Extracting Information Content within Noisy, Sampled Profile Data from Charged Particle Beams*

Christopher K. Allen
Willem Blokland
Sarah Cousineau
John Galambos

HB2008 Workshop
Nashville Tennessee
August 25 – 29, 2008

*This work was supported by SNS through UT-Battelle, LLC, under contract DE-AC05-00OR22725 for the U.S. DOE.

Outline

- **Profile Data**
- **The Problem**
- **Model of Measurement Random Process**
- **Computations of Beam Position μ and Size σ**
- **Conclusions**
- **Open Questions**

Profile Data
1D Projections of the Beam Distribution

Say $f(x,y)$ is the transverse beam distribution.

The projection, or *profile*, of f in the horizontal plane is

$$f_x(x) = \int_{-L/2}^{+L/2} f(x,y)\, dy.$$

When measuring the projection f_x is *sampled* at axis locations

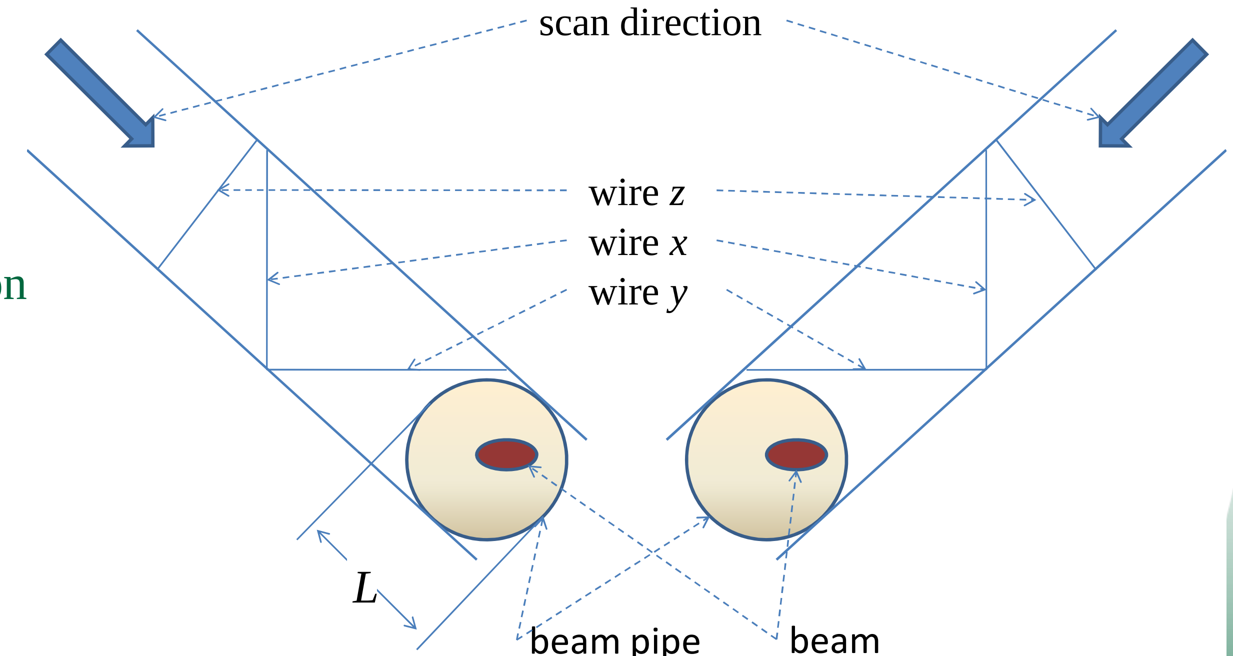
$$x_k = kh$$

with constant sampling intervals h, and N samples.

Thus, the *sampled profile* is given as the discrete set

$$\{f_{x,k}\} = \{f_x(x_1), f_x(x_2), \ldots, f_x(x_N)\}$$

Prototypical Profile Device – The Wire Scanner



We drop the subscript x from here out

Profile Data
Objectives: What Do We Want?

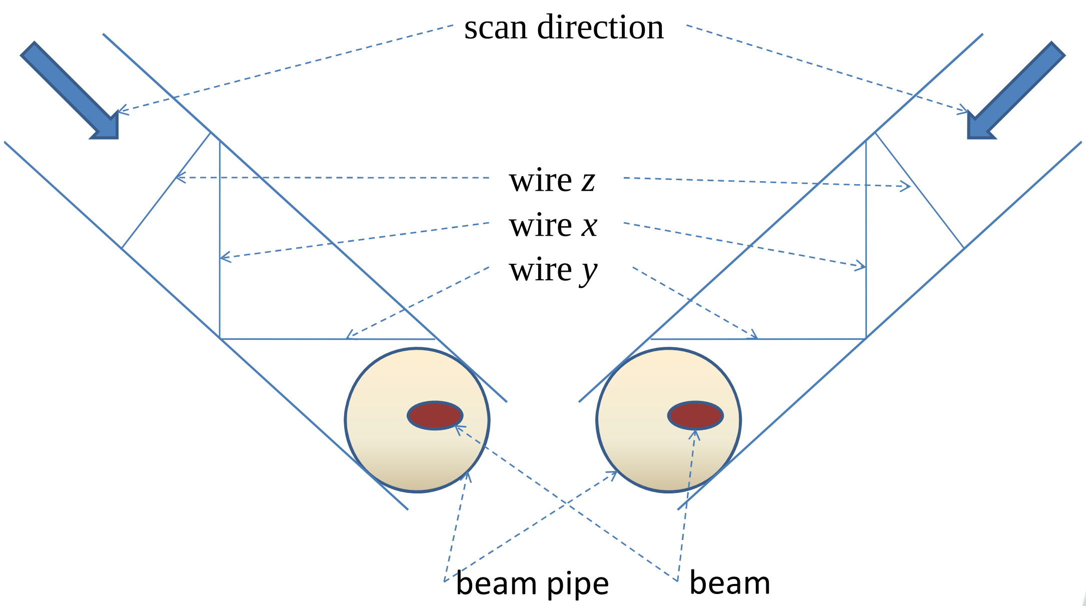
At this point, we only want two quantities from the measured data

