliers have later to be incorporated in the observed data set $\mathcal{Y}$ to investigate the robustness of the proposed algorithms to such perturbations. By denoting $\{y_i^o\}_{i=1,\ldots,n^o}$ the set of outlying events, following a uniform distribution over the interval $[y_1, y_n]$, the data set from which the chip time must be extracted will then take the form $\mathcal{Y} = \{y_i\}_{i=1,\ldots,n'-n^o} \cup \{y_i^o\}_{i=1,\ldots,n^o}$. Missing observations can be similarly considered.

B. A novel cost function

As shown in a few earlier studies [3], [5], [17], the problem of period estimation can be formulated as the minimization of a cost function relying on ToA folding through quantization operations. In this way, knowledge of coefficients k_i is not necessary and, if ToA differences are considered, the effect of time offset is eliminated.

A novel cost function is proposed here, whose highly oscillatory behavior results from a mixing of partial functions operating on punctured observation sets. In order to eliminate the phase term ϕ, an adjacent pair differencing as proposed by Sidiropolous *et al.* [5], is processed. Hence, a new set of observations $\boldsymbol{t}$, with $t_i = y_{i+1} - y_i = (N_c + \Delta c_i) T_c + \delta_i,\ i \in \{1,\ 2,\ ...,\ n'-1\}$, is obtained. Here $\Delta c_i = c_{i+1} - c_i$ and $\delta_i = \eta_{i+1} - \eta_i$. The latter term represents a correlated random variable (r.v.) with distribution $\mathcal{N}\left(0,\ 2\sigma_\eta^2\right)$. The proposed cost function is defined from the set of time data $\boldsymbol{t} = \{t_i\}_{i\in\{1,\ 2,\ ...,\ n\}}$ as[1]

$$f\left(\boldsymbol{t}, \widetilde{T}\right) = \left\| \frac{\mathbf{1}_n \otimes S - \boldsymbol{t}}{\widetilde{T}} - \text{round}\left(\frac{\mathbf{1}_n \otimes S - \boldsymbol{t}}{\widetilde{T}}\right)\right\|_1 . \quad (1)$$

where $S = \sum_{i=1}^n t_i$, $\mathbf{1}_n$ is a nth dimensional vector of ones, $\otimes$ denotes the Kronecker product, $n = n' - 1$ and $\|.\|_1$ stands for the ℓ_1 matrix norm. Due to the operations involved, the variation of the cost function becomes very fast in the nearness of T_c. This characteristic translates into pseudo-periodical oscillations. In the jitter free case, (1) is a sum of periodical functions $h_m(x) = |(S - \Delta c_m)x - \text{round}((S - \Delta c_m)x)|$ of respective frequencies $(S - \Delta c_m)$, where $x = T_c/\widetilde{T}$ and $m \in \{1,\ 2,\ ...,\ n\}$. Cost function's oscillating property will be exploited in the next sections for a restricted nearby of $x = 1$, where $(x + 1/x) \cong 1$. Hence $h_m(1/x)$ has also an approximate oscillating behavior with the same frequency as $h_m(x)$. As Δc_m is a symetrical r.v., with 0 mean, it can be shown that the local minima of the cost function $f(1/x)$ will be located with a pseudo-frequency $S \cong nN_cT_c$. Consequently, the approximate pseudo-frequency of oscillation for $f\left(\boldsymbol{t}, \widetilde{T}\right)$ is $f_o \simeq nN_c/T_c$.

III. Proposed Method for Chip Time Estimation

A. Multiple Data Sets - Adjacent algorithm

The procedure proposed here combines multiple hops (MH) controlled by the pseudo-frequency estimate and a Golden Section Search (GSS). Owing to this limited complexity processing, the cost function envelope is precisely evaluated over the search space while avoiding the problem of staying in a particular local minimum. A multidimensional approach will be developed to deal with many data sets of various dimensions.

Now, let us describe our search scheme, considering a n-dimensional vector of adjacent observations $\boldsymbol{t}$. To launch the procedure, it should be assumed first that a confidence interval I_{T_c} is known for an initial chip time estimate $\hat{T}_c^{'}$, as a result of a preprocessing of the time data with a coarse estimation method. A rough estimate of the cost function pseudo-frequency is also required, which can be computed utilizing relation $\hat{f}_o^{(0)} = n\hat{T}_f / \left(\hat{T}_c^{'}\right)^2$, where the approximate value of the frame time is $\hat{T}_f = (1/n)\sum_{i=1}^n t_i$. For example, if the initial estimate $\hat{T}_c^{'}$ results from the SLS2-ADJ algorithm proposed by Sidiropoulos *et al.*, $I_{T_c} = [0.98 \times T_c,\ 1.02 \times T_c]$ can be obtained for our refining method to apply, with $n = 40$ and a jitter smaller than 30%. We define the percent jitter as $(3\sigma_\eta/T_c) \times 100$ to reflect the measurement noise corresponding to the ToA model defined in section II.A. The oscillation pseudo-frequency depends upon the number of observations taken into account. Hence, processing many data sets of different lengths n_k will lead to various groups of local minima "matching only" in the nearby of T_c. The proposed algorithm, denoted as Multiple Data Sets - Adjacent (MDS-ADJ), is now proposed by mixing various minima positions resulting from the processing of multiple data sets.

