THE BARANKIN BOUND AND THRESHOLD BEHAVIOR IN FREQUENCY ESTIMATION

Luc Knockaert

Abstract

This correspondence presents the Barankin bound as a fundamental statistical tool for the understanding of the threshold effect associated with the estimation of the frequency of a sinusoid in additive white Gaussian noise. It is shown that the threshold effect takes hold whenever the Barankin bound departs significantly from the Cramer-Rao bound. In terms of the signal-to-noise ratio (SNR) and the data length $T$, the quantity $\text{SNR} \times T / \ln T$ is shown to be a good indicator for deciding whether the SNR is above threshold or not.

1 INTRODUCTION

The problem of estimating the frequency of a sinusoid in additive white Gaussian noise is one of considerable interest. In most cases [1], [2] the maximum likelihood (ML) procedure is utilized to obtain what one expects to be a sufficiently unbiased and efficient estimator of the frequency. Due to the nonlinear nature of the frequency estimation problem, the so-called threshold effect [2], [3] takes hold whenever the SNR drops below a critical data-length dependent level $\text{SNR}_c(T)$. The threshold effect can be characterized by an almost instant and drastic deterioration of the frequency estimator variance with respect to the Cramer-Rao bound (CRB) below this critical SNR level. In [2] the threshold effect was related to the existence of highly probable outliers, farly removed from the exact frequency, and in [3] a more technical device, related to the phase locked loop, was proposed to explain the phenomenon. The aim of this correspondence is to provide a more fundamental approach to the understanding of the threshold effect. Our starting point is the fact that the CRB, although being the best lower bound in the linear Gaussian case, is a less appropriate tool when dealing with non-linear problems such as frequency estimation. For non-linear problems, the Barankin bound (BRB) [4], [5] is a stronger lower bound for the variances of unbiased estimators, including the CRB as a limit case. The threshold effect region
can therefore, in the sense of Barakin, be defined as the region where the BRB suddenly departs from the CRB. This is fully exploited in the sequel, resulting in a simple indicator quantity for threshold behavior in frequency estimation.

2 THE BARANKIN BOUND

The simplest form of the Barankin bound for the estimation of a scalar real parameter \( a \) can be stated as follows [5]. Let \( p(x|a) \) be the probability density of the vector \( x \), given \( a \). Let \( h \) be a real number, independent of \( x \) such that \( a + h \) ranges over all possible values of \( a \). Then for any unbiased estimator \( \xi(x) \) we have

\[
\text{var}(\xi) \geq \text{BRB},
\]

where

\[
\text{BRB} = \sup_h h^2 \int p(x|a+h)^2/p(x|a)dx - 1 \geq \text{CRB},
\]

and the CRB is given by

\[
\text{CRB} = \lim_{h \to 0} h^2 \int p(x|a+h)^2/p(x|a)dx - 1 = \frac{1}{\text{var}[\partial \ln p(x|a)/\partial a]}.
\]

To avoid theoretical complications we assume that the integral in the denominator of (2) exists and that the support of \( p(x|a) \) and its partial derivative with respect to \( a \) is \( R^T \) for almost all \( a \).

A natural way to measure the deviation of the BRB from the CRB is the ratio

\[
Q = \frac{\text{BRB}}{\text{CRB}} = \sup_h V(h),
\]

where

\[
V(h) = \frac{h^2 \text{var}[\partial \ln p(x|a)/\partial a]}{\int p(x|a+h)^2/p(x|a)dx - 1}.
\]

When \( Q = 1 \), then the supremum in (5) is obtained for \( h = 0 \), and in that case we say that there is no Barankin threshold effect. This does not imply that there is no threshold effect whatsoever, since there exist still stronger bounds than the above BRB [4], [5]. When \( Q > 1 \), then the supremum is obtained for \( h \neq 0 \), and in that case there surely exists a threshold effect, since the BRB and hence the variance of the estimator then depart from the CRB.
To show what this means in practice we apply this to a simple, but frequently occurring non-linear problem. Let the observed data vector $x$ be given by

$$x = f(a) + n,$$

where $n$ is $N(0, \sigma^2 I)$ Gaussian noise, and $f(a)$ is a function, in general non-linear, mapping the parameter $a$ into the data space.

After some elementary calculations, we obtain

$$V(h) = \frac{h^2 |f'(a)|^2}{\sigma^2 \left( e^{\frac{f(a+h)-f(a)^2}{\sigma^2}} - 1 \right)}.$$  

(7)

Note that when the problem is linear, i.e. when $f'(a)$ is a constant vector, $V(h)$ is a strictly decreasing function of $h^2$, which implies that $Q = 1$. This is easily understood, since linear problems in additive Gaussian noise never exhibit threshold effect. Note also that when we have $M$ independent realizations of the same process, the above formula remains valid after replacing $\sigma^2$ with $\sigma^2 / M$.

