LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-414706

Calculating the Errors on the Feynman Variance, Y_{2F}

Peter D. Barnes, Jr.

July 14, 2009

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Calculating the Errors on the Feynman Variance, Y_{2F}

In a multi-channel detector we record the time of arrival of detected counts. We divide the timeline into separate gates of width T, and ask how many counts arrived in each gate. Finally, we form the count distribution histogram, **B**, a tally of how many gates had zero counts, how many had one count, . . . For illustration, we'll use the gamma block data from NTS Run 20071114224434, RTO 1.2, which was RF1-24, 13.7 kg HEU with 2" Pb shielding.

The count distribution for a gate width of 416 μs is shown in Figure 1. Note that **B** is a vector across multiplicity, n. **B** also depends on the gate time, $\mathbf{B} = \mathbf{B}(T) = \mathbf{B}(T_i)$, so we could just as well think of **B** as a rectangular matrix. For the moment we'll work with a fixed T, and use a single column of the full count distribution matrix. We'll leave the T dependence implicit.

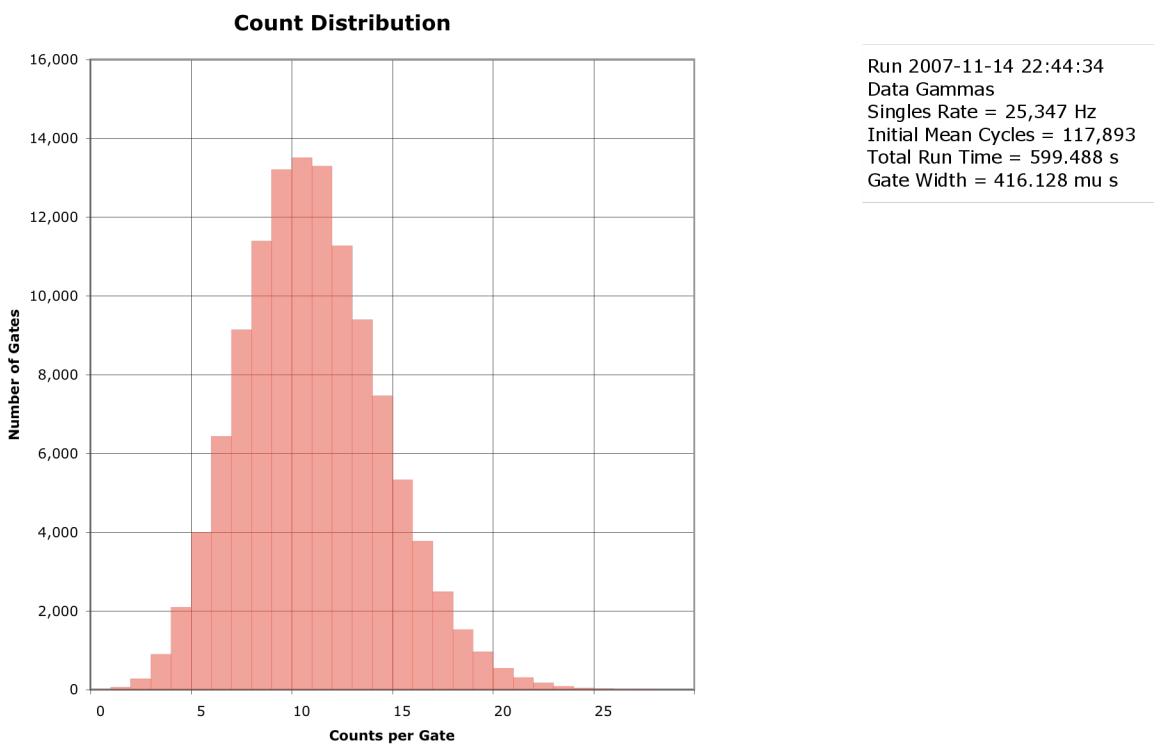

Figure 1. Count distribution from NTS Run 20071114224434, RTO 1.2, RF1-24, 13.7 kg HEU with 2" Pb shielding, gamma blocks only, for a gate width of 416 μs.

The Feynman variance is a scalar random variable which depends on the vector count distribution:

$$Y_{2\mathrm{F}} = Y_{2\mathrm{F}}\left(B_n\right) = Y_{2\mathrm{F}}\left(\mathbf{B}\right) \tag{1}$$

Using Sean's formulary, the *Idiot's Guide*,[1] it's straight-forward to calculate $Y_{2\mathrm{F}}$ "by hand" from **B**:

$$\begin{aligned} N &= \sum_{n=0}^{\infty} B_n \\ n_{\mathrm{tot}} &= \sum_{n=0}^{\infty} nB_n \\ \bar{c} &= \frac{n_{\mathrm{tot}}}{N} \end{aligned} \qquad \begin{aligned} \mathcal{M}_2 &= \frac{1}{N}\sum_{n=2}^{n_{\mathrm{tot}}} \binom{n}{2} B_n \\ Y_{2\mathrm{F}} &= \frac{\mathcal{M}_2}{\bar{c}} - \frac{\bar{c}}{2!} \end{aligned} \tag{2}$$

where N is the total number of gates in the timeline, n_{tot} is the total number of observed counts, $\bar{c}$ is the average number of counts per gate, and $\mathcal{M}_2$ is the second combinatorial moment.

My calculation is shown in Figure 2. For comparison, the result of analysis pass `V490CD0` is shown in Figure 3. They give the same results, so I haven't made a gross error. Note that this run yields about 118 K independent gates for each gate width.

Again, using Sean's formulary, it is straightforward to calculate the variance as well.

$$\sigma_{Y_{2\mathrm{F}}}^2 = \frac{1}{N\bar{c}^2}\sum_{n=2}^{n_{\mathrm{tot}}}\left[\binom{n}{2} - n\left(Y_{2\mathrm{F}} + \bar{c}\right)\right]^2 b_n\left(1 - b_n\right) \tag{3}$$

Again, we agree, but let's look at this in more detail.