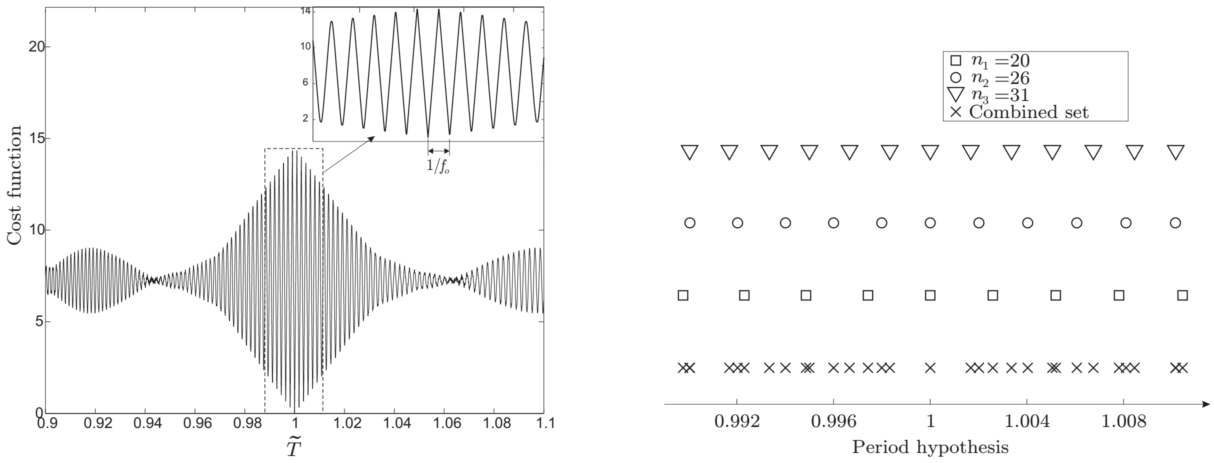
Let $\boldsymbol{t}^{(k)},\ k = 1, 2, ..., K$ be a series of observation sets with different increasing lengths defined by $\boldsymbol{n} = [n_1,\ n_2,\ ...,\ n_K]$. Each of these data sets leads to a group $\boldsymbol{v}^{(k)}$ of local minima owing to an iterative process (Multi-Hop/local GSS), that will be described later in this section. A vector $\boldsymbol{u}$, defined as the increasingly rearranged version of the combined vector $\boldsymbol{v} = \bigcup_{k=1}^K \left(\boldsymbol{v}^{(k)}\right)$, is then created. Roughly speaking, the estimation will be based on evaluating the dispersion of each K consecutive terms from $\boldsymbol{u}$, considering that the minimum spread corresponds to the maximal matching in terms of local minima placement. An example is given in figure 1 (b).

An essential step of the algorithm is the choice of $\boldsymbol{n}$. The simplest case $K = 2$ is considered now for the purpose of exemplification. In this case, two pseudo-periods of oscillation θ_1 and θ_2 are obtained, depending on n_1 and n_2, respectively. It is considered $n_1 < n_2$ so naturally $\theta_1 > \theta_2$, and, in the noise-free case, the elements of $\boldsymbol{v}^{(1)}$ and $\boldsymbol{v}^{(2)}$ can be expressed as $v_1^{(p)} \simeq T_c + p\theta_1$ and $v_2^{(r)} \simeq T_c + r\theta_2$, respectively, where $(p, r) \in \mathbb{Z}^2$. In order to evaluate the matching between the two sets of minima, the following distance is computed for each entry $v_1^{(p)}$, $a_p = \min\{|p\theta_1 - r\theta_2|,\ r \in \{l_1,\ l_1+1,\ ...,\ l_2\}\}$, where $l_1 = \left\lceil \frac{(T_c - \min(\boldsymbol{v}^{(2)}))}{\theta_2} \right\rceil$ and $l_2 = \left\lfloor \frac{(\max(\boldsymbol{v}^{(2)}) - T_c)}{\theta_2} \right\rfloor$. For close values n_1 and n_2, the matching between $\boldsymbol{v}^{(1)}$ and $\boldsymbol{v}^{(2)}$ is perfect for $(p, r) = (0, 0)$ ($\widetilde{T} = T_c$), with a slow but constant variation, whereas if the difference between n_1 and n_2 is increased, the distance a_p exhibits some oscillations with possible small magnitudes for time values far from T_c. This introductory example reveals that a large difference between consecutive values n_k is likely to lead to a performance degradation if the chip time estimate relies on measuring the dispersion of values in the different groups of discovered local minima.

