Blind Carrier Frequency Offset estimation and Mean Square Error Lower bounds

Philippe Ciblat

École Nationale Supérieure des Télécommunications, Paris, France

Outline

1. Blind Carrier Frequency Offset synchronization
   - Harmonic retrieval in multiplicative noise
   - Design of powerful estimates
   - Asymptotic analysis
   - Probability of outliers
2. Mean Square Error Lower bound
   - Standard Cramer-Rao bound
   - Cramer-Rao bound with nuisance parameter
   - Bayesian Cramer-Rao bound
   - Other bounds
     - Deterministic approach : Battacharya, Barankin
     - Random approach : Ziv-Zakai

Harmonic retrieval (I)

We assume

$$y(n) = a(n)e^{2i\pi f_0 n} + b(n), \quad n = 0, \ldots, N-1$$

with

- $y(n)$: the received signal
- $a(n)$: a zero-mean *random process* or a *time-varying amplitude*.
- $b(n)$: circular white Gaussian stationary additive noise.

Goal : Estimating the frequency f_0 in multiplicative and additive noise

Outline :

1. Short review on some estimates
2. Derivations of asymptotic performance and non-asymptotic performance
3. MSE lower bounds associated with this problem

Harmonic retrieval (II)

Previous model holds for

Digital Communications : Non-data-aided/Blind synchronization.

$$a(n) = \sum_{l=0}^{L} h_l s_{n-l}$$

$\rightsquigarrow$ *circular/noncircular* complex-valued MA process
$\rightsquigarrow$ non-Gaussian process

Radar : Jakes model
$\rightsquigarrow$ $a(n)$ *circular* complex-valued Gaussian process.

Direction of Arrival (DOA) : Frequency domain.
$\rightsquigarrow$ $a(n)$ *circular* complex-valued Gaussian process.

Literature on estimator design

Digital Communications community (COM)

- A. Viterbi, U. Mengali, M. Moeneclaey
  - ⇝ Ad hoc algorithms based on modulation properties (Gaussian channel)

Signal Processing community (SP)

- P. Whittle, D. Brillinger, E. Hannan, A. Walker (1950-1970)
  - ⇝ Constant amplitude and periodogram analysis.
- O. Besson, P. Ciblat, M. Ghogho, G.B. Giannakis, H. Messer, E. Serpedin, P. Stoica (1990-present)
  - ⇝ Time-varying amplitude
  - ⇝ Notion of non-circularity
  - ⇝ Notion of cyclostationary
  - ⇝ Asymptotic performance analysis

Definition of circularity

Circularity (strict sense)

Let Z be a zero-mean complex random variable. Z is said circular in strict sense iff

$$Z \quad \text{and} \quad Ze^{i\theta}$$

have the same distribution for any θ.

Property

$$\mathbb{E}[\underbrace{Z \cdots Z}_{p \ \text{times}} \underbrace{\overline{Z} \cdots \overline{Z}}_{q \ \text{times}}] = 0$$

as soon as $p \neq q$.

Remark Z is M-order noncircular/ $M-1$-order circular random variable if only the moments of order $(M-1)$ or less satisfy the previous property

Second order circular case (I)

Assumptions

- $a(n)$ is second order circular (= circular in wide sense)

$$\mathbb{E}[a(n)^2] = 0$$

- $a(n)$ is Gaussian
- $a(n)$ is colored
- $a(n)$ obeys the Jakes model

$$r_a(\tau) = J_0(2\pi f_d \tau)$$

  and so $r_a(\tau)$ is real-valued.

$\rightsquigarrow$ Applications : Radar
$\rightsquigarrow$ SP community

Second order circular case (II)

We get

$$r_y(\tau) = \mathbb{E}\left[y(n+\tau)\overline{y(n)}\right] = r_a(\tau)e^{2i\pi f_0 \tau}, \quad \forall \tau \neq 0$$

As $r_a(\tau)$ is real-valued (as in Jakes model), we obtain

$$\hat{f}_N = \frac{1}{2\pi\tau} \angle \hat{r}_N(\tau)$$

where $\hat{r}_N(\tau)$ is the empirical estimate of $r_y(\tau)$ when N samples are available.

Remark

Estimating frequency boils down to estimating constant phase.

Non-circular case (I)

Assumptions

- $a(n)$ is M-order noncircular

$$\mathbb{E}[a(n)^M] \neq 0$$

- $a(n)$ is Gaussian or not
- $a(n)$ is colored or not
- $a(n)$ is a MA process

$$a(n) = \sum_{l=0}^{L} h_l s_{n-l}$$

where $\{h_l\}$ is the impulse channel response and s_n is the unknown M-order noncircular data

Non-circular case (II)

Remark

Any usual constellation is rotationally symmetric over $2\pi/M$.

Constellation	M
P-PAM	2
P-PSK	P
P-QAM	4

One can prove that any usual constellation is M-order noncircular

$\rightsquigarrow$ Applications : Digital communications
$\rightsquigarrow$ COM and SP community

Second order non-circular case (I)

Deterministic ML based method : Besson 1998

$$\{\hat{\mathbf{a}}_N, \hat{f}_N\} = \arg\min_{\mathbf{a},f} \mathbf{K}_N(\mathbf{a}, f) = \frac{1}{N}\sum_{n=0}^{N-1} |y(n) - a(n)e^{2i\pi fn}|^2$$

Non-linear least square (NLLS) asymptotically equivalent to maximization of periodogram of $y^2(n)$

$$\hat{f}_N = \arg\min_{f} \mathbf{J}_N(f) = \left| \frac{1}{N}\sum_{n=0}^{N-1} y^2(n)e^{-2i\pi(2f)n} \right|^2$$

$\rightsquigarrow$ Traditional Square-Power estimate in COM community for BPSK

Second order non-circular case (II)

Remark

As $u_a(0) = \mathbb{E}[a^2(n)] \neq 0$, then

$$z(n) = y^2(n) = r_a(0)e^{2i\pi(2f_0)n} + e(n)$$

where $e(n)$ is a *non-Gaussian* and *non-stationary* additive noise.