Variance I – Ensemble

Imagine that we have an ensemble of n_{exp} experiments for $Y_{2\mathrm{F}}$. That is, a vector **B** from each experiment. The standard definition for the variance is

$$\mathrm{var}\, Y_{2\mathrm{F}} = \sigma_{Y_{2\mathrm{F}}}^2 = \left\langle (Y_{2\mathrm{F}} - \overline{Y_{2\mathrm{F}}})^2 \right\rangle_{n_{\mathrm{exp}}} = \left\langle (\Delta Y_{2\mathrm{F}})^2 \right\rangle_{n_{\mathrm{exp}}} \tag{4}$$

where the ensemble average is defined as:

$$\left\langle \ldots \right\rangle_{n_{\mathrm{exp}}} \equiv \frac{1}{n_{\mathrm{exp}}}\sum_{k=0}^{n_{\mathrm{exp}}-1} \ldots_k \tag{5}$$

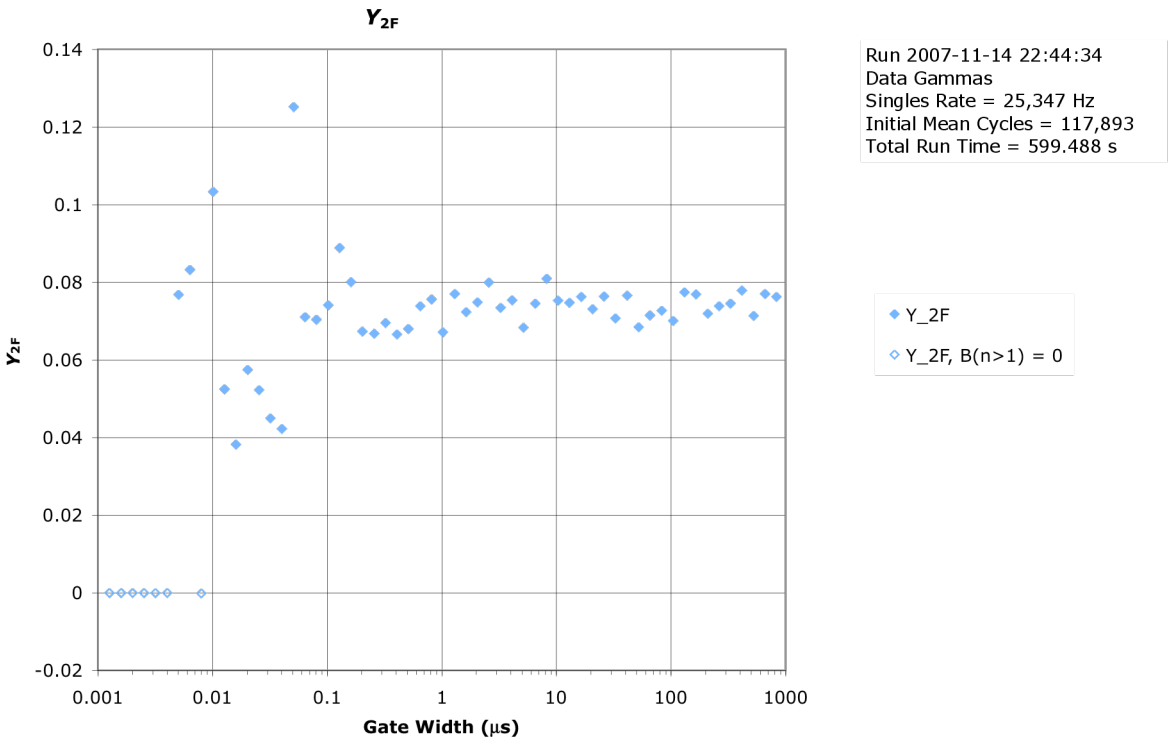

Figure 2. Y_{2F} as a function of gate width from the data, using this calculation. Open symbols represent gate widths where $B_{n\geq 2} = 0$ and Y_{2F} is undefined.

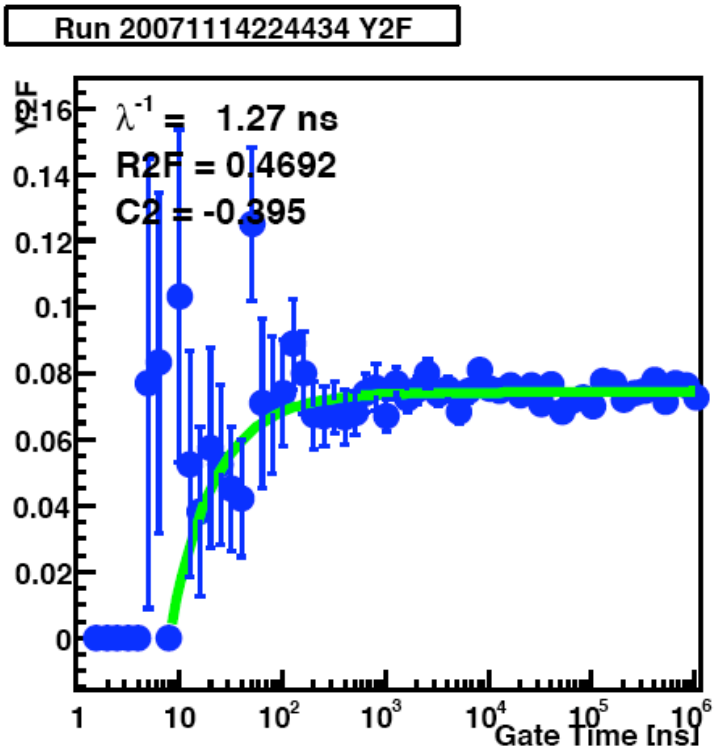

Figure 3. Y_{2F} as a function of gate width from NTS Run 20071114224434, RTO 1.2 (RF1-24 + 2" Pb), gamma blocks only, using the standard analysis package, version and pass `V490CD0`.