- Beam Position μ
- Beam Size σ

This seemed like a reasonable expectation, however…

- The data are noisy
- Beam jitter
- Missing data points
- Many data sets

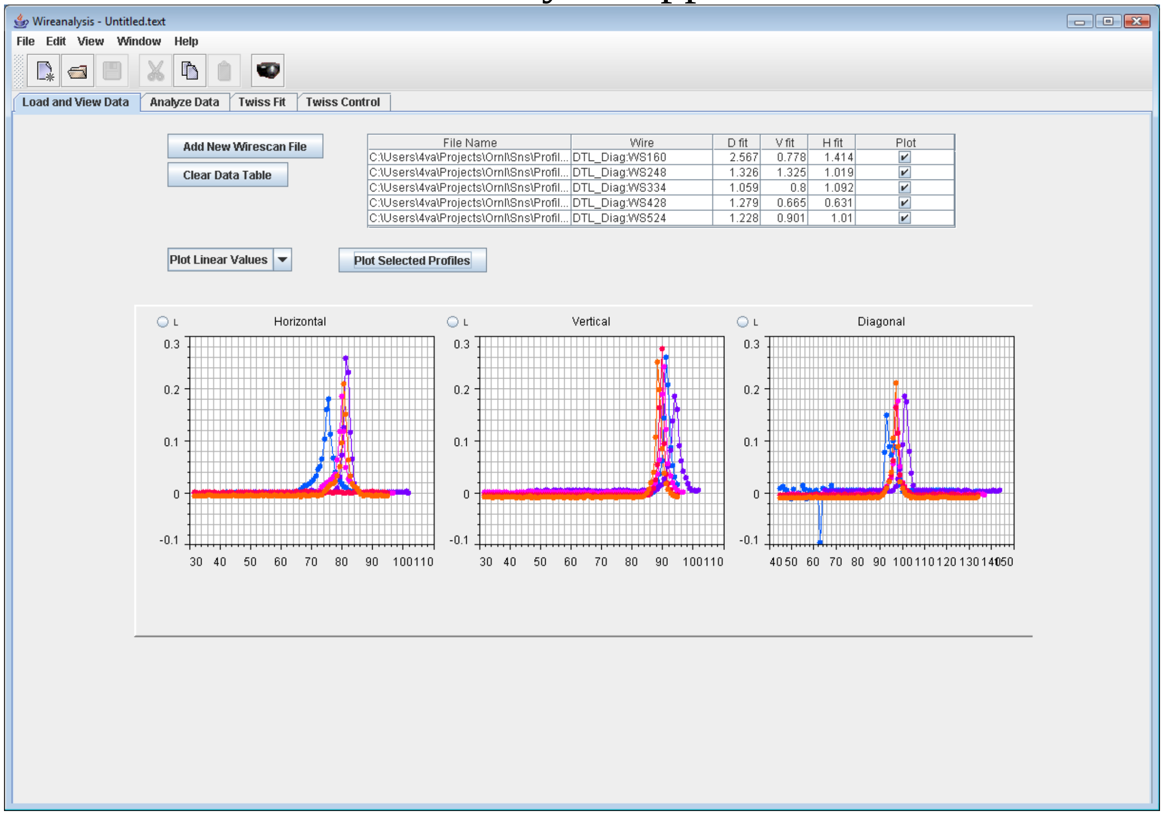
Wire Scanner



The Problem

Processing many data sets for Simple Parameters

SNS Wire Analysis Application



Original Goal: Estimate Twiss parameters

Within SNS CCL:

- First compute beam sizes
  - 5 wire scanners with 3 wires
  - 15 data sets of ~150 samples each
- Most effort is manual data processing
  - Looking for bad data sets
  - Removing errant data points
  - Clipping noise baseline
  - Reject bad fits, Etc.
- We just want 10 numbers !

Can computation of the beam position and size from profile data be automated?

Beam Properties and Measurement Model

Computing Beam Position μ and Size σ

- If we know the sampled profile f_k exactly, normalizing by the step length h the position μ and size σ are approximated*

$$\mu = \frac{1}{S}\sum_{k=1}^{N} k\, f_k\,, \quad \sigma = \left[\frac{1}{S}\sum_{k=1}^{N}(k-\mu)^2 f_k\right]^{1/2}, \quad \text{where} \quad S = \sum_{k=1}^{N} f_k$$

- However, we do not know the $\{f_k\}$.

The Measurement Model

- Each measurement m_k contains noise from electronics, jitter, etc.
- Model as Gaussian white-noise process W with mean B and variance V**

$$m_k = f_k + W_k$$

measurement random process

- We **must** account for this noise when approximating μ and σ.

* That is, μ and σ are in units of step length h – not necessarily integers

**The noise can be characterized by a calibration experiment (w/o beam)

Measurement Random Process

- Gaussian noise process p.d.f. is $P(W=w)=\dfrac{1}{\sqrt{2\pi V}}e^{-\frac{(w-B)^2}{2V^2}}$
  - Then probability that measurement process M_k has value m_k is the same as the probability that noise process W has value $m_k - f_k$

$$P(M_k = m_k) = \frac{1}{\sqrt{2\pi V}} e^{-\frac{(m_k - f_k - B)^2}{2V^2}}$$

  - Assuming independent events, probability (p.d.f.) of the data set $\{m_k\}$ is

$$P\left(\{M_k\} = \{m_k\}\right) = \frac{1}{(2\pi)^{N/2} V^N} e^{-\frac{1}{2V^2}\sum_{k=1}^{N}(m_k - f_k - B)^2}$$

This is the p.d.f. of our *measurement random process*

Technique #1
Direct Computation with Measurement Data

- Inspecting $P(\{m_k\})$, the sample set $\{f_k\}$ that maximizes the probability of obtaining measurement set $\{m_k\}$ is $\boldsymbol{f_k = m_k - B}$ **for all** $\boldsymbol{k}$
  - Compute position μ and size σ directly from measurement data $\{m_k - B\}$
  - However, $\{m_k\}$ is a sampling from a random process, we must characterize statistical properties of computations involving these samples…

Defining computations*

$$S_n(\bar{k}) \equiv \sum_{k=1}^{N} (k-\bar{k})^n f_k$$

$$\tilde{S}_n(\bar{k}) \equiv \sum_{k=1}^{N} (k-\bar{k})^n (m_k - B)$$

*Recall $\mu = S_1(0)/S_0(0)$
and $\sigma^2 = S_2(\mu)/S_0(0)$

We get

$$\text{Mean}[\tilde{S}_n(\bar{k})] = S_n(\bar{k})$$

$$\text{Var}[\tilde{S}_n(\bar{k})] = N_n(\bar{k})V$$

where $N_n(\bar{k}) \equiv \sum_{k=1}^{N} (k-\bar{k})^n$

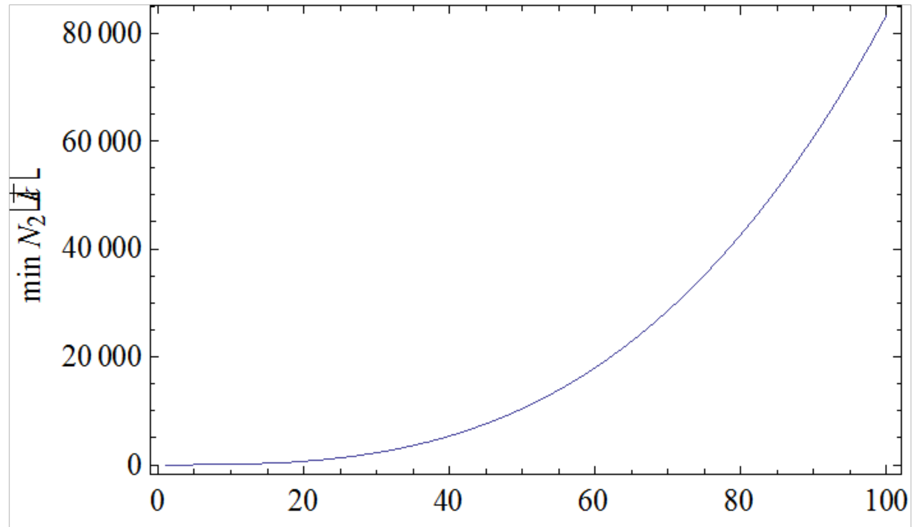
Technique #1
Direct Computation with Measurement Data

