The Information Limit in Clutter: CRLB in the Presence of False Measurements and the ML-PDA Estimator

Yaakov Bar-Shalom, Distinguished IEEE AESS Lecturer

ESP (Estimation and Signal Processing) Lab
University of Connecticut, ECE Dept.
Box U-2157, Storrs, CT 06269-2157

Phone: 860-486-4823 Fax: 860-486-5585
E-mail: ybs@ee.uconn.edu

May 27, 2011

Fitting a Straight Line Through n Points

Theorem 0

One can pass a straight line through any n points on a sheet of paper.

Proof

Use a thick enough pencil.

Fitting a Straight Line Through n Points

Theorem 0

One can pass a straight line through any n points on a sheet of paper.

Proof

Use a thick enough pencil.

Model

Assume each point is a noisy observation (with additive zero-mean white noise with known variance) of the true points which lie on a straight line

$$z(i) = ay(i) + b + w(i) \qquad i = 1, \ldots, n \tag{1}$$

with the noises

$$E[w(i)] = 0 \qquad E[w(i)w(k)] = \sigma^2 \delta_{ik} \tag{2}$$

This is the simplest linear regression problem with known error statistics.

Solution

The solution is obtained with *LS* and is the same as with the *ML* criterion under the (additional) Gaussian assumption on the noises.

Fitting a Straight Line Through n Points (cont'd)

Model

Assume each point is a noisy observation (with additive zero-mean white noise with known variance) of the true points which lie on a straight line

$$z(i) = ay(i) + b + w(i) \qquad i = 1, \ldots, n \tag{1}$$

with the noises

$$E[w(i)] = 0 \qquad E[w(i)w(k)] = \sigma^2 \delta_{ik} \tag{2}$$

This is the simplest linear regression problem with known error statistics.

Solution

The solution is obtained with *LS* and is the same as with the *ML* criterion under the (additional) Gaussian assumption on the noises.

The Cramer-Rao Lower Bound

For an unbiased estimate $\hat{x}$ the error covariance is bounded from below

$$E[(x - \hat{x})(x - \hat{x})'] \geq J^{-1} \tag{3}$$

where J is the *Fisher information matrix*.

Under the Gaussian assumption, the MLE of the parameter vector $[a, b]$ in the above problem is efficient — it meets the CRLB.

There is an implicit assumption in the above problem formulation: No measurement origin uncertainty, i.e., for each $y(i)$ there is a single $z(i)$ that obeys the (linear-Gaussian) model.

Mathematically this amounts to

(1) No extraneous measurements (no false alarms or clutter): $P_{FA} = 0$

(2) Correct measurement always available (target detected): $P_D = 1$.

The Cramer-Rao Lower Bound

For an unbiased estimate $\hat{x}$ the error covariance is bounded from below

$$E[(x - \hat{x})(x - \hat{x})'] \geq J^{-1} \tag{3}$$

where J is the *Fisher information matrix*.

Under the Gaussian assumption, the MLE of the parameter vector $[a,\, b]$ in the above problem is efficient — it meets the CRLB.

There is an implicit assumption in the above problem formulation: No measurement origin uncertainty, i.e., for each $y(i)$ there is a single $z(i)$ that obeys the (linear-Gaussian) model.

Mathematically this amounts to

(1) No extraneous measurements (no false alarms or clutter): $P_{FA} = 0$

(2) Correct measurement always available (target detected): $P_D = 1$.

The Cramer-Rao Lower Bound

For an unbiased estimate $\hat{x}$ the error covariance is bounded from below

$$E[(x - \hat{x})(x - \hat{x})'] \geq J^{-1} \tag{3}$$

where J is the *Fisher information matrix*.

Under the Gaussian assumption, the MLE of the parameter vector $[a, b]$ in the above problem is efficient — it meets the CRLB.

There is an implicit assumption in the above problem formulation: No measurement origin uncertainty, i.e., for each $y(i)$ there is a single $z(i)$ that obeys the (linear-Gaussian) model.

Mathematically this amounts to

(1) No extraneous measurements (no false alarms or clutter): $P_{FA} = 0$

(2) Correct measurement always available (target detected): $P_D = 1$.

The Cramer-Rao Lower Bound

For an unbiased estimate $\hat{x}$ the error covariance is bounded from below

$$E[(x - \hat{x})(x - \hat{x})'] \geq J^{-1} \quad (3)$$

where J is the *Fisher information matrix*.

Under the Gaussian assumption, the MLE of the parameter vector $[a, b]$ in the above problem is efficient — it meets the CRLB.

There is an implicit assumption in the above problem formulation: No measurement origin uncertainty, i.e., for each $y(i)$ there is a single $z(i)$ that obeys the (linear-Gaussian) model.

Mathematically this amounts to

(1) No extraneous measurements (no false alarms or clutter): $P_{FA} = 0$

(2) Correct measurement always available (target detected): $P_D = 1$.

In remote sensing problems (with radar, sonar or EO sensors), for low observable (LO) targets — with low SNR — the two assumptions

$$P_{FA} = 0 \tag{4}$$

$$P_D = 1 \tag{5}$$

do not hold.

The detection threshold must be low enough to detect low SNR targets and this causes unavoidable FAs (clutter) that look like the correct measurements.

The Real World

In remote sensing problems (with radar, sonar or EO sensors), for low observable (LO) targets — with low SNR — the two assumptions

$$P_{FA} = 0 \quad (4)$$

$$P_D = 1 \quad (5)$$

do not hold.

The detection threshold must be low enough to detect low SNR targets and this causes unavoidable FAs (clutter) that look like the correct measurements.