Conclusion

- Frequency estimation in multiplicative and additive noise

  $\Updownarrow$

  Frequency estimation in additive noise *but non-standard noise*
- Periodogram based on $y^2(n)$ instead of $y(n)$.
- If $a(n)$ colored, periodogram not exhaustive.

Cyclostationary based method

- Let $u_y(n,\tau) = \mathbb{E}[y(n+\tau)y(n)]$ be the *pseudo-correlation*

Definition

$y(n)$ is cyclostationary w.r.t. its pseudo-correlation iff $n \mapsto u_y(n,\tau)$ is periodic of period $1/\alpha_0$. Then

$$u_y(n,\tau) = \sum_k u_y^{(k\alpha_0)}(\tau) e^{2i\pi k\alpha_0 n}$$

with

- $k\alpha_0$: k^{th} cyclic frequency
- $u_y^{(k\alpha_0)}(\tau)$: cyclic pseudo correlation
- $c_a^{(k\alpha_0)}(e^{2i\pi f}) = \text{F.T.}(\tau \mapsto u_y^{(k\alpha_0)}(\tau))$: cyclic pseudo spectrum

- $n \mapsto u_y(n,\tau)$ is periodic of period $1/\alpha_0$ with $\alpha_0 = 2f_0$
- Ciblat & Loubaton 2000

Contrast function

Remark

Estimating frequency in multiplicative and additive noise boils down to estimating a cyclic frequency

$$f_0 = \arg\max_f \mathbf{J}_{\mathbf{W}}(f) = \mathbf{u}_y^{(2f)^{\mathrm{H}}} \mathbf{W} \mathbf{u}_y^{(2f)} = \left\| \mathbf{u}_y^{(2f)} \right\|_{\mathbf{W}}^2$$

with $\mathbf{u}_y^{(\alpha)} = [u_y^{(\alpha)}(-T), \cdots, u_y^{(\alpha)}(T)]^{\mathrm{T}}$.

In practice, $\mathbf{u}_y^{(2f)}$ is not available and needs to be estimated

$$\begin{aligned} u_y^{(\alpha)}(\tau) &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} u_y(n,\tau) e^{-2i\pi\alpha n} \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}[y(n+\tau)y(n)] e^{-2i\pi\alpha n} \end{aligned}$$

Contrast process

$$\hat{f}_N = \arg\max_f \mathbf{J}_{N,\mathbf{w}}(f) = \left\| \hat{\mathbf{u}}_N^{(\alpha)} \right\|_{\mathbf{w}}^2$$

with $\hat{\mathbf{u}}_y^{(\alpha)} = [\hat{u}_y^{(\alpha)}(-T), \cdots, \hat{u}_y^{(\alpha)}(T)]^{\mathrm{T}}$ and

$$\hat{u}_N^{(\alpha)}(\tau) = \frac{1}{N}\sum_{n=0}^{N-1} y(n+\tau)y(n)e^{-2i\pi\alpha n}.$$

Then

$$\hat{f}_N = \arg\max_f \mathbf{J}_{N,\mathbf{w}}(f) = \left\| \frac{1}{N}\sum_{n=0}^{N-1} \mathbf{z}(n)e^{-2i\pi(2f)n} \right\|_{\mathbf{w}}^2$$

with $\mathbf{z}(n) = [z_{-T}(n), \ldots, z_T(n)]^{\mathrm{T}}$ and
$z_\tau(n) = y(n+\tau)y(n) = u_y^{(\alpha_0)}(\tau)e^{2i\pi\alpha_0 n} + e_\tau(n)$.

Remarks

- Multi-variate periodogram
- Weighted periodogram
- Extended Square-Power algorithm
- Asymptotic performance
  - Giannakis & Zhou
    1995 : cyclostationarity approach and CRB bounds
  - Besson & Stoica
    1999 : deterministic NLS with white real-valued multiplicative noise
  - Ghogho & Swami
    1999 : deterministic NLS with white real-valued multiplicative noise
  - Ciblat & Loubaton
    2000 : weighted multi-variate periodogram and analysis with colored complex-valued multiplicative noise

High-order noncircular case

***P*-PSK** : Viterbi 1983.

$$\mathbb{E}[a(n)^P] \neq 0 \Leftrightarrow \hat{f}_N = \arg\max_f \left\| \frac{1}{N} \sum_{n=0}^{N-1} y^P(n) e^{-2i\pi(Pf)n} \right\|^2$$

Tutorial done by Morelli-Mengali in 1998.

***P*-QAM** : Moeneclaey 2001 & Serpedin 2004

$$\mathbb{E}[a(n)^4] \neq 0 \Leftrightarrow \hat{f}_N = \arg\max_f \left\| \frac{1}{N} \sum_{n=0}^{N-1} y^4(n) e^{-2i\pi(4f)n} \right\|^2$$

$\Rightarrow$ *The so-called M-power estimate*

Asymptotic analysis

- Consistency

$$\hat{f}_N - f_0 \stackrel{p.s.}{\rightarrow} 0$$

- Asymptotic normality : it exists p such that

$$N^p(\hat{f}_N - f_0) \stackrel{\mathcal{D}}{\rightarrow} \mathcal{N}(0, \gamma)$$

  with
  - p the so-called convergence speed
  - γ the so-called asymptotic covariance
- Asymptotic covariance

$$\mathrm{MSE} = \mathbb{E}[(\hat{f}_N - f_0)^2] \sim \frac{\gamma}{N^{2p}}$$

Convergence analysis

- Consistency
- Asymptotic normality (with $p = 3/2$)

are proven in Ciblat & Loubaton for

$$\hat{\alpha}_N = \arg\max_{\alpha} \mathbf{J}_{N,\mathbf{W}}(\alpha) = \left\| \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{z}(n) e^{-2i\pi\alpha n} \right\|_{\mathbf{W}}^2$$

where

$$\mathbf{z}(n) = \mathbf{u} e^{2i\pi\alpha_0 n} + \mathbf{e}(n)$$

whatever the noise process $\mathbf{e}(n)$ satisfying standard mixing conditions

Remarks

- Analysis valid for second order and high order noncircular case
- Derivations of the asymptotic covariance need still to be done