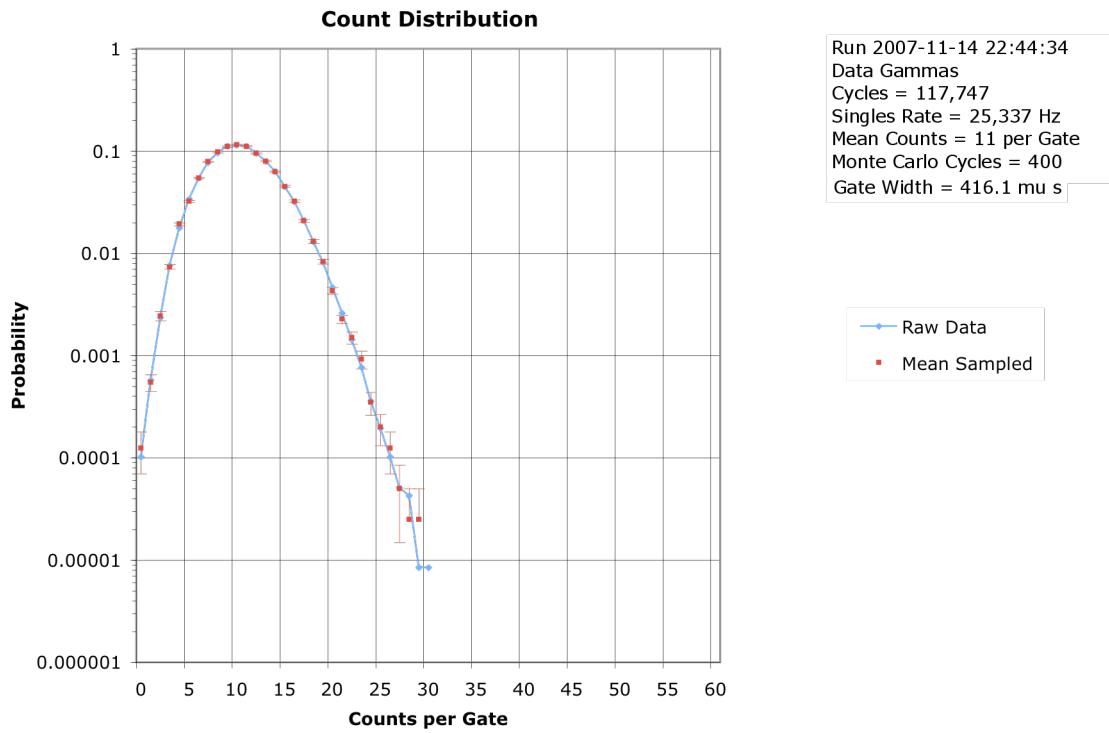

Figure 4. (Blue curve) Measured count probability distribution from Run 20071114224434 using 416 µs gate width. (Red points) Average Monte Carlo sampled count distribution for 100 experiments using 400 gates.

The mean value of Y_{2F} is

$$\overline{Y_{2F}} = \langle Y_{2F} \rangle_{n_{\exp}} = \frac{1}{n_{\exp}} \sum_{k=0}^{n_{\exp}-1} \left(Y_{2F} \right)_k = \frac{1}{n_{\exp}} \sum_{k=0}^{n_{\exp}-1} Y_{2F}\left(\mathbf{B_k} \right) \tag{6}$$

Using the data, we can pick a gate time and obtain a good measurement of the count probability distribution, **b**, shown as the solid line in Figure 4. This differs from Figure 1 only in the normalization; Figure 1 shows the raw number of gates, **B**; Figure 4 is the same distribution divided by the total number of gates, $\mathbf{b} = \frac{1}{N}\mathbf{B}$, to obtain the (normalized) probability distribution.

Using this well-measured count probability distribution, we can simulate the result for a different live time, just by generating a count distribution **B'**, drawn according to the measured **b**, with the specified total number of gates (cycles). In this case I simulate $n_{exp} = 100$ experiments with $N = 400$ gates. The results are summarized in Table 1; the quantities are described in Table 2. Then, across the 100 experiments, I calculate the average, standard deviation, and the standard error on the average, for each of the defined quantities. Figure 4 also shows the count probability distribution for all experiments taken together, 40,000 gates total, which agrees with the source probability distribution, **b**.

Table 1. Summary of $n_{exp} = 100$ experiments each with $N = 400$ gates. See Table 2 for the definition of each quantity.

Index	0	99	Average	±SE Mean	Stdev
N	400	400	400.000	0.000	0
$N \geq 2$	400	400	399.700	0.048	0.482
$\bar{c}$	10.693	10.438	10.533	0.017	0.171
n_{tot}	4277	4175	4213.160	6.859	64.589
$\mathcal{M}_2$	57.788	54.968	56.292	0.188	1.882
$Y_{2\text{F}}$	0.058	0.048	0.076	0.004	0.044
var $Y_{2\text{F}}$	0.066	0.063	0.065	0.000	0.002
σ	0.257	0.251	0.254	0.000	0.004
r	25,695.240	25,082.447	25,311.704	41.207	412.068

Table 2. Definitions of quantities in Table 1.

Quantity	Definition
N	Number of gates sampled
$N \geq 2$	Number of gates with 2 or more counts, $N \geq 2 = N - B_0 - B_1$
$\bar{c}$	Average number of counts per gate
n_{tot}	Total number of counts in the experiment
$\mathcal{M}_2$	Second combinatorial moment of the count distribution
$Y_{2\text{F}}$	Feynman's second moment
var $Y_{2\text{F}}$	Variance computed using "*Idiot's Guide*", Eq. 3 above
σ	Square root of the variance
r	Singles rate in Hz

Notice the highlighted numbers in Table 1. The *Idiot's Guide* variance formula gives $\bar{\sigma} = 0.254$, whereas the Monte Carlo sample has a standard deviation of only 0.043. Where does this difference come from?

Variance II – Covariance

The variance formula derived in the *Idiot's Guide* is based on the usual expression for propagating errors,

$$\sigma_{Y_{2\text{F}}}^2 = \text{var}\, Y_{2\text{F}} = \frac{1}{n_{tot}} \sum_{n=0}^{n_{tot}-1} \left(\frac{\partial Y_{2\text{F}}}{\partial B_n} \right)^2 \sigma_{B_n}^2 \tag{7}$$

This is actually just a sum across the diagonal elements of the full covariance matrix. This simplification is appropriate when there are no correlations, or off-diagonal elements.

The complete expression is

$$\sigma_{Y_{2F}}^2 = \left(\nabla Y_{2F}\right)^{\mathrm{T}} \cdot \boldsymbol{\Sigma} \cdot \nabla Y_{2F} \tag{8}$$

Here ∇ is the usual vector gradient,

$$\nabla \equiv \left\{ \frac{\partial}{\partial B_n} \right\} \tag{9}$$

$\left(\ \right)^{\mathrm{T}}$ denotes the transpose (in this case from column to row vector), and $\boldsymbol{\Sigma}$ is the covariance matrix. In our case,

$$\nabla Y_{2F} = \left\{ \frac{1}{N\overline{c}} \left[\binom{n}{2} - n\left(Y_{2F} + \overline{c}\right) \right] \right\} \tag{10}$$

For an ensemble, the covariance matrix is defined as

$$\begin{aligned} \boldsymbol{\Sigma} = \left\{ \sigma_{n,m}^2 \right\} &= \left\{ \left\langle \Delta B_n \cdot \Delta B_m \right\rangle_{n_{\exp}} \right\} \\ &= \left\{ \left\langle \left(B_n - \overline{B_n}\right) \cdot \left(B_m - \overline{B_m}\right) \right\rangle_{n_{\exp}} \right\} \\ &= \left\langle \Delta \mathbf{B} \cdot \Delta \mathbf{B}^{\mathrm{T}} \right\rangle_{n_{\exp}} \end{aligned} \tag{11}$$

(Note the outer product in the last line, resulting in a square matrix with n_{tot} rows/columns.)

Correlations in the Count Distribution

It's high time to be really clear on the question we want to answer. We take some data, an "experiment." We compute Y_{2F}. Based on this single experiment, we want to know what variation we could expect had we run an ensemble of identical experiments. That means, suppose we knew the true count probability distribution, **b**, and we took a series of runs each with the same number of gates, N, what variation should we expect on the measured Y_{2F}?

