

[aprendtech.com](#) >> [blog](#) >> [this post](#)

If you have trouble viewing this, try [the pdf of this post](#). You can [download the code](#) used to produce the figures in this post. [Edited May 12, 2012 to clarify that constant covariance log data and photon count data give same formula for the CRLB.](#)

SNR with PHA vs. number of bins

In this post, I will discuss the signal to noise ratio (SNR) of a photon counting detector with pulse height analysis (PHA). I will show that it approaches the ideal full-spectrum SNR as the number of bins gets large. I will also show that we can get quite close to the ideal value even with a small number of bins, which was one of the main points in [my paper](#).

I will begin by re-visiting a result from my [last post](#) that the Cramèr-Rao lower bound (CRLB) with multivariate normal log of photon count data assuming a constant covariance is nearly equal to the accurate matrix that includes the variation of the covariance. The derivation becomes problematic with a large number of bins since the mean value in each bin may be small enough so the normal approximation to the counts is not valid. I will show an alternate derivation that uses the Poisson model of the count data, which is valid for small counts. The result is formally the same as with the normal approximation so the CRLB formula can be applied for any number of bins. This result is new and was not included in [my paper](#).

Imaging task for SNR with A-space data

Summarizing my previous development, to determine the SNR with energy spectrum information I analyzed the imaging task shown in Fig. 1. I introduced the A-space description by using the attenuation coefficients of the background μ_b and feature μ_f materials as the basis function. With this basis set, the a vectors for the feature and background materials are simply

$$\mathbf{a}_f = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{a}_b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

The A-vector in the background region is $\mathbf{A}_b = t_b \mathbf{a}_b$ while in the feature region it is $\mathbf{A}_f = (t_b - t_f) \mathbf{a}_b + t_f \mathbf{a}_f$. We can apply the statistical detection theory model described in [my previous post](#) by assuming that the A-vector data are described by a multinormal distribution with expected values under two hypotheses: $\langle \mathbf{A}_{H0} \rangle = \mathbf{A}_b$ and $\langle \mathbf{A}_{H1} \rangle = \mathbf{A}_f$ and a covariance \mathbf{C}_A . From my last post, the performance will depend only on the signal to noise ratio

$$SNR^2 = \delta \mathbf{A}^T \mathbf{C}_A^{-1} \delta \mathbf{A} \quad (1)$$

where $\delta \mathbf{A} = \langle \mathbf{A}_{H1} \rangle - \langle \mathbf{A}_{H0} \rangle = t_f (\mathbf{a}_f - \mathbf{a}_b)$.

$$\begin{aligned} \delta \mathbf{A} &= \langle \mathbf{A}_{H1} \rangle - \langle \mathbf{A}_{H0} \rangle \\ &= t_f (\mathbf{a}_f - \mathbf{a}_b) \\ &= t_f \begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{aligned}$$

In general, the covariance of the A-space data depends on the estimator used. But we know that the 'minimum' covariance (suitably defined for matrix data) is the CRLB, so we can determine the optimal performance by using it for the covariance \mathbf{C}_A . In [previous post](#), I showed that we could use a linearized model with multivariate normal noise data. With this model, the CRLB is

$$\mathbf{C}_A = \left(\mathbf{M}^T \mathbf{C}_L^{-1} \mathbf{M} \right)^{-1} \quad (2)$$

where $\mathbf{M} = \frac{\partial \mathbf{L}}{\partial \mathbf{A}}$ is the gradient of log of the x-ray measurements in A-space evaluated at the operating point of the model and \mathbf{C}_L is the covariance of the log data with noise. The SNR with the imaging task in Fig. 1 is

$$SNR^2 = \delta \mathbf{A}^T \mathbf{C}_A^{-1} \delta \mathbf{A} = \delta \mathbf{A}^T \mathbf{M}^T \mathbf{C}_L^{-1} \mathbf{M} \delta \mathbf{A} \quad (3)$$