Asymptotic covariance (I)

Second-order noncircular case :
whatever the second-order noncircular process $a(n)$, Ciblat & Loubaton (*IEEE SP 2002*) have proven that

1. $\mathbf{W}_{opt} = \mathbf{Id}_{2T+1}$
2. $T_{opt} = L$ with L the memory size of $a(n)$
3. $$\text{MSE} \sim \frac{3}{4\pi^2 N^3} \cdot \frac{\int_0^1 |c_a(e^{2i\pi f})|^2 \mathcal{X}(e^{2i\pi f}) df}{\left(\int_0^1 |c_a(e^{2i\pi f})|^2 df\right)^2}.$$

with
$\mathcal{X}(e^{2i\pi f}) = (s_a(e^{2i\pi f}) + \sigma^2)(\overline{s_a(e^{-2i\pi f})} + \sigma^2) - c_a(e^{2i\pi f})\overline{c_a(e^{-2i\pi f})}$

- if $a(n)$ is a white real-valued process, then asymptotic covariance also available in Ghogho and in Besson

Asymptotic covariance (II)

High-order noncircular case :
Serpedin (*IEEE TCOM 2003* and *IEEE TWIRELESS 2003*) has proven that

$$\mathrm{MSE}_{P\text{-PSK}} \sim \frac{24}{\pi^2 P N^3} \frac{B-D}{C^2}$$

with

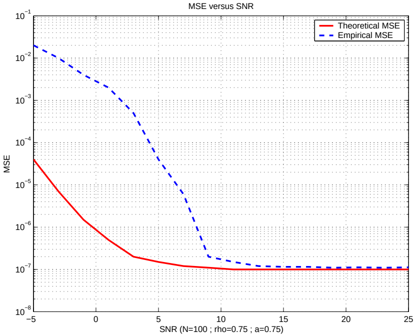
$$\begin{aligned} B &= \sum_{q=0}^{P} (C_P^q)^2 q! \sigma_b^{2q} \\ C &= e^{-1/\sigma_b^2} {}_2F_1(2P+1, 2P+1, 1/\sigma_b^2) \\ D &= \frac{P!}{(2P)!} e^{-1/\sigma_b^2} {}_2F_1(P+1, P+1, 1/\sigma_b^2) \end{aligned}$$

- Similar equations for P-QAM constellation

Numerical illustrations

Set-up :

- $a(n) = s(n) + 0.75s(n-1)$ with $s(n)$ white Gaussian process
- Performance of "weighted periodogram-based estimate" vs. SNR



Questions :

⇝ How far away from Cramer-Rao Bound we are ?

⇝ Irrelevancy of MSE at low SNR (*outliers effect*).

Outliers effect

We focus on the following M-power estimate

$$\hat{f}_N = \frac{1}{M} \arg \max_{\alpha \in]-1/2, 1/2]} \left| \frac{1}{N} \sum_{n=0}^{N-1} y(n)^M e^{-2i\pi\alpha n} \right|^2$$

with

$$y(n)^M = u e^{2i\pi M f_0 n} + e(n)$$

This periodogram is maximizing by proceeding into two steps

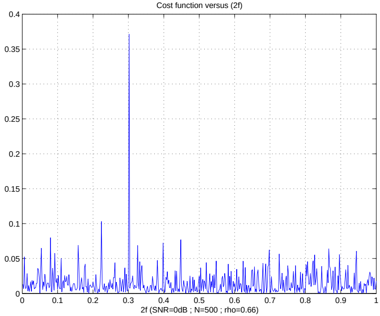
- a "coarse" step detecting the peak
- a "fine" step refining the estimation around the peak

Remark

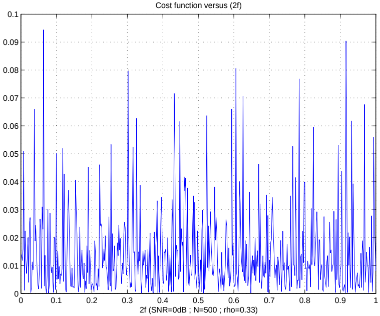
At *low SNR* and/or when *few samples are available*, the coarse step may fail. This leads to the so-called *outliers effect*.

Example

- $a(n)$ is a complex-valued white zero-mean Gaussian process with unit-variance and pseudo-variance $u = \mathbb{E}[a(n)^2]$
  $\rightsquigarrow |u|$ refers to non-circularity rate.
- $SNR = 0\text{dB}$ and $N = 500$



Cost function with $|u| = 2/3$



Cost function with $|u| = 1/3$

Mean Square Error

True MSE

$$\mathrm{MSE} = \frac{p}{12} + (1-p)\mathrm{MSE}_{\mathrm{o.f.}}$$

where

- p is the probability of coarse step failure
- $\mathrm{MSE}_{\mathrm{o.f.}}$ is the standard "outliers effect"-free MSE

Available Results :

1. $\mathrm{MSE}_{\mathrm{o.f.}}$ seen in previous slides
2. p recently derived (Ciblat & Ghogho submitted to TCOM)

Failure probability p (I)

Let Y_k (resp. E_k) be the N-FFT of $y(n)^M$ (resp. $e(n)$)

$$|Y_k| = \begin{cases} |ue^{2i\pi M\phi_0} + E_0| & \text{si} \quad k = 0 \\ |E_k| & \text{si} \quad k \neq 0 \end{cases}, (f_0 = 0)$$

