

Notes on Stein's Ambiguity Function Processing

Chris Hall

August 20, 2022

1 Introduction

These notes are intended to document my thoughts on Seymour Stein's 1981 paper titled "Algorithms for ambiguity function processing."

Section two of these notes contains the missing derivation for the Time of Arrival estimator variance stated in Stein's Algorithms for Ambiguity Function processing.

The remaining sections deal with Stein's definition of effective SNR, and how effective SNR changes depending on the detection method employed. To correctly calculate the effective SNR, the signal power and noise power of the received signals must be correctly propagated through every step of a detection system.

Three detection methods will be analyzed.

$$r = \sum s_1(t)s_2^*(t) \quad (1)$$

$$r = |\sum s_1(t)s_2^*(t)| \quad (2)$$

$$r = \sum s_1(t)s_1(t-1)s_2^*(t)s_2^*(t-1) \quad (3)$$

2 Derivation of Stein's Time of Arrival Accuracy and the Maximum Likelihood Estimator

Most of this derivation is pieced together from chapters 6 and 10 of Robert McDonough's "Detection of Signals in Noise" second edition.

We start by defining a noisy received signal:

$$r(t) = s_r(t|x) + n(t) \quad (4)$$

For the case where $n(t)$ is additive white Gaussian noise, the probability distribution of $r(t)$ is

$$p_i(r) = \frac{(2\pi)^{\frac{m}{2}}}{\sqrt{\text{Det}(C)}} \exp \left[-\frac{1}{2}(r - s_i)^T C^{-1}(r - s_i) \right] \quad (5)$$

C is the covariance matrix for the random noise component, which will be diagonal: $\frac{2\Delta t}{N_0}$. The sample period is Δt .

$$p(r) = \left(\frac{\Delta t}{\pi N_0} \right)^{\frac{m}{2}} \exp \left[-\frac{\Delta t}{N_0} \sum (r - s_r)^2 \right] \quad (6)$$

Where m is the number of samples in the signal and the summation is taken over the m samples. To get the continuous time signal model, we will take the limit as sample period, Δt , goes to zero.

$$p(r) = \exp \left[- \left(\frac{1}{N_0} \right) \int_0^T [r(t) - s_r(t)]^2 dt \right] \quad (7)$$

Here x represents any unknown parameters in the signal, s . These unknown parameters could be amplitude, phase, or time offset. In our case we are interested in examining signals with an unknown time offset.

$$\bar{r}(t) = E[r(t)] = s_r(t|x) \quad (8)$$

The likelihood function will be

$$p(r|x) = \alpha \exp \left\{ - \left(\frac{1}{N_0} \right) \int_0^T [r(t) - s_r(t|x)]^2 dt \right\} \quad (9)$$

The maximum likelihood estimate comes from setting the derivative of the log likelihood function equal to zero.

$$\frac{\partial [\ln p(r|x)]}{\partial x} = 0 \quad (10)$$

$$\frac{2}{N_0} \int_0^T [r(t) - s_r(t|x)] \left[\frac{\partial s_r(t|x)}{\partial x} \right] dt = 0 \quad (11)$$

Note that $r(t) - s_r(t|x) = n(t)$ and $E[n(t)n(t')] = R_n(t - t') = (N_0/2)\delta(t - t')$.

$$s_r(t|x) = s_r(t|\tau) = s_r(t - \tau) \quad (12)$$

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} [r(t) - s_r(t - \tau)] \left[\frac{\partial s_r(t - \tau)}{\partial x} \right] dt = 0 \quad (13)$$

The second term reduces to zero.

$$- \int_{\frac{T}{2}}^{\frac{T}{2}} s_r(t - \tau) \left[\frac{\partial s_r(t - \tau)}{\partial \tau} \right] dt \quad (14)$$

$$= - \int_{-\frac{T}{2}-\tau}^{\frac{T}{2}-\tau} s_r(u) \left[\frac{\partial s_r(u)}{\partial u} \right] du \quad (15)$$

$$= \frac{1}{2} \int_{-\frac{T}{2}-\tau}^{\frac{T}{2}-\tau} \left[\frac{\partial s_r^2(u)}{\partial u} \right] du \quad (16)$$

$$= \frac{1}{2} [s_r^2(T/2 - \tau) - s_r^2(-T/2 - \tau)] = \frac{1}{2} [s_r^2(T/2) - s_r^2(-T/2)] = 0 \quad (17)$$

The maximum likelihood estimate, $\hat{\tau}$, is the value of τ for which

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} r(t) \left[\frac{\partial s_r(t - \tau)}{\partial \tau} \right] dt = 0 \quad (18)$$

2.1 Fisher Information and Estimator Variance

The Cramer Rao lower bound is a lower bound on the variance of an estimator, given by:

$$\text{Cov}(\hat{x}(y)) \geq \left[I + \frac{\partial b(x)}{\partial x} \right]^T F^{-1}(x) \left[I + \frac{\partial b(x)}{\partial x} \right] \quad (19)$$