For this we need the generalization of the binomial distribution, the multinomial distribution. In a binomial distribution, there are only two possible outcomes, "success" and "failure," with a constant probability, p, of success across trials. In a multinomial distribution each of N trials results in one of a finite number, n_{tot}, of outcomes, each with a constant probability, b_n, across trials. Not surprisingly, the marginal distribution for any one of the outcomes is a binomial. In a multinomial the outcomes are (anti-)correlated because the total number of trials is fixed; a positive fluctuation in the number of trials with a given outcome has to be compensated by negative fluctuations for other outcomes.

The multinomial probability distribution is

$$\mathcal{P}(\mathbf{B};N,\mathbf{b}) = \begin{cases} \dfrac{N!}{\prod_n B_n!}\prod_n b_n^{B_n}, & \text{when } \sum_{\mathrm{n}} B_n = N \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

and the total probability for all outcomes is unity:

$$\sum_{\mathrm{n}} b_n = 1 \tag{13}$$

The covariance matrix for the multinomial distribution can be written in this simple form:

$$\mathbf{\Sigma}_{\text{Multinomial}} = \left\{\Sigma_{i,j}\right\} = \begin{cases} Nb_i\left(1-b_i\right) & i = j \\ -Nb_ib_j & i \neq j \end{cases} \tag{14}$$

The diagonal entries are just the variance of the binomial distribution.

So how does this apply to us?

Figure 5 shows the covariance matrix calculated for the ensemble of experiments, using Eq. 11. Clearly, there are anti-correlations, especially where **b** has strong support. Not surprisingly, this is somewhat noisy, since each experiment has only 400 gates. Figure 6 shows the covariance computed using Eq. 14 and the count probability distribution **b** measured in the full data set with 118K gates.

Finally, I computed the error using Eq 8, and obtained the results shown in Table 3. There is still considerable variation run to run for the numerical covariance. To understand the size of this variation, I computed each of the methods for ten iterations at a sample of gate widths. These results are summarized in Table 4. I conclude that the Monte Carlo and all covariance calculations are in agreement; the *Idiot's Guide* disagrees, especially for long gate widths.

As a last note, the live-time dependence is

$$\begin{aligned} \sigma^2_{Y_{2\mathrm{F}}} &= \left(\nabla Y_{2\mathrm{F}}\right)^{\mathrm{T}} \cdot \mathbf{\Sigma} \cdot \nabla Y_{2\mathrm{F}} \\ &\approx \frac{1}{N} \cdot N \cdot \frac{1}{N} \\ &\approx \frac{1}{N} \end{aligned} \tag{15}$$

So the error falls as the square root of the number of gates.

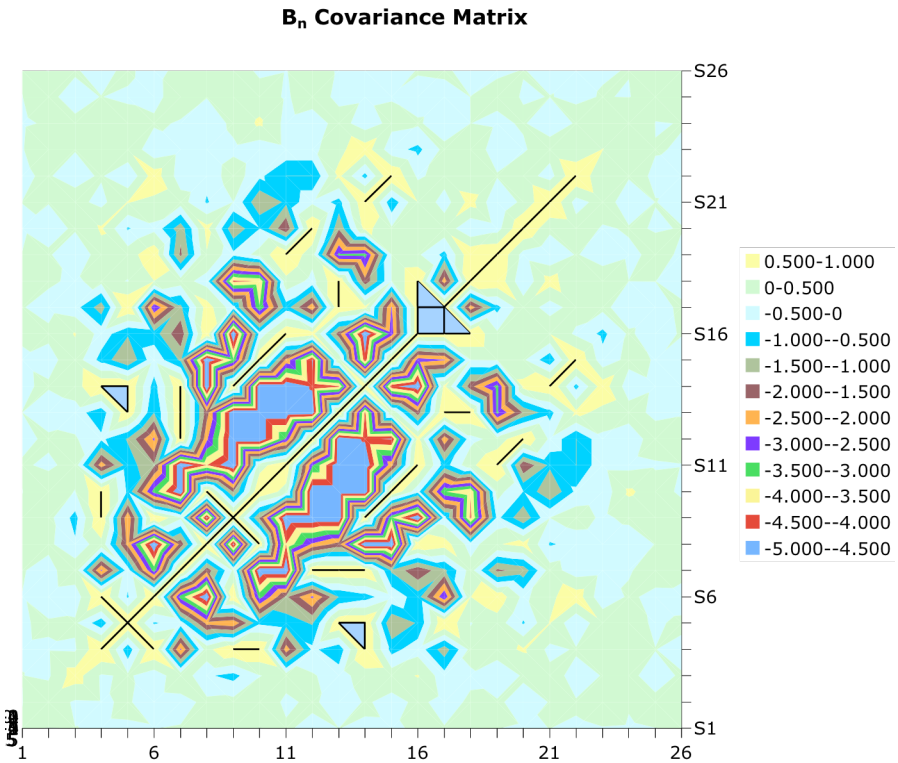

Figure 5. Covariance matrix for the ensemble, computed using Eq. 11.

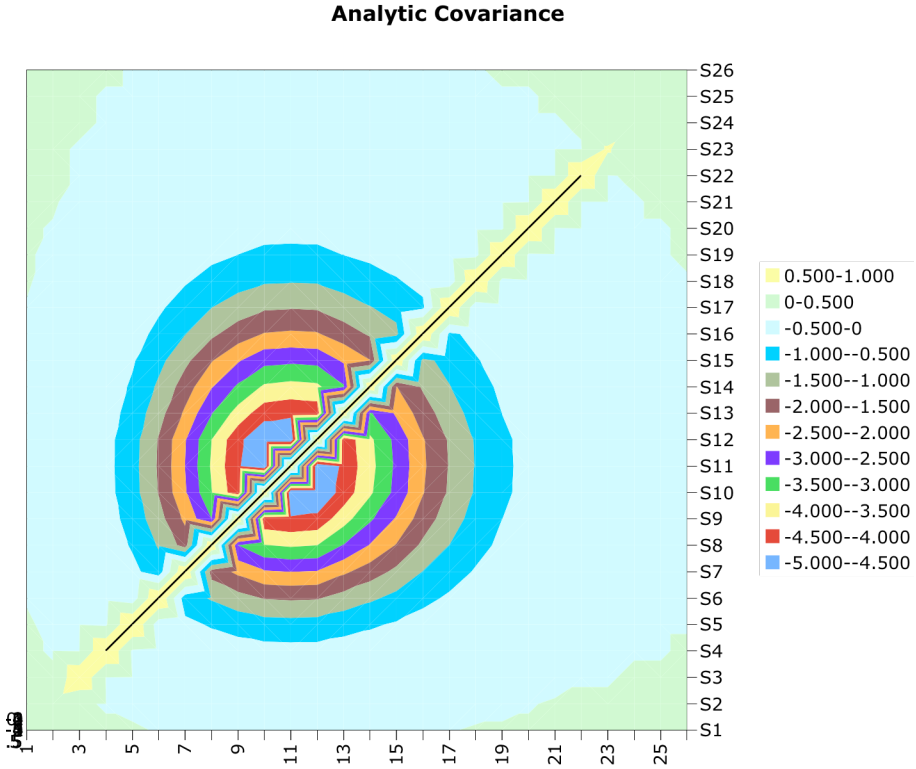

Figure 6. Covariance computed using Eq. 14.