The failure probability may write as follows

$$p = 1 - \text{Pb}(\forall k \neq 0, |Y_k| < |Y_0|) = 1 - \int p_1(x)p_2(x)dx$$

where

$$\begin{aligned} p_1(x) &= \text{Pb}(\forall k \neq 0, |Y_k| < x) \\ &= \int_{-\infty}^{x} \cdots \int_{-\infty}^{x} p_{|Y_1|,\cdots,|Y_{N-1}|}(y_1, \cdots, y_N - 1)dy_1 \cdots dy_{N-1} \\ p_2(x) &= p_{|Y_0|}(x) \end{aligned}$$

$\Rightarrow$ *The distribution of FFT points are needed*

Failure probability p (II)

Constant-amplitude multiplicative noise :

- $a(n) = a, \quad \forall n$
- $M = 1$
- Rife & Boorstyn (*IEEE IT 1974*)

$\rightsquigarrow$ $e(n)$ is white circular Gaussian process

Time-varying multiplicative noise :

- $a(n)$ is white and belongs to an usual constellation

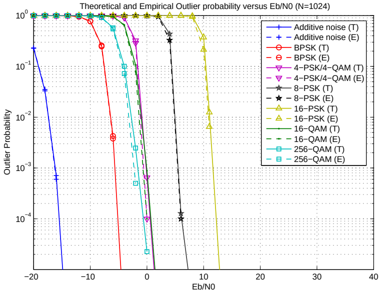
$\rightsquigarrow$ $e(n)$ is white *noncircular* and *non-Gaussian* process

Result

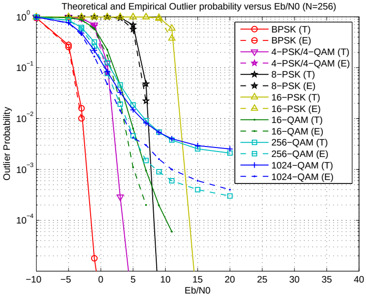
Under Gaussian assumption, a closed-form expression for p can be addressed which strongly depends on

$$\sigma_e^2 = \mathbb{E}[|a(n)|^{2M}] - |\mathbb{E}[a(n)^M]|^2 + \sum_{m=0}^{M-1} (C_M^m)^2 \mathbb{E}[|a(n)|^{2m}] \mathbb{E}[|b(n)|^{2(M-m)}]$$

Simulations : p versus SNR



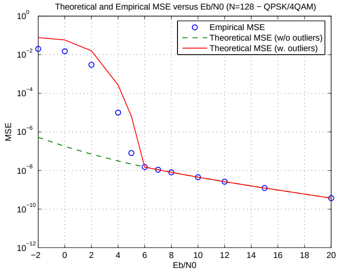
$N = 1024$



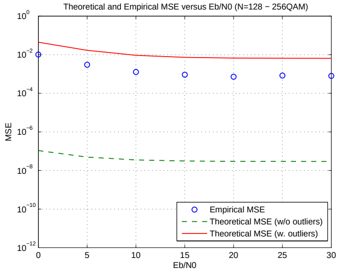
$N = 256$

- p strongly depends on P for P-PSK
- p slightly depends on P for P-QAM
- *Self-noise* for QAM due to $\sigma_e^2 = \mathbb{E}[|a(n)|^8] - |\mathbb{E}[a(n)^4]|^2 \neq 0$ in noiseless case

Simulations : MSE versus SNR



4-QAM and $N = 128$

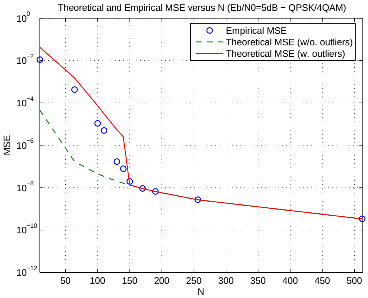


256-QAM and $N = 128$

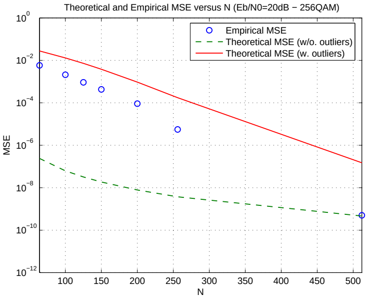
Threshold analysis

- For 4-QAM, $\mathrm{SNR}_{\mathrm{th.}} = 6$dB if $N = 128$
- For P-QAM (with $P > 4$), floor effect for $p \Rightarrow$ no threshold

Simulations : MSE versus N



4-QAM and $E_b/N_0 = 5$dB



256-QAM and $E_b/N_0 = 20$dB

- When N increases, p decreases (without floor effect)
- Any MSE is reachable BUT sometimes with very large N

Remarks

Estimation accuracy

- Data-aided context can improve the performance but outliers effect still exists (Mengali IEEE TCOM 2000)
- Cramer-Rao bound (CRB) with coded scheme is less than CRB without coded scheme (Moeneclaey IEEE COML 2003)
- Turbo-estimation is an appropriate solution (Vandendorpe & al. EURASIP JWCN 2005)

Questions

1. MSE value : is it far away from the lower bound (Cramer-Rao Bound) ?
2. Outliers effect : is it intrinsic to M power estimate or to any estimate ?