MDS-ADJ algorithm

0) Define a confidence interval I_{T_c} according to a preprocessing approach such as SDS-ADJ;
1) Choose the length values $\{n_k\}$ of the various observation subsets $\boldsymbol{t}^{(k)},\ k = 1, 2, ..., K$;
2) For $k = 1$ compute the starting pseudofrequency $\hat{f}_o^{(1,0)} = n_1\hat{T}_f / \left(\hat{T}_c^{'}\right)^2$;
3) Determine through GSS the first local minimum $v_{1,0}$ within $I_{T_c}^{(1,0)} = \left[\min(I_{T_c}),\ \min(I_{T_c}) + \frac{1}{\hat{f}_o^{(1,0)}}\right]$;

[1] round(.) denotes rounding to the nearest integer



(a) (b)

Figure 1. Cost function properties: (a) Evaluation for $T_c = 1$, $n = 30$ and $N_c = 20$; (b) Local minima positions for $T_c = 1$, $N_c = 20$ and various sets of observations having distinct dimensions.

4) For $i \geqslant 1$, an interval $I_{T_c}^{(1,i)}$ is defined for a local search as

$$I_{T_c}^{(1,i)} = \left[v_{1,(i-1)} + \frac{5}{6\hat{f}_o^{(1,i-1)}},\, v_{1,(i-1)} + \frac{7}{6\hat{f}_o^{(1,i-1)}} \right], \tag{2}$$

where $v_{1,(i-1)}$ stands for the local minimum found at $(i-1)$th step;

5) A local GSS is performed and the minimizer $v_{1,i}$ evaluations are necessary to determine the localwithin $I_{T_c}^{(1,i)}$ is stored for the next iteration;

6) The frequency estimate is updated according to $\hat{f}_o^{(1,i)} = 1/\left(v_{1,i} - v_{1,(i-1)}\right)$;

7) Steps 2, 3 and 4 are repeated until the upper bound of the search interval I_{T_c} is reached;

8) A vector $\boldsymbol{v}^{(1)} = [v_{1,1}, v_{1,2}, ..., v_{1,L_1}]$ is obtained; repeat steps 2 - 7 for $k \in \{2, \ldots, K\}$;

9) Form $\boldsymbol{u}$ as the increasingly rearranged version of $\boldsymbol{v} = \bigcup_{k=1}^{K}\left(\boldsymbol{v}^{(k)}\right)$ and compute $s_i = (1/K)\sum_{j=i}^{i+K}(u_j - m_i)^2$, where $m_i = (1/K)\sum_{j=i}^{i+K} u_j$, for $i \in \mathcal{I} = \{1, ,2, ..., \text{card}\,(\boldsymbol{u}) - K\}$;

10) Estimate the chip time as $\hat{T}_c = \underset{v_{K,i} \in \boldsymbol{v}^{(K)}}{\arg\min} |v_{K,i} - m_{i_m}|$, where $i_m = \underset{i \in \mathcal{I}}{\arg\min}\,(s_i)$.

As can be seen, the MH-GSS processing is of central importance in this generalized estimation algorithm. It must be emphasized that the computation of the standard deviations s_i for each K successive elements of $\boldsymbol{u}$ is achieved to find the nearby of T_c only, the final estimate being derived from the larger observation set $\boldsymbol{t}^{(K)}$, relatively to the i_m-th group of candidate values, as it provides the best accuracy about the global minimum location.

Theorem 1: (Loose performance bound for chip time eevaluations are necessary to determine the localstimation) Let $\boldsymbol{t}^{(K)}$ be the longest set of ToA differences for MDS-ADJ. Then, the obtained estimate $\hat{T}_c$ verifies the following inequality, for $n_K \gg \pi/2$:

$$\mathbb{E}\left(\left(\hat{T}_c - T_c\right)^2\right) \geqslant B_L \cong \frac{2\sigma_\eta^2}{(n_K N_c)^2}. \tag{3}$$

where $\mathbb{E}(\times)$ denotes the expectation of $\times$.

Proof: See Appendix A. ∎

We derive also in the next theorem a closed-form expression of the ratio between our loose performance bound B_L and the Cramér-Rao Bound (CRB) B_{T_d} coresponding to TOA differences model. Owing to this result, we can assess the efficiency of our estimator with respect to the number of observations.

Theorem 2: Let $\boldsymbol{t}$ be a n-dimensional vector of recorded TOA differences. Denoting B_T and B_{T_d}, the CRB of the ML estimate of T_c, for ToA model, and respectively for ToA differences model, the performance ratio $\eta_L = B_L/B_{T_d}$ can be expressed, for $n' \gg 1$, as

$$\eta_L \simeq \frac{n'^3}{12\,(n'-1)^2}. \tag{4}$$

Proof: First, we know from [14] that the CRB regarding the period estimation for the ToA statistical model can be expressed as $B_T = \mathbb{E}\left(\hat{B}_T\right)$, where $\hat{B}_T = n'\sigma_\eta^2 / \left(n' \sum_{j=1}^{n'} k_j^2 - \left(\sum_{j=1}^{n'} k_j\right)^2\right)$ reflects the randomness of the coefficients k_j (n' being the number of time observations). If a set of time differences $\boldsymbol{t}$ is considered to eliminate the influence of the phase term, the bound to take into account is $\hat{B}_{T_d} = 2\hat{B}_T$ [14]. Hence, the performance ratio η_L becomes