Table 3. Summary of error estimates.

Method		Sources		σ
1	***"Idiot's Guide"***	Eq 3	**b** from data, resampled for $N = 400$, averaged over Monte Carlo ensemble	0.2542
2	**Monte Carlo**	Eq 4		0.0440
3	**Covariance**	$\mathbf{\Sigma}$ ∇Y_{2F}	Eq. 11, Monte Carlo ensemble averaged over Monte Carlo ensemble	0.0441
4		$\mathbf{\Sigma}$ ∇Y_{2F}	Eq. 14, **b** from data, scaled to $N = 400$, averaged over Monte Carlo ensemble	0.0416
5		$\mathbf{\Sigma}$ ∇Y_{2F}	Eq. 14, both using **b** from data, scaled to $N = 400$	0.416

Table 4. Error estimates for each method, as a function of gate time, *T*. Missing entries occur when at least one iteration of the Monte Carlo fails to generate a single gate with more than one count. Method numbers refer to lines in Table 3.

T (μs)	Method				
	1	**2**	**3**	**4**	**5**
1.024		0.09 ± 0.01	0.082 ± 0.009	0.092 ± 0.004	0.0789 ± 0.0000
3.25		0.057 ± 0.004	0.058 ± 0.006	0.057 ± 0.001	0.0546 ± 0.0000
10.32	0.034 ± 0.001	0.037 ± 0.003	0.045 ± 0.004	0.040 ± 0.001	0.0389 ± 0.0000
32.77	0.0333 ± 0.0005	0.033 ± 0.003	0.039 ± 0.003	0.0358 ± 0.0003	0.0356 ± 0.0000
104.0	0.0670 ± 0.0005	0.049 ± 0.003	0.041 ± 0.003	0.0490 ± 0.0001	0.0490 ± 0.0000
330.3	0.2018 ± 0.0005	0.044 ± 0.003	0.043 ± 0.003	0.0415 ± 0.0001	0.0415 ± 0.0000
1049	0.6459 ± 0.0006	0.040 ± 0.003	0.040 ± 0.003	0.0407 ± 0.0000	0.0407 ± 0.0000

Implementation Results

I implemented the full covariance method for calculating the errors in an analysis framework based on the Root package.[2] Figure 7 shows the full count distribution, **B**, as a function of multiplicity, b_n, *vs.* gate width index. Figure 8 shows the covariance matrix for gate index 55, $T = 416$ µs; compare to Figure 6. Finally, Figure 9 shows the gate time dependence of Y_{2F} with a fit using the full errors; compare to Figure 2-3.

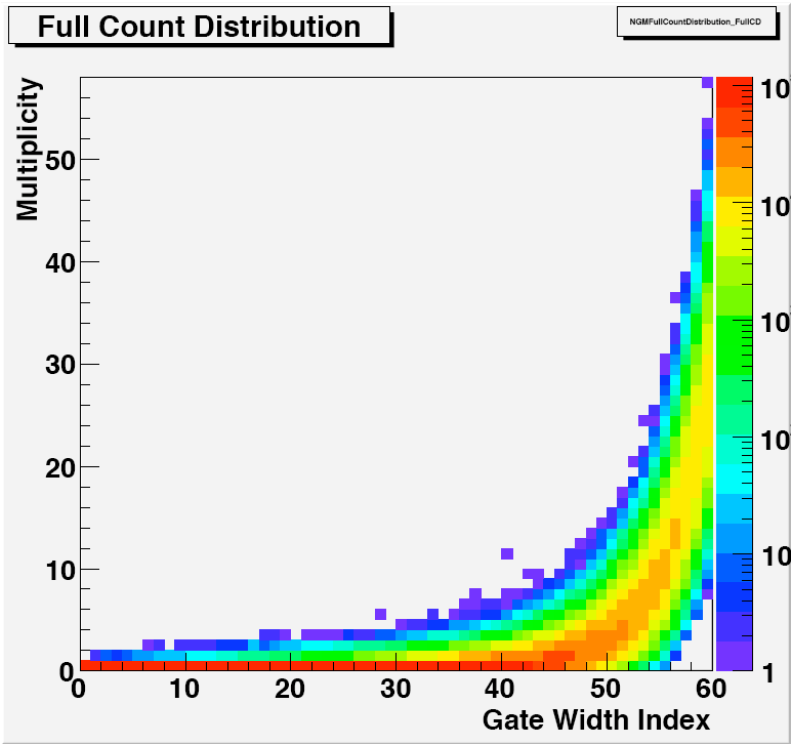

Figure 7. Full count distribution, **B**, showing multiplicity, b_n, *vs.* gate width index.

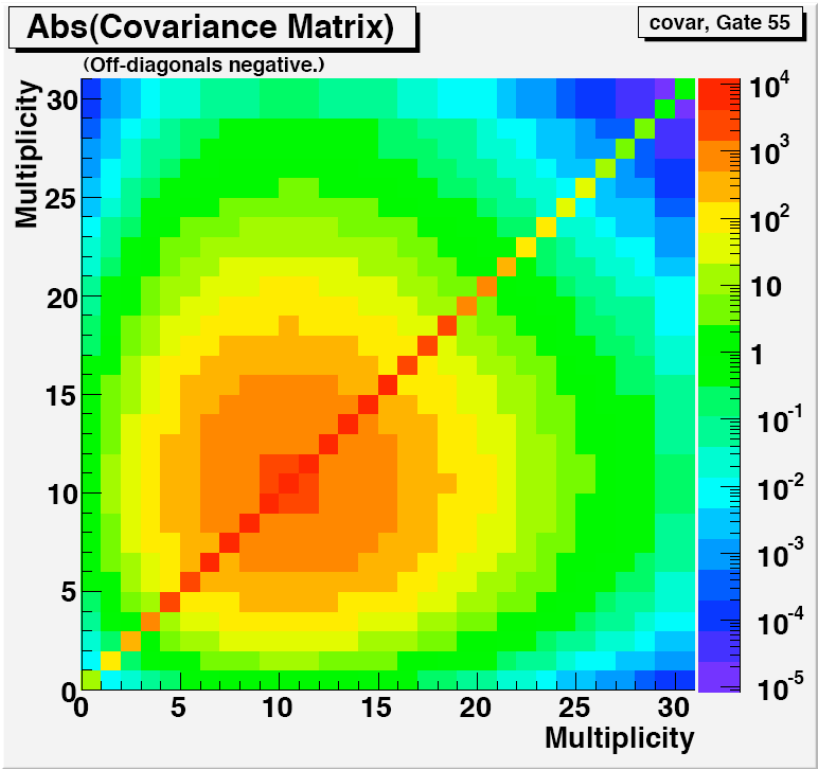

Figure 8. Covariance matrix for gate index 55, corresponding to $T = 416$ µs.

