Pade order investigation

Asym vs. FESEM thickness or Rate
-0.5σ to +2 σ, bkg subtract
Run 1 data
x-error bars turned into y-errors
26Jan 2016

Pade approximates

In mathematics a **Padé approximant** is the "best" approximation of a function by a rational function of given order.

Given a function *f* and two integers *m* ≥ 0 and *n* ≥ 1, the *Padé approximant* of order [m/n] is the rational function

$$y = \frac{a_n x^n + a_{n-1} x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0}{b_m x^m + b_{m-1} x^{m-1} + \ldots + b_2 x^2 + b_1 x + b_0},$$

Taylor series expansions are one example of Pade' (Pade (1,0), Pade (2,0), Pade(3,0)…

The typical fitting function $A = \frac{Ao}{1+\gamma T}$ is also Pade' (0,1)

F testing

- The goodness of a fit is typically found by looking at reduced χ^2 or reduced R^2, which show how far the fit is from the data
- It is possible to overfit functions looking only at these “goodness of fit” tests
- An “F-test” can be used to see, to a given degree of confidence, if adding the next order term in an expansion is justified. If the F-test fails, there is a n% chance that the term isn’t needed

$$F = \frac{S_{j-1} - S_j}{S_j}(N-j-1)$$

is distributed as a Fisher–Snedecor $F(1, N-j-1)$ variable if the j^{th} degree is not justified.

From the tabulated value of the F distribution one can then give the prescription in Table 10.2.

Table 10.2. Maximum degree needed in polynomial approximation.

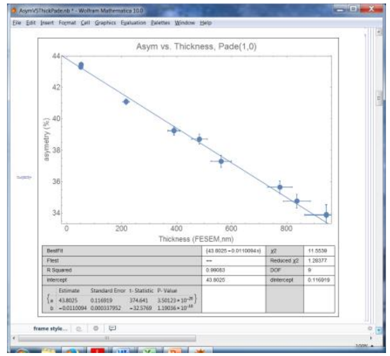
$N-j-1$	2	3	4	6	8	12	20	60	120
Reject j^{th} order to 95% confidence level if F is smaller than	18.5	10.1	7.7	6	5.3	4.7	4.3	4	3.9

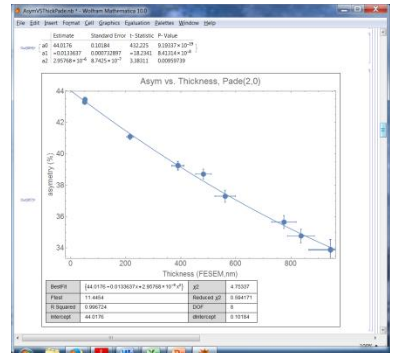
Frederick James,
Statistical methods in
experimental physics 2nd ed.

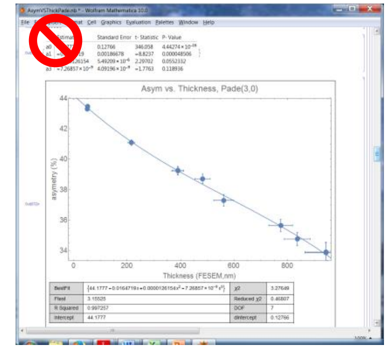
Comparison of fitting functions for asymmetry zero thickness extrapolation

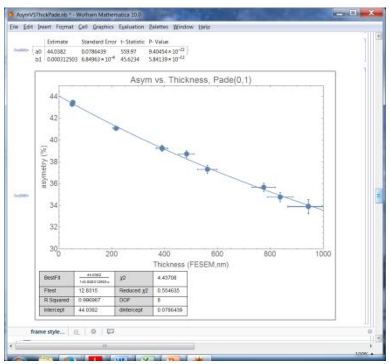
- Two ways to look at data
  - Asymmetry vs. Thickness
    - Asymmetry using Daniel's best data: -0.5σ - +2.0 σ, background subtracted
    - FESEM thickness, 500 nm point fixed to best average
  - Asymmetry vs. Rate