Where F is the Fisher information matrix. For an unbiased estimator, the bias term $b(x)$ will be zero, and the equation simplifies to

$$\text{Cov}(\hat{x}(y)) \geq F^{-1}(x) \quad (20)$$

In our case the Fisher information matrix consists of only one element since there is only one unknown variable to be estimated, τ . The Fisher information matrix is calculated as

$$F(x) = E \left[\frac{\partial \ln[p(y|x)]}{\partial x} \right]^T \left[\frac{\partial \ln[p(y|x)]}{\partial x} \right] \quad (21)$$

$$\frac{\partial \ln[p(y|x)]}{\partial x} = \int_{-\frac{T}{2}}^{\frac{T}{2}} [r(t) - s_r(t)] \left[\frac{\partial s_r(t - \tau)}{\partial \tau} \right] dt \quad (22)$$

Note that $r(t) - s_r(t|x) = n(t)$ and $E[n(t)n(t')] = R_n(t - t') = \frac{N_0}{2} \delta(t - t')$.

The Fisher information matrix has elements

$$F_{ij} = \left(\frac{2}{N_0} \right)^2 \int_{-\frac{T}{2}}^{\frac{T}{2}} \int_{-\frac{T}{2}}^{\frac{T}{2}} \left(\frac{N_0}{2} \right) \delta(t - t') \left[\frac{\partial s_r(t|x)}{\partial x_i} \right] \left[\frac{\partial s_r(t|x)}{\partial x_j} \right] dt dt' \quad (23)$$

$$F_{ij} = \frac{2}{N_0} \int_{-\frac{T}{2}}^{\frac{T}{2}} \left[\frac{\partial s_r(t|x)}{\partial x_i} \right] \left[\frac{\partial s_r(t|x)}{\partial x_j} \right] dt \quad (24)$$

There is only one unknown, τ , that we are interested in estimating, so our Fisher information matrix will only have one element, $\tau = x_i = x_j$. Now we can substitute τ for x since we are trying to estimate the time delay.

The Cramer Rao Lower Bound is

$$\text{Var}(\hat{t} - \tau) \geq \left[\frac{2}{N_0} \int_{-\frac{T}{2}}^{\frac{T}{2}} \left[\frac{\partial s_r(t - \tau)}{\partial \tau} \right]^2 dt \right]^{-1} \quad (25)$$

$$= \left[\frac{2}{N_0} \int_{-\frac{T}{2}}^{\frac{T}{2}} \left[\frac{\partial s_r(t)}{\partial t} \right]^2 dt \right]^{-1} \quad (26)$$

Applying Parseval's theorem, we get

$$\text{Var}(t - \tau) \geq \left[\frac{2}{N_0} \int_{-\infty}^{\infty} \omega^2 |S_r(i\omega)|^2 df \right]^{-1} \quad (27)$$

Received signal energy is

$$E_r = \int_{-\frac{T}{2}}^{\frac{T}{2}} s_r^2(t) dt = \int_{-\infty}^{\infty} |S_r(i\omega)|^2 df \quad (28)$$

$$\beta^2 = \frac{1}{E_r} \int_{-\infty}^{\infty} \omega^2 |S_r(i\omega)|^2 df \quad (29)$$

$$\text{Var}(\hat{\tau} - \tau) \geq \frac{1}{(2E_r/N_0)\beta^2} \quad (30)$$

The noise power is given as $\sigma_n^2 = \frac{N_0 B}{2}$. Signal power is $P_r = \frac{E_r}{T}$. SNR will then be $\gamma = \frac{2E_r}{N_0 B T}$

$$\boxed{\text{Var}(\hat{\tau} - \tau) \geq \frac{1}{\gamma B T \beta^2}} \quad (31)$$

3 Matched Filter Detection

Stein defines the output of a correlation at a time lag, $\tau = 0$, as

$$r(t) = \sum s_1(t) s_2^*(t) \quad (32)$$

We want to determine the SNR of the correlated signal. First let's start with something simple: a single sample. Note this assumption does not affect our final results. For sequences of a generalized length we would substitute $E_s = \sum s^2$ for s^2 in the derivations outlined below. The correlation then becomes

$$y = (s_1 + n_1)(s_2^* + n_2) \quad (33)$$

$$y = s_1 s_2^* + s_1 n_2 + s_2 n_1 + n_1 n_2 \quad (34)$$

The first term is the signal term and the remaining three terms are noise terms. The SNR will be given as

$$\gamma = \frac{|s_1 s_2|^2}{\sigma_{n1}^2 |s_1|^2 + \sigma_{n2}^2 |s_2|^2 + \sigma_1^2 \sigma_2^2} \quad (35)$$