Mean Square Error Lower Bounds

Signal Model

$$y(n) = a(n)e^{2i\pi f_0 n} + b(n), \quad n = 0, \ldots, N-1 \Leftrightarrow \mathbf{y} = \mathbf{D}(f_0)\mathbf{a} + \mathbf{b}$$

where

- $\mathbf{y} = [y(0), \cdots, y(N-1)]^{\mathrm{T}}$
- $\mathbf{D}(f_0) = \mathrm{diag}([1, \cdots, e^{2i\pi f_0 (N-1)}])$
- Noise variance assumed to be known (for sake of simplicity)

f_0 : (deterministic) parameter of interest
$\{a(0), \cdots, a(N-1)\}$: parameters of nuisance

Each assumption on the parameters of nuisance (deterministic/random, etc.) leads to ONE Cramer-Rao-type bound

Unconditional CRB

We consider the likelihood for parameters $\{f_0, \mathbf{a}\}$:

$$\Lambda(f,\mathbf{a}) \quad \left(\propto e^{\frac{-\|\mathbf{y}-\mathbf{D}(f)\mathbf{a}\|^2}{2N_0}}\right)$$

$a(n)$ are viewed as real nuisance $\rightsquigarrow$ *stochastic*

Unconditional CRB or True CRB or Stochastic CRB

Unconditional Likelihood is equal to True-Likelihood

$$\Lambda_u(f) = \mathbb{E}_{\mathbf{a}}[\Lambda(f,\mathbf{a})] = \int \Lambda(f,\mathbf{a})p(\mathbf{a})d\mathbf{a}$$

$$\Rightarrow \mathrm{UCRB}(f) = \frac{1}{\mathbb{E}_{\mathbf{y}}\left[\left|\frac{\partial}{\partial f}\ln\Lambda_u(f)\right|^2\right]} = \frac{1}{\mathbb{E}_{\mathbf{y}}\left[\left|\frac{\partial}{\partial f}\ln\mathbb{E}_{\mathbf{a}}[\Lambda(f,\mathbf{a})]\right|^2\right]}$$

$\rightsquigarrow$ Often *untractable*

$\rightsquigarrow$ UCRB mainly analysed by Moeneclaey

$\rightsquigarrow$ Approximation at low SNR ($e^x = 1 + x + x^2/2$ if x small)

Conditional CRB

$a(n)$ are viewed as parameters of interest $\rightsquigarrow$ *deterministic*

Conditional CRB or Deterministic CRB

Conditional Likelihood is equal to Deterministic Likelihood

$$\Lambda_c(f) = \Lambda(f, \hat{\mathbf{a}}_f) \quad \text{where} \quad \frac{\partial \Lambda(f, \mathbf{a})}{\partial \mathbf{a}}_{|\hat{\mathbf{a}}_f} = 0$$

$$\Rightarrow \mathrm{CCRB}(f) = \frac{1}{\mathbb{E}_{\mathbf{y}}\left[\left|\frac{\partial}{\partial f}\ln\Lambda_c(f)\right|^2\right]}$$

Average CCRB or Asymptotic CCRB

$$< \mathrm{CCRB} >(f) = \frac{1}{\mathbb{E}_{\mathbf{y},\mathbf{a}}\left[\left|\frac{\partial}{\partial f}\ln\Lambda_c(f)\right|^2\right]}$$

$\rightsquigarrow$ CCRB not used although CML well spread

$\rightsquigarrow$ CCRB mainly analysed by Stoica and Vazquez

Modified CRB

$a(n)$ are viewed as *known* parameters

Modified CRB

$$\Rightarrow \mathrm{MCRB}(f) = \frac{1}{\mathbb{E}_{\mathbf{y},\mathbf{a}}\left[\left|\frac{\partial}{\partial f}\ln\Lambda(f,\mathbf{a})\right|^2\right]}$$

⇝ Closed-form expressions *tractable*
⇝ MCRB introduced by Mengali
⇝ MCRB very often used in COM/SP community

Gaussian CRB

$a(n)$ are viewed as *Gaussian process*

Gaussian CRB

Gaussian Likelihood

$$\Lambda_g(f) = \mathbb{E}_{\mathbf{a}}[\Lambda(f, \mathbf{a})]$$

where **a** is a Gaussian vector.

$$\Rightarrow \mathrm{GCRB}(f) = \frac{1}{\mathbb{E}_{\mathbf{y}}\left[\left|\frac{\partial}{\partial f}\ln\Lambda_g(f)\right|^2\right]}$$

⇝ Closed-form expressions *tractable*
⇝ Not valid for digital communications but this is still a bound for all the consistent estimates based on data sample covariance matrix
⇝ GCRB developed in SP community (Giannakis, Ghogho, Ciblat)

Bayesian CRB

f_0 is also viewed as stochastic variable with an a priori pdf $p(f)$

Let $\hat{\boldsymbol{\theta}}$ be an unbiased estimate of $\boldsymbol{\theta}_0 = [f_0, \mathbf{a}]$. Then

$$\mathrm{MSE}_{|\mathrm{Bayesian}} = \mathbb{E}_{\mathbf{y},\boldsymbol{\theta}}[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^{\mathrm{T}}] \geq \mathbf{J}^{-1} = \mathrm{BCRB}$$

with

$$\mathbf{J} = \mathbb{E}_{\mathbf{y},\boldsymbol{\theta}}\left[\frac{\partial \ln \Lambda(\mathbf{y}, f, \mathbf{a})}{\partial \boldsymbol{\theta}} \frac{\partial \ln \Lambda(\mathbf{y}, f, \mathbf{a})}{\partial \boldsymbol{\theta}}^{\mathrm{T}}\right]$$

and $\Lambda(\mathbf{y}, f, \mathbf{a})$ is the joint density of $[\mathbf{y}, f, \mathbf{a}]$.

Remarks

No link in the literature between xCRB and BCRB

Bayesian Algorithm

Deterministic approach :

- Optimal unbiased estimate does not always exist (except ML in asymptotic regime)

Stochastic/Bayesian approach :

- Optimal unbiased estimate always exists : the so-called MMSE estimator

$$\hat{\boldsymbol{\theta}} = \mathbb{E}_{\boldsymbol{\theta}|\mathbf{y}}[\boldsymbol{\theta}] = \int \boldsymbol{\theta} p(\boldsymbol{\theta}|\mathbf{y}) d\boldsymbol{\theta}$$

Remarks

- *The MMSE is the mean of the a posteriori density*
- $p(\boldsymbol{\theta})$ must be differentiable
- SP community (Van Trees)

Link between xCRB (I)

All these bounds (except GCRB) lower-bound the mean square error !