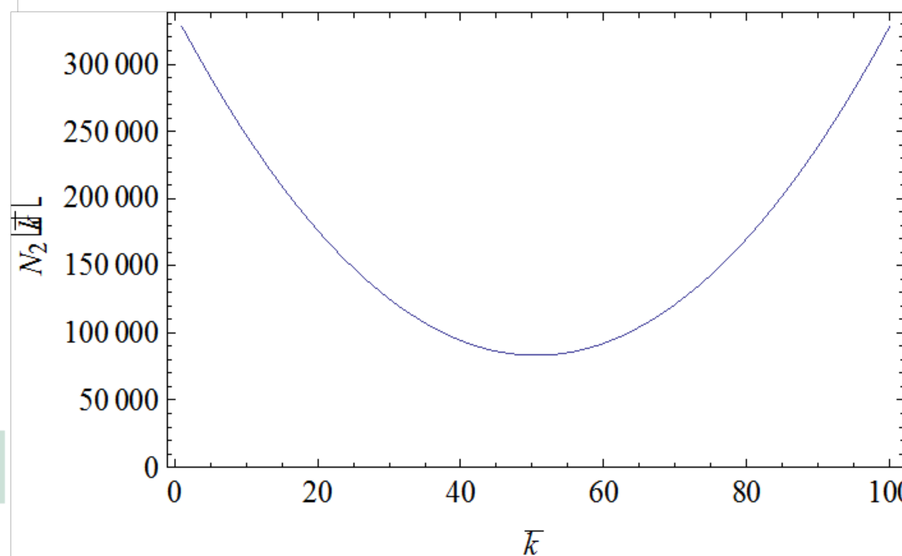
$$N_n(\bar{k}) \equiv \sum_{k=1}^{N} (k-\bar{k})^n$$

- Approximate μ and σ with measured

$$\mu \approx \tilde{S}_1(0)/\tilde{S}_0(0)$$
$$\sigma^2 \approx \tilde{S}_2(\mu)/\tilde{S}_0(0)$$

  which are the expected values
  - If W is ergodic these approximations get better as $N \to \infty$
- The variances in these values are dominated by $N_1(0)V$ and $N_2(\mu)V$
  - N_n is exponentially increasing as $N \to \infty$
  - N_n is huge for typical measurements
- Although the expected values are exactly μ and σ, the variances become enormous as $N \to \infty$.
  - $V < \sigma \times 10^{-7}$ for ~10% accuracy





Is there any way around this??

Technique #2
Assuming a Known Profile for f_k

- **Assume** a profile for $f(x)$ which is parameterized by μ and σ
  - Apply Bayesian techniques to estimate parameters μ and σ
  - **Example**: Take f as a **Gaussian** – must add amplitude parameter A

$$f(x;A,\mu,\sigma) = Ae^{-\frac{(x-h\mu)^2}{2(h\sigma)^2}} \quad \text{then} \quad f_k(A,\mu,\sigma) = Ae^{-\frac{(k-\mu)^2}{2\sigma^2}}$$

  - We want to know (A,μ,σ) given $\{m_k\}$ - Bayes says that

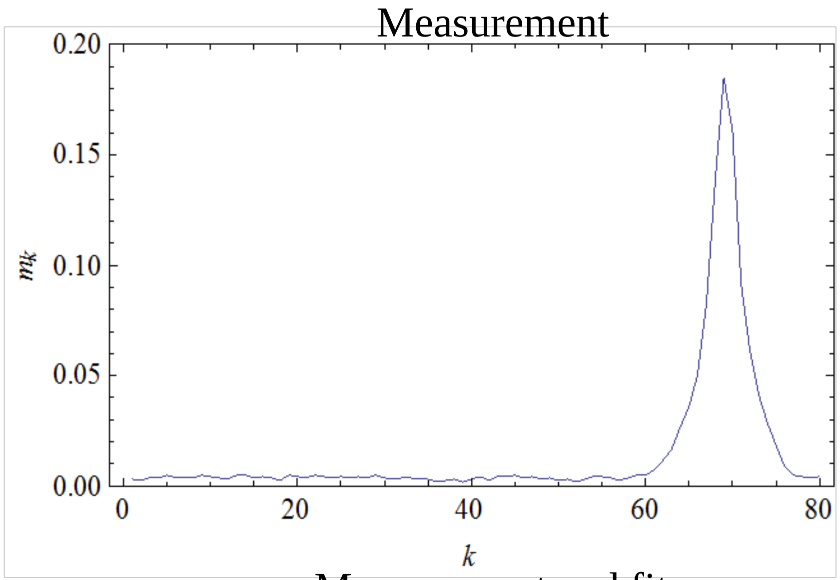
$$P\left(A,\mu,\sigma \mid \{m_k\},B,V\right) \propto P\left(\{m_k\} \mid A,\mu,\sigma,B,V\right)P(A,\mu,\sigma)$$

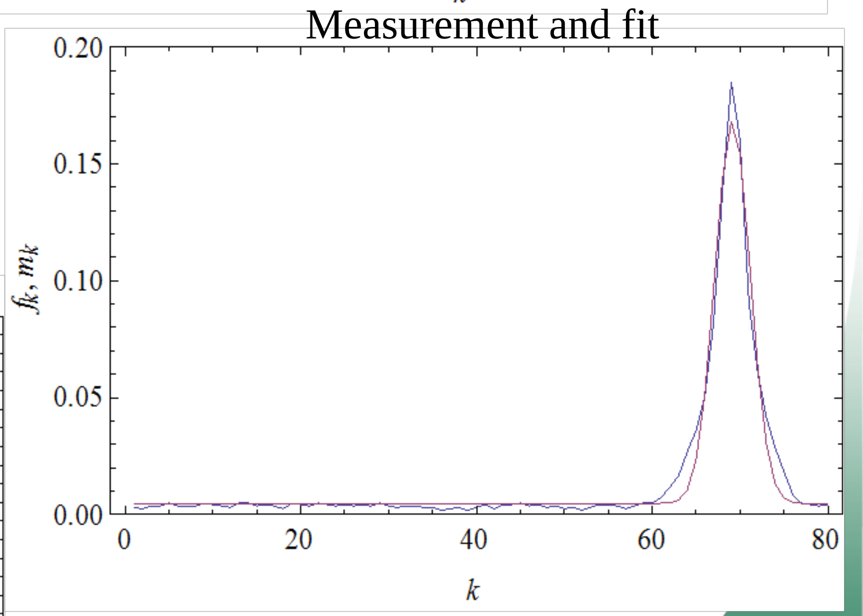
  - Look for A, μ, and σ that maximize $P(\{m_k\}|A,\mu\sigma,B,V)P(A,\mu,\sigma)$
    - We know $P(\{m_k\}|A,\mu\sigma,B,V)$
    - The *prior distribution* $P(A,\mu,\sigma) = P(A,\sigma)\ P(\mu)$ can be shown to be uniform because A and σ are related by $A\sigma \propto Q$, the beam charge
    - **The result is a χ-squared maximization of $P(\{m_k\}|A,\mu\sigma,B,V)$**

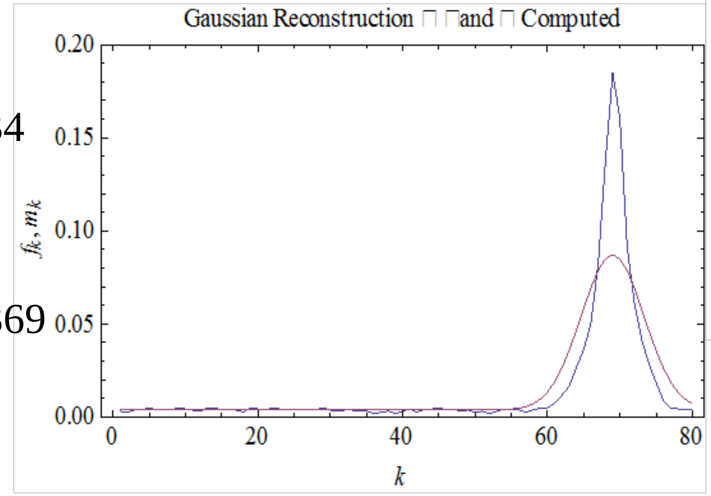
We can also eliminate the need for noise characterization by including *B* as a parameter