The third term in the denominator is not obvious. I'm claiming that the product of two random variables has a variance equal to the product of the variances of the two random variables.

$$\text{var}(n_1 n_2) = E[n_1^2 n_2^2] - E[n_1 n_2]^2 \quad (36)$$

If both n_1 and n_2 have a mean of zero then $E[n_1 n_2] = 0$.

$$\text{var}(n_1 n_2) = E[n_1^2 n_2^2] = E[n_1^2] E[n_2^2] = \sigma_1^2 \sigma_2^2 \quad (37)$$

Returning to the original problem, we can invert the fraction and decompose it into three terms.

$$\frac{1}{\gamma} = \frac{\sigma_{n1}^2 |s_1|^2 + \sigma_{n2}^2 |s_2|^2 + \sigma_1^2 \sigma_2^2}{|s_1|^2 |s_2|^2} \quad (38)$$

$$\frac{1}{\gamma} = \frac{\sigma_1^2 |s_1|^2}{|s_1|^2 |s_2|^2} + \frac{\sigma_2^2 |s_2|^2}{|s_1|^2 |s_2|^2} + \frac{\sigma_1^2 \sigma_2^2}{|s_1|^2 |s_2|^2} \quad (39)$$

$$\frac{1}{\gamma} = \frac{\sigma_1^2}{|s_2|^2} + \frac{\sigma_2^2}{|s_1|^2} + \frac{\sigma_1^2 \sigma_2^2}{|s_1|^2 |s_2|^2} \quad (40)$$

$$\boxed{\frac{1}{\gamma} = \frac{1}{\gamma_1} + \frac{1}{\gamma_2} + \frac{1}{\gamma_1\gamma_2}} \quad (41)$$

If $s_1(t) = \alpha s_2(t)$ then

$$r = \sum s_1(t)s_2^*(t) \quad (42)$$

will be purely real. The noise will still be complex, but only the real component of the noise will degrade our autocorrelation measurement. The real component of complex Gaussian noise contains half of the power of the complex Gaussian noise. This affects our effective SNR.

$$\boxed{\frac{1}{\gamma} = \frac{1}{2} \left(\frac{1}{\gamma_1} + \frac{1}{\gamma_2} + \frac{1}{\gamma_1\gamma_2} \right)} \quad (43)$$

This one half factor should only be used if we are dealing with complex signals and we know the phase relationship between them. If the original signals are real, we get no gain from looking at only the real component, and the one half factor should not be used. If we have signals with an unknown phase relationship, and our detection function involves taking a magnitude, then we can not benefit from coherent detection and the one half factor should not be used.

4 Non-Coherent Matched Filter Detection

$$r(t) = \sum |s_1(t)s_2(t)| \quad (44)$$

$$\gamma = \frac{|s_1s_2|^2}{\text{var}(|s_1n_2 + s_2n_1 + n_1n_2|)} \quad (45)$$

The first two terms in the denominator are Rayleigh distributed with variance

$$\text{var}(|s_1n_2|) = \frac{4-\pi}{2} |s_1|^2 \sigma_{n_2}^2 \quad (46)$$

$$\text{var}(|s_1n_2 + s_2n_1|) = \frac{4-\pi}{2} (|s_1|^2 \sigma_{n_2}^2 + |s_2|^2 \sigma_{n_1}^2) \quad (47)$$

For a minute let's claim that the third term in the denominator is insignificant compared to the other two, due to some assumption about the relative powers of n_1 and n_2 .

$$\gamma = \frac{|s_1|^2 |s_2|^2}{\left(\frac{4-\pi}{2}\right) (|s_1|^2 \sigma_{n_2}^2 + |s_2|^2 \sigma_{n_1}^2)} \quad (48)$$

$$\gamma = \frac{2|s_1|^2 |s_2|^2}{(4-\pi) (|s_1|^2 \sigma_{n_2}^2 + |s_2|^2 \sigma_{n_1}^2)} \quad (49)$$

$$\frac{1}{\gamma} = \frac{(4-\pi)(|s_1|^2 \sigma_{n_2}^2)}{2|s_1|^2 |s_2|^2} + \frac{(4-\pi)(|s_2|^2 \sigma_{n_1}^2)}{2|s_1|^2 |s_2|^2} \quad (50)$$

$$\frac{1}{\gamma} = \frac{(4-\pi)\sigma_{n_2}^2}{|s_2|^2} + \frac{(4-\pi)\sigma_{n_1}^2}{2|s_1|^2} \quad (51)$$

$$\boxed{\frac{1}{\gamma} = \frac{4 - \pi}{2} \left(\frac{1}{\gamma_1} + \frac{1}{\gamma_2} \right)} \quad (52)$$

The solution above is valid for the case where one of our correlated signals is clean, or has infinite SNR. Now let's go back to the case where the third term is not negligible.