Results

$$\mathrm{UCRB} \geq \mathrm{MCRB}$$

and

$$<\mathrm{CCRB}> \geq \mathrm{MCRB}$$

- *At high SNR :* $\mathrm{UCRB} = \mathrm{MCRB}$ (if the values of the parameters of nuisance belongs to a discrete set)
- *For large samples :* $\mathrm{CCRB} \overset{N\to\infty}{\longrightarrow} <\mathrm{CCRB}>$ (ergodism)
- *Under Gaussian assumption :* $\mathrm{UCRB} = \mathrm{GCRB}$

$\rightsquigarrow$ MCRB usually too optimistic
$\rightsquigarrow$ GCRB unable to take into acocunt high order information

Link between xCRB (II)

Application to blind synchronization

$a(n)$ belongs to a constellation and thus to a discrete set

- At high SNR,

$$\mathrm{UCRB} = \mathrm{MCRB}$$

  MCRB is of interest in digital communications

- At low SNR,

$$\mathrm{UCRB} \gg \mathrm{MCRB}$$

  Let M be the order of non-circularity (Moeneclaey IEEE COML 2001).

$$\mathrm{UCRB} = \mathcal{O}(1/\mathrm{SNR}^M) \quad \text{and} \quad \mathrm{MCRB} = \mathcal{O}(1/\mathrm{SNR})$$

- GCRB likely useful for BPSK but not for other constellations

Example (I)

Harmonic retrieval where $a(n)$ is complex-valued white (discrete) process with $\mathbb{E}[|a(n)|^2] = 1$ and $\mathbb{E}[a(n)^2] = u$.

$$\mathrm{MCRB} = \frac{3\sigma^2}{2\pi^2 N^3} \quad \text{and} \quad \mathrm{GCRB} = \frac{3\left[(1-|u|^2) + 2\sigma^2 + \sigma^4\right]}{4\pi^2 |u|^2 N^3}$$

$$\mathrm{UCRB}_{|\text{low SNR}} = \frac{3\sigma^4}{4\pi^2 |u|^2 N^3}$$

and

$$\mathrm{UCRB}_{|\text{high SNR}} = \mathrm{MCRB} = \frac{3\sigma^2}{2\pi^2 N^3}$$

$\rightsquigarrow$ MCRB quite relevant BUT does not depend on non-circularity rate.
$\rightsquigarrow$ At low SNR, second-order noncircularity leads to GCRB=UCRB
$\rightsquigarrow$ If $|u| \neq 1$, floor error with GCRB not with MCRB and UCRB

Example (II)

We consider $u = 1$ (e.g. $a(n) \in \mathrm{BPSK}$)

$$\mathrm{MCRB} = \frac{3\sigma^2}{2\pi^2 N^3} \quad \text{and} \quad \mathrm{GCRB} = \frac{3\sigma^2}{2\pi^2 N^3} + \frac{3\sigma^4}{4\pi^2 N^3}$$

and

$$\mathrm{UCRB}_{|\text{high SNR}} = \frac{3\sigma^2}{2\pi^2 N^3} \quad \text{and} \quad \mathrm{UCRB}_{|\text{low SNR}} = \frac{3\sigma^4}{4\pi^2 N^3}$$

At high SNR

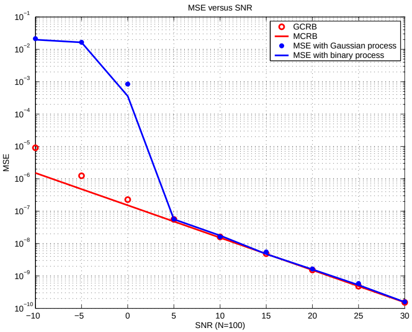
$$\mathrm{UCRB} = \mathrm{MCRB} = \mathrm{GCRB}$$

At low SNR

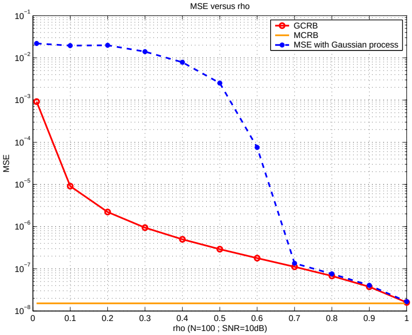
$$\mathrm{UCRB} = \mathrm{GCRB}$$

$\rightsquigarrow$ GCRB relevant for BPSK

Example (III)



MSE versus SNR



MSE versus $|u|$

$\rightsquigarrow$ For BPSK signal, we are lucky (GCRB$\approx$MCRB) !

Asymptotic Gaussian CRB (I)

⇝ Several works for obtaining asymptotic (large sample) expressions for GCRB.

- Circular case : Ghogho 2001 (based on Whittle's theorem)
- Real-valued case : Ghogho 1999
- Non-circular case : Ciblat 2003 (large Toeplitz matrices)

	White	Colored
Circular	∞	$O(1/N)$
Real-valued	$O(1/N^3)$ No floor error Reached by Square Power	$O(1/N^3)$ No floor effect
Non-circular	$O(1/N^3)$ No floor error Reached by Square Power	$O(1/N^3)$ Floor effect

Asymptotic Gaussian CRB (II)

Second-order noncircular case : Ciblat (EURASIP SP 2005)

$$\mathrm{GCRB} \sim \frac{3}{4\pi^2 \xi N^3} \quad \text{with} \quad \xi = \int_0^1 \frac{c_a(e^{2i\pi f})\overline{c_a(e^{-2i\pi f})}}{\mathcal{X}(e^{2i\pi f})} df$$

$$\mathrm{MSE} \sim \frac{3\eta}{4\pi^2 N^3} \quad \text{with} \quad \eta = \frac{\int_0^1 |c_a(e^{2i\pi f})|^2 \mathcal{X}(e^{2i\pi f}) df}{\left(\int_0^1 |c_a(e^{2i\pi f})|^2 df\right)^2}$$