Problem 1
Estimate the parameter vector x of the (possibly nonlinear) relationship

$$z_j(i) = \begin{cases} f(x, t_i) + w_j(i) & \text{if origin is "target"} \\ u_j(i) & \text{if origin is "false"} \end{cases} \tag{6}$$

where $i = 1, \ldots, n$, $j = 1, \ldots, m_i$, and

- i is the index of time t_i
- m_i is the number of measurements at t_i with at most one of them being from the "target", with probability P_D (and with a Gaussian error w)
- the remaining measurements (a priori, a random number) are false (clutter), with their values u uniformly distributed in the measurement space
- all the random variables and detection events are mutually independent.

Problem 1

Estimate the parameter vector x of the (possibly nonlinear) relationship

$$z_j(i) = \begin{cases} f(x, t_i) + w_j(i) & \text{if origin is “target"} \\ u_j(i) & \text{if origin is “false"} \end{cases} \tag{6}$$

where $i = 1, \ldots, n$, $j = 1, \ldots, m_i$, and

- i is the index of time t_i
- m_i is the number of measurements at t_i with at most one of them being from the “target", with probability P_D (and with a Gaussian error w)
- the remaining measurements (a priori, a random number) are false (clutter), with their values u uniformly distributed in the measurement space
- all the random variables and detection events are mutually independent.

Generalized LS Fitting with $P_D < 1$ and $P_{FA} > 0$ (cont'd)

Problem 2

- Find the CRLB for this problem accounting for the measurement origin uncertainty
- Determine if the estimator is statistically efficient — it meets the CRLB, i.e., it extracts all the information from the data
- Find the lowest SNR for which one can have efficiency.

Generalized LS Fitting with $P_D < 1$ and $P_{FA} > 0$ (cont'd)

Problem 2

- Find the CRLB for this problem accounting for the measurement origin uncertainty
- Determine if the estimator is statistically efficient — it meets the CRLB, i.e., it extracts all the information from the data
- Find the lowest SNR for which one can have efficiency.

Generalized LS Fitting with $P_D < 1$ and $P_{FA} > 0$ (cont'd)

Problem 2

- Find the CRLB for this problem accounting for the measurement origin uncertainty
- Determine if the estimator is statistically efficient — it meets the CRLB, i.e., it extracts all the information from the data
- Find the lowest SNR for which one can have efficiency.

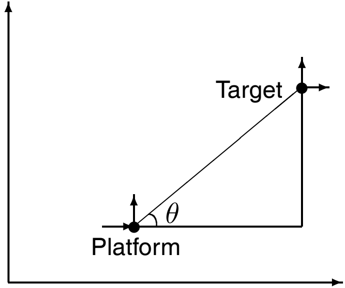
Motivation — Target Motion Analysis with Passive Sonar

TMA: Estimation of a target's initial position and its constant velocity from bearings-only measurements (and possibly Doppler) corrupted by noise, in the presence of clutter/FA.

Bearings-only TMA scenario

Motivation — Target Motion Analysis with Passive Sonar

TMA: Estimation of a target's initial position and its constant velocity from bearings-only measurements (and possibly Doppler) corrupted by noise, in the presence of clutter/FA.



Bearings-only TMA scenario

Applicability

This constant velocity motion model is widely used in underwater passive target tracking since it is a good approximation for actual scenarios, at least for a while.

TMA is an example of parameter estimation — initial condition estimation — for an object with deterministic motion.

Another example: exoatmospheric motion of ballistic objects — this motion is fully determined by the initial position and velocity.

For motion affected by randomness, state estimation is needed. Similar results are available for state estimation in presence of measurement origin uncertainty.

Applicability

This constant velocity motion model is widely used in underwater passive target tracking since it is a good approximation for actual scenarios, at least for a while.

TMA is an example of parameter estimation — initial condition estimation — for an object with deterministic motion.

Another example: exoatmospheric motion of ballistic objects — this motion is fully determined by the initial position and velocity.

For motion affected by randomness, state estimation is needed. Similar results are available for state estimation in presence of measurement origin uncertainty.

Applicability

This constant velocity motion model is widely used in underwater passive target tracking since it is a good approximation for actual scenarios, at least for a while.

TMA is an example of parameter estimation — initial condition estimation — for an object with deterministic motion.

Another example: exoatmospheric motion of ballistic objects — this motion is fully determined by the initial position and velocity.

For motion affected by randomness, state estimation is needed. Similar results are available for state estimation in presence of measurement origin uncertainty.

Applicability

This constant velocity motion model is widely used in underwater passive target tracking since it is a good approximation for actual scenarios, at least for a while.

TMA is an example of parameter estimation — initial condition estimation — for an object with deterministic motion.

Another example: exoatmospheric motion of ballistic objects — this motion is fully determined by the initial position and velocity.

For motion affected by randomness, state estimation is needed. Similar results are available for state estimation in presence of measurement origin uncertainty.

Motivation for Using Feature Information

In addition to kinematic measurements (position, Doppler) the estimation algorithm can also use feature measurements to help reduce the data (measurement) association uncertainty.

In all systems Amplitude Information (AI) is used *implicitly* to determine whether there is a valid measurement — thresholding (a minimal use).

If fully taken advantage of, i.e., with statistical models (e.g., Swerling fluctuation models) AI can be used in the estimation process itself to enhance the performance:

- Tracking under very low SNR conditions becomes more accurate
- The algorithm with AI is efficient (meets the CRLB) for *lower SNR than the one without AI*.

Motivation for Using Feature Information

In addition to kinematic measurements (position, Doppler) the estimation algorithm can also use feature measurements to help reduce the data (measurement) association uncertainty.

In all systems Amplitude Information (AI) is used *implicitly* to determine whether there is a valid measurement — thresholding (a minimal use).