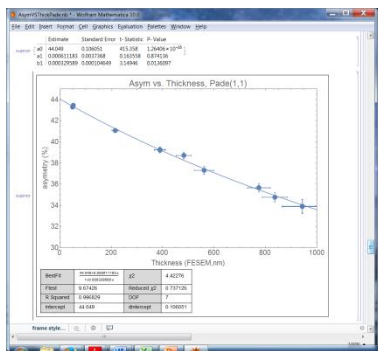
Asymmetry vs. Thickness

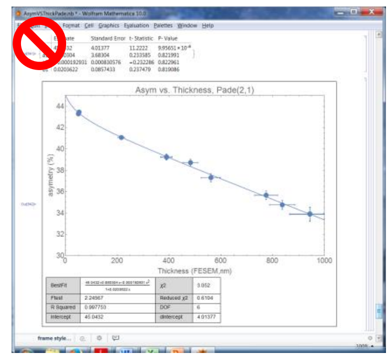


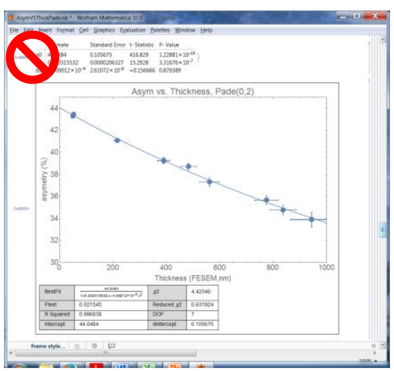


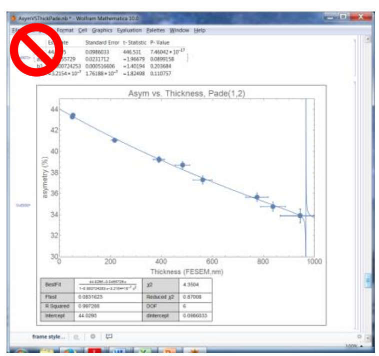










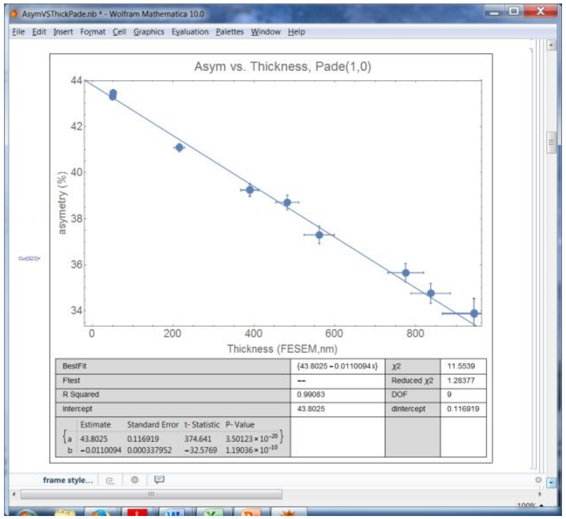


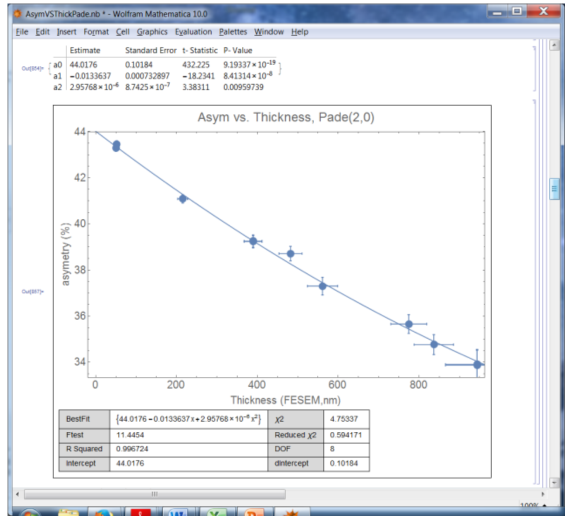
Y error bars have been manipulated to have the x uncertainty included since mathematical typically only fits with y uncertainty. Pade (0,1) (typical fit) used to transform error bars