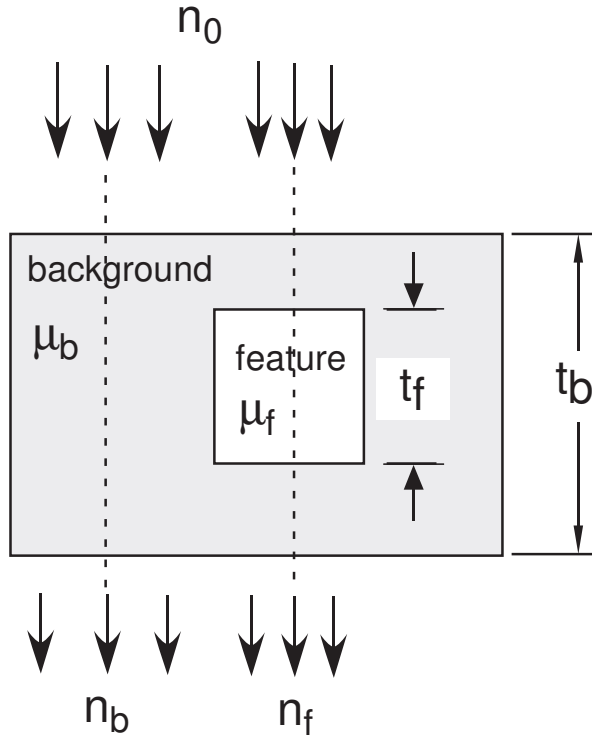


Figure 1: Imaging task for SNR computation

Table 1: Notation

photon counts	$\mathbf{N} = \begin{bmatrix} n_1 \\ n_2 \\ \dots \\ n_K \end{bmatrix}$	mean counts	$\langle \mathbf{N} \rangle = \begin{bmatrix} \langle n_1 \rangle \\ \langle n_1 \rangle \\ \dots \\ \langle n_K \rangle \end{bmatrix} = \begin{bmatrix} g_1(\mathbf{A}) \\ g_2(\mathbf{A}) \\ \dots \\ g_K(\mathbf{A}) \end{bmatrix} = \mathbf{g}(\mathbf{A})$
covariance	$\mathbf{C} = \begin{bmatrix} \langle n_1 \rangle & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & \langle n_K \rangle \end{bmatrix}$	atten. coeff.	$\mathbf{M} = \begin{bmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \\ \dots & \dots \\ \mu_{K1} & \mu_{K2} \end{bmatrix}$

CRLB with complete energy information

I would like to apply this model to show that the SNR approaches the full spectrum as the number of bins becomes large. However, as mentioned, with a large number of bins the data may not satisfy the assumptions of a large number of counts. In that case, I will use a Poisson model, which is valid for small mean counts, to show that the same formula (2) can be used in all cases.

I will derive the CRLB for independent multivariate Poisson data. As I have discussed in a [previous post](#), this is a good model for the PHA counts if the deadtime is negligible compared to the mean inter-arrival time of the photons. Table 1 lists the notation I use in this post.

It is convenient to use the Fisher information matrix \mathbf{F} , which is the inverse of the CRLB. As shown in Theorem 3.2 of Kay[1], \mathbf{F} has components which are the negative of the expected value of the second derivative of the log of the likelihood $\mathcal{L} = \log [Pr(\mathbf{N}; \mathbf{A})]$

$$F_{ij} = - \left\langle \frac{\partial^2 \mathcal{L}}{\partial A_i \partial A_j} \right\rangle$$