If fully taken advantage of, i.e., with statistical models (e.g., Swerling fluctuation models) AI can be used in the estimation process itself to enhance the performance:

- Tracking under very low SNR conditions becomes more accurate
- The algorithm with AI is efficient (meets the CRLB) for *lower SNR than the one without AI.*

Motivation for Using Feature Information

In addition to kinematic measurements (position, Doppler) the estimation algorithm can also use feature measurements to help reduce the data (measurement) association uncertainty.

In all systems Amplitude Information (AI) is used *implicitly* to determine whether there is a valid measurement — thresholding (a minimal use).

If fully taken advantage of, i.e., with statistical models (e.g., Swerling fluctuation models) AI can be used in the estimation process itself to enhance the performance:

- Tracking under very low SNR conditions becomes more accurate
- The algorithm with AI is efficient (meets the CRLB) for *lower SNR than the one without AI*.

The target parameter is the 5-dimensional vector (includes emitted frquency)

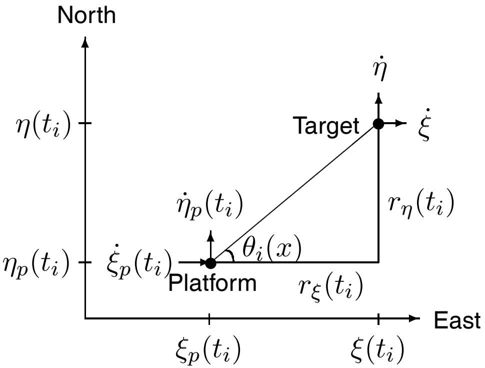
$$x \triangleq \left[\begin{array}{ccccc} \xi(t_0) & \eta(t_0) & \dot{\xi} & \dot{\eta} & \gamma \end{array} \right]'$$

The scenario at t_i

Target Model for Passive Narrowband Sonar

The target parameter is the 5-dimensional vector (includes emitted frquency)

$$x \triangleq \left[\begin{array}{ccccc} \xi(t_0) & \eta(t_0) & \dot{\xi} & \dot{\eta} & \gamma \end{array} \right]'$$



The scenario at t_i

The (noisy) measurements are bearing (DOA), frequency and amplitude

$$z_j(i) \triangleq \begin{bmatrix} \beta_{ij} & f_{ij} & a_{ij} \end{bmatrix}', \qquad j = 1, \ldots, m_i; \quad i = 1, \ldots, n$$

i.e., n scans, with m_i measurements in scan i.

The following assumptions are made:

- A target-originated measurement is received by the sensor only once during a scan with known probability P_D, independently across scans
- Target-originated angle and frequency measurements are corrupted by independent additive zero-mean white Gaussian noise sequences
- False measurements occur according to a spatial Poisson process with known spatial density λ
- The fluctuating amplitudes are independent from scan to scan and have a known pdf for the target and the false measurements (Swerling model with known average SNR = d).

Measurement Model for Passive Narrowband Sonar

The (noisy) measurements are bearing (DOA), frequency and amplitude

$$z_j(i) \triangleq \begin{bmatrix} \beta_{ij} & f_{ij} & a_{ij} \end{bmatrix}', \qquad j = 1, \ldots, m_i; \quad i = 1, \ldots, n$$

i.e., n scans, with m_i measurements in scan i.

The following assumptions are made:

- A target-originated measurement is received by the sensor only once during a scan with known probability P_D, independently across scans
- Target-originated angle and frequency measurements are corrupted by independent additive zero-mean white Gaussian noise sequences
- False measurements occur according to a spatial Poisson process with known spatial density λ
- The fluctuating amplitudes are independent from scan to scan and have a known pdf for the target and the false measurements (Swerling model with known average SNR = d).

The ML-PDA Parameter Estimation Approach

The Maximum Likelihood (ML) estimator combined with the Probabilistic Data Association (PDA) technique is obtained as follows:

- The exact joint pdf of the entire set of measurements is obtained using the PDA approach — (*without* approximations)
- The total log-likelihood ratio (LLR) of the target parameter vector is derived
- The maximization of the LLR is done numerically using a quasi-Newton (variable metric) method to yield the maximum likelihood estimate $\hat{x}$ of the parameter x
- The LLR may have many local maxima — to overcome this, a multi-pass approach is used.

The ML-PDA Parameter Estimation Approach

The Maximum Likelihood (ML) estimator combined with the Probabilistic Data Association (PDA) technique is obtained as follows:

- The exact joint pdf of the entire set of measurements is obtained using the PDA approach — (*without* approximations)
- The total log-likelihood ratio (LLR) of the target parameter vector is derived
- The maximization of the LLR is done numerically using a quasi-Newton (variable metric) method to yield the maximum likelihood estimate $\hat{x}$ of the parameter x
- The LLR may have many local maxima — to overcome this, a multi-pass approach is used.

The ML-PDA Parameter Estimation Approach

The Maximum Likelihood (ML) estimator combined with the Probabilistic Data Association (PDA) technique is obtained as follows:

- The exact joint pdf of the entire set of measurements is obtained using the PDA approach — (*without* approximations)
- The total log-likelihood ratio (LLR) of the target parameter vector is derived
- The maximization of the LLR is done numerically using a quasi-Newton (variable metric) method to yield the maximum likelihood estimate $\hat{x}$ of the parameter x
- The LLR may have many local maxima — to overcome this, a multi-pass approach is used.