Gaussian RMS Fit
Gaussian-Like Profile

- Measurement
  - $N = 80$ sample points
  - Noise floor $B \sim 0.00369$
  - $A \sim \max\{m_k\} - B = 0.180$
- Gaussian Fit
  - $A = 0.164$
  - $\mu = 69.2$
  - $\sigma = 1.99$
  - $B = 0.00478$
- Computed
  - $A = 0.0834$
  - $\mu = 69.0$
  - $\sigma = 4.33.$
  - $B = 0.00369$



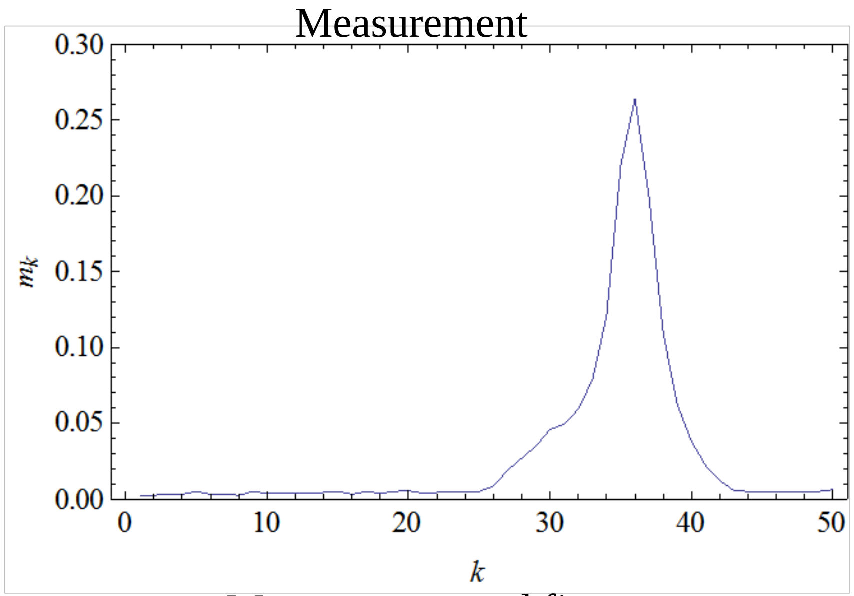


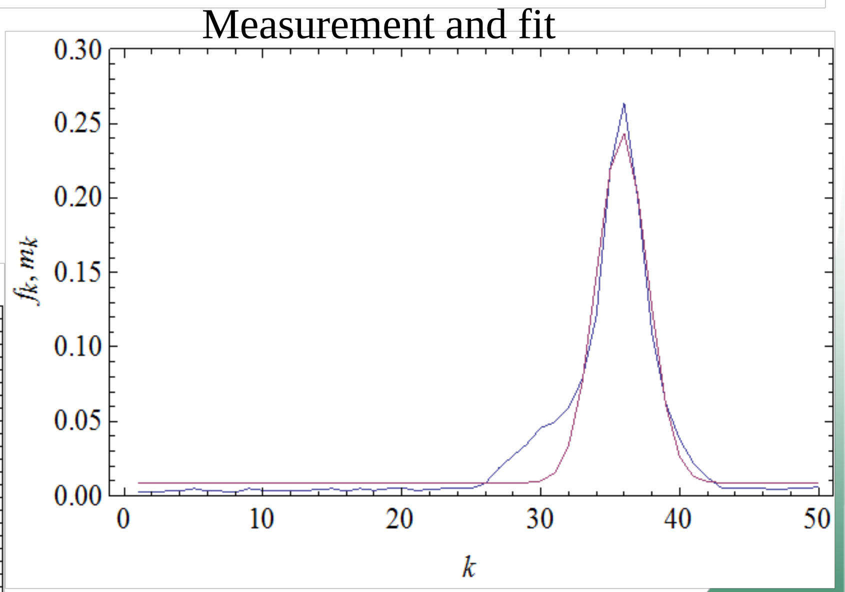


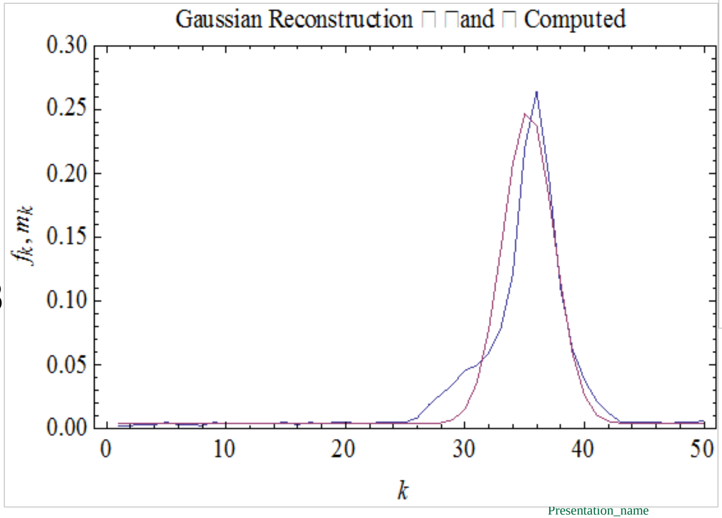
80 points

Gaussian RMS Fit
Profile with Halo

- Measurement
  - N = 50 sample points
  - Noise floor $B \sim 0.00387$
  - $A \sim \max\{m_k\} - B = 0.260$
- Gaussian Fit
  - $A = 0.236$
  - $\mu = 35.9$
  - $\sigma = 1.81$
  - $B = 0.00874$
- Computed
  - $A = 0.245$
  - $\mu = 35.3$
  - $\sigma = 2.14$
  - $B = 0.0038$