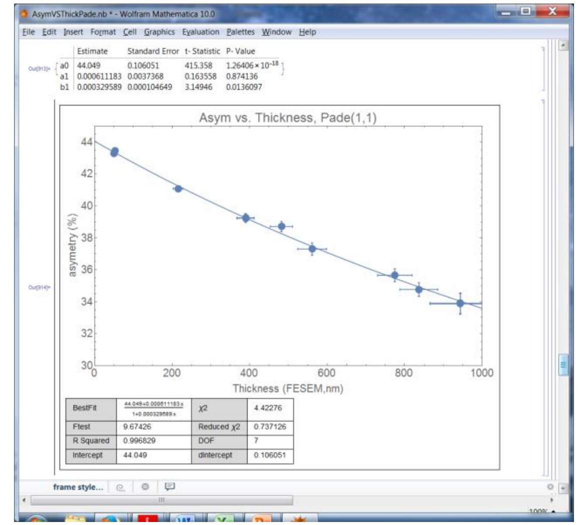
Pade(n,m) orders: Asy vs. Thick

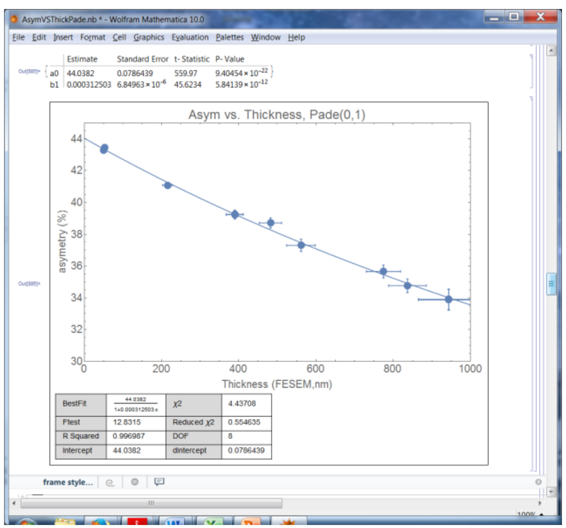
Pade(n,m)	intercept	dA	R^2	red. χ^2	d.o.f.	Ftest
(1,0)	43.8025	0.1169	0.991	1.28	9	--
(2,0)	44.0176	0.1018	0.997	0.594	8	11.45
(3,0)	~~44.1777~~	0.128	0.997	0.546	7	3.15 (rej F test)
(0,1)	44.0382	0.0786	0.997	.554	8	11.23
(0,2)	~~44.0484~~	0.1057	0.997	0.737	7	0.022 (rej ftest)
(1,1)	44.049	0.1061	0.997	0.737	7	9.67
(1,2)	~~44.0295~~	0.0986	0.997	0.870	6	0.083 (rej. Ftest)
(2,1)	~~45.0432~~	4.014	0.9977	0.6104	6	2.25 (rej. Ftest)