$$\eta_L = \mathbb{E}\left[\frac{B_L}{\hat{B}_{T_d}}\right] = \frac{\mathbb{E}\left[\sum_{j=1}^{n'} k_j^2\right] - \frac{1}{n'}\mathbb{E}\left[\left(\sum_{j=1}^{n'} k_j\right)^2\right]}{(n'-1)^2 N_c^2}.$$

with $k_i = iN_c + c_i$, $i \in \{1, 2, ..., n'\}$, according to the time-hopping format of the signal. For the first term in the numerator, we get $\mathbb{E}\left[\sum_{j=1}^{n'} k_j^2\right] = N_c^2 \sum_{j=1}^{n'} j^2 + 2N_c x_c^{(1)} + y_c^{(1)}$, where $x_c^{(1)} = \mathbb{E}\left[\sum_{j=1}^{n'} j c_j\right]$ and $y_c^{(1)} = \mathbb{E}\left[\sum_{j=1}^{n'} c_j^2\right]$. As random variables $\sum_{j=1}^{n'} jc_j$ and $\sum_{j=1}^{n'} c_j^2$ become normal as n' increases (central limit theorem), this first term finally takes the following expression $\mathbb{E}\left[\sum_{j=1}^{n'} k_j^2\right] = N_c^2\left(n'^3 + 3n'^2 + 3n'\right)/3 + N_c\left(n'^2 + n' - 1\right)/2$. For the second term in the numerator we obtain $\mathbb{E}\left[\left(\sum_{j=1}^{n'} k_j\right)^2\right] = N_c^2 n'^2\left(n'+1\right)^2/4 + 2N_c n'\left(n'+1\right)x_c^{(2)}/2 + y_c^{(2)}$, where $x_c^{(2)} = \mathbb{E}\left[\sum_{j=1}^{n'} c_j\right]$ and $y_c^{(2)} = \mathbb{E}\left[\left(\sum_{j=1}^{n'} c_j\right)^2\right]$. For large values of n', random variable $\left(\sum_{j=1}^{n'} c_j\right)^2$ has a non-central χ^2 distribution, with one degree of freedom, whose mean is significantly larger than its variance; hence, $y_c^{(2)} \simeq \left(x_c^{(2)}\right)^2$. Consequently, we get $E\left[\left(\sum_{j=1}^{n'} k_j\right)^2\right]/n' \simeq N_c^2\left(n'^3 + 4n'^2 + 4n'\right)/4 - N_c n'^2/2 + n'/4$. Assuming $n' \gg 1$ concludes the proof. ∎

B. Complexity

For a given length n of the input set $\boldsymbol{t}$, each evaluation of (1) has $O(n)$ complexity. If a confidence interval is considered, with a width $\Delta t = \gamma T_c$, $\gamma < 1$, then, the number l of local minima to be searched is $f_o \Delta t$, yielding $l \simeq \lfloor \gamma n N_c \rfloor$. An average number of iterations i_{GSS} is considered [20], $i_{GSS} = 1 + [(\ln(L) - \ln(\epsilon)) / \ln(\tau)]$, where $L = \frac{T_c}{nN_c}$ is the average width of the search interval (the width decreases at each iteration due to frequency f_o increase), ϵ is the desired accuracy and τ stands for the golden ratio[2]. Hence, $l \cdot i_{GSS}$ of $O(n)$ complexity function evaluations are required. Now, the complexity of MDS-ADJ is addressed. Considering a set of lengths $\boldsymbol{n} = \{n_1, n_2, ..., n_K\}$, $l_k \simeq \lfloor \gamma n_k N_c \rfloor$ minima are to be searched for each n_k, $k \in \{1, 2, ..., K\}$. Hence, $\sum_{k=1}^{K} l_k \cdot i_{GSS}^{(k)}$ evaluations of the cost function are needed to determine the elements of $\boldsymbol{u}$; $i_{GSS}^{(k)} = 1 + [(\ln(L_k) - \ln(\epsilon)) / \ln(\tau)]$ represents the average number of iterations corresponding to n_k and $L_k = \frac{T_c}{n_k N_c}$. One must take note that for different values of k we have different complexities, hence, in the numerical results section we will present as indicator the average number of evaluation for per length. Then, card$(\boldsymbol{u}) - K$ supplementary operations are necessary, where card$(\boldsymbol{u}) = \sum_{k=1}^{K} l_k$, for computing the mean and standard deviation for each K consecutive elements from $\boldsymbol{u}$.