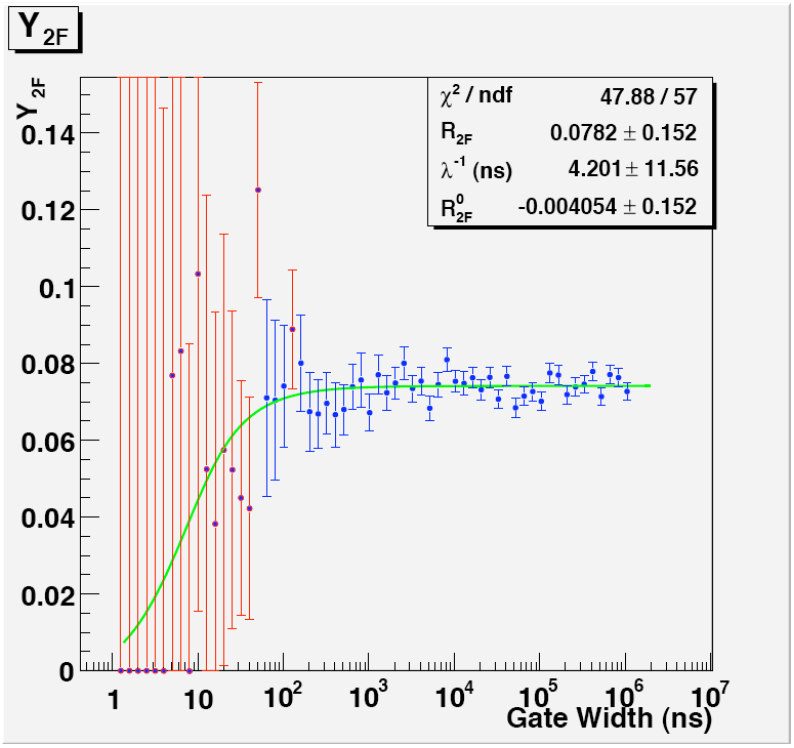

Figure 9. Y_{2F} as a function of gate width from the data, with a fit, calculated by the analysis framework.

Gate Time Dependence

With the full covariance machinery in hand, I recomputed the error as a function of gate time. Figure 10 shows the Feynman variance, the error computed using the full covariance, Eq. 14, and the conventional error using the *Idiot's Guide*, Eq. 3. Correlations clearly become important in reducing the errors at long gate times. Figure 11 shows the same two errors, along with the range of non-zero counts per gate present in **B**, as a function of gate width. The range of non-zero counts is the dimension of the covariance matrix. Evidently correlations become important when this range is larger than of order 10.

With the live-time dependence from Eq. 15, we can estimate the errors we would have obtained had we run for a different live-time, or allocated our live time differently between gates of different widths. Figure 12 shows the scaled error for three different strategies. "Uniform number of gates" allocates the same number of gates, N, for each gate width. This is essentially the strategy we use now. "Uniform Y_{2F} / Sigma" allocates live-time to obtain uniform significance. "Uniform Live Time" allocates more gates for shorter gate widths, so that all gate widths receive the same live time. All of these cases assume a fixed total live time.

Figure 13 shows the significance, Y_{2F}/σ, for a uniform one second live-time for each gate. This type of plot is useful for estimating the live-time that will be required to reach a stated level of significance. For example, at 1 μs we have ~40 σ at 1 s live-time. We can estimate that we need $\left(\frac{6}{40}\right)^2 \times 1\text{ s} = 22.5\text{ ms}$ to achieve a 6-σ measurement.

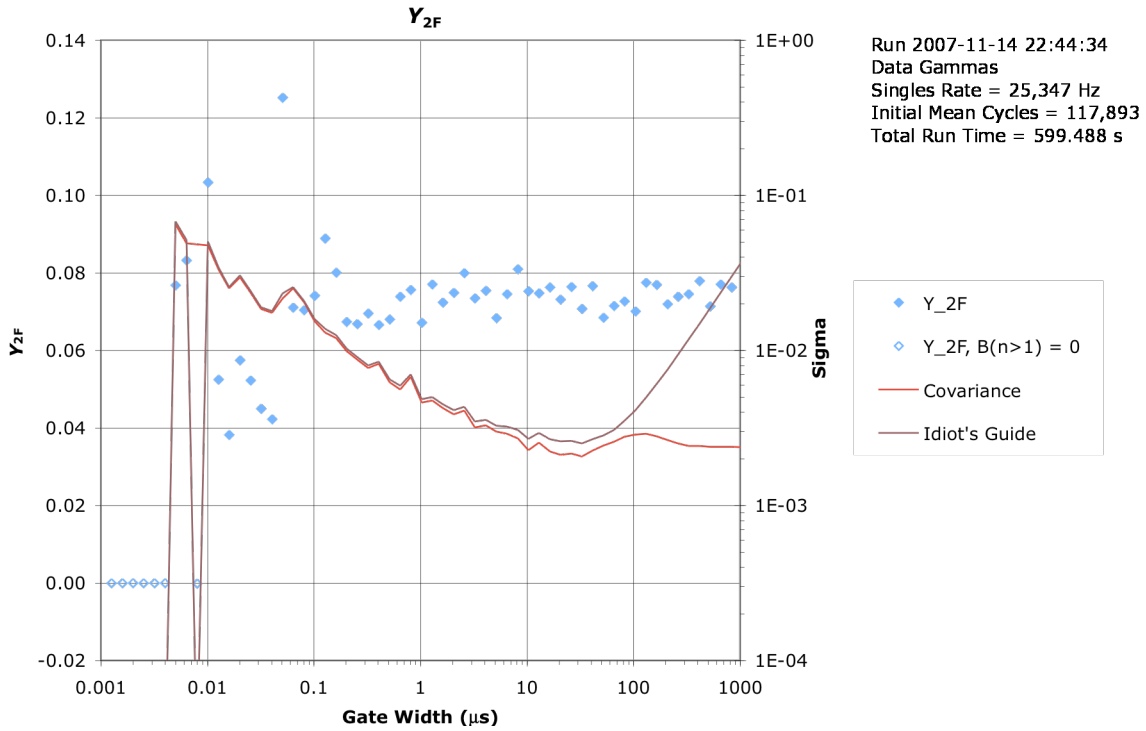

Figure 10. Feynman variance and error, computed using full covariance, Eq. 14, and the *Idiot's Guide*, Eq. 3.