One can proven that (Cauchy-Schwartz inequality)

$$\mathrm{GCRB} = \mathrm{MSE} \text{ iff } a(n) \text{ white process}$$

Other types of bound

Remark

xCRB unable to predict and analyze the outliers effect

Solutions

Introducing other tighter lower bounds

- Deterministic approach
  - $\rightsquigarrow$ Battacharyya bound
  - $\rightsquigarrow$ Barankin bound
- Stochastic approach
  - $\rightsquigarrow$ Ziv-Zakai bound
  - $\rightsquigarrow$ Weiss-Weinstein bound

Battacharyya bound (I)

Review on CRB : consider the vector $\mathbf{z}$,

$$\mathbf{z} = \left[\begin{array}{c} \boldsymbol{\theta} - \boldsymbol{\theta}_0 \\ \frac{\partial \ln(p(\mathbf{y}|\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} \end{array} \right]$$

By construction, $\mathbb{E}[\mathbf{z}\mathbf{z}^{\mathrm{T}}]$ is nonnegative matrix. This implies that

$$\left[\begin{array}{cc} \mathrm{MSE} & 1 \\ 1 & \mathrm{FIM} \end{array} \right] \geq 0$$

and

$$\mathrm{MSE} \geq \mathrm{FIM}^{-1} = \mathrm{CRB}$$

Battacharyya bound (II)

consider the vector $\mathbf{z}_N$,

$$\mathbf{z}_N = \begin{bmatrix} \boldsymbol{\theta} - \boldsymbol{\theta}_0 \\ \frac{\partial \ln(p(\mathbf{y}|\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} \\ \vdots \\ \frac{\partial^N \ln(p(\mathbf{y}|\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}^N} \end{bmatrix}$$

Once again $\mathbb{E}[\mathbf{z}_N \mathbf{z}_N^{\mathrm{T}}]$ is nonnegative matrix and this leads to

$$\mathrm{MSE} \geq \mathrm{BaB} = \mathrm{CRB} + \text{one positive term}$$

Barankin bound (I)

We consider "test-points" $\mathcal{E}_n = [\boldsymbol{\theta}^{(1)} - \boldsymbol{\theta}_0, \ldots, \boldsymbol{\theta}^{(n)} - \boldsymbol{\theta}_0]$.
Furthermore $\mathbf{B}_n = (B_{k,l})_{1 \leq k,l \leq n}$ is the following $n \times n$ matrix

$$B_{k,l} = \mathbb{E}_{\mathbf{y}} \left[\frac{p(\mathbf{y}|\boldsymbol{\theta}^{(k)}) p(\mathbf{y}|\boldsymbol{\theta}^{(l)})}{p(\mathbf{y}|\boldsymbol{\theta}_0)^2} \right]$$

Definition

Barankin bound of order $n \rightsquigarrow \mathrm{BB}_n(\boldsymbol{\theta}_0) = \sup_{\mathcal{E}_n} \underbrace{\mathcal{E}_n (\mathbf{B}_n(\mathcal{E}_n) - \mathbf{1}_n \mathbf{1}_n^{\mathrm{T}})^{-1} \mathcal{E}_n^{\mathrm{T}}}_{S_n(\mathcal{E}_n)}$

with $\mathbf{1}_n = \mathrm{ones}(n, 1)$

$\rightsquigarrow$ *MSE of any unbiased estimator is greater than any* BB_n
$\rightsquigarrow$ *As* $n \to \infty$, BB_∞ *becomes even the tightest lower bound*

Barankin bound (II)

- BB_1 used (one test-point)
- Main task : closed-form expression for matrix **B**

Remark

$$\mathrm{CRB} = \lim_{\mathcal{E}\to 0} S_1(\mathcal{E})$$

- ⇝ CRB inspects the likelihood only around the true point
- ⇝ CRB and BaB unable to observe outliers

$$\mathrm{BB} = \sup_{\mathcal{E}} S_1(\mathcal{E})$$

- ⇝ BB scans all the research interval
- ⇝ BB takes into account outliers effect in lower bound.

- Pure harmonic retrieval : Knockaert in 1997
- Circular multiplicative noise : Messer in 1992 for DOA issue

Derivations

- Let $y(n) = ae^{2i\pi f_0 n} + b(n) \rightsquigarrow$ Information in *mean* of $y(n)$.
- Let $y(n) = a(n)e^{2i\pi f_0 n} + b(n) \rightsquigarrow$ Information in *variance* of $y(n)$.

Closed-form expression (Ciblat EURASIP SP 2005)

$$B_{k,l} = \begin{cases} \frac{1}{\sqrt{\det(\mathbf{Q}_{k,l})}} & \text{if } \mathbf{Q}_{k,l} > 0 \\ +\infty & \text{otherwise} \end{cases},$$

where

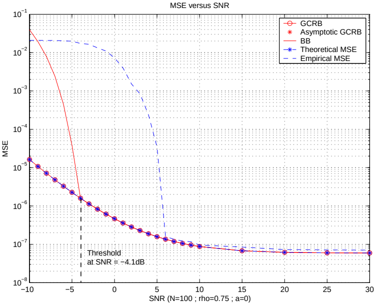
$$\mathbf{Q}_{k,l} = (\widetilde{\mathbf{R}}_{f(k)}^{-1} + \widetilde{\mathbf{R}}_{f(l)}^{-1})\widetilde{\mathbf{R}}_{f_0} - \mathbf{Id}_{2N}$$

and

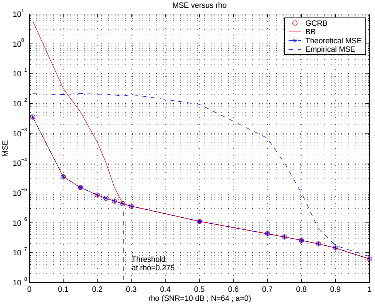
$$\widetilde{\mathbf{R}}_f = \left[\begin{array}{cc} \frac{\mathbf{E}[\mathbf{y}_N \mathbf{y}_N^{\mathrm{H}}]}{\mathbf{E}[\mathbf{y}_N \mathbf{y}_N^{\mathrm{T}}]} & \frac{\mathbf{E}[\mathbf{y}_N \mathbf{y}_N^{\mathrm{T}}]}{\mathbf{E}[\mathbf{y}_N \mathbf{y}_N^{\mathrm{H}}]} \end{array} \right].$$

