Empirical Bayes for data fusion at the National Ignition Facility (NIF)

CASIS 2017

May 17, 2017

Florin Munteanu

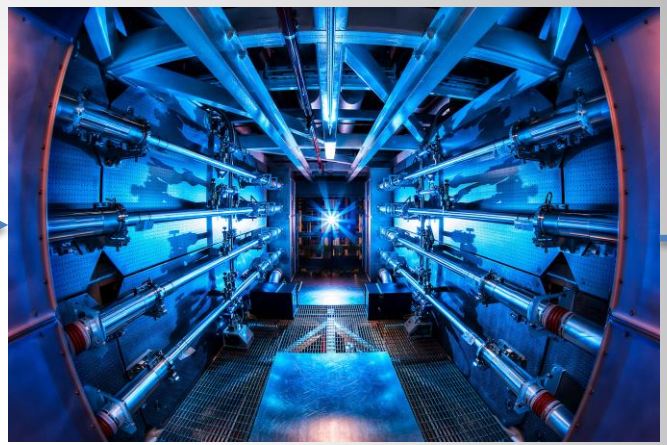
Lawrence Livermore National Laboratory

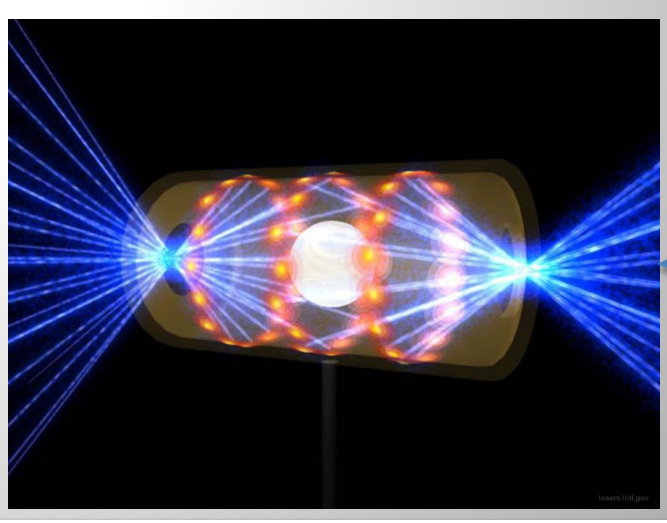
LLNL-PRES-728328

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

National Ignition Facility

World's most energetic laser (192 laser beams, a total of 1.8 MJ Energy)





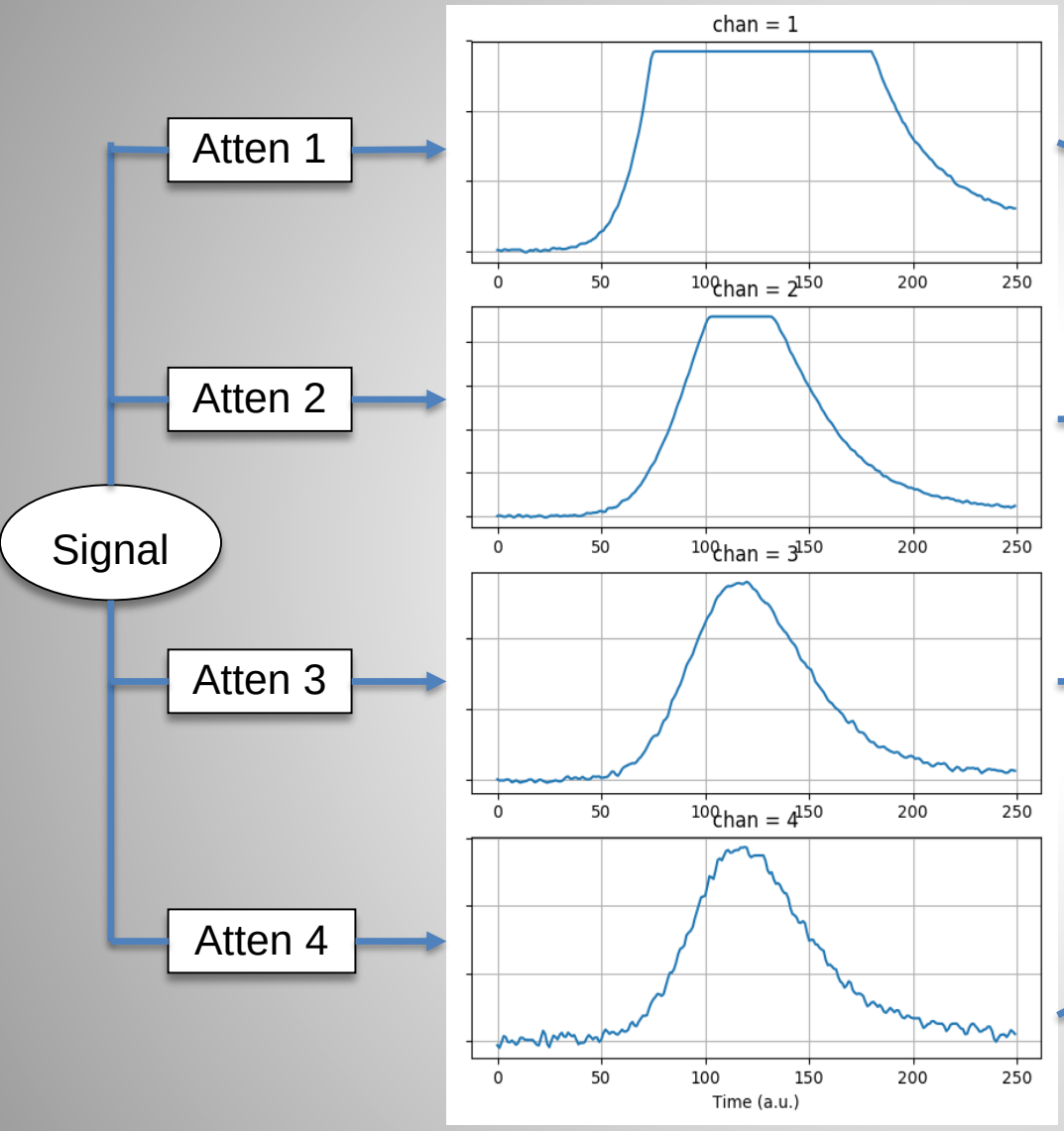
A large amount of radiation is produced and recorded.