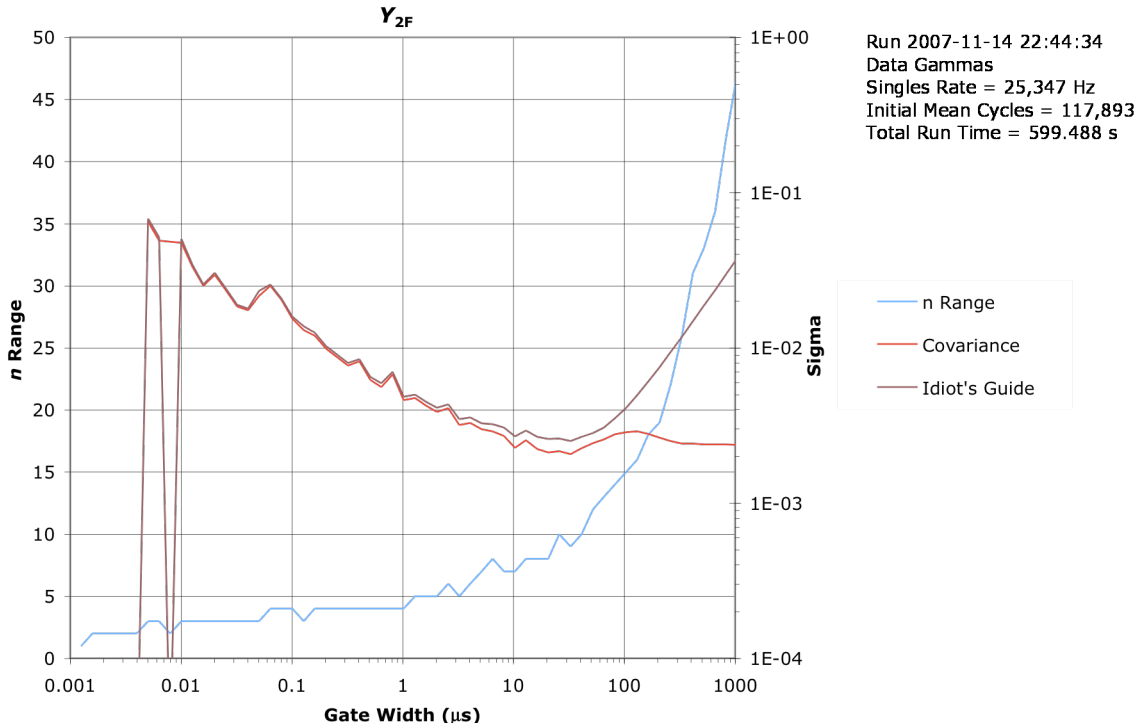

Figure 11. Error on Feynman variance (right axis) compared to the range of counts per gate present as a function of gate time. Correlations become important when the count range is of order 10.

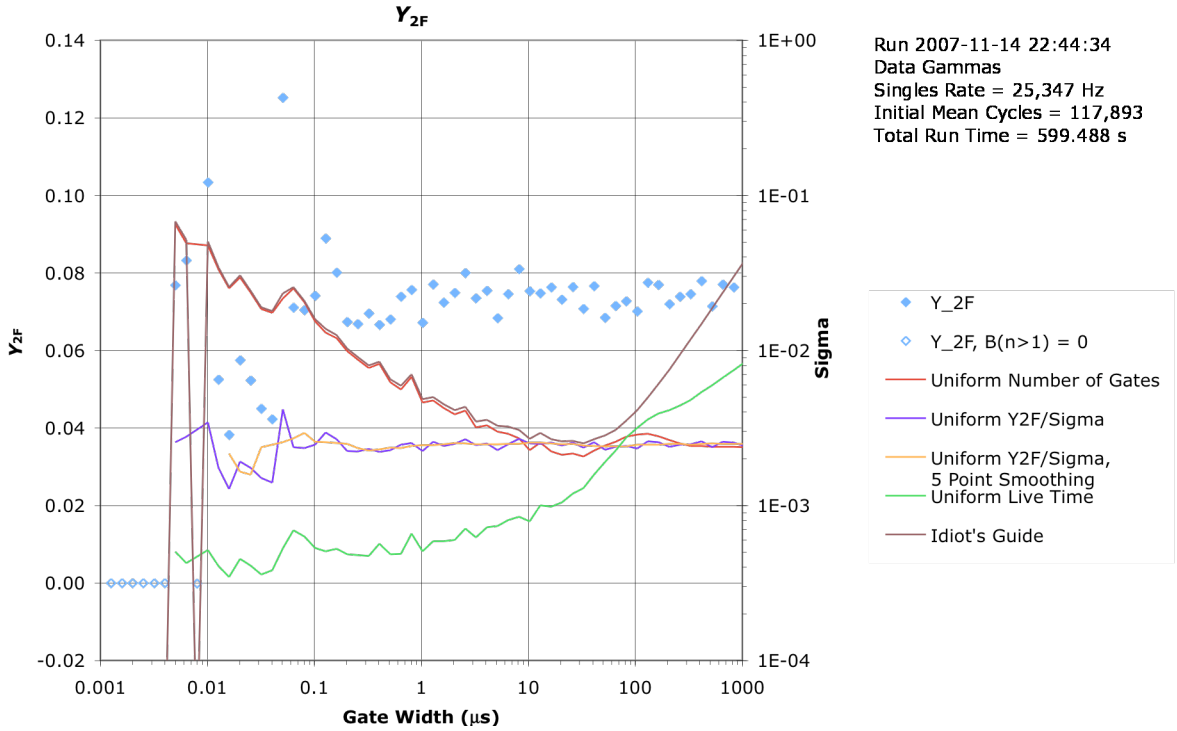

Figure 12. Error estimates for different live-time strategies. See the text for definitions of each strategy.