IV. Numerical Results

We now study the performance of MDS-ADJ chip time estimator as a function of the measurement noise level and in presence of missing pulses or outliers. A few alternate estimation methods will be considered for comparison purposes :

- *Periodogram*, which yields a ML period estimate for a given record length n [14], was first proposed by Fogel and Gavish in [15] and reconsidered in [16]. This approach requires a very fine sampling of the frequency domain due to the very narrow peak corresponding to the solution.
- *Separate Least Squares Line Search - Adjacent (SLS2-ADJ)*, developed in [5], relies on a simple *round* operator-based cost function using ToA differences, with $O(n)$ complexity. The initial phase at the receiver side is eliminated with the expense of a 3 dB noise amplification.
- *Separate Least Squares Line Search - All (SLS2-ALL)* is an extension of *SLS2-ADJ*, which exploits all the possible positive differences between ToAs, with the same *round*-based cost function. The complexity is $O(n^2)$, a ML estimate being obtained for a n-length observation set [5].

One must take note that periodogram, SLS2-ADJ and SLS2-ALL are based on a line search over an interval of hypotheses, with performances highly depending on the sampling step. On the other hand, our method does not depend on this parameter at all as the exploration of the cost function is conducted through GSS and successive moves.

Monte Carlo (MC) simulations have been conducted to evaluate the performance of the various estimation algorithms. A first set of results evaluates the statistical efficiency of different methods through 10000 MC runs against jitter only, with no missing data or outlier. The record length $n' = 40$ has been considered to evaluate the performance. The parameters considered for the transmitted signal were $T_c = 1$ and $N_c = 20$, the uniformly distributed code elements c_i being changed for each run. The search interval was restricted to $[0.55\,, 1.95]$ in order to avoid ambiguities $m_1 T_c / m_2$, where $m_1, m_2 \in \mathbb{N}$ [16]. A sampling step $\tau_s = 5 \cdot 10^{-5}$ was chosen to achieve the estimation using periodogram, SLS2-ADJ and SLS2-ALL. For the MH-GSS search procedure we used in simulation $\epsilon = 5 \cdot 10^{-5}$ accuracy. It should be mentioned that, due to observation diversity enabled by MDS-ADJ, a careful design of $\boldsymbol{n} = \{n_1, n_2, ..., n_K\}$ is required to get the best results with this method. The largest observation set was selected so that $n_K = n = n' - 1$, where n' is the length of the unique input set of ToA differences $\boldsymbol{t} = \{t_1, t_2, ..., t_n\}$ used by other methods. For MDS-ADJ algorithm $\boldsymbol{n} = \{2k | \, k \in \{10, 11, ..., 14\}\} \cup \{30, 31, ..., 39\}$ and a search interval $I_{T_c} = [0.98, 1.02]$ were used. Figure 2 (a) gives the statistical performance as a function of jitter for the five considered methods. As pointed out in previous papers [5], [16], we can observe that periodogram and SLS2-ALL are optimal, as their errors reach the CRB B_T. Despite MDS-ADJ non-optimality, this approach achieves very good performance and it can be observed that the results are consistent with the approximate bound stated in theorem 1. Concerning the bias, MDS-ADJ yields similar performance to periodogram and SLS2-ALL at low jitter. A key feature of our approach is to provide a tradeoff between statistical performances and complexity. About this latter aspect, MDS-ADJ requires less computations than SLS2-ALL which relies also on the *round* operator. For the case considered in this section ($n = 40$ observations), a total number of $0.04/\tau_s = 800$ evaluations of complexity $O(n^2)$ are requested for SLS2-ALL whereas an average number of 200 operations of complexity $O(i)$, for each $i \in \{2k | \, k \in \{10, 11, ..., 14\}\} \cup \{30, 31, ..., 39\}$ are involved for MDS-ADJ; 351 supplementary operations are needed in our algorithm to get the means m_i and standard deviations s_i . The periodogram is computed in 800 points over I_{T_c}, each one with complexity $O(n)$; this approach involves lower complexity than SLS2-ALL using trigonometric functions instead.

In order to reflect more realistically the propagation channel uncertainties, missing observations and outliers were considered in a second set of trials, to evaluate the robustness of MDS-ADJ. The same simulation setup was considered with two distinct settings: first, some missing events have been inserted in the observed time datas, the positions of the missing data being modeled via a uniform distribution. By inspecting the standard deviation curve, we can conclude that missing observations produce a virtual increase of N_c, resulting in a standard deviation below the bound. However, the curve leaves the estimation bound earlier, the performance of MDS-ADJ being degraded. In a second situation, we analyzed the estimation errors in presence of erroneous datas due to false alarms; again, a uniform distribution has been used to generate these time events. It can be seen that MSE curve degrades as the rate of outliers increases, the performance reduction being more pronounced when the noise level is low. Again, no significant effect on bias is noticed. These results show that the detection stage must be designed in order to limit the false alarm rate if a MDS-ADJ estimation is achieved afterwards.

From these numerical simulations, we remark that the proposed MDS-ADJ approach is a good alternative to reference methods such as periodogram or SLS2-ALL. Its statistical accuracy at low to moderate noise levels is rather close to the CRB, while avoiding both search space sampling and evaluation of trigonometric functions. By diversification of the observed time data, an efficient tradeoff can be achieved between estimation errors and computational complexity. However, a few drawbacks can be noticed : first, a restricted search interval is needed to launch the MDS-ADJ algorithm, hence a coarse estimate has to be found first (e.g. through SLS2-ADJ, which requires relatively few calculations); second, no solution for an optimal setting of observation lengths $\boldsymbol{n}$ is available yet; and third, MDS-ADJ is sensitive to outliers.