Data analysis offers insight into the state of the matter.

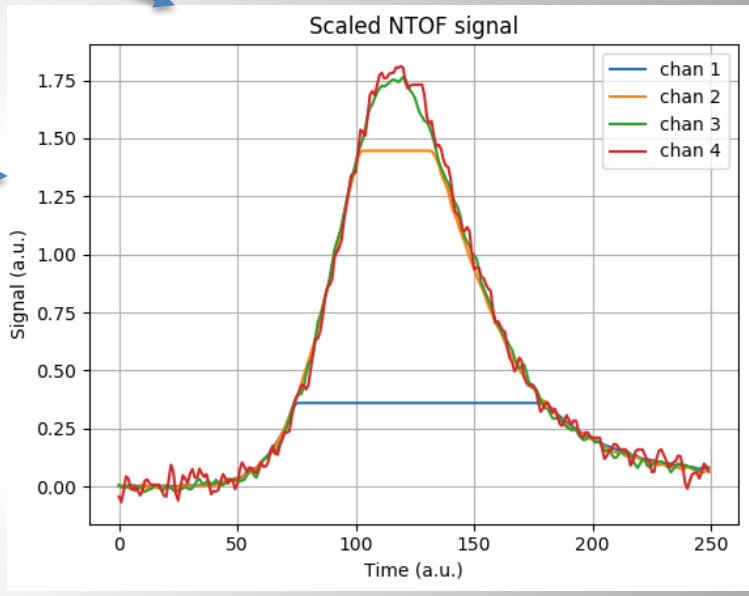
Data is corrupted by noise and instrumentation mis-calibration.

Inferring the true original signal is key to correct physical interpretation.