The ML-PDA Parameter Estimation Approach

The Maximum Likelihood (ML) estimator combined with the Probabilistic Data Association (PDA) technique is obtained as follows:

- The exact joint pdf of the entire set of measurements is obtained using the PDA approach — (*without* approximations)
- The total log-likelihood ratio (LLR) of the target parameter vector is derived
- The maximization of the LLR is done numerically using a quasi-Newton (variable metric) method to yield the maximum likelihood estimate $\hat{x}$ of the parameter x
- The LLR may have many local maxima — to overcome this, a multi-pass approach is used.

The ML-PDA Parameter Estimation Approach

The Maximum Likelihood (ML) estimator combined with the Probabilistic Data Association (PDA) technique is obtained as follows:

- The exact joint pdf of the entire set of measurements is obtained using the PDA approach — (*without* approximations)
- The total log-likelihood ratio (LLR) of the target parameter vector is derived
- The maximization of the LLR is done numerically using a quasi-Newton (variable metric) method to yield the maximum likelihood estimate $\hat{x}$ of the parameter x
- The LLR may have many local maxima — to overcome this, a multi-pass approach is used.

The ML-PDA Estimator

If there are m_i detections at t_i we have the following mutually exclusive and exhaustive events (PDA approach):

$$\varepsilon_j(i) \triangleq \begin{cases} \{\text{measurement } z_j(i) \text{ is from the target}\} & j = 1, \ldots, m_i \\ \{\text{all measurements are false}\} & j = 0 \end{cases}$$

The pdf of the measurements in scan i conditioned on the above events can be written as

$$p\left[Z(i)|\varepsilon_j(i), x\right] = \begin{cases} V^{1-m_i} p(\beta_{ij}) p(f_{ij}) \rho_{ij} \prod_{k=1}^{m_i} p_0^{\tau}(a_{ik}) & j = 1, \ldots, m_i \\ V^{-m_i} \prod_{k=1}^{m_i} p_0^{\tau}(a_{ik}) & j = 0 \end{cases}$$

where V is the volume of the surveillance region and

$$\rho_{ij} = p_1^{\tau}(a_{ij}) / p_0^{\tau}(a_{ij})$$

is the amplitude likelihood ratio — a ratio of two truncated Rayleigh pdfs.

The ML-PDA Estimator

If there are m_i detections at t_i we have the following mutually exclusive and exhaustive events (PDA approach):

$$\varepsilon_j(i) \triangleq \begin{cases} \{\text{measurement } z_j(i) \text{ is from the target}\} & j = 1, \ldots, m_i \\ \{\text{all measurements are false}\} & j = 0 \end{cases}$$

The pdf of the measurements in scan i conditioned on the above events can be written as

$$p\left[Z(i)|\varepsilon_j(i), x\right] = \begin{cases} V^{1-m_i} p(\beta_{ij}) p(f_{ij}) \rho_{ij} \prod_{k=1}^{m_i} p_0^\tau(a_{ik}) & j = 1, \ldots, m_i \\ V^{-m_i} \prod_{k=1}^{m_i} p_0^\tau(a_{ik}) & j = 0 \end{cases}$$

where V is the volume of the surveillance region and

$$\rho_{ij} = p_1^\tau(a_{ij}) / p_0^\tau(a_{ij})$$

is the amplitude likelihood ratio — a ratio of two truncated Rayleigh pdfs.

The ML-PDA Estimator (cont'd)

Using the total probability theorem one can write the likelihood function (LF) of the set of measurements at t_i as a uniform-Gaussian-Rayleigh mixture.

After some lengthy manipulations, the total log-likelihood ratio (LLR) of the parameter x based on the entire data set Z^n without knowing the origin of the measurements is obtained as

$$\ell[Z^n, x] = \sum_{i=1}^{n} \ell_i[Z(i), x] = \sum_{i=1}^{n} \ln \Bigg[(1 - P_D) + \frac{P_D}{\lambda} \sum_{j=1}^{m_i} \frac{\rho_{ij}}{2\pi\sigma_\theta\sigma_\gamma} \cdot \exp\left(-\frac{1}{2}\left[\frac{\beta_{ij} - \theta_i(x)}{\sigma_\theta}\right]^2 - \frac{1}{2}\left[\frac{f_{ij} - \gamma_i(x)}{\sigma_\gamma}\right]^2 \right) \Bigg]$$

The LLR ℓ is preferable to the LF since it is a (physically) dimensionless quantity.

The maximum likelihood estimate is obtained by finding the parameter $\hat{x}$ that maximizes the above total log-likelihood ratio.

Using the total probability theorem one can write the likelihood function (LF) of the set of measurements at t_i as a uniform-Gaussian-Rayleigh mixture.

After some lengthy manipulations, the total log-likelihood ratio (LLR) of the parameter x based on the entire data set Z^n without knowing the origin of the measurements is obtained as

$$\ell[Z^n,x] = \sum_{i=1}^{n} \ell_i[Z(i),x] = \sum_{i=1}^{n} \ln \Bigg[(1-P_D) + \frac{P_D}{\lambda} \sum_{j=1}^{m_i} \frac{\rho_{ij}}{2\pi\sigma_\theta\sigma_\gamma} \cdot \exp\left(-\frac{1}{2}\left[\frac{\beta_{ij}-\theta_i(x)}{\sigma_\theta}\right]^2 - \frac{1}{2}\left[\frac{f_{ij}-\gamma_i(x)}{\sigma_\gamma}\right]^2 \right)\Bigg]$$

The LLR ℓ is preferable to the LF since it is a (physically) dimensionless quantity.

The maximum likelihood estimate is obtained by finding the parameter $\hat{x}$ that maximizes the above total log-likelihood ratio.

The ML-PDA Estimator (cont'd)

Using the total probability theorem one can write the likelihood function (LF) of the set of measurements at t_i as a uniform-Gaussian-Rayleigh mixture.