Potential fits: not statistically rejected





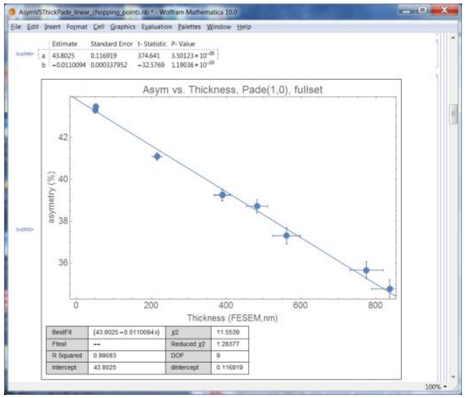


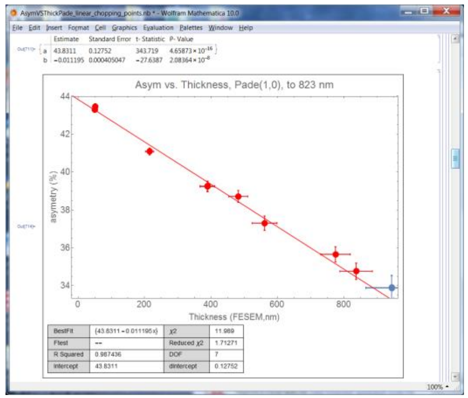


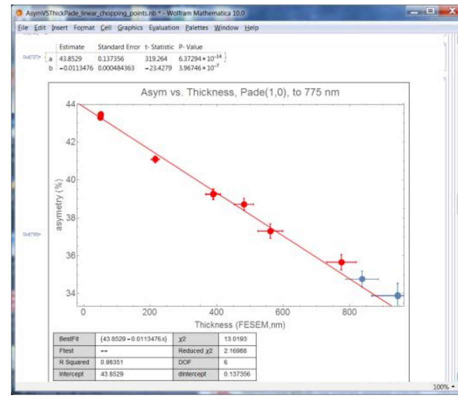
Pade(n,m)	Asym(%)	dA	
(1,0)	43.8025	0.1169	linear
(2,0)	44.0176	0.1018	
(0,1)	**44.0382**	**0.0786**	**Normal fit**
(1,1)	44.049	0.1061	
averaged	**44.0352**		

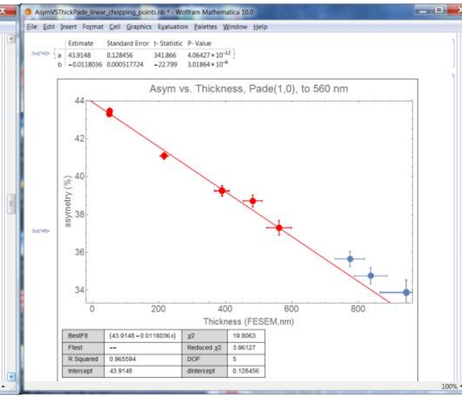
Zero thickness extrapolation largely independent of fit function used, assuming statistically reasonable fits
Error bars shown in y include x errors – need to fix graphs

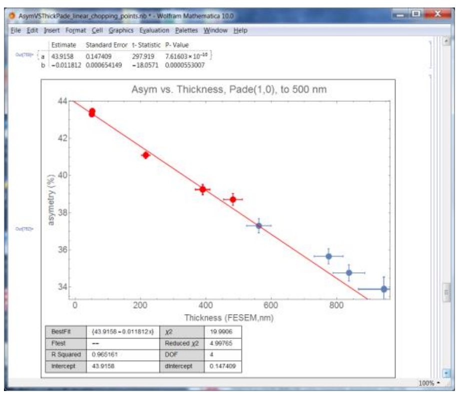
Linear fit: how many points?