### 3 APPLICATION TO FREQUENCY ESTIMATION

Consider the single tone frequency estimation problem

$$z(t) = Ae^{j(at+\alpha)} + \nu(t) \quad t = 0,1,\ldots,T-1.$$  

(8)

The $\nu(t)$ are white complex independent Gaussian random variables with the same noise variance $\sigma^2$ and the amplitude $A > 0$ and initial phase $\alpha$ are assumed known. Of primary interest is the estimation of the angular frequency $-\pi \leq a < \pi$.

In the terms of the preceding section, we have a $2T$–dimensional data space

$$x_k = \Re\{z(k-1)\},$$

$$x_{k+T} = \Im\{z(k-1)\} \quad k = 1,2,\ldots,T$$

(9)

and

$$f_k(a) = A \cos((k-1)a + \alpha),$$

$$f_{k+T}(a) = A \sin((k-1)a + \alpha) \quad k = 1,2,\ldots,T.$$  

(10)
Also

\[ f'_k(a) = -A(k - 1) \sin((k - 1)a + \alpha), \]
\[ f'_{k+T}(a) = A(k - 1) \cos((k - 1)a + \alpha) \quad k = 1, 2, \ldots, T \]  

(11)

and

\[ |f'(a)|^2 = \sum_{k=1}^{T} \left\{ |f_k'(a)|^2 + |f'_{k+T}(a)|^2 \right\} \]
\[ = A^2 \sum_{k=1}^{T} (k - 1)^2 \]
\[ = A^2 T^6 (2T - 1)(T - 1). \]  

(12)

In the same vein we have

\[ |f(a + h) - f(a)|^2 = 2A^2 \left[ T - 1 - \sum_{k=1}^{T-1} \cos(kh) \right] \]
\[ = A^2 \left[ 2T - 1 - \frac{\sin(T - \frac{1}{2})h}{\sin \frac{h}{2}} \right], \]  

(13)

where the explicit result for the sum of cosinuses can be found in [6], p. 73. Note that there is no dependence on the initial phase \( \alpha \). The objective function \( V(h) \) can be written as

\[ V(h) = \frac{\Gamma h^2 T(T - 1)/6}{e \left[ 1 - \sin(T - \frac{1}{2})h/(2T - 1) \sin \frac{h}{2} \right] - 1}, \]  

(14)

where

\[ \Gamma = (2T - 1) \frac{A^2}{\sigma^2}. \]  

(15)

In order to find \( Q = \sup_{h} V(h) \), we need to know the range of \( h \). Since \( V(h) \) is an even function of \( h \), and since \( a + h \) has to range over all possible values of \( a \) modulo \( 2\pi \), we are enabled to take \( 0 \leq h \leq \pi \) and hence

\[ Q = \sup_{0 \leq h \leq \pi} V(h). \]  

(16)

In order not to have immediate threshold behavior we must require that \( h = 0 \) is at least a local maximum of the function \( V(h) \). For \( h \) sufficiently small we have

\[ V(h) \approx 1 - h^2 \frac{T(T - 1)}{12} \left[ \Gamma - \left( \frac{3}{5} - \frac{1}{5T(T - 1)} \right) \right] + O(h^4). \]  

(17)
For $h = 0$ to be a local maximum a necessary condition is therefore

$$\Gamma > \frac{3}{5} - \frac{1}{5T(T-1)}. \tag{18}$$

The above inequality is always satisfied if we simply take $\Gamma > \frac{3}{5}$. Since the SNR (not in decibels) is defined as $[2], [3]$ SNR = $A^2/2\sigma^2$, this implies that we should have

$$\text{SNR} \times T > \frac{3}{20}. \tag{19}$$

When condition (18) is not satisfied, $h = 0$ is a local minimum of $V(h)$ and $Q$ will certainly always largely exceed unity. This means that the threshold effect will always be active in that case, resulting in large variances for any unbiased estimator of the angular frequency, and thereby making efficient frequency estimation almost impossible. This confirms partly, but not completely, as we shall see in the sequel, the conclusions of $[3]$, where the indicator quantity $3\sigma^2/A^2T$ was utilized, which quantity should therefore be smaller than 10. Note also that when we have $M$ independent data realizations at our disposition, condition (19) remains valid after replacing $T$ with $MT$, the total number of data points.