After some lengthy manipulations, the total log-likelihood ratio (LLR) of the parameter x based on the entire data set Z^n without knowing the origin of the measurements is obtained as

$$\begin{aligned}\ell[Z^n, x] &= \sum_{i=1}^{n} \ell_i[Z(i), x] = \sum_{i=1}^{n} \ln \Bigg[(1 - P_D) + \frac{P_D}{\lambda} \sum_{j=1}^{m_i} \frac{\rho_{ij}}{2\pi\sigma_\theta\sigma_\gamma} \\ &\cdot \exp\left(-\frac{1}{2}\left[\frac{\beta_{ij} - \theta_i(x)}{\sigma_\theta}\right]^2 - \frac{1}{2}\left[\frac{f_{ij} - \gamma_i(x)}{\sigma_\gamma}\right]^2 \right)\Bigg]\end{aligned}$$

The LLR ℓ is preferable to the LF since it is a (physically) dimensionless quantity.

The maximum likelihood estimate is obtained by finding the parameter $\hat{x}$ that maximizes the above total log-likelihood ratio.

Using the total probability theorem one can write the likelihood function (LF) of the set of measurements at t_i as a uniform-Gaussian-Rayleigh mixture.

After some lengthy manipulations, the total log-likelihood ratio (LLR) of the parameter x based on the entire data set Z^n without knowing the origin of the measurements is obtained as

$$\ell[Z^n, x] = \sum_{i=1}^{n} \ell_i[Z(i), x] = \sum_{i=1}^{n} \ln \left[(1 - P_D) + \frac{P_D}{\lambda} \sum_{j=1}^{m_i} \frac{\rho_{ij}}{2\pi\sigma_\theta\sigma_\gamma} \cdot \exp\left(-\frac{1}{2}\left[\frac{\beta_{ij} - \theta_i(x)}{\sigma_\theta}\right]^2 - \frac{1}{2}\left[\frac{f_{ij} - \gamma_i(x)}{\sigma_\gamma}\right]^2 \right) \right]$$

The LLR ℓ is preferable to the LF since it is a (physically) dimensionless quantity.

The maximum likelihood estimate is obtained by finding the parameter $\hat{x}$ that maximizes the above total log-likelihood ratio.

Cramer-Rao Lower Bound in the Presence of False Measurements

Theorem 1. For an unbiased estimate, from measurements with Gaussian errors in the presence of Poisson distributed false measurements, the estimation error covariance is bounded from below as follows

$$E\left\{(x-\hat{x})(x-\hat{x})'\right\} \geq J^{-1}$$

with the Fisher information matrix (FIM) *in clutter*, J, given by

$$J = q_2 J_0$$

where

- J_0 is the FIM in the absence of clutter ($P_{FA} = 0$, $P_D = 1$)
- q_2, **a scalar**, is the **information reduction factor (IRF)** that accounts for the loss of information due to false measurements ($P_{FA} > 0$) and imperfect target detection ($P_D < 1$)

Cramer-Rao Lower Bound in the Presence of False Measurements

Theorem 1. For an unbiased estimate, from measurements with Gaussian errors in the presence of Poisson distributed false measurements, the estimation error covariance is bounded from below as follows

$$E\{(x - \hat{x})(x - \hat{x})'\} \geq J^{-1}$$

with the Fisher information matrix (FIM) *in clutter*, J, given by

$$J = q_2 J_0$$

where

- J_0 is the FIM in the absence of clutter ($P_{FA} = 0$, $P_D = 1$)
- q_2, **a scalar**, is the **information reduction factor (IRF)** that accounts for the loss of information due to false measurements ($P_{FA} > 0$) and imperfect target detection ($P_D < 1$)

The Information Reduction Factor

The information reduction factor q_2 involves a multi-dimensional integration (evaluated numerically).

For narrowband passive sonar (bearing and frequency measurements) with amplitude information (modeled by Rayleigh fluctuations with SNR=d), the factor q_2 is given by

$$q_2\left(P_D, \lambda v_g, g\right) = \frac{1}{1+d} \sum_{m=1}^{\infty} \frac{2^{m-1} \mu_f(m-1)}{\left(g^2 P_{FA}\right)^{m-1}} I_2\left(m, P_D, g\right)$$

where

- $I_2\left(m, P_D, g\right)$ is an m-fold integral
- μ_f is the Poisson pmf of the number of false alarms
- g is the “number of sigmas” of the “validation gate” (region in the measurement space where the true measurement will be with high probability if detected; typically, $g = 5$).

The Information Reduction Factor

The information reduction factor q_2 involves a multi-dimensional integration (evaluated numerically).

For narrowband passive sonar (bearing and frequency measurements) with amplitude information (modeled by Rayleigh fluctuations with SNR=d), the factor q_2 is given by

$$q_2\left(P_D, \lambda v_g, g\right) = \frac{1}{1+d} \sum_{m=1}^{\infty} \frac{2^{m-1} \mu_f (m-1)}{\left(g^2 P_{FA}\right)^{m-1}} I_2\left(m, P_D, g\right)$$

where

- $I_2\left(m, P_D, g\right)$ is an m-fold integral
- μ_f is the Poisson pmf of the number of false alarms
- g is the “number of sigmas” of the “validation gate” (region in the measurement space where the true measurement will be with high probability if detected; typically, $g = 5$).

Track Acceptance Test

Each track estimate is tested whether it can be used as an acceptable track — we should reject noise-only tracks.

This is also necessary due to the multimodal nature of the log-likelihood ratio since the numerical maximization might converge to a wrong peak.