The bin counts n_k are independent Poisson random variables, which have a distribution function

$$Pr(n_k; \mathbf{A}) = \frac{e^{-\langle n_k \rangle} \langle n_k \rangle^{n_k}}{n_k!} = \frac{e^{-g_k(\mathbf{A})} g_k(\mathbf{A})^{n_k}}{n_k!}$$

where the expected value $\langle n_k \rangle = g_k(\mathbf{A})$ is assumed to be a function of the A-vector. Since the n_k are independent the joint probability is the product of the distributions and the log of the likelihood is

$$\mathcal{L} = \sum_{k=1}^K [-g_k + n_k \log(g_k) - \log(n_k!)].$$

To evaluate the Fisher information, we need the derivatives. The first derivative can be expressed as a matrix product of the gradient of \mathbf{g} , the inverse of the photon count covariance \mathbf{C}^{-1} , and the difference of the measured data and the expected values. Since the count data are independent, their covariance matrix and its inverse are

$$\mathbf{C} = \begin{bmatrix} g_1 & & 0 \\ & \dots & \\ 0 & & g_K \end{bmatrix}, \quad \mathbf{C}^{-1} = \begin{bmatrix} 1/g_1 & & 0 \\ & \dots & \\ 0 & & 1/g_K \end{bmatrix}$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial A_j} &= \sum \left[\frac{\partial g_k}{\partial A_j} \left(\frac{n_k - g_k}{g_k} \right) \right] \\ &= \frac{\partial \mathbf{g}^T}{\partial A_j} \begin{bmatrix} 1/g_1 & & 0 \\ & \dots & \\ 0 & & 1/g_K \end{bmatrix} (\mathbf{N} - \mathbf{g}) \\ &= \frac{\partial \mathbf{g}^T}{\partial A_j} \mathbf{C}^{-1} (\mathbf{N} - \mathbf{g}) \end{aligned}$$

Note that this covariance is valid even with small number of counts per bin.

The second derivative is

$$\frac{\partial^2 \mathcal{L}}{\partial A_i \partial A_j} = \frac{\partial}{\partial A_i} \left[\frac{\partial \mathbf{g}^T}{\partial A_j} \mathbf{C}^{-1} \right] (\mathbf{N} - \mathbf{g}) - \frac{\partial \mathbf{g}^T}{\partial A_j} \mathbf{C}^{-1} \frac{\partial \mathbf{g}}{\partial A_i}.$$

The expected value of the first term is zero since the term in the bracket does not depend on the data and $\langle \mathbf{N} \rangle = \mathbf{g}$. The second term does not depend on the data so

$$\begin{aligned} F_{ij} &= - \left\langle \frac{\partial^2 \mathcal{L}}{\partial A_i \partial A_j} \right\rangle \\ &= \frac{\partial \mathbf{g}^T}{\partial A_j} \mathbf{C}^{-1} \frac{\partial \mathbf{g}}{\partial A_i} \end{aligned}$$

The full matrix is

$$\mathbf{F} = \frac{\partial \mathbf{g}^T}{\partial \mathbf{A}} \mathbf{C}^{-1} \frac{\partial \mathbf{g}}{\partial \mathbf{A}^T}. \quad (4)$$

where I have used the notation of van den Bos[2] Appendix D so

$$\frac{\partial \mathbf{g}}{\partial \mathbf{A}^T} = \begin{bmatrix} \frac{\partial g_1}{\partial A_1} & \frac{\partial g_1}{\partial A_2} & \dots \\ \frac{\partial g_2}{\partial A_1} & \frac{\partial g_2}{\partial A_2} & \dots \\ \dots & \dots & \dots \end{bmatrix}$$

and $\left(\frac{\partial \mathbf{g}}{\partial \mathbf{A}^T} \right)^T = \frac{\partial \mathbf{g}^T}{\partial \mathbf{A}}$.

If the number of bins is large so the energy width is small, we can use the Beer's law exponential model of the expected value of the counts

$$g_k(\mathbf{A}) = \langle n_k \rangle = n_{k0} \exp[-\mu_{k1} A_1 - \mu_{k2} A_2]$$

where n_{k0} is the mean of the number of photons incident on the object in the $k'th$ energy bin. Differentiating,

$$\frac{\partial g_k}{\partial A_i} = -\mu_{ki}g_k.$$

We can express all the components as a matrix multiplication

$$\frac{\partial \mathbf{g}}{\partial \mathbf{A}^T} = -\mathbf{M}^T \mathbf{C}$$

and the Fisher matrix (4) is

$$\mathbf{F} = \mathbf{M}^T \mathbf{C} \mathbf{M} \quad (5)$$

where I have used the fact that all covariance matrices are symmetrical so $\mathbf{C}^T = \mathbf{C}$.