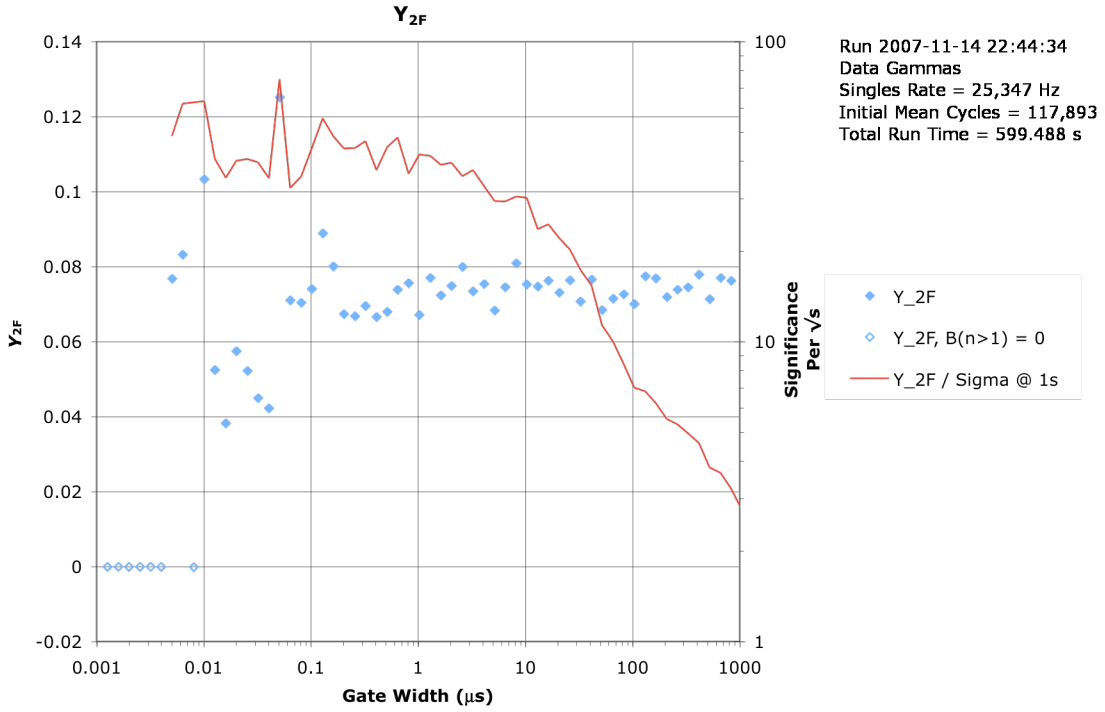

Figure 13. Significance *vs.* gate width for one second of live-time for each gate width.

Appendix A: Formulary

Here are the complete set of formulas needed to compute all the moments using the covariance.

First, some definitions:

(From Eq. 8) $$\sigma_X^2 = \left(\nabla X\right)^{\mathrm{T}} \cdot \boldsymbol{\Sigma} \cdot \nabla X \tag{A1}$$

(From Eq. 9) $$\nabla \equiv \left\{ \frac{\partial}{\partial B_n} \right\} \tag{A2}$$

(From Eq. 12) $$\boldsymbol{\Sigma}_{\text{Multinomial}} = \left\{ \Sigma_{i,j} \right\} = \begin{cases} Nb_i\left(1 - b_i\right) & i = j \\ -Nb_i b_j & i \neq j \end{cases} \tag{A3}$$

The count distribution, **B,** will be updated as new data arrives, so we have ready access to **B**, not the probability distribution, $\mathbf{b} = \frac{1}{N}\mathbf{B}$. In addition, we'll extract common factors as much as possible. to compute things efficiently. Here are the forms we'll use:

$$\mathbf{K} = -N\boldsymbol{\Sigma} = \begin{cases} -NB_i + B_i B_i & i = j \\ B_i B_j & i \neq j \end{cases} \tag{A4}$$

$$\boldsymbol{D} \equiv N\nabla = \mathrm{N}\left\{ \frac{\partial}{\partial B_n} \right\} \tag{A5}$$

$$\begin{aligned} \sigma_X^2 &= \left(\nabla X\right)^{\mathrm{T}} \cdot \boldsymbol{\Sigma} \cdot \nabla X \\ &= \left(\frac{1}{N}\boldsymbol{D}X\right)^T \left(-\frac{1}{N}\mathbf{K}\right)\left(\frac{1}{N}\boldsymbol{D}X\right) \\ &= -\frac{1}{N^3}\left(\boldsymbol{D}X\right)^T \mathbf{K}\left(\boldsymbol{D}X\right) \end{aligned} \tag{A6}$$

Moment, *m*	**Moment Gradient *Dm***
$N = \sum B_n$ (A7)	0 (independent variable) (A8)
$\mathcal{M}_1 = \frac{1}{N}\sum nB_n$ (A9)	$\boldsymbol{D}\mathcal{M}_1 = \{n\}$ (A10)
$\mathcal{M}_2 = \frac{1}{N}\sum \binom{n}{2} B_n$ (A11)	$\boldsymbol{D}\mathcal{M}_2 = \left\{\binom{n}{2}\right\}$ (A12)
$\mathcal{M}_3 = \frac{1}{N}\sum \binom{n}{3} B_n$ (A13)	$\boldsymbol{D}\mathcal{M}_3 = \left\{\binom{n}{3}\right\}$ (A14)
$Y_1 = \mathcal{M}_1$ (A15)	$\boldsymbol{D}Y_1 = \boldsymbol{D}\mathcal{M}_1$ (A16)
$Y_2 = \mathcal{M}_2 - \frac{Y_1^2}{2}$ (A17)	$\boldsymbol{D}Y_2 = \boldsymbol{D}\mathcal{M}_2 - Y_1\boldsymbol{D}Y_1$ (A18)
$Y_3 = \mathcal{M}_3 - Y_2Y_1 - \frac{1}{3!}Y_1^3$ (A19)	$\boldsymbol{D}Y_3 = \boldsymbol{D}\mathcal{M}_3 - \left(Y_1\boldsymbol{D}Y_2 + Y_2\boldsymbol{D}Y_1\right) - \frac{1}{2}Y_1^2\boldsymbol{D}Y_1$ (A20)
$Y_{2F} = \frac{Y_2}{Y_1}$ (A21)	$\boldsymbol{D}Y_{2F} = \frac{1}{Y_1^2}\left(Y_1\boldsymbol{D}Y_2 - Y_2\boldsymbol{D}Y_1\right)$ (A22)
$Y_{3F} = \frac{Y_3}{Y_1}$ (A23)	$\boldsymbol{D}Y_{2F} = \frac{1}{Y_1^2}\left(Y_1\boldsymbol{D}Y_3 - Y_3\boldsymbol{D}Y_1\right)$ (A24)
$\overline{c} = \mathcal{M}_1$ (A25)	$\boldsymbol{D}\overline{c} = \boldsymbol{D}\mathcal{M}_1$ (A26)
$n_{Tot} = N\overline{c}_1$ (A27)	$\boldsymbol{D}n_{Tot} = N\boldsymbol{D}\overline{c}$ (A28)

References

[1] S. Walston, *The Idiot's Guide to the Statistical Theory of Fission Chains*, August 11, 2008, Technical Report (Lawrence Livermore National Laboratory, 2008).

[2] R. Brun, et al., *ROOT: An object oriented analysis framework*, (CERN, Geneva, Switzerland).