The test is formulated as a Neyman-Pearson hypothesis testing problem where the false track rejection power of the test is maximized for a given true track miss probability:

- The test statistic is ℓ (the LLR) for which the first two moments can be calculated under the "target present" hypothesis (H_1) with a certain SNR.
- Since the statistic is the sum of (typically) tens of independent random variables, the test threshold τ_ℓ is obtained assuming the test statistic obeys the CLT, i.e., it is Gaussian with the above moments
- Gaussian model for $p(\ell)$: For 5% miss (false rejection) probability it yields 3–6 misses per 100 runs.

Track Acceptance Test

Each track estimate is tested whether it can be used as an acceptable track — we should reject noise-only tracks.

This is also necessary due to the multimodal nature of the log-likelihood ratio since the numerical maximization might converge to a wrong peak.

The test is formulated as a Neyman-Pearson hypothesis testing problem where the false track rejection power of the test is maximized for a given true track miss probability:

- The test statistic is ℓ (the LLR) for which the first two moments can be calculated under the "target present" hypothesis (H_1) with a certain SNR.
- Since the statistic is the sum of (typically) tens of independent random variables, the test threshold τ_ℓ is obtained assuming the test statistic obeys the CLT, i.e., it is Gaussian with the above moments
- Gaussian model for $p(\ell)$: For 5% miss (false rejection) probability it yields 3–6 misses per 100 runs.

False Track Acceptance

We also want to evaluate the false track rejection power of the test.

The false track acceptance probability $P\{\ell > \tau_\ell | H_0\}$ under the "target absent" hypothesis (H_0)should be (very) small — the tail of $p(\ell|H_0)$.

The first model used was a Gaussian for $p(\ell|H_0)$ — very inaccurate for the tail.

Under H_0 the LLR surface has thousands of peaks, from which we want the pdf of the highest (the global maximum) — the tail of the tail.

Approach

- Use extreme value theory, specifically a Gumbel distribution;
- Use ML estimation to obtain the parameters of the Gumbel distribution from a set of samples;
- Results: $\approx 10^{-3}$ verified by simulations (Gaussian model yields $\approx 10^{-5}$).

False Track Acceptance

We also want to evaluate the false track rejection power of the test.

The false track acceptance probability $P\{\ell > \tau_\ell | H_0\}$ under the “target absent” hypothesis (H_0)should be (very) small — the tail of $p(\ell|H_0)$.

The first model used was a Gaussian for $p(\ell|H_0)$ — very inaccurate for the tail.

Under H_0 the LLR surface has thousands of peaks, from which we want the pdf of the highest (the global maximum) — the tail of the tail.

Approach

- Use extreme value theory, specifically a Gumbel distribution;
- Use ML estimation to obtain the parameters of the Gumbel distribution from a set of samples;
- Results: $\approx 10^{-3}$ verified by simulations (Gaussian model yields $\approx 10^{-5}$).

False Track Acceptance

We also want to evaluate the false track rejection power of the test.

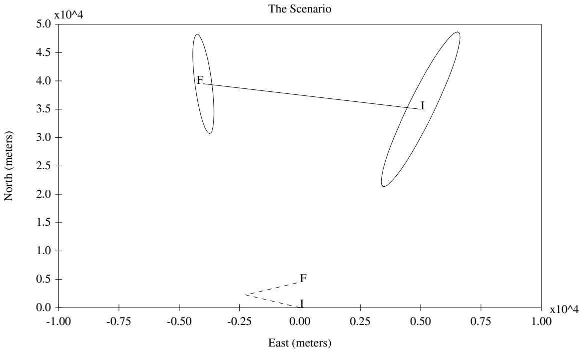
The false track acceptance probability $P\{\ell > \tau_\ell | H_0\}$ under the "target absent" hypothesis (H_0)should be (very) small — the tail of $p(\ell|H_0)$.

The first model used was a Gaussian for $p(\ell|H_0)$ — very inaccurate for the tail.

Under H_0 the LLR surface has thousands of peaks, from which we want the pdf of the highest (the global maximum) — the tail of the tail.

Approach

- Use extreme value theory, specifically a Gumbel distribution;
- Use ML estimation to obtain the parameters of the Gumbel distribution from a set of samples;
- Results: $\approx 10^{-3}$ verified by simulations (Gaussian model yields $\approx 10^{-5}$).



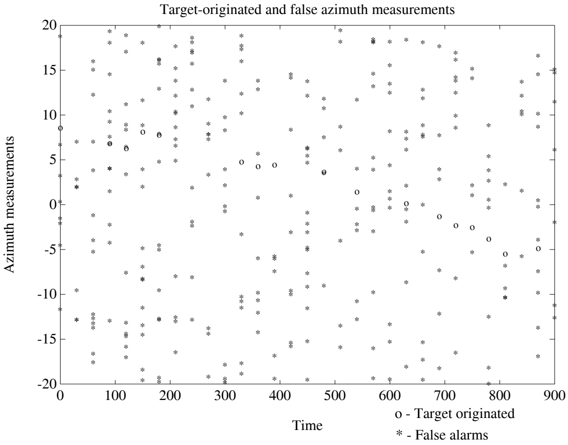
Trajectories of the target and platform with CRLB based uncertainties for initial and final position

Simulation Scenario (cont'd)

Scenario Parameters:

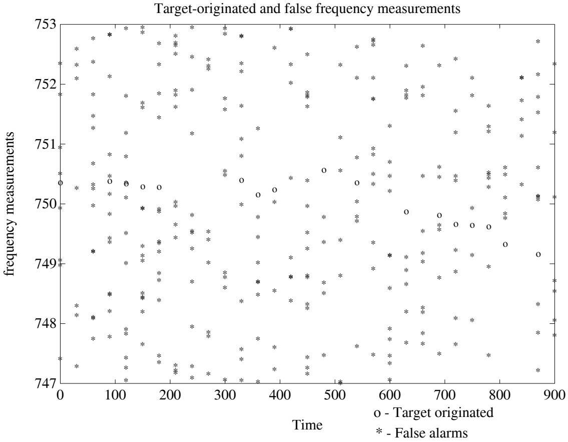
- The SNR in a resolution cell ($3^{\mathrm{o}} \times 0.25$Hz) is 6dB (0dB/Hz) — average power of target signal at detector is 4$\times$ the power of noise
- The probability of target signal detection is 0.5 per scan
- These values give, on average, 10 false alarms in the entire surveillance region per scan.