The Fisher information matrix derived in my [previous post](#) assuming log data is $\mathbf{F} = \mathbf{M}^T \mathbf{C}_L^{-1} \mathbf{M}$ (see Eq. 8 of my previous post, which I reproduce here)

$$\text{cov}(\hat{\mathbf{A}}_{MLE}) = (\mathbf{M}^T \mathbf{C}_L^{-1} \mathbf{M})^{-1} \quad (6)$$

Recall that the CRLB is the inverse of \mathbf{F} , so the Fisher information matrix from my previous post is the term in the parentheses of Eq. 6. With photon count data, the variance is the inverse of the number of counts and since \mathbf{C}_L is a diagonal matrix, $\mathbf{C}_L^{-1} = \mathbf{C}$. Thus Eq. 5 has the same form as the Fisher information of constant covariance multivariate normal data, derived in my previous post but here it is derived without the assumptions that the covariance is constant and the mean count values are large. Therefore, we can use the same formula (2) for the covariance of the A-space data for any number of bins.

SNR for K bin PHA

To derive the SNR with K-bin data, we need to evaluate Eq. 3 with

$$\delta \mathbf{A} = t_f \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} M_{11} & M_{12} \\ \dots & \dots \\ M_{K1} & M_{K2} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} \langle n_1 \rangle & & 0 \\ & \dots & \\ 0 & & \langle n_K \rangle \end{bmatrix}.$$

This is straight-forward but tedious, so I used Maxima to automate the matrix computation. Maxima is an open-source symbolic math program that is available for download [here](#). A good introduction is available [here](#). I develop the code by writing a batch script with a text editor and then running it using the Maxima *batch* command. For example, I copy this command to the clipboard and then paste it into the WxMaxima GUI. With modern computers, the evaluation is very fast and we can see the results almost immediately

```
batch ("E:\\Projects\\Blog\\Posts\\P29D2kbin\\code\\SNRkbin.max")$
```

A Maxima script to compute the SNR with a 3 bin detector is

```

/*
SNRkbin.max
Maxima batch file to derive SNR^2 for K bin PHA
to invoke copy and paste into Maxima:
batch("E:\\Projects\\Blog\\Posts\\P29D2kbin\\code\\SNRkbin.max")$
REA 2012-Feb-14 11:38
COPYRIGHT: © Aprend Technology, 2012. All rights reserved.
*/
"clear all workspace variables"$
kill(all)$
"The SNR^2 USING 3 bin PHA"$
"the M matrix with general coeffs and its transpose"$
M: matrix ([m11,m12],[m21,m22],[m31,m32])$
MT:transpose(M);
"the covariance matrix"$
C:diagmatrix(3,0);
"fill the diagonal"$
S: matrix ([n1,n2,n3]);
for k:1 thru 3 do C[k,k]:S[1,k]$
"display c"$
C;
"The delta_A vector for basis materials feature and background"$
dA: matrix ([1],[ - 1])$
dAT:transpose(dA);
"the SNR^2"$
d2:ratsimp(dAT.MT.C.M.dA);
"factor the individual terms"$
d2simp:n1*(m12-m11)^2 + n2*(m22-m21)^2 + n3*(m32-m31)^2;
"verify correct"$
ratsimp(d2-d2simp);

```

The initial result was (we can use edit|"copy as Latex" in the Maxima GUI to copy expressions and then paste them into a Latex or Lyx file)

$$(m_{32}^2 - 2m_{31}m_{32} + m_{31}^2) n_3 + (m_{22}^2 - 2m_{21}m_{22} + m_{21}^2) n_2 + (m_{12}^2 - 2m_{11}m_{12} + m_{11}^2) n_1.$$