Numerical illustrations

$a(n)$ white Gaussian process with unit-variance and $\mathbb{E}[a(n)^2] = u$.



MSE versus SNR



MSE versus u

- Threshold analysis : $\mathrm{BB} = \max(\mathrm{GCRB}, S(1/4))$
- Important gap between BB and standard Square-Power estimate

Ziv-Zakai bound (I)

- Bayesian bound : random parameter
- Two classes :
  - Hölder inequality :
    - Bayesian Battacharyya
    - Bobrovsky-Zakai (1976)
    - Weiss-Weinstein bound (1985)
  - Kotelnikov inequality :
    - Ziv-Zakai (1969)
    - Bellini-Tartara (1975)

State-of-the-Art

Ziv-Zakai bound (ZZB) derivations

1. bearing estimation and additive noise (Bell IEEE IT 1997)
2. time-delay estimation (Weiss IEEE SP 1983)

Ziv-Zakai bound (II)

Definition

The mean square error (MSE) for φ_1 is bounded by

$$\text{MSE} \geq \int_0^{\infty} h_1 \left(\max_{h_0} g(h_0, h_1) \right) dh_1.$$

where

- $g(h_0, h_1) = \int \min(p(\varphi), p(\varphi + \mathbf{h})) P_e(\varphi, \varphi + \mathbf{h}) d\varphi$
- $\varphi = [\phi_0, f_0]$ and $\mathbf{h} = [h_0, h_1]$
- $p(.)$ is the *a priori* density function of φ
- $P_e(\varphi, \varphi + \mathbf{h})$ is the error probability when the optimal detector decides between the following two equally likely hypotheses

$$\begin{cases} H_0 : & y(n) = a(n)e^{2i\pi(\phi_0 + f_0 n)} + b(n) \\ H_1 : & y(n) = a(n)e^{2i\pi((\phi_0 + h_0) + (f_0 + h_1)n)} + b(n) \end{cases}$$

$\rightsquigarrow$ *Detection theory with multiplicative noise*

Derivations

Result

$$\mathrm{MSE}_1 \geq \int_0^{1/2} (1/2 - h_1) h_1 (\max_{h_0} (1/2 - h_0) P_e(h_0, h_1)) dh_1$$

with

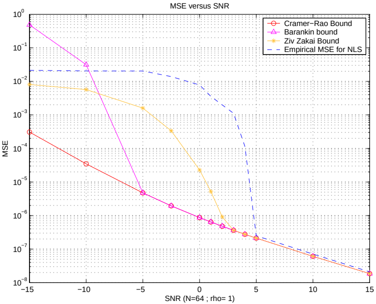
$$P_e(h_0, h_1) = \frac{(\theta_1/\theta_2)^{\alpha_1}}{\alpha_1} B(\alpha_1, \alpha_2) {}_2F_1(\alpha_1 + \alpha_2, \alpha_1, \alpha_1 + 1 - \theta_1/\theta_2)$$

where

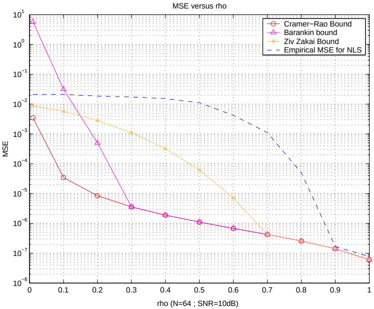
- $B(\alpha_1, \alpha_2) = \Gamma(\alpha_1 + \alpha_2)/\Gamma(\alpha_1)$ is called either the Euler's first integral or the Beta function
- ${}_2F_1(\alpha, \beta, \gamma; x)$ is the hyper-geometric function
- Closed-form expressions of θ_1, θ_2, α_1, α_2 depend on $\widetilde{\mathbf{R}}_{\mathbf{h}}$ and $\widetilde{\mathbf{R}}_0$

Numerical illustrations

$a(n)$ white Gaussian process with unit-variance and $\mathbb{E}[a(n)^2] = u$.



MSE versus SNR



MSE versus u

- Small gap between ZZB and standard Square-Power estimate

Bibliography

- *BCRB and CRB* : H.L. Van Trees,"Detection, Estimation, and Modulation Theory", Part 1, 1968.
- *xCRB* :
  - G. Vazquez, "Non-Data-Aided Digital Synchronization" in the book "Signal Processing Advances in Communications" edited by G.Giannakis et al., 2000.
  - M. Moeneclaey, "On the True and the Modified CRB for the estimation of a scalar paramter in the presence of nuisance parameter", IEEE TCOM, Nov. 1998.
- *Asymptotic analysis* :
  - P. Ciblat, "Asymptotic analysis of blind cyclic correlation based symbol rate estimation", IEEE IT, Jul. 2002.
  - M. Ghogho, "Frequency estimation in the presence of Doppler spread : performance analysis", IEEE SP, Apr. 2001.
- *ZZB* : J. Ziv, "Some lower bounds on Signal Processing, IEEE IT, May 1969.
- *Outliers effect* : D.C. Rife, "Single-tone parameter estimation from discrete-time observations", IEEE IT, Sep. 1974.