Neutron Time of Flight (NTOF) signal



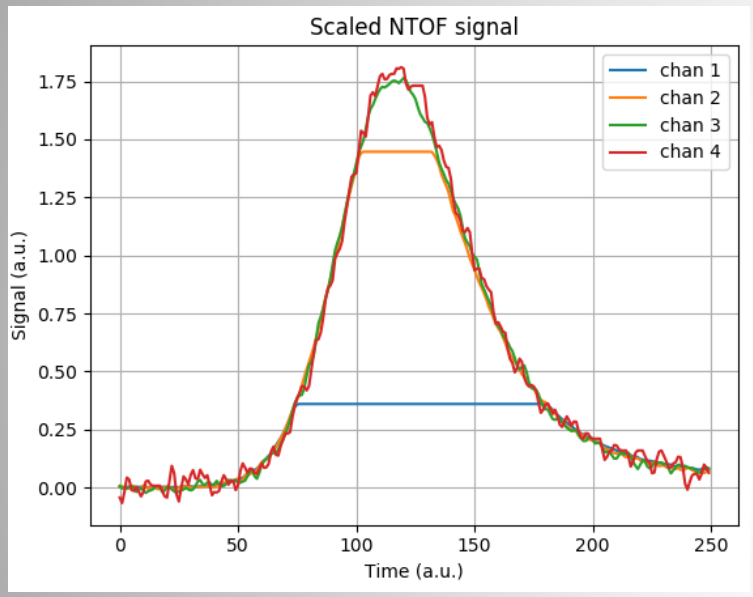
The signal is recorded by multiple scopes at different attenuations.

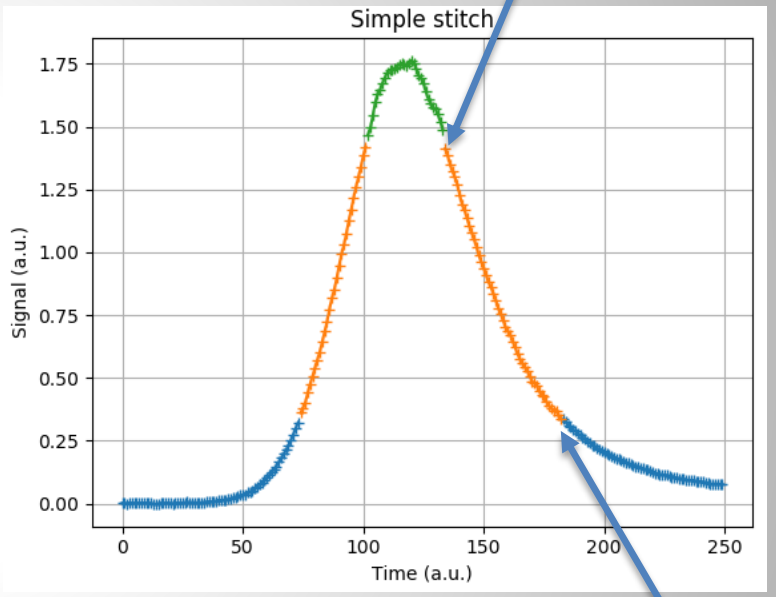


- Goal: Retrieve the original signal!
- The fiducial ensures good registration.
- Attenuator values measured with errors.

NTOF simple stitch

For each portion select the trace with best signal to noise ratio (SNR).





Gap

Overlap

Main issues:

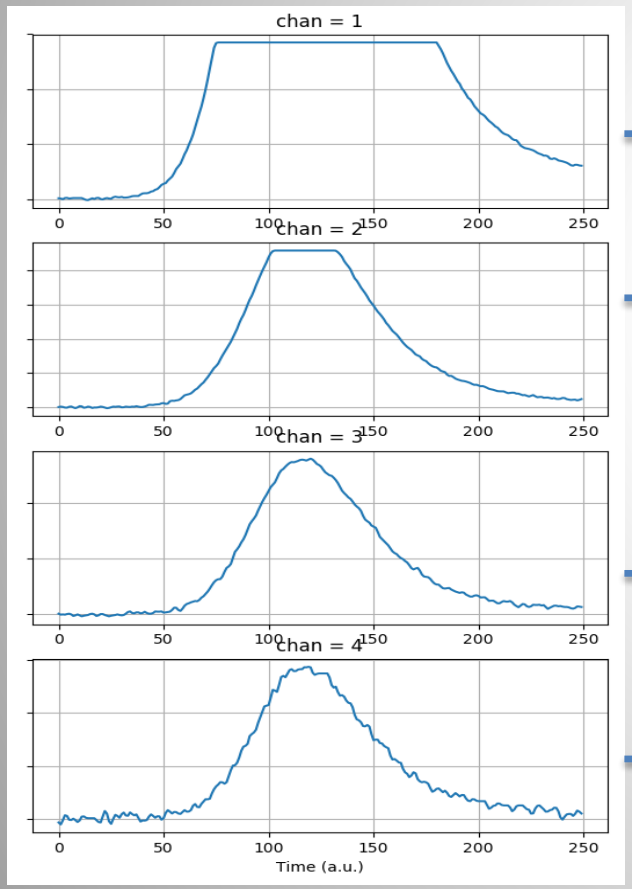
1. Many available points from lower SNR traces not used -> noise.
2. The Attenuator mis-calibrations are evident in the gaps/overlaps between different traces.