V. Conclusions

The issue of blind chip time estimation of time-hopping signals has been considered in this paper, which is directly related to the

[2]The golden ratio is $\tau = \frac{1+\sqrt{5}}{2}$

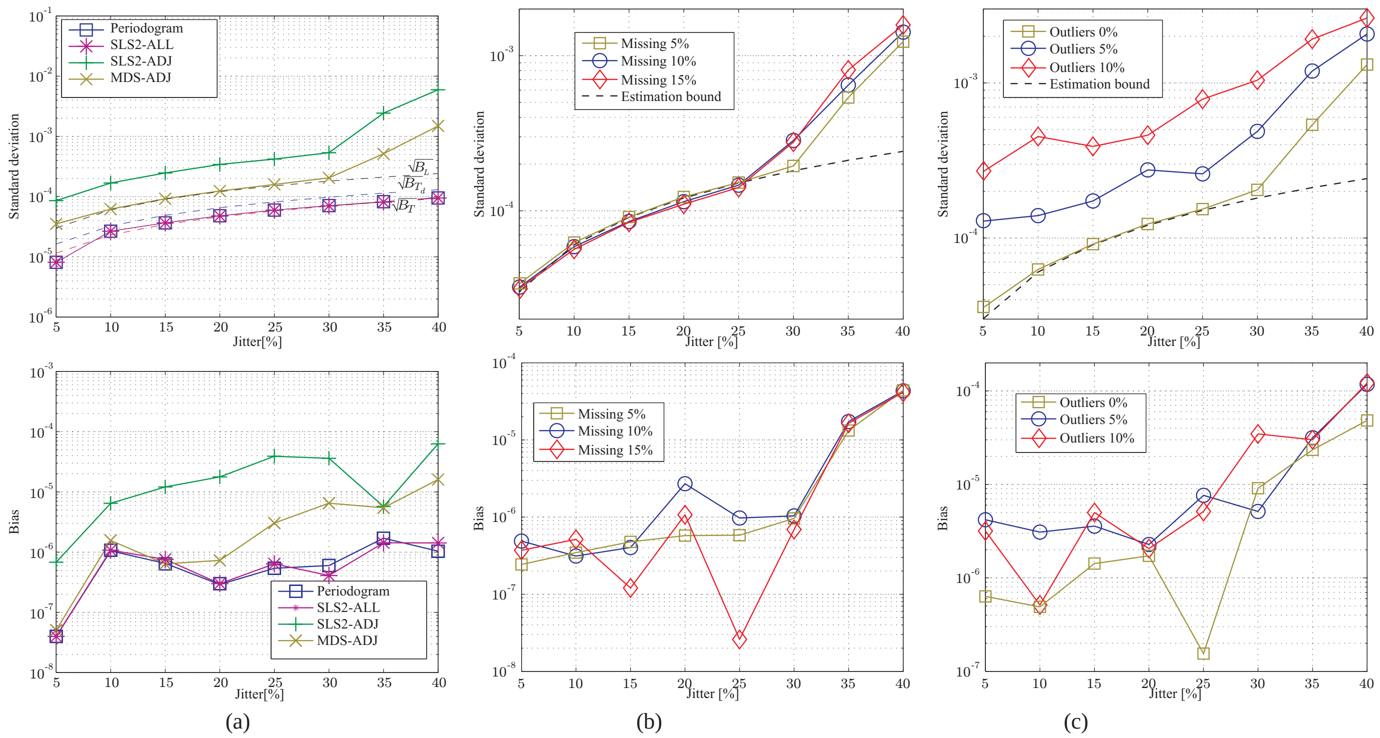


Figure 2. Chip time estimation errors against jitter: (a) Different methods comparison, (b) Missing observations and (c) Outliers effect on MDS-ADJ estimation

general problem of period estimation from sparse, noisy timing data. Following recent results of Sidiropoulos *et al.*, the role of the *round* operator has been investigated to derive a period estimate. A novel cost function has been introduced, which combines various incomplete observation sets. The highly oscillatory behavior of this function has been examined through theoretical developments. An algorithm for chip time estimation taking advantage of this property has then been proposed. Thanks to the pseudo-periodic nature of the cost function, which involves only basic operations, a good statistical performance is achieved at limited computational cost. Moreover, compared to methods like periodogram or SLS2-ALL, no sampling is required.