The program did not factor the quadratics in the parentheses but this is easily done. The SNR with 3 bins generalizes to K bins as

$$SNR_K^2 = \frac{t_f^2}{2} \sum_{k=1}^K \langle n_k \rangle \langle \mu_{k2} - \mu_{k1} \rangle_k^2 \quad (7)$$

As explained in my paper, the factor of $1/2$ is introduced to make the imaging task SNR comparable to the detection theory SNR, which assumes only one measurement. The notation $\langle \mu \rangle_k$ means the weighted mean of the function $\mu(E)$ in the spectrum of bin k . For example, if bin k corresponds to energies from E_k to $E_k + \Delta E$ and $S(E)$ is the spectrum incident on the detector, then the weighted mean is

$$\langle \mu \rangle_k = \frac{\int_{E_k}^{E_k + \Delta E} \mu(E) S(E) dE}{\int_{E_k}^{E_k + \Delta E} S(E) dE}$$

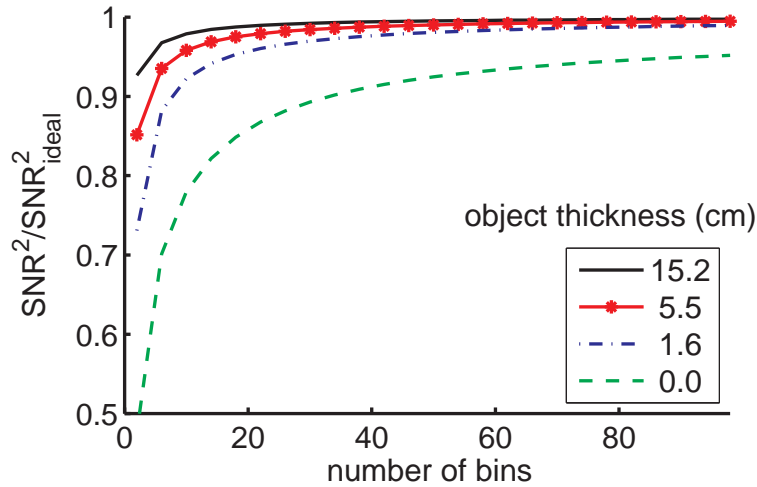


Figure 2: SNR vs. number of bins and object thickness. The SNR is normalized by dividing by the ideal with full energy spectrum information.

Using the notation $\delta\mu = \mu_2(E) - \mu_1(E)$, for the difference in attenuation coefficient functions, Eq. 7 is the same as Eq. 35 of my paper

$$SNR_K^2 = \frac{t_f^2}{2} \sum_{k=1}^K \langle n_k \rangle \langle \delta\mu \rangle_k^2.$$

Simulate K-bin SNR

I wrote Matlab code to compute the SNR as a function of the number of PHA bins. I normalized the result by dividing by the ideal, full-spectrum SNR. I also computed the results as a function of object thickness. Recall that the A-vectors for this kind of object lie along a straight line in the A-plane. The results are shown in Fig. 2.

Conclusions

There are two interesting features of Fig. 2. First, as we expected, the SNR approaches the ideal result as the number of bins gets larger. This is shown mathematically in my paper using the mean value theorem. The second interesting aspect is that the performance with a small number of bins depends strongly on the object thickness. My interpretation of this result is that as the object thickness increases the spectral width also decreases due to beam hardening. Therefore, there is less energy information and it is easier for a low energy-resolution detector to extract it. The improvement occurs for relatively thin objects. From the figure, the three-bin SNR with a 5.5 cm object is about 93% of the ideal value.

Last edited May 12, 2012

©2012 by Aprend Technology and Robert E. Alvarez

Linking is allowed but reposting or mirroring is expressly forbidden.

References

- [1] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, 1st ed., Prentice Hall PTR, 1993.
- [2] A. van den Bos, *Parameter Estimation for Scientists and Engineers*, Hoboken, New Jersey: Wiley-Interscience, 2007.