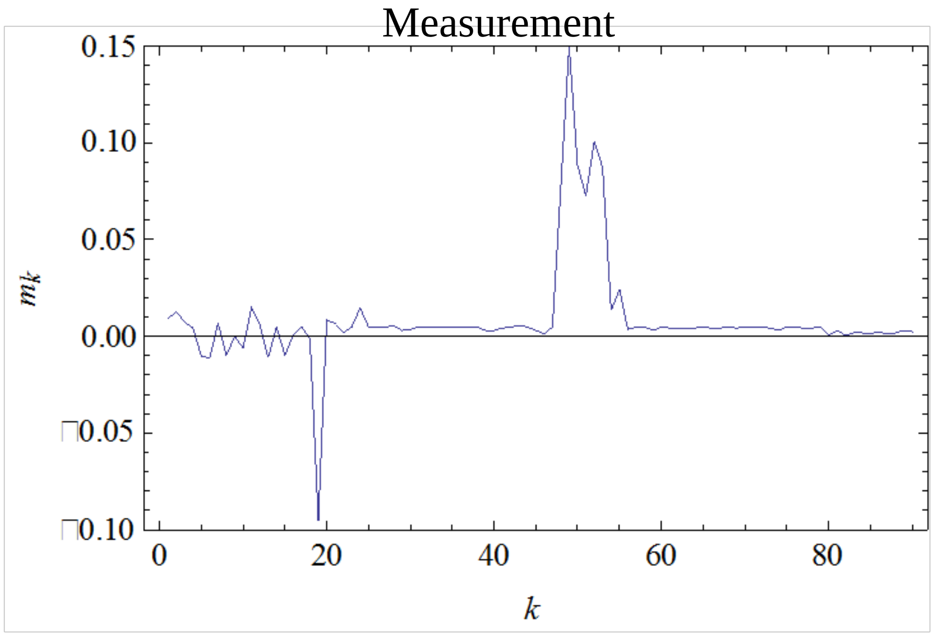


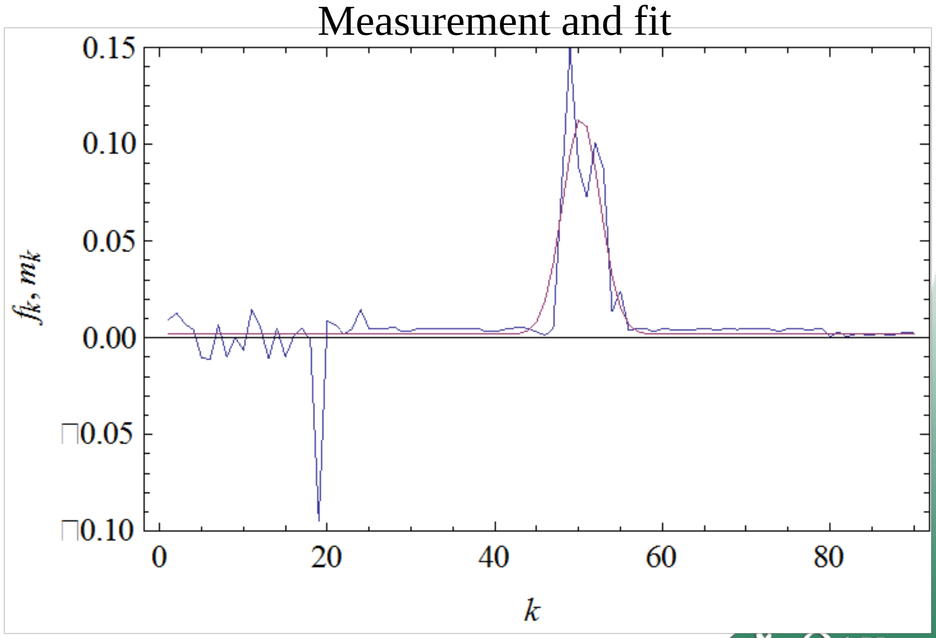


Only 50 points

Gaussian RMS Fit
Extremely Noisy Profile

- Measurement
  - N = 90 sample points
  - Noise floor $B \sim 0.00107$
  - $A \sim \max\{m_k\} - B = 0.149$
- Gaussian Fit
  - $A = 0.112$
  - $\mu = 50.3$
  - $\sigma = 2.26$
  - $B = 0.00181$
- Computed
  -





Conclusions

- Direct Computation of μ and σ from Measurements
  - Highly sensitive to noise and thus dubious
  - Requires calibration measurement (twice as long)
- Gaussian Fits
  - Direct RMS data fit is the most probable from Bayesian standpoint
  - Work well without halo
  - Good noise rejection
  - Seems to prefer core of the beam
  - Include noise baseline as parameter to avoid calibration (faster)
- Data Smoothing (not covered)
  - Significant loss of original signal
- Data Sampling (not covered) – Spectral power loss $\propto \exp[-\sigma^2/h]$
  - An h providing > 3 samples per σ gives good signal reconstruction
  - An h with < 1.5 samples per σ gives poor signal reconstruction

The Crux

- A primary motivation for determining μ and σ is **halo mitigation**
  - A primary cause of **halo formation** is **poor matching** between accelerating structures
  - We originally wanted μ and σ to compute Twiss parameters in order to re-adjustment quadrupole strengths for a good match (automated matching?)
  - Gaussian fits are suspect when halo is present

- Gaussian Fitting: You need a good match in order to match

Open Questions

Without Visual Inspection (That is, Automatically…)

- How do we recognize corrupted data?
  - Reject it if we find it?
- How do we recognize halo?
  - If we can recognize halo how do we compute μ and σ?
- Is there a better assumed profile than Gaussian?
  - Maxwell-Boltzmann is known to be stationary but no analytic form exists
- More fundamentally – is it possible to automate matching?
  - If so, how?

Thank You !

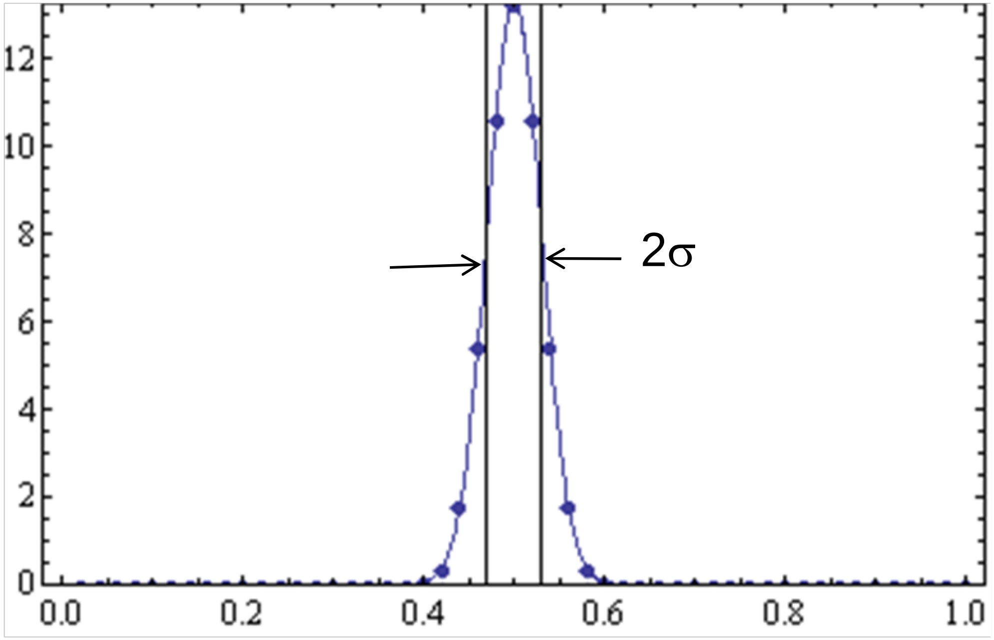
Any ideas, suggestions, comments welcome!

Sampling Intervals
Resolving Information Content

Choosing number of samples per scan (in order to maintain information content)

- Assume Gaussian profile
- Fourier transform of Gaussian with std = σ is Gaussian with std = $1/\sigma$
- Nyquist says when sampling at interval of h the highest frequency is $1/2h$.