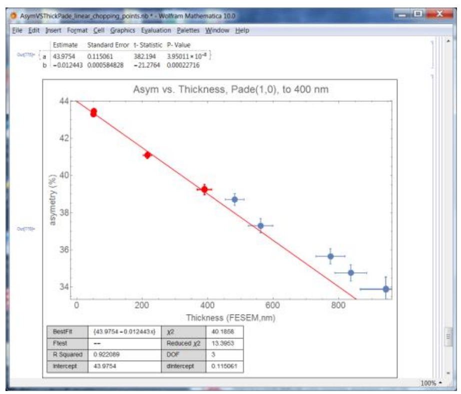








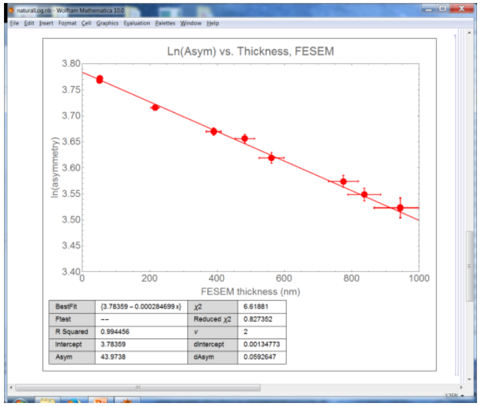


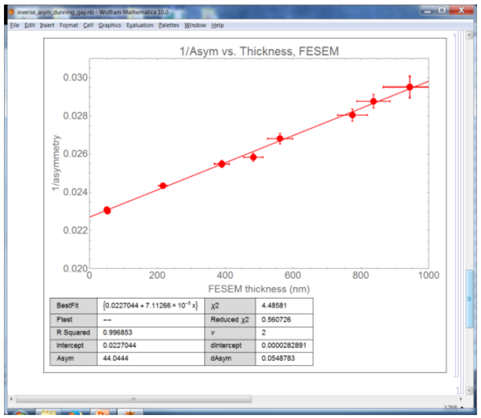


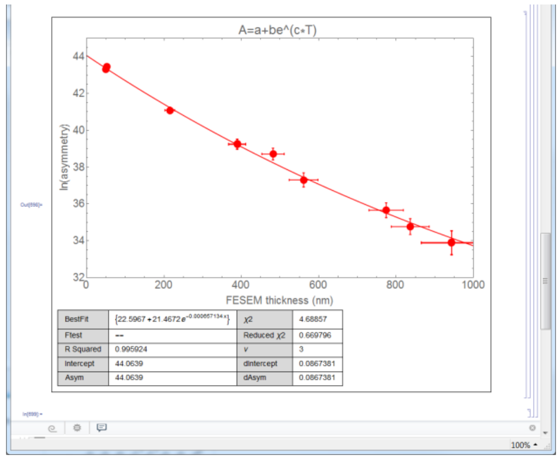
Points kept (of 11)	Asym	dA
5	43.9754	0.115
6	43.916	0.147
7	43.915	0.128
8	43.853	0.137
9	43.833	0.128
all	43.803	0.117

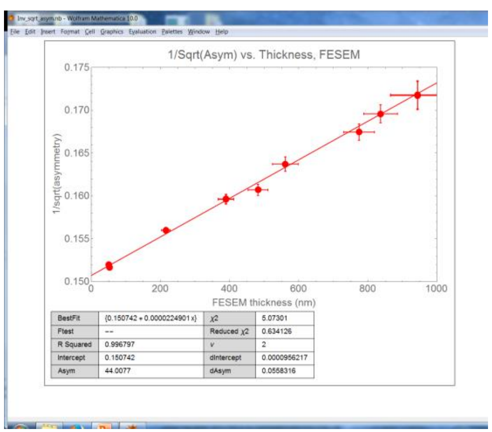
Other functional forms have been used historically to fit asym. vs. thickness

- $\ln(A) = a + b * T$
- $\frac{1}{A} = a + b * T$ (similar to inverting standard)
- $\frac{1}{\sqrt{A}} = a + b * T$
- A=a+b*e^(c*T)