Appendix

Proof of Theorem 1

A set of n realizations $\{x_1, x_2, ..., x_n\} \in X^n$, where $X \sim \mathcal{N}(\mu, \sigma^2)$, is considered. For each x_i, $i \in \{1, 2, ..., n\}$, and a given value β, three possible events can be distinguished: $E_i^{=} = \{x_i = \beta\}$, $E_i^{<} = \{x_i < \beta\}$ and $E_i^{>} = \{x_i > \beta\}$. Our interest is to establish, for a given value β, which is the probability of event $\left\{y_{\frac{n-1}{2}} = \beta\right\}$, where $\{y_1, y_2, ..., y_n\}$ is the ascending order version of $\{x_1, x_2, ..., x_n\}$. Here, we only take in consideration the case of n odd. It can be easily shown that $\{E_i^{<} \cap E_j^{>}\} \cap \{E_i^{>} \cap E_j^{<}\} = \varnothing$, the sets $\{E_i^{=}\} \cup \{E_j^{>} | j \in \{1, 2, ..., n\}, j \neq i\}$ and respectively $\{E_i^{=}\} \cup \{E_j^{<} | j \in \{1, 2, ..., n\}, j \neq i\}$ contain elements that are independent one to each other and $\forall\, i \in \{1, 2, ..., n\}$, $P\left(E_i^{<}\right) = \left(1 + \text{erf}\left((\beta - \mu) / \left(\sigma\sqrt{2}\right)\right)\right)/2$ and $P\left(E_i^{>}\right) = \left(1 - \text{erf}\left((\beta - \mu) / \left(\sigma\sqrt{2}\right)\right)\right)/2$. We define,

$$E_i^{'} = \bigcup_{k=1}^{K} \left\{ \left\{ \bigcap_{j \in \mathbf{J}_{i,k}^{>}} E_j^{>} \right\} \cap \left\{ \bigcap_{j \in \mathbf{J}_{i,k}^{<}} E_j^{<} \right\} \right\},$$

where, $\mathbf{J}_{i,k}^{>}$ and $\mathbf{J}_{i,k}^{<}$ have the same length $(n-1)/2$, $\{\mathbf{J}_{i,k}^{>} \cap \mathbf{J}_{i,k}^{<}\} = \varnothing$ and $\{\mathbf{J}_{i,k}^{>} \cup \mathbf{J}_{i,k}^{<}\} = \{j \in \{1, 2, ..., n\} \,|\, j \neq i\}$. $K_n = \binom{\frac{n-1}{2}}{n-1}$ represents the total number of combinations $(\mathbf{J}_{i,k}^{>}, \mathbf{J}_{i,k}^{<})$. Thus, $P\left(y_{\frac{n+1}{2}} = \beta\right) = P\left(\bigcup_{i=1}^{n} E_i^{'}\right)$, which can be approximated as [21]

$$P\left(y_{\frac{n+1}{2}} = \beta\right) \simeq C_n \cdot P(x = \beta) \cdot \exp\left(-\frac{(n-1)(\beta - \mu)^2}{\pi\sigma^2}\right),$$

where $C_n = nK_n 4^{\frac{1-n}{2}}$. The corresponding probability density function is:

$$f(\beta) \simeq \frac{C_n}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(\beta - \mu)^2}{\sigma\sqrt{2}}\left(1 + \frac{2(n-1)}{\pi}\right)\right). \quad (5)$$

We particularize now the above relation to our problem. To evaluate the estimation error, we express our cost function (1) as a sum of partial functions $h_m\left(\boldsymbol{t}, \widetilde{T}\right) = \left|(S - t_m)/\widetilde{T} - \text{round}\left((S - t_m)/\widetilde{T}\right)\right|$; $m = 1, 2, ..., n$. It can be easily shown that $(S - t_m) = (C - \Delta c_m) T_c + \Gamma - \delta_m$, where $C \cong nN_c$ and $\Gamma = \eta_{n+1} - \eta_1$. In the nearby of T_c, each function h_m equals zero at $\widetilde{T}_m = \varphi_m / (C - \Delta c_m) \cong \varphi_m / C$, where $\varphi_m = \Gamma - \delta_m$. The random variables δ_m and Γ have the probability distribution $\mathcal{N}(0, \sigma^2)$, where $\sigma^2 = 2\sigma_\eta^2$. For a given set of observations $\boldsymbol{t}$, only two terms of $\{\delta_m\}$ exhibit a correlation with Γ which has a constant value. Hence, we can assume that each r.v. φ_m is approximately distributed as $\mathcal{N}(\Gamma, \sigma^2)$. For the sake of simplicity of the presentation we denote $\Psi = \{\varphi_m';\, m = 1, 2, ..., n\}$ as the set of increasingly ordered values of φ_m. As partial functions h_m are convex in the nearby of T_c, with the same form but differently delayed, the global minimum of the cost function, corresponding to estimate $\hat{T}_c$, is $\xi = \varphi'_{\frac{n+1}{2}}$. From (5), the corresponding probability density function for a single realization can be expressed as

$$f(\xi, \Gamma) \simeq \frac{C_n}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(\xi - \Gamma)^2}{\sigma\sqrt{2}}\left(1 + \frac{2(n-1)}{\pi}\right)\right).$$