On the other hand, when condition (18) is satisfied, the function $V(h)$ drops sharply in the vicinity of $h = 0$ and afterwards behaves as a parabola proportional to $h^2$, slightly modulated due to the presence of the factor $\sin((T - 0.5)h)/(2T - 1) \sin(h/2)$. Hence it appears that the local maximum, closest to $h = \pi$, of the latter factor determines the global behavior of $V(h)$ in the vicinity of $\pi$. For $T$ odd, this happens when $h = \pi$, and for $T$ even, this happens when $h \approx \pi - \pi/(T - 0.5)$. This is illustrated in Figure 1, where $V(h)$ is plotted for SNR $\times T = 2.6, 2.4$ and $T = 33$. This allows us to infer that

$$Q \approx \max \left\{ 1, \frac{\Gamma \pi^2 T(T-1)/6}{e^{\Gamma(2T-2)/(2T-1)} - 1} \right\}, \tag{20}$$

where the approximation is exact for $T$ odd. Hence we conclude that we are certainly in the threshold region when

$$\frac{\Gamma \pi^2 T(T-1)/6}{e^{\Gamma(2T-2)/(2T-1)} - 1} > 1, \tag{21}$$

since in that case no unbiased estimator can achieve the exact CRB. On the other hand we have chances of not being in the threshold region when

$$\frac{\Gamma \pi^2 T(T-1)/6}{e^{\Gamma(2T-2)/(2T-1)} - 1} \ll 1. \tag{22}$$
Defining $\gamma = \Gamma(2T - 2)/(2T - 1)$, equation (21) can be written as

$$\frac{e^\gamma - 1}{\gamma} < (2T - 1)T \frac{\pi^2}{12}.$$ (23)

Taking advantage of the fact that

$$e^{\gamma/2} \leq \frac{e^\gamma - 1}{\gamma}, \quad \forall \gamma \geq 0,$$ (24)

equation (23) can be strengthened to

$$\gamma < 2 \ln \left( (2T - 1)T \frac{\pi^2}{12} \right).$$ (25)

For sufficiently large $T$, this means that we are in the threshold region when

$$\text{SNR} \times \frac{T}{\ln T} < 1.$$ (26)

This is further illustrated in Figure 2, where $Q$ is plotted as a function of $\Theta$ for fixed $T = 33$. By contrast, the property of being outside the threshold region may be described by the approximate inequality

$$\text{SNR} \times \frac{T}{\ln T} \gg 1.$$ (27)

Hence it appears that, at the level of the Barankin bound, the indicator quantity

$$\Theta = \text{SNR} \times \frac{T}{\ln T}$$ (28)

is crucial to the understanding of threshold behavior.

Based on the data in [3], we construct the following table, pertaining to the onset of the threshold effect for the maximum likelihood estimator of the angular frequency:

<table>
<thead>
<tr>
<th>$T$</th>
<th>SNR(dB)</th>
<th>$\Theta$</th>
<th>$\text{SNR} \times T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>0</td>
<td>9.2</td>
<td>32.0</td>
</tr>
<tr>
<td>64</td>
<td>-2.75</td>
<td>8.2</td>
<td>34.0</td>
</tr>
<tr>
<td>128</td>
<td>-5.5</td>
<td>7.4</td>
<td>36.1</td>
</tr>
<tr>
<td>256</td>
<td>-8.25</td>
<td>6.9</td>
<td>38.3</td>
</tr>
<tr>
<td>512</td>
<td>-11</td>
<td>6.5</td>
<td>40.7</td>
</tr>
</tbody>
</table>

It is seen that the indicator quantity $\text{SNR} \times T$ is not a very good one, since it clearly forms an increasing sequence, and is therefore not likely to possess an upper bound. On the other hand,
the quantity $\Theta$ forms a decreasing sequence, and therefore always has an upper bound. Since $T/\ln T$ is an increasing function of $T$ whenever $T > e$, the condition, say $\Theta \geq 70$, will therefore always pull the ML estimator out of the threshold region. The reason why the $\Theta$ sequence is not approximately constant may be explained by the biasedness which typically affects ML estimators, especially when the number of data points is small [7] p. 426. In the above context, the author would like to thank one of the reviewers for pointing out that in [8] a hybrid Barankin-Bhattacharyya bound was developed, with the possibility of incorporating biasedness corrections and a generalization to the multiple harmonics problem.

4 CONCLUSION

A fundamental approach to the understanding of the threshold effect in frequency estimation, based on the Barankin bound, has been proposed. It is shown that threshold behavior is a typical non-linear effect due to the departure of the Barankin bound from the Cramer-Rao bound at low SNR levels. A simple indicator quantity, with better behavior than the one proposed in [3], to characterize the onset of the threshold effect, has been derived. The problem of including a possible biasedness corrector in the Barankin bound in the sense of [8], in order to come up with an even better indicator quantity, is the subject of ongoing research.

References