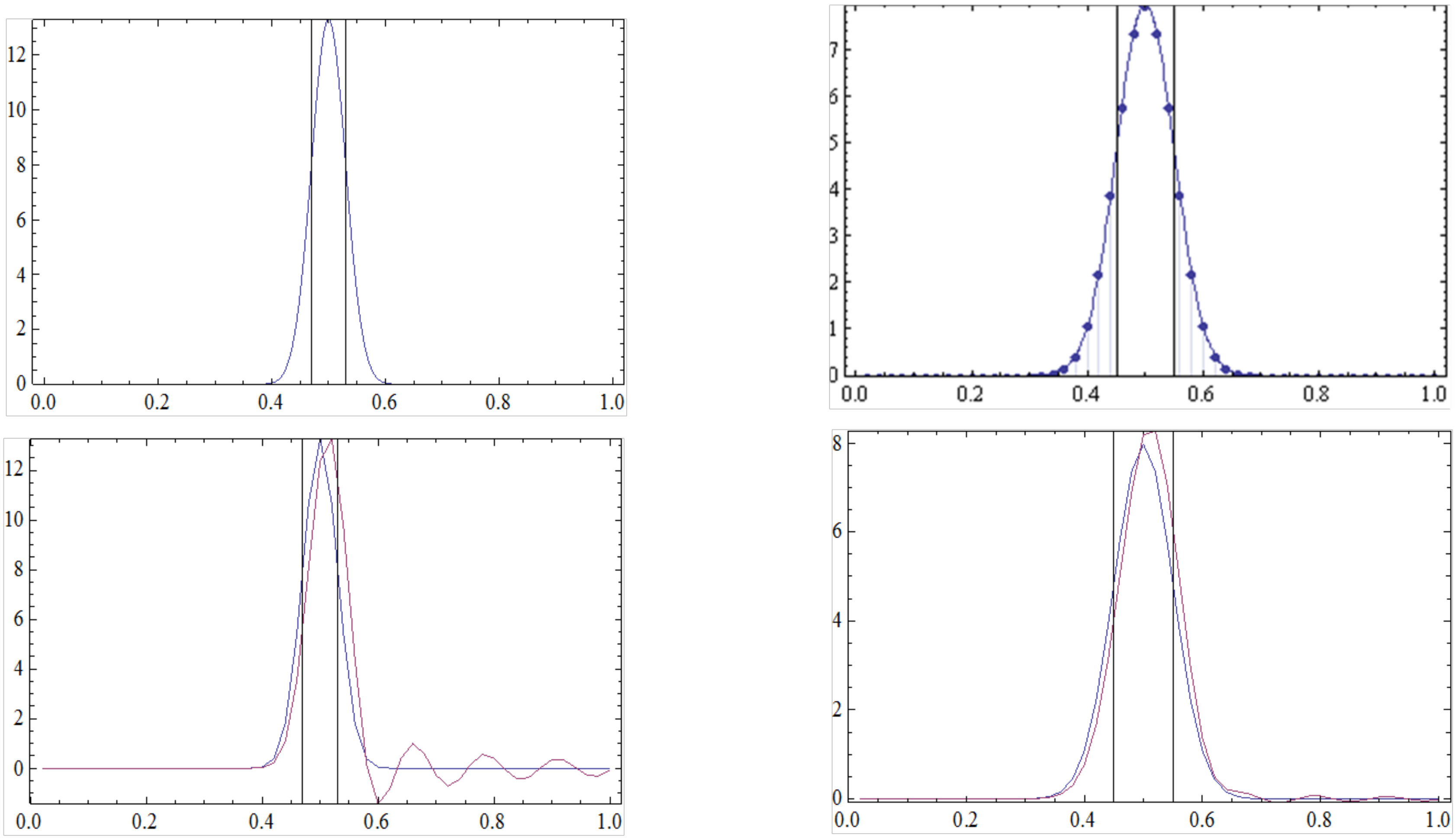
$\Rightarrow \sigma/h > 3$ is reasonable
$\Rightarrow \sigma/h < 1.5$ is dubious



Gaussian signal with
- $N = 50$ samples
- $\sigma/h = 1.5$

Sampling Interval
“Perfect*” (D/A) Reconstruction from Samples

$\sigma/h = 1.5$, $N = 50$ $\sigma/h = 2.5$, $N = 50$

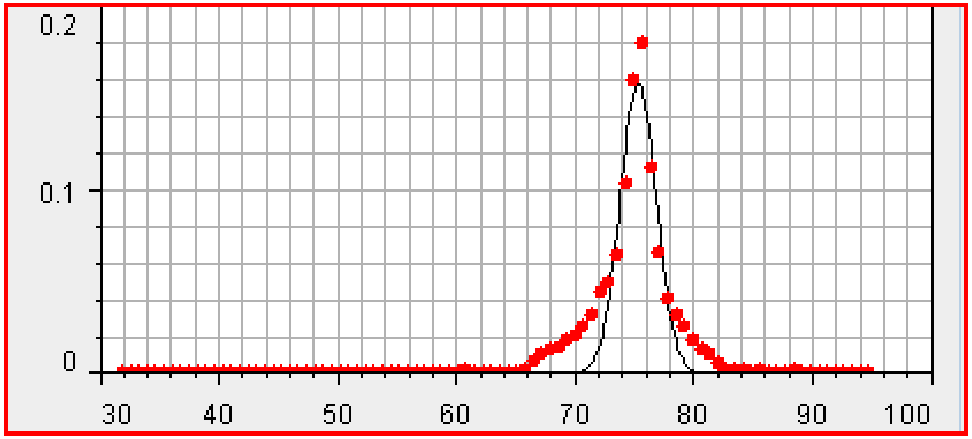


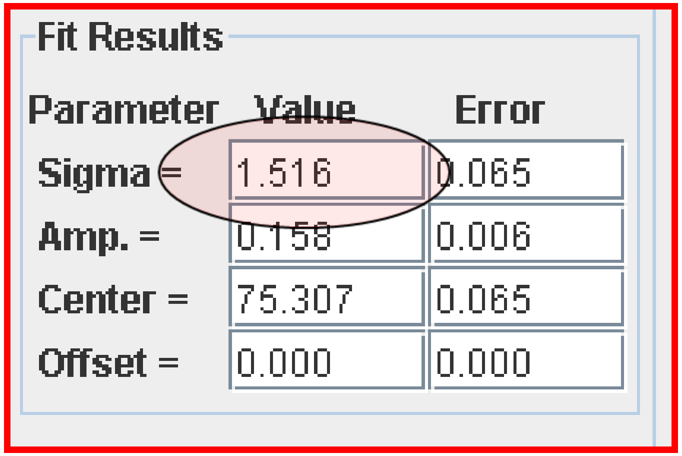
Red - reconstruction Red - reconstruction

*via Shannon sampling theorem

Gaussian χ-Squared Fit

- Significant shoulders
  - Gaussian fit does not accurately represent the signal
  - Beam size (sigma) is too small





Fit Results

Parameter	Value	Error
Sigma =	1.516	0.065
Amp. =	0.158	0.006
Center =	75.307	0.065
Offset =	0.000	0.000