Considering that $f(\xi) \simeq \int_{-\infty}^{\infty} f(\xi,\Gamma) f(\Gamma)\, d\Gamma$, we then obtain the following closed form expression:

$$f(\xi) \simeq C_n \sqrt{\frac{\pi}{\pi + 2(n-1)}} \cdot \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\xi^2}{2\sigma^2} \frac{\pi + 2(n-1)}{2\pi + 2(n-1)}\right). \quad (6)$$

If we look at argument of the exponential in the relation above, ξ is a normal, unbiased variable with the approximate variance $\sigma^2 (2\pi + 2(n-1)) / (\pi + 2(n-1))$. For large values of n the variance of ξ can be further approximated to σ^2. As the chip time estimate is $\hat{T}_c = \xi / C$, we finally get

$$\mathrm{var}\left(\hat{T}_c\right) = \frac{\mathrm{var}(\xi)}{C^2} \simeq \frac{2\sigma_n^2}{(nN_c)^2}.$$

ACKNOWLEDGMENT*:* The authors are grateful to the reviewers for their comments which significantly helped to improve this paper.

REFERENCES

[1] R. G. Wiley, *Electronic Intelligence: The Analysis of Radar Signals*, 2nd edition, Norwood, MA: Artech House, 1993.

[2] E. Fogel, M. Gavish, “Performance evaluation of zero-crossing-based bit synchronizers,” *IEEE Trans. On Communications*, 37 (6), pp. 663-665, 1989.

[3] D.H. Staelin, "Fast folding algorithm for detection of periodic pulse trains," *Proc. IEEE*, Vol. 57, pp. 724-725, 1969.

[4] D. R. Brillinger, “Maximum Likelihood Analysis of Spike Trains of Interacting Nerve Cells,” *Biol. Cyber.*, 59, pp. 189-200, 1988.

[5] N. D. Sidiropoulos, A. Swami, B. M. Sadler, "Quasi-ML period estimation from incomplete timing data," *IEEE Trans. on Signal Processing*, Vol. 53, No. 2, pp. 733-739, 2005.

[6] W. A. Gardner, "Signal interception : a unifying theoretical framework for feature detection," *IEEE Trans. On Communications*, Vol. 36, No. 8, pp. 897-906, 1988.

[7] R. A. Dillard, G. M. Dillard, *Detectability of Spread-Spectrum Signals*, Artech House Publishers, Norwood, MA, USA, 1989.

[8] H. Arslan, Z. N. Chen, M.-G. Di Benedetto, *Ultra Wideband Wireless Communication*, Wiley-Interscience, Oct. 2006.

[9] U. S. Federal Communications Commission, FCC 02-48: First Report and Order, Feb. 2002.

[10] M.Z. Win and R.A. Scholtz, “Impulse radio: How it works,” *IEEE Commun. Lett.*, vol.2, no.2, pp.36–38, 1998.

[11] K. Mardia, "New techniques for the deinterleaving of repetitive sequences," *IEE Proceedings*, 136, Pt. F, 4 (Aug. 1989), pp. 149-154.

[12] D. A. Gray, B. J. Slocumb, S. D. Elton, "Parameter estimation for periodic discrete event processes," *Proc. IEEE ICASSP'94* , Vol IV, 1994.

[13] T. L. Conroy and J. B. Moore, "On the estimation of interleaved pulse train phases," *IEEE Trans. on Signal Processing*, vol. 48, issue 12, pp. 3420-3425, 2000.

[14] B. M. Sadler, S. D. Casey, “On periodic pulse interval analysis with outliers and missing observations,” *IEEE Trans. on Signal Processing*, vol. 46, issue 11, pp. 2990-3002, 1998.

[15] E. Fogel, M. Gavish, "Parameter estimation of quasi-periodic sequences," *Proc. of the Int. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, vol. 4, pp. 2348-2351, April 1988.

[16] I. V. L. Clarkson, "Approximate maximum-likelihood period estimation from sparse, noisy timing data," *IEEE Trans. Signal Processing*, Vol. 56, No. 5, pp. 1779-1787, May 2008.

[17] S. D. Elton, D.A. Gray, "The application of circular statistics to specific radar pulse train detection," *Proc. of the 7th European Signal Processing Conf. (EUSIPCO)*, vol. 1, pp. 284-287, Sept. 1994.

[18] R. Jackson, D. Pycock, M. Xu, M. Salous, M. Knowles and S. Harman, "Event detection and period extraction using multi-scale symmetry and entropy," *Signal Processing*, vol. 85 , No. 3, pp. 591 - 605, March 2005.

[19] I. Guvenc, Z. Sahinoglu, and P. Orlik, “TOA estimation for IR-UWB systems with different transceiver types,” *IEEE Trans. on Microwave Theory and Techniques*, 54(4), 1876-1886, 2006.

[20] R. P. Brent, “Algorithms for minimization without derivatives,” Dover Publications, ISBN : 0-486-41998-3, 2002.

[21] S. Winitzki, "Uniform approximations for transcendental functions," in *Proc. Int. Conf. Computational Science and Its Applications (ICCSA 2003)*, Montreal, Canada, May 2003.