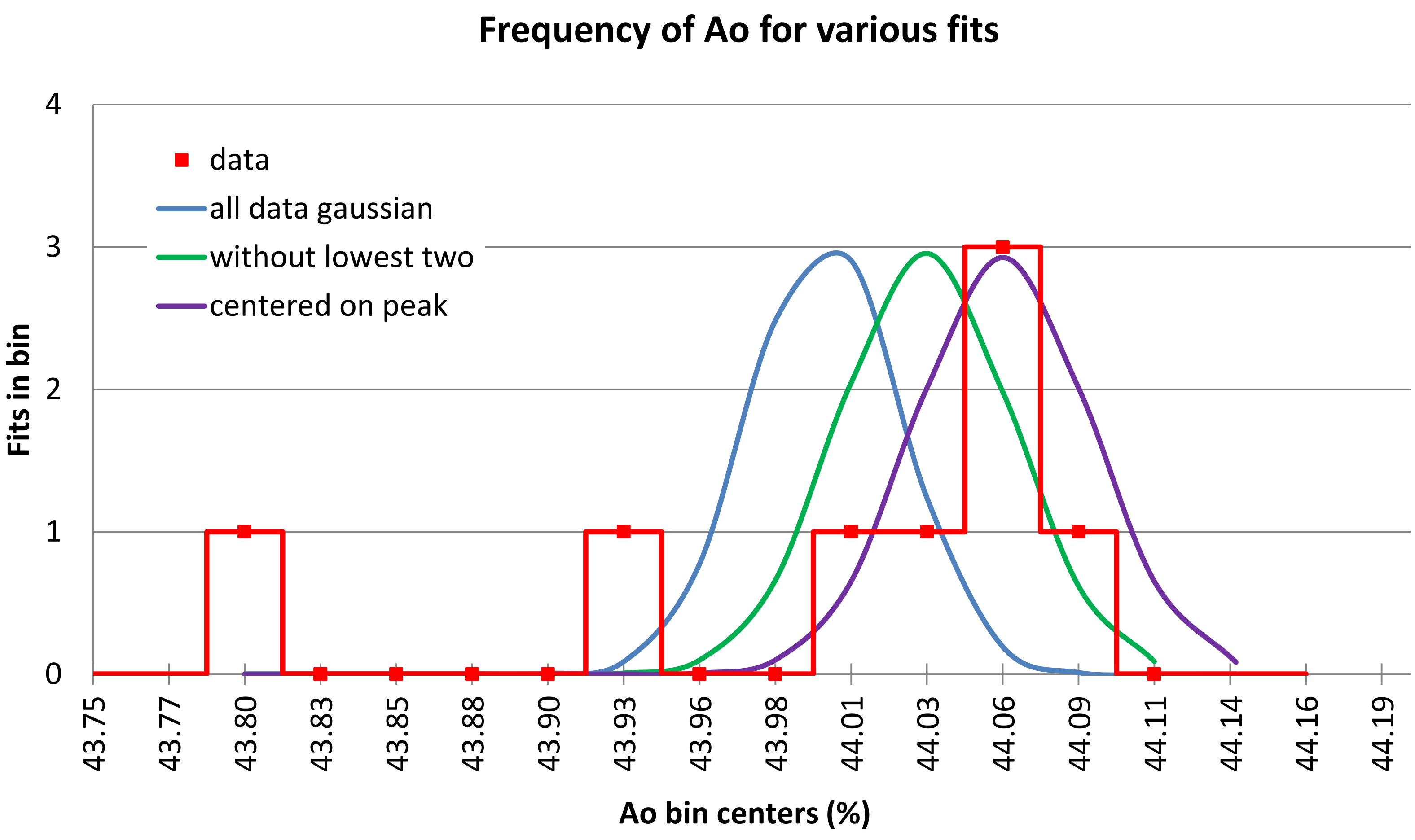






Form	Asym(%)	dA	Red. Chi sq	DOF
ln(A)=a+bT	43.914	0.059	0.827	9
1/A=a+bT	44.044	0.0549	0.561	9
1/√A=a+bT	44.008	0.0558	0.634	9
A=a+b*e^c*T	44.064	0.0867	0.670	8

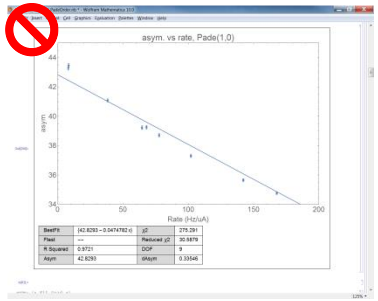
Variation of Ao with fitting function

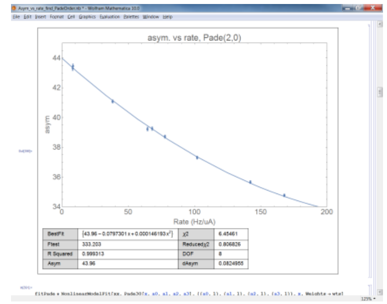


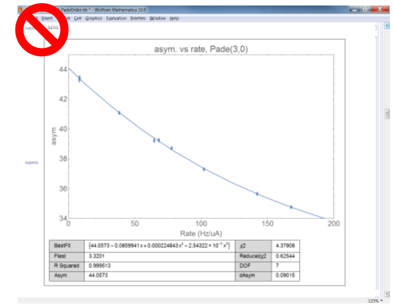
Consider Asym vs. Rate instead?

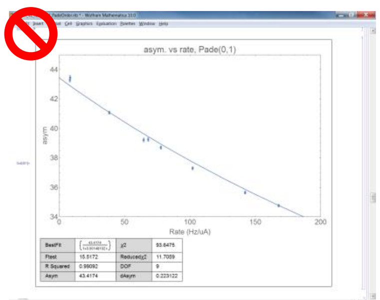
- Plot Asymmetry vs. average detector rate
- Run one data only thus far, “gold” cuts
  - -0.5σ to +2 σ, bkg subtract
  - x-error bars turned into y-errors (using Pade (1,1))
- Fitted Pade(n,m) orders until F test started failing

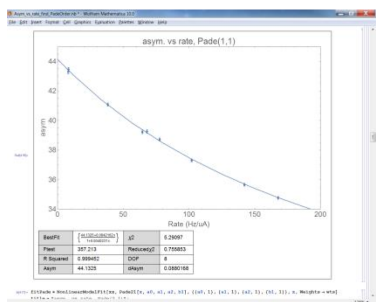
Pade orders: Asym vs. Rate

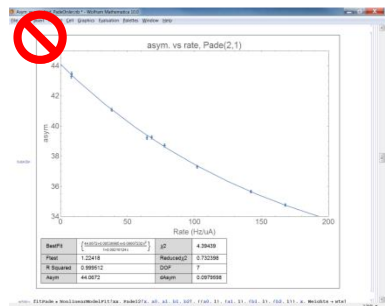


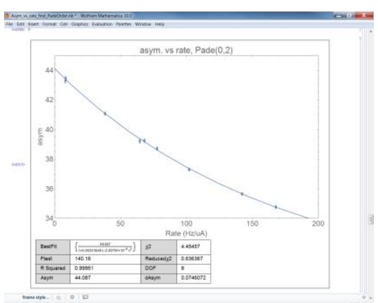


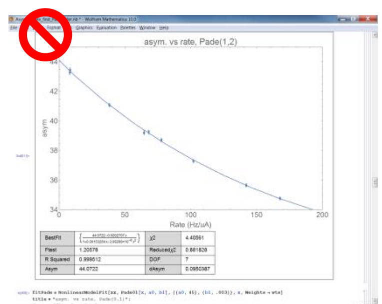


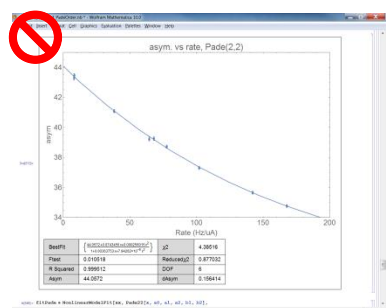


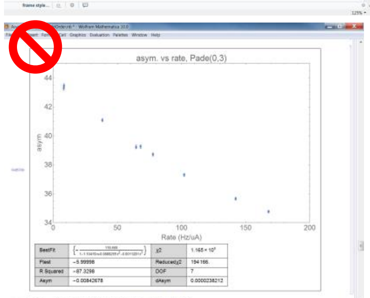












Again, rate uncertainty extrapolated to asym uncertainty for fitting. Rate uncertainties much smaller percentage than thickness uncertainties

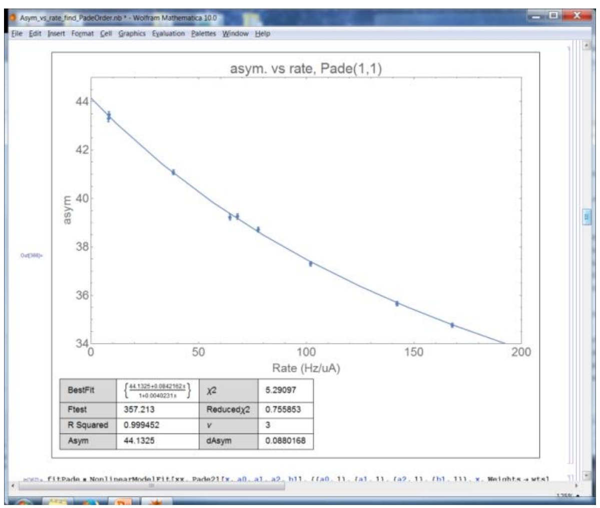
Pade(n,m) orders: A vs rate

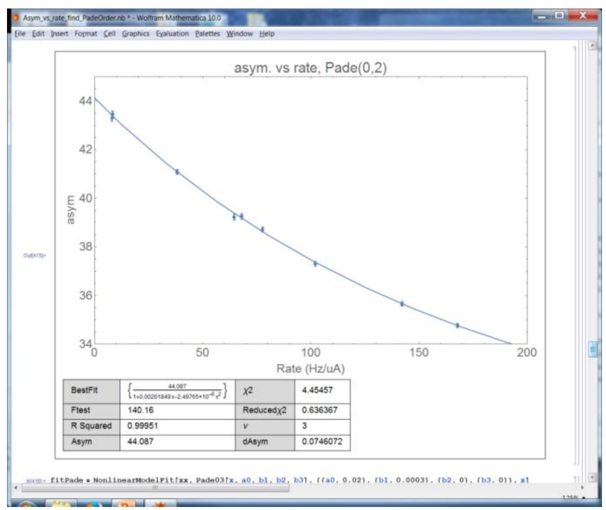
Pade(n,m)	intercept	dA	R^2	red. χ^2	Ftest	D.o.F	
(1,0)	~~42.8~~	.335	.97	31	--	9	Reject chi
(2,0)	43.96	.082	.999	0.807	333	8	
(3,0)	~~44.06~~	.090	.999	0.930	2.84	7	Reject F
(1,1)	44.133	.088	.999	0.756	357	8	
(2,1)	~~44.067~~	.098	.999	0.732	1.22	7	Reject F
(1,2)	~~44.072~~	.095	.999	0.882	1.21	7	Reject F
(0,1)	~~43.42~~	.223	.991	11.7	15.51	9	Reject chi
(0,2)	44.087	0.075	0.999	0.636	140.16	8	
(0,3)	Doesn't converge					7	
(2,2)	~~44.057~~	.156	.999	0.87	.0105	6	Reject F

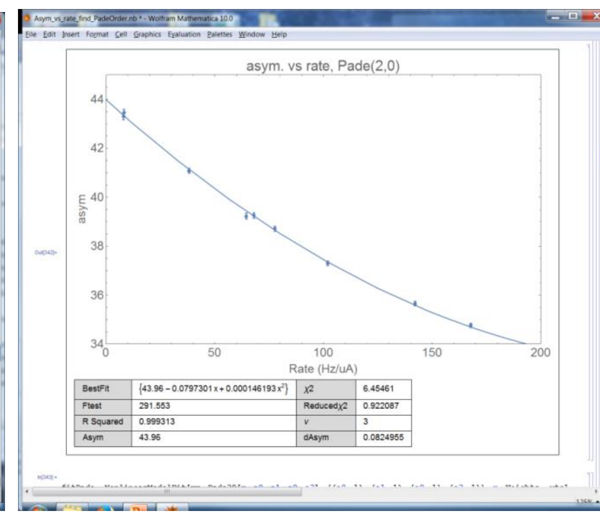
Viable fits: A vs. R

Pade(n,m)	intercept	dA
(2,0)	43.96	.082
(1,1)	44.133	.088
(0,2)	44.087	0.075
average	44.058	

Haven't yet run other forms: square roots, ln, exponential with A vs. R data







Summary of non-excluded fits

vs. R

Pade(n,m)	intercept	dA	R^2	red. χ^2	Dof	Ftest
(2,0)	43.96	.082	.999	0.807	8	333
(1,1)	44.133	.088	.999	0.756	8	357
(0,2)	44.087	0.075	0.999	0.636	8	140.16

vs. T

Pade(n,m)	intercept	dA	R^2	red. χ^2	d.o.f.	Ftest
(1,0)	43.8025	0.1169	0.991	1.28	9	--
(2,0)	44.0176	0.1018	0.997	0.594	8	11.45
(0,1)	44.0382	0.0786	0.997	.554	8	11.23
(1,1)	44.049	0.1061	0.997	0.737	7	9.67

vs. T

Form	Asym(%)	dA	Red. Chi sq	DOF
ln(A)=a+bT	43.914	0.059	0.827	9
1/A=a+bT	44.044	0.0549	0.561	9
1/√A=a+bT	44.008	0.0558	0.634	9
A=a+b*e^c*T	44.064	0.0867	0.670	8

Conclusions

- Fitting A vs. T: std. fit form gives lowest uncertainties
- Use Pade analysis, F-testing to determine other viable functional forms
- Fitting A vs. Rate: 3 forms have viable fits, uncertainties all comparable to best in A vs. T
- Translating x uncertainties to y axis (done by root, this mathematica analysis) requires model dependence, likely not a large error factor.