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

*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 $\mu$ and Size $\sigma$

• 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

\[
fx(x) = \int_{-L/2}^{+L/2} f(x,y) \, dy.
\]

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

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

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 $\mu$
- Beam Size $\sigma$

This seemed like a reasonable expectation, however…

- The data are noisy
- Beam jitter
- Missing data points
- Many data sets
**The Problem**
Processing many data sets for Simple Parameters

**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 $\mu$ and Size $\sigma$

- If we know the sampled profile $f_k$ exactly, normalizing by the step length $h$ the position $\mu$ and size $\sigma$ 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 $V**$

$$
m_k = f_k + W_k \quad \text{measurement random process}
$$

- We must account for this noise when approximating $\mu$ and $\sigma$.

* That is, $\mu$ and $\sigma$ 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 \( P(W = w) = \frac{1}{\sqrt{2\pi V}} e^{-\frac{(w-B)^2}{2V^2}} \)

  - 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

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

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 \( f_k = m_k - B \) for all \( k \)
  - Compute position \( \mu \) and size \( \sigma \) 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 \( N_n(\bar{k}) \equiv \sum_{k=1}^{N} (k - \bar{k})^n \)
Approximate $\mu$ and $\sigma$ 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 $\mu$ and $\sigma$, the variances become enormous as $N \to \infty$.

- $V < \sigma \times 10^{-7}$ for $\sim 10\%$ 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 $\mu$ and $\sigma$
  - Apply Bayesian techniques to estimate parameters $\mu$ and $\sigma$

- **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 $(A, \mu, \sigma)$ given $\{m_k\}$ - Bayes says that

\[
P(A, \mu, \sigma \mid \{m_k\}, B, V) \propto P(\{m_k\} \mid A, \mu, \sigma, B, V)P(A, \mu, \sigma)
\]

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

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 \( B \approx 0.00369 \)
  - \( A \approx \max \{m_k\} - B = 0.180 \)

- **Gaussian Fit**
  - \( A = 0.164 \)
  - \( \mu = 69.2 \)
  - \( \sigma = 1.99 \)
  - \( B = 0.00478 \)

- **Computed**
  - \( A = 0.0834 \)
  - \( \mu = 69.0 \)
  - \( \sigma = 4.33 \)
  - \( B = 0.00369 \)
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**
  - $A = 0.245$
  - $\mu = 35.3$
  - $\sigma = 2.14$
  - $B = 0.00387$
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 \( \mu \) and \( \sigma \) 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 \( \sigma \) gives good signal reconstruction
  - An \( h \) with < 1.5 samples per \( \sigma \) gives poor signal reconstruction
The Crux

• A primary motivation for determining $\mu$ and $\sigma$ is **halo mitigation**
  – A primary cause of **halo formation** is **poor matching** between accelerating structures
  – We originally wanted $\mu$ and $\sigma$ 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 $\mu$ and $\sigma$?

• 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 = $\sigma$ 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
Sampling Interval
“Perfect*” (D/A) Reconstruction from Samples

\[ \sigma/h = 1.5, \; N = 50 \]

\[ \sigma/h = 2.5, \; N = 50 \]

*via Shannon sampling theorem
Gaussian $\chi$-Squared Fit

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

$\chi^2$ minimization
- $f(x)$ is a Gaussian at location $\bar{x}$ with standard deviation $\sigma$
- $\{m_k\}$ are measurements
- $\{x_k\}$ are measurement locations

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

![Fit Results Table]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sigma</td>
<td>1.516</td>
<td>0.065</td>
</tr>
<tr>
<td>Amp.</td>
<td>0.158</td>
<td>0.006</td>
</tr>
<tr>
<td>Center</td>
<td>75.307</td>
<td>0.065</td>
</tr>
<tr>
<td>Offset</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Direct RMS Size Calculation

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

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

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

Fit Results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sigma</td>
<td>4.154</td>
<td>0.000</td>
</tr>
<tr>
<td>Amp.</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Center</td>
<td>74.748</td>
<td>0.000</td>
</tr>
<tr>
<td>Offset</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Noise amplifying term
Observations
What I Have Seen So Far

Gaussian Fit*

<table>
<thead>
<tr>
<th></th>
<th>Noise unknown</th>
<th>Noise charact.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>Halo</td>
<td>Bad</td>
<td>Bad</td>
</tr>
<tr>
<td>Noisy data</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>Jittery data</td>
<td>Good</td>
<td>Good</td>
</tr>
</tbody>
</table>

Statistical Calculation

<table>
<thead>
<tr>
<th></th>
<th>Noise unknown</th>
<th>Noise charact.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>Bad</td>
<td>Good</td>
</tr>
<tr>
<td>Halo</td>
<td>Bad</td>
<td