The Measurements



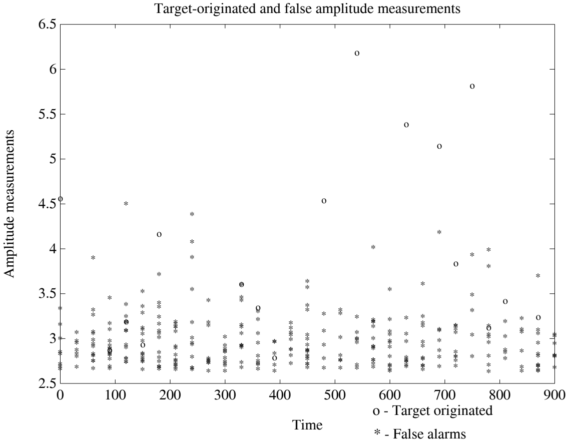
Azimuth (bearing) measurements

The Measurements (cont'd)



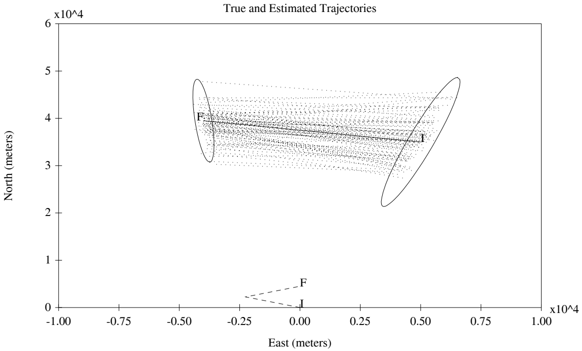
Frequency measurements

The Measurements (cont'd)



Amplitude measurements

Estimated Tracks



Estimated tracks from 100 Monte Carlo runs

Estimated Tracks (cont'd)

- In 94–97 runs out of 100 the estimated trajectory endpoints fall in the corresponding 95% uncertainty ellipses based on the CRLB — s.d. from 100 MC runs is

$$\sqrt{0.95 \cdot 0.05/100} \approx 2\%$$

- The normalized estimation error squared is 5.37, which lies within the 95% probability region $[4.4, 5.64]$ — the estimator is **efficient**
- The track acceptance test was carried out with a theoretical miss probability of 5% (In 1000 runs no false track was ever accepted)
- Coarse search — grid — is used to start the fine search.

Estimated Tracks (cont'd)

- In 94–97 runs out of 100 the estimated trajectory endpoints fall in the corresponding 95% uncertainty ellipses based on the CRLB — s.d. from 100 MC runs is

$$\sqrt{0.95 \cdot 0.05/100} \approx 2\%$$

- The normalized estimation error squared is 5.37, which lies within the 95% probability region $[4.4, 5.64]$ — the estimator is **efficient**
- The track acceptance test was carried out with a theoretical miss probability of 5% (In 1000 runs no false track was ever accepted)
- Coarse search — grid — is used to start the fine search.

Estimated Tracks (cont'd)

- In 94–97 runs out of 100 the estimated trajectory endpoints fall in the corresponding 95% uncertainty ellipses based on the CRLB — s.d. from 100 MC runs is

$$\sqrt{0.95 \cdot 0.05/100} \approx 2\%$$

- The normalized estimation error squared is 5.37, which lies within the 95% probability region $[4.4, 5.64]$ — the estimator is **efficient**
- The track acceptance test was carried out with a theoretical miss probability of 5% (In 1000 runs no false track was ever accepted)
- Coarse search — grid — is used to start the fine search.

Estimated Tracks (cont'd)

- In 94–97 runs out of 100 the estimated trajectory endpoints fall in the corresponding 95% uncertainty ellipses based on the CRLB — s.d. from 100 MC runs is

$$\sqrt{0.95 \cdot 0.05/100} \approx 2\%$$

- The normalized estimation error squared is 5.37, which lies within the 95% probability region $[4.4, 5.64]$ — the estimator is **efficient**
- The track acceptance test was carried out with a theoretical miss probability of 5% (In 1000 runs no false track was ever accepted)
- Coarse search — grid — is used to start the fine search.

Accuracy of Estimates with AI

Unit	x_{true}	x_{init}	$\overline{x}$	σ_{CRLB}	$\hat{\sigma}$
m	5000	–12000 to 12000	4991	667	821
m	35000	49000 to 50000	35423	5576	5588
m/s	–10	–16 to 5	–9.96	0.85	0.96
m/s	5	–4 to 9	4.87	4.73	4.99
Hz	750	747 to 751	749.52	2.371	2.531

Results of 100 Monte Carlo runs with AI ($\mathrm{SNR} = 6.1\mathrm{dB}$)

Accuracy of Estimates without AI

Unit	x_{true}	x_{init}	$\overline{x}$	σ_{CRLB}	$\hat{\sigma}$
m	5000	–12000 to 12000	6395	689	8653
m	35000	49000 to 50000	41370	5759	23094
m/s	–10	–17 to 5	–9.86	0.88	1.21
m/s	5	–5 to 10	3.55	4.89	7.36
Hz	750	747 to 751	749.03	2.448	2.751