$$\gamma = \frac{|s_1 s_2|^2}{\text{Var}(|s_1 n_2 + s_2 n_1 + n_1 n_2|)} \quad (53)$$

Using the triangle inequality

$$\gamma \geq \frac{|s_1 s_2|^2}{\text{Var}(|s_1 n_2 + s_2 n_1|) + \text{Var}(|n_1 n_2|)} \quad (54)$$

$$\gamma \geq \frac{|s_1 s_2|^2}{\left(\frac{4-\pi}{2}\right) (|s_1|^2 \sigma_{n_2}^2 + |s_2|^2 \sigma_{n_1}^2) + 0.38 \sigma_{n_1}^2 \sigma_{n_2}^2} \quad (55)$$

$$\frac{1}{\gamma} \leq \frac{(4 - \pi) |s_1|^2 \sigma_{n_2}^2}{2 |s_1|^2 |s_2|^2} + \frac{(4 - \pi) |s_2|^2 \sigma_{n_1}^2}{2 |s_1|^2 |s_2|^2} + \frac{0.38 \sigma_{n_1}^2 \sigma_{n_2}^2}{|s_1|^2 |s_2|^2} \quad (56)$$

$$\frac{1}{\gamma} \leq \frac{(4 - \pi) \sigma_{n_2}^2}{2 |s_2|^2} + \frac{(4 - \pi) \sigma_{n_1}^2}{2 |s_1|^2} + \frac{0.38 \sigma_{n_1}^2 \sigma_{n_2}^2}{|s_1|^2 |s_2|^2} \quad (57)$$

$$\boxed{\frac{1}{\gamma} \leq \frac{4 - \pi}{2 \gamma_1} + \frac{4 - \pi}{2 \gamma_2} + \frac{0.38}{\gamma_1 \gamma_2}} \quad (58)$$

5 Differential Matched Filter Detection

Work in progress.

$$r = \sum s_1(t) s_1(t-1) s_2^*(t) s_2^*(t-1) \quad (59)$$

$$\gamma = \frac{P_s}{P_n} \quad (60)$$

$$P_s = |s_1(t) s_1(t-1) s_2(t) s_2(t-1)|^2 \quad (61)$$

$$P_n = \text{var}(s_{11} s_{12} s_{21} n_{22} + s_{11} s_{12} s_{22} n_{21} + s_{11} s_{12} n_{21} n_{22} \quad (62)$$

$$+ s_{11} n_{12} s_{21} s_{22} + s_{11} n_{12} s_{21} n_{22} + s_{11} n_{12} s_{22} n_{21} + s_{11} n_{12} n_{21} n_{22} \quad (63)$$

$$+ s_{12} n_{11} s_{21} s_{22} + s_{12} n_{11} s_{21} n_{22} + s_{12} n_{11} s_{22} n_{21} + s_{12} n_{11} n_{21} n_{22} \quad (64)$$

$$+ n_{11} n_{12} s_{21} s_{22} + n_{11} n_{12} s_{21} n_{22} + n_{11} n_{12} s_{22} n_{21} + n_{11} n_{12} n_{21} n_{22}) \quad (65)$$

Where $s_{11} = s_1(t)$, $s_{12} = s_1(t-1)$, $s_{21} = s_2(t)$, $s_{22} = s_2(t-1)$, $n_{11} = n_1(t)$, $n_{12} = n_1(t-1)$, $n_{21} = n_2(t)$, and $n_{22} = n_2(t-1)$.

This is probably easier to figure out experimentally. It is tempting to start aggressively canceling out terms that appear in both the numerator and the denominator. However, this would require invoking the Cauchy-Schwarz inequality.

6 Concluding Thoughts

Although these results allow us to calculate a post detection SNR, or at least a lower bound on SNR, we should be careful about using this value to calculate probability of detection or probability of false alarm metrics. The noise statistics are not purely Gaussian, or complex Gaussian, or Rayleigh distributed. When selecting a detection threshold, the actual probability distribution of the post detection data should be used.

In our results we make an implicit assumption that signal two and signal one are scaled versions of each other. That is $S_1(t) = \alpha s_2(t)$. This allows us to avoid the Cauchy-Schwarz inequality by guaranteeing it is actually an equality. The following simplification is made based on this assumption: $|s_1(t)s_2(t)|^2 = |s_1(t)|^2|s_2(t)|^2$. This assumption falls apart if $s_2(t)$ is not a scaled and rotated copy of $s_1(t)$.

The assumption also falls apart if there is a lag between $s_1(t)$ and $s_2(t)$. The work above deals with computing the output SNR for the time instant where both signals align perfectly. The implication of this is that Signal to Noise Ratio is not constant over a CAF. This should make intuitive sense given that the post detection signal magnitude is not constant over the CAF.