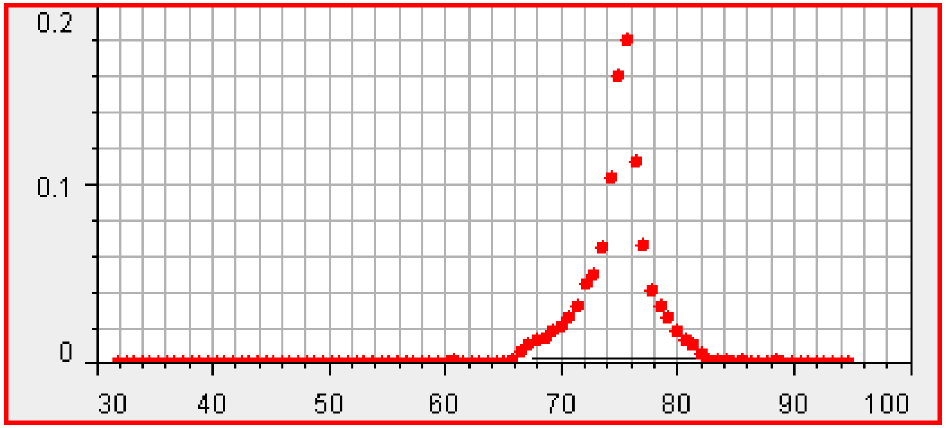
<u>χ^2 minimization</u>

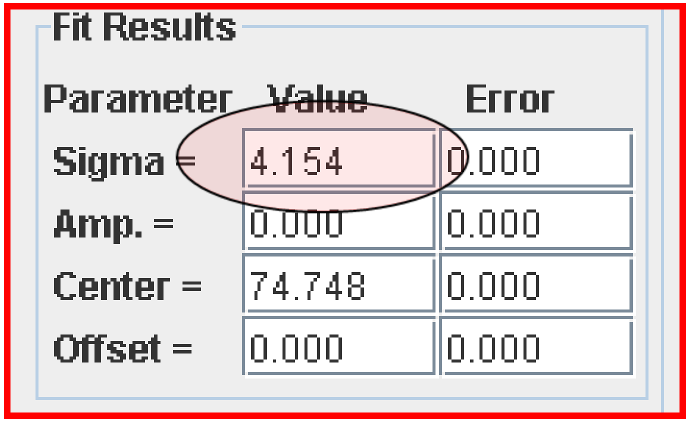
- $f(x)$ is a Gaussian at location $\bar{x}$ with standard deviation σ
- $\{m_k\}$ are measurements
- $\{x_k\}$ are measurement locations

$$\sigma = \arg\min_{\bar{x},\sigma,B} \chi^2(\bar{x},\sigma,B) = \sum_{k=1}^{N} \left[m_k - f(x_k;\bar{x},\sigma,B)\right]^2$$

Direct RMS Size Calculation

- Highly sensitive to background noise
  - Direct RMS calculation does not accurately produce beam size
  - Beam size is two large




Fit Results

Parameter	Value	Error
Sigma =	4.154	0.000
Amp. =	0.000	0.000
Center =	74.748	0.000
Offset =	0.000	0.000

<u>Standard Deviation of Measured Data</u>

- h step length
- $\bar{k}$ is (discrete) mean value
- $\{m_k\}$ are measurements
- $\{x_k\}$ are measurement locations

$$\left\langle x^2 \right\rangle^{1/2} = \left[\frac{1}{L} \sum_{k=1}^{N} x_k^2 m_k \right]^{1/2} = \left[\frac{h}{N} \sum_{k=1}^{N} \left(k - \bar{k} \right)^2 m_k \right]$$

Noise amplifying term

Observations
What I Have Seen So Far

Gaussian Fit*

	Noise unknown	Noise charact.
Gaussian	Good	Good
Halo	Bad	Bad
Noisy data	Good	Good
Jittery data	Good	Good

Statistical Calculation

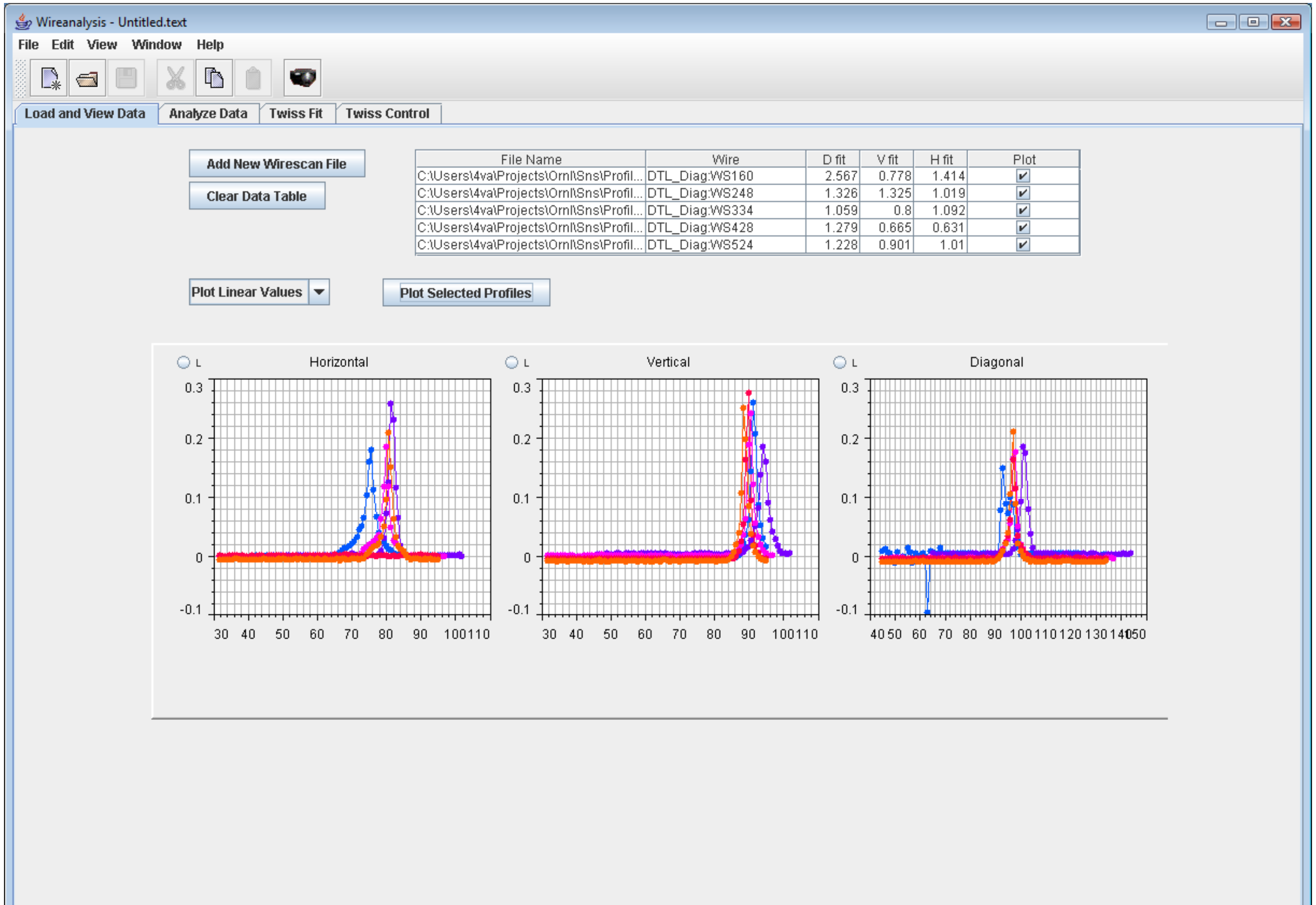
	Noise unknown	Noise charact.
Gaussian	Bad	Good
Halo	Bad	Good
Noisy data	Bad	marginal
Jittery data	Bad	Bad

- Additional Bayesian analysis (i.e., most probable) gives marginal return
- Critical to know the noise offset for direct statistical calculation

*RMS, or most likely, fits

Computations Involving Profile Data

- **Beam Position**
- **Beam Size**
- **Twiss Parameters**



Measurement Model

- Measurement process
  - Each measurement mk is taken during one macro-pulse
  - A stepper motor advances the profile device step length h after which the next measurement is made
  - We assume the beam is reproducible, that is, each beam pulse is identical to the previous.
  - Gaussian white noise process with mean M and variance V.