Alternative: NTOF weighted stitch

$$\hat{s} = \sum_{i=1}^{M} w_i y_i = \sum_{i=1}^{M} w_i (A_i s + n_i)$$

$$E[\hat{s}] = s \rightarrow \sum_{i=1}^{M} w_i A_i = 1$$

$$Var[\hat{s}] \rightarrow min$$



X

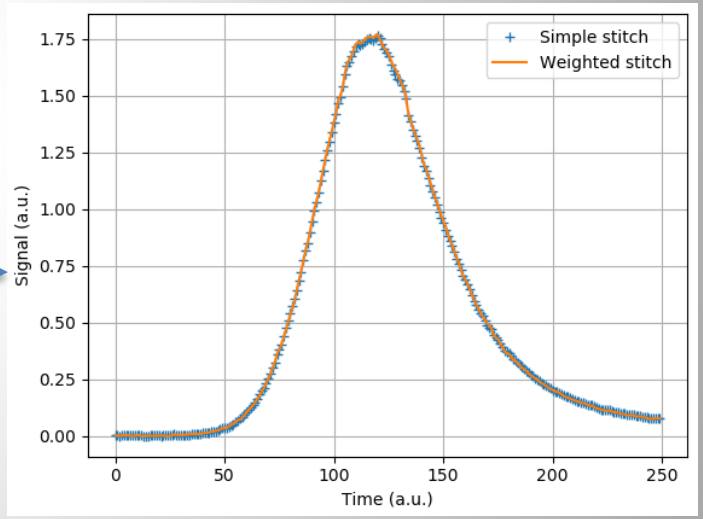
X

X

X

$$w_i = \frac{A_i}{\sigma_i^2} \frac{1}{\sum_{k=1}^{p} \left[\frac{A_k}{\sigma_k}\right]^2}$$

+



Makes use of all available data

Still does not solve the attenuator calibration problem.

Proposed: Data fusion

Signal at time “k” from the scope “i” (N data, M signals)

$$y_i(k) = A_i s(k) + n_i(k)$$

$$n_i(k) = N(0, \sigma_i^2)$$

$$y_i \equiv [y_i(1), y_i(2), ..., y_i(N)]^T$$

$$n_i \equiv [n_i(1), n_i(2), ..., n_i(N)]^T$$

$$\underset{NM\times 1}{\begin{bmatrix} y_1 \\ y_2 \\ .... \\ y_M \end{bmatrix}} = \underset{NM\times N}{\begin{bmatrix} A_1\mathbf{I_N} \\ A_2\mathbf{I_N} \\ ..... \\ A_M\mathbf{I_N} \end{bmatrix}} \underset{N\times 1}{s} + \underset{NM\times 1}{\begin{bmatrix} n_1 \\ n_2 \\ .... \\ n_M \end{bmatrix}}$$

Known values:

- measured traces - y_i
- measurement noise - σ_i
- initial A_i^0 estimates

Variables to determine:

- true signal - s(k)
- true attenuator values - A_i

Signal distribution

Bayes theorem

$$P(s|y_1,..,y_M,\sigma_1,..,\sigma_M,A_1,..,A_M) \propto P(y_1,..,y_M|s,\sigma_1,..,\sigma_M,A_1,..,A_M)P(s)\prod_{i=1}^{M}P(A_i)\prod_{i=1}^{M}P(\sigma_i)$$

$$\begin{cases} P(y_1,..,y_M|s,\sigma_1,..,\sigma_M,A_1,..,A_M) & = \frac{1}{(2\pi)^{\frac{NM}{2}}\prod_{i=1}^{M}\sigma_i^N}e^{-\sum_{i=1}^{M}\frac{(y_i-A_is)^T(y_i-A_is)}{2\sigma_i^2}} \\ P(A_i) & = \frac{1}{\sqrt{2\pi}\sigma_A}e^{-\frac{(A_i-A_i^0)^2}{2\sigma_A^2}} \\ P(\sigma_i) & = 1 \end{cases}$$

A_i^0 measured attens

σ_A unknown!

Separate estimation compounds the errors in Ai

$$P(s) \propto e^{-\frac{1}{2\lambda^2}\sum_{k=1}^{N}\left[s(k)-s^Th_k\right]^2}$$

'h_k' suitable filter (Savitzky–Golay, for example)

$$\hat{s},\hat{A}_i,\hat{\sigma}_A = \underset{s,A_i,\sigma_A}{argmin}[-\log(P(s|y_i,\sigma_i,A_i,\sigma_A))]$$

Empirical Bayes for σ_A determination

Calculate the marginal distribution over A_i

$$m(s|y_i,\sigma_i,\sigma_A) = \prod_{i=1}^{M} \int P(s|y_i,\sigma_i,A_i,\sigma_A)dA_i$$

$$P(s|y_i,\sigma_i,A_i,\sigma_A) \propto \frac{1}{(2\pi\sigma_A^2)^{\frac{M}{2}}} \prod_{i=1}^{M} e^{-\left[\sum_{k=1}^{N} \frac{[y_i(k)-A_i s(k)]^2}{2\sigma_i^2} + \frac{(A_i-A_i^0)^2}{2\sigma_A^2}\right]}$$

Minimize with respect σ_A

$$-\log\left[m(s|y_i,\sigma_i,\sigma_A)\right] = M\log(\sigma_A) + \sum_{i=1}^{M} \frac{{A_i^0}^2}{2\sigma_A^2} - \sum_{i=1}^{M} \frac{\left[\sum_{k=1}^{N} \frac{s(k)y_i(k)}{2\sigma_i^2} + \frac{A_i^0}{2\sigma_A^2}\right]^2}{\sum_{k=1}^{N} \frac{s(k)^2}{2\sigma_i^2} + \frac{1}{2\sigma_A^2}}$$

Iterative algorithm

1. Start with the best weighted stitched signal

2. Estimate the unknown signal ‘s’

$$\hat{s} = \underset{s}{argmin}[-\log(P(s|y_i, \sigma_i, A_i))]$$

3. Estimate the value of σ_A by minimizing the marginal

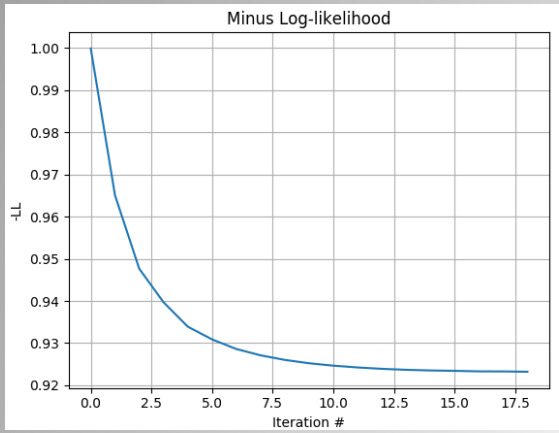
$$\hat{\sigma_A} = \underset{\sigma_A}{argmin}[-\log(m(s|y_i, \sigma_i, A_i, \sigma_A))]$$

4. Estimate the attenuations

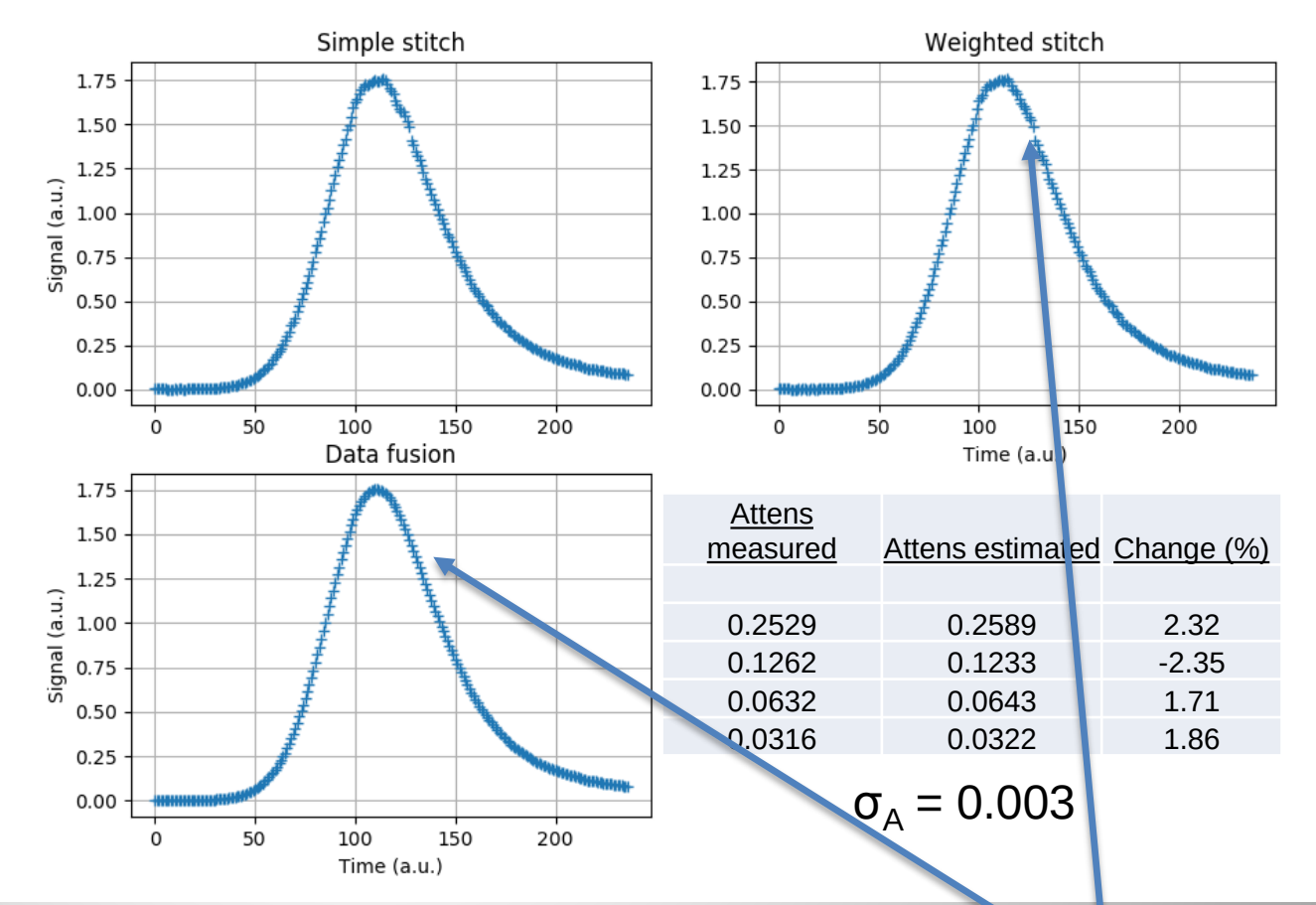
$$\hat{A}_i = \underset{A_i}{argmin}[-\log(P(A_i|s, y_i, \sigma_i, \sigma_A))] \Longleftrightarrow A_i = \frac{\sum_{k=1}^{N} \frac{s(k)y_i(k)}{\sigma_i^2} + \frac{A_i^0}{\sigma_A^2}}{\sum_{k=1}^{N} \frac{s(k)^2}{\sigma_i^2} + \frac{1}{\sigma_A^2}} \xrightarrow[\sigma_A \to 0]{} A_i^0$$

5. Repeat 2-4

Results – experimental data



Converged after 12 iterations

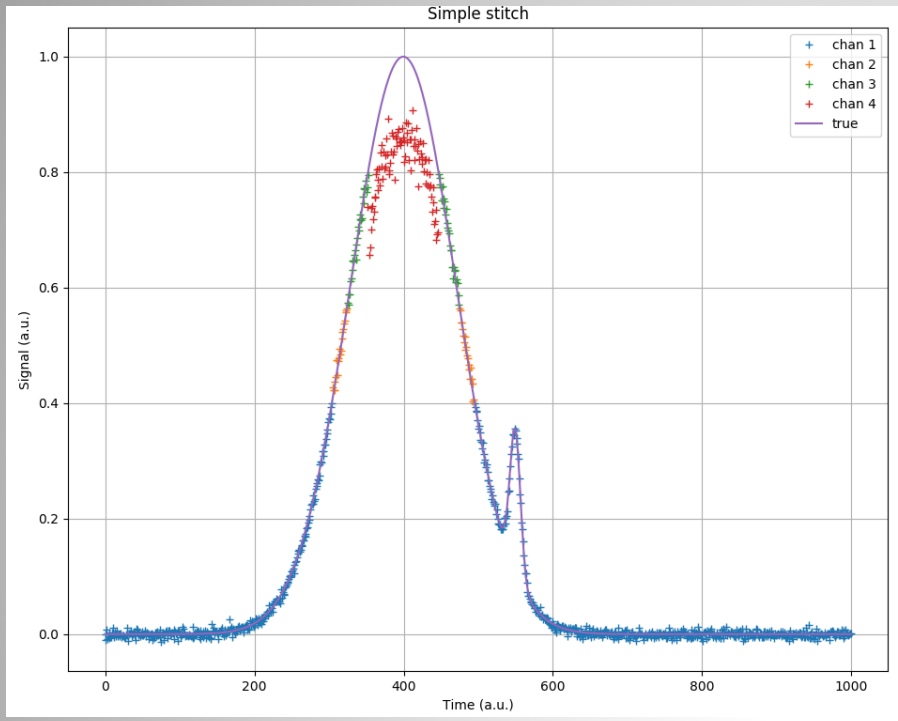


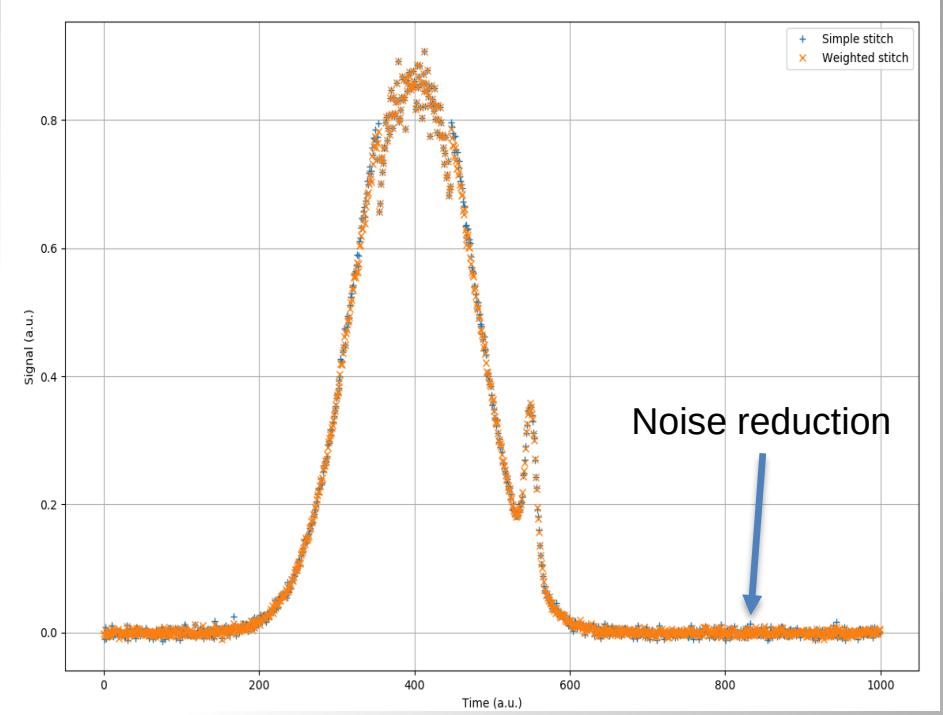
Attens measured	Attens estimated	Change (%)
0.2529	0.2589	2.32
0.1262	0.1233	-2.35
0.0632	0.0643	1.71
0.0316	0.0322	1.86

$\sigma_A = 0.003$

Finding the correct attenuator values removed the gaps in data

Result – synthetic data

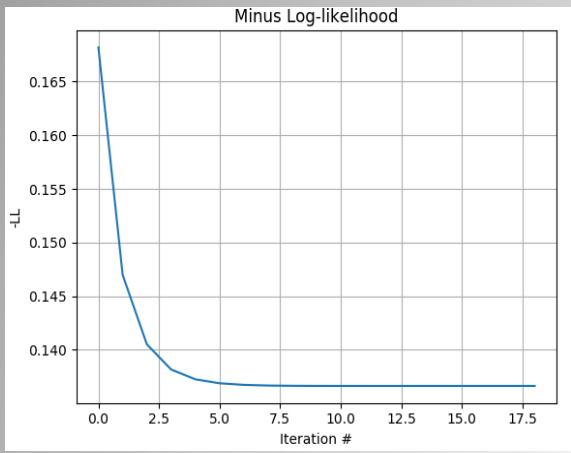




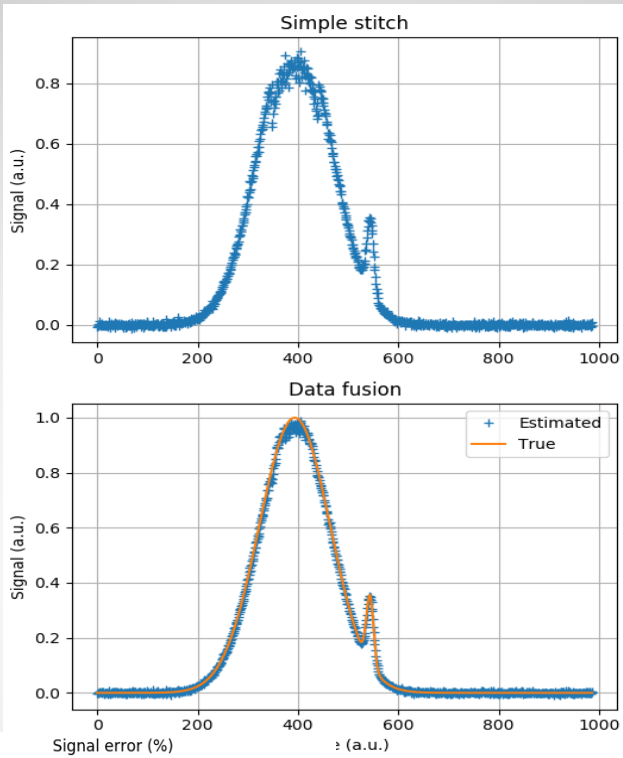
Simple stitch is badly corrupted when attenuator mis-calibrations are present (red trace lower than its true value!).

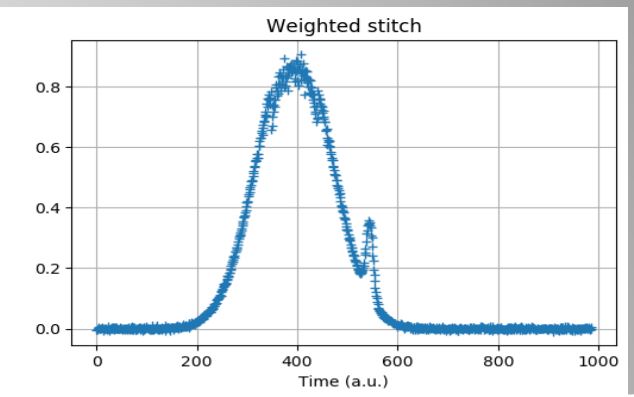
Weighted stitch reduces the random noise but is still affected by attenuator mis-calibration.

Result – synthetic data



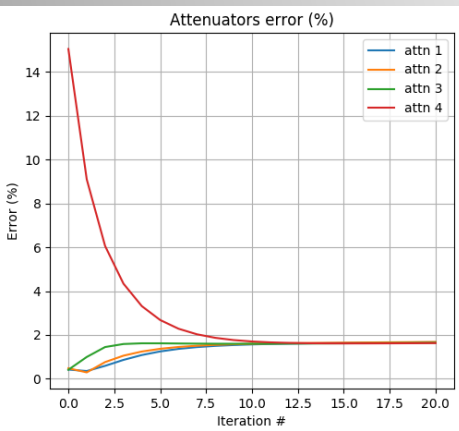
Converged after 5 iterations

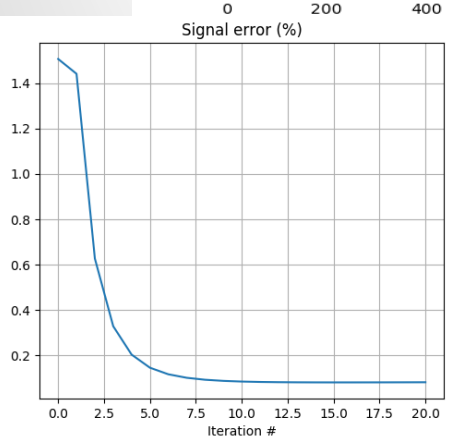




Attens true	Attens measured	Attens determined	Error (%)
0.9958	1	1.0124	-1.67
0.7034	0.7	0.7151	-1.66
0.5021	0.5	0.5103	-1.63
0.1738	0.2	0.1767	-1.67

σA true	σA determined
0.01	0.012





Finding the correct attenuator values removed the gaps/overlap in data

Summary

- Hardware mis-calibration and instrumentation noise affects negatively the quality of the original signal estimation.
- Simultaneous estimation of the signal and hardware calibration values is necessary.
- Empirical Bayes approach offers a valid alternative for solving a difficult optimization problem.

Lawrence Livermore
National Laboratory