Results of 100 Monte Carlo runs without AI ($\text{SNR} = 6.1\text{dB}$)

Computation Times

K	with AI accepted tracks	with AI average time taken (s)	without AI accepted tracks	without AI average time taken (s)
3	95	3.19	71	3.57
2	91	2.68	60	2.63
1	82	1.97	18	1.69

Performances of estimators for different number of passes (K)

Conclusions

- The CRLB in clutter is characterized by the **scalar information reduction factor (IRF)**.
- The use of AI gives a moderate increase in the information reduction factor (reduction in the CRLB), which is significant under low SNR conditions (13% at 6dB cell SNR — 0.45 vs. 0.4).
- The cell SNR limit down to which the estimator with AI is efficient is 6dB. This is 3–4dB lower than the limit without AI.
- The percentage of accepted tracks with AI is substantially higher than that without AI (95% vs. 70%).
- This technique is (probably) the most powerful for detecting LO tracks for nonmaneuvering targets (via parameter estimation).
- **Recent result:** In nonlinear dynamic systems the **Bayesian CRLB for state estimation** (Tichavsky-Muravchik-Nehorai), the effect of the clutter is quantified by a similar **(scalar) IRF** (Zhang-Willett-BarShalom).

Conclusions

- The CRLB in clutter is characterized by the **scalar information reduction factor (IRF)**.
- The use of AI gives a moderate increase in the information reduction factor (reduction in the CRLB), which is significant under low SNR conditions (13% at 6dB cell SNR — 0.45 vs. 0.4).
- The cell SNR limit down to which the estimator with AI is efficient is 6dB. This is 3–4dB lower than the limit without AI.
- The percentage of accepted tracks with AI is substantially higher than that without AI (95% vs. 70%).
- This technique is (probably) the most powerful for detecting LO tracks for nonmaneuvering targets (via parameter estimation).
- **Recent result:** In nonlinear dynamic systems the **Bayesian CRLB for state estimation** (Tichavsky-Muravchik-Nehorai), the effect of the clutter is quantified by a similar **(scalar) IRF** (Zhang-Willett-BarShalom).

Conclusions

- The CRLB in clutter is characterized by the **scalar information reduction factor (IRF)**.
- The use of AI gives a moderate increase in the information reduction factor (reduction in the CRLB), which is significant under low SNR conditions (13% at 6dB cell SNR — 0.45 vs. 0.4).
- The cell SNR limit down to which the estimator with AI is efficient is 6dB. This is 3–4dB lower than the limit without AI.
- The percentage of accepted tracks with AI is substantially higher than that without AI (95% vs. 70%).
- This technique is (probably) the most powerful for detecting LO tracks for nonmaneuvering targets (via parameter estimation).
- **Recent result:** In nonlinear dynamic systems the **Bayesian CRLB for state estimation** (Tichavsky-Muravchik-Nehorai), the effect of the clutter is quantified by a similar **(scalar) IRF** (Zhang-Willett-BarShalom).

Conclusions

- The CRLB in clutter is characterized by the **scalar information reduction factor (IRF)**.
- The use of AI gives a moderate increase in the information reduction factor (reduction in the CRLB), which is significant under low SNR conditions (13% at 6dB cell SNR — 0.45 vs. 0.4).
- The cell SNR limit down to which the estimator with AI is efficient is 6dB. This is 3–4dB lower than the limit without AI.
- The percentage of accepted tracks with AI is substantially higher than that without AI (95% vs. 70%).
- This technique is (probably) the most powerful for detecting LO tracks for nonmaneuvering targets (via parameter estimation).
- **Recent result:** In nonlinear dynamic systems the **Bayesian CRLB for state estimation** (Tichavsky-Muravchik-Nehorai), the effect of the clutter is quantified by a similar **(scalar) IRF** (Zhang-Willett-BarShalom).

Conclusions

- The CRLB in clutter is characterized by the **scalar information reduction factor (IRF)**.
- The use of AI gives a moderate increase in the information reduction factor (reduction in the CRLB), which is significant under low SNR conditions (13% at 6dB cell SNR — 0.45 vs. 0.4).
- The cell SNR limit down to which the estimator with AI is efficient is 6dB. This is 3–4dB lower than the limit without AI.
- The percentage of accepted tracks with AI is substantially higher than that without AI (95% vs. 70%).
- This technique is (probably) the most powerful for detecting LO tracks for nonmaneuvering targets (via parameter estimation).
- **Recent result:** In nonlinear dynamic systems the **Bayesian CRLB for state estimation** (Tichavsky-Muravchik-Nehorai), the effect of the clutter is quantified by a similar **(scalar) IRF** (Zhang-Willett-BarShalom).

Conclusions

- The CRLB in clutter is characterized by the **scalar information reduction factor (IRF)**.
- The use of AI gives a moderate increase in the information reduction factor (reduction in the CRLB), which is significant under low SNR conditions (13% at 6dB cell SNR — 0.45 vs. 0.4).
- The cell SNR limit down to which the estimator with AI is efficient is 6dB. This is 3–4dB lower than the limit without AI.
- The percentage of accepted tracks with AI is substantially higher than that without AI (95% vs. 70%).
- This technique is (probably) the most powerful for detecting LO tracks for nonmaneuvering targets (via parameter estimation).
- **Recent result:** In nonlinear dynamic systems the **Bayesian CRLB for state estimation** (Tichavsky-Muravchik-Nehorai), the effect of the clutter is quantified by a similar **(scalar) IRF** (Zhang-Willett-BarShalom).