Profile Data
Processing and Data Analysis

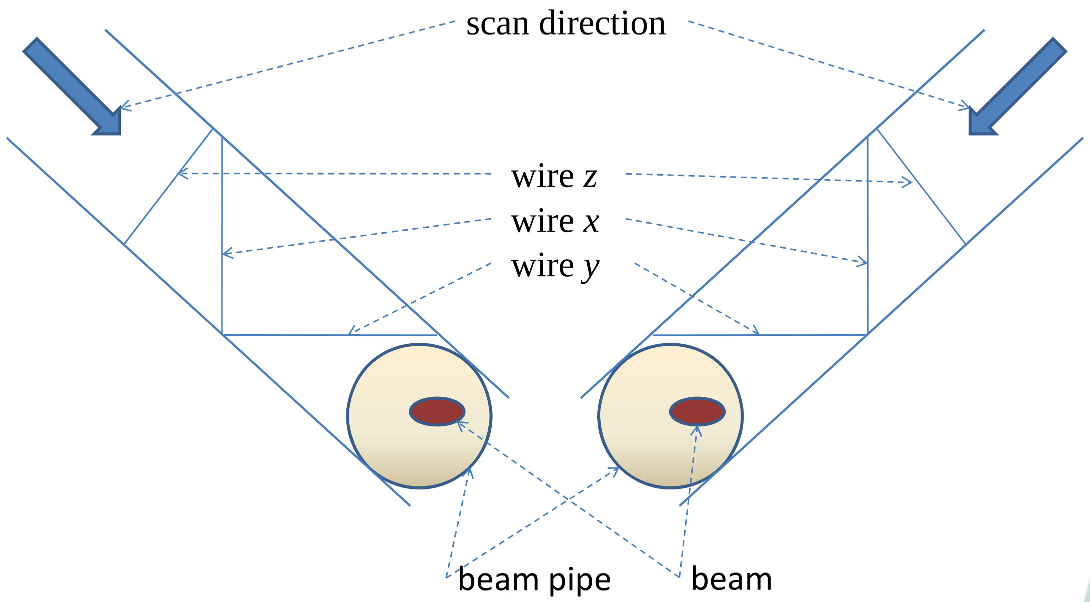
We wish to infer beam properties from collected profile data.

However – can think of profile data as 3-view, 1-dimension tomography

⇒Data contain limited amount of information

⇒Profile data have noise, jitter, missing data points, etc.

Wire Scanner



We want to recover…

- Beam Position μ
- Beam Size σ

This is a reasonable expectation.

The difficulty arises because we have so many data, and it's noisy ….

OAK RIDGE National Laboratory

Measurement Model

- Each sample contains noise from
  - Electronics
  - Jitter, etc.
- If the jitter is minimal, then it is reasonable to model the noise as a Gaussian white noise process W with mean B and variance V*.
  - Each measurement m_k will be composed of the (actual) sampled projection** f_k and a noise component W_k

$$m_k = f_k + W_k$$

- The white noise assumption implies
  - $W_k = W$ for all k (i.e., the noise is position independent)

* The noise can be characterized by a calibration experiment (no beam)

**This assumes that the beam is pulse reproducible

Centroid Location (Beam Position)

- **Let μ be the beam centroid position (i.e., beam position)**

$$\mu \equiv \frac{\int_{-b/2}^{+b/2} x f(x)\,dx}{\int_{-b/2}^{+b/2} f(x)\,dx} \approx h \frac{S_1(N)}{S_0(N)}$$

where the S_n are the sampled summations

$$S_n(N) \equiv \sum_{k=1}^{N} k^n f_k$$

OAK RIDGE National Laboratory

Expected Beta (Beam Size)

- **Let σ be the beam size**

$$\sigma^2 \equiv \frac{\int_{-b/2}^{+b/2} x^2 f(x)\,dx}{\int_{-b/2}^{+b/2} f(x)\,dx} - \left[\frac{\int_{-b/2}^{+b/2} x f(x)\,dx}{\int_{-b/2}^{+b/2} f(x)\,dx}\right]^2 \approx h^2 \frac{S_2(N)}{S_0(N)} + h^2 \left[\frac{S_1(N)}{S_0(N)}\right]^2$$

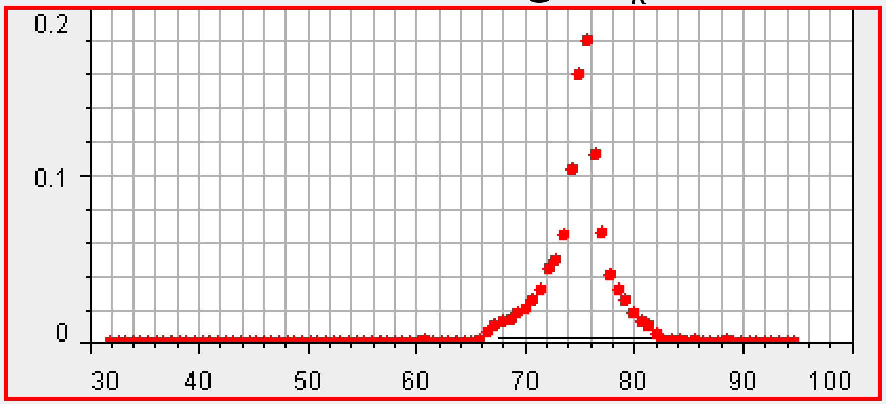
- **Once again we include the noise process and from our measurements {m_k} compute**

$$\tilde{\sigma}(N) = h^2 \frac{\tilde{S}_2(N)}{\tilde{S}_0(N)} - h^2 \left[\frac{\tilde{S}_1(N)}{\tilde{S}_0(N)}\right]^2 \qquad \text{where} \qquad \tilde{S}_n(N) = \sum_{k=1}^{N} k^2 m_k$$

The Problem – Halo
What is the Beam Size?

Determining beam size can be a very subjective process

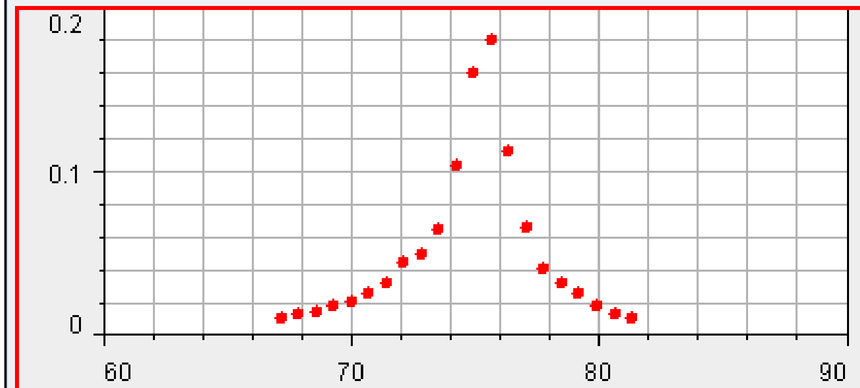
Direct calculation using $\{m_k\}$



Fit Results

Parameter	Value	Error
Sigma =	4.154	0.000
Amp. =	0.000	0.000
Center =	74.748	0.000
Offset =	0.000	0.000

Direct calculation with manual processing



Fit Results

Parameter	Value	Error
Sigma =	2.695	0.000
Amp. =	0.000	0.000
Center =	74.852	0.000
Offset =	0.000	0.000

Labor Intensive!

Fit Results

Parameter	Value	Error
Sigma =	1.516	0.065
Amp. =	0.158	0.006
Center =	75.307	0.065
Offset =	0.000	0.000

1.516

Gaussian fit result

The Problem - Jittery Data

- How to compute beam size
  - Do we trust a Gaussian fit?
  - Data smoothing?
- Reject measurement altogether?
  - How to automatically identify bad data

Fit Results

Parameter	Value	Error
Sigma =	2.291	0.245
Amp. =	0.112	0.008
Center =	94.401	0.245
Offset =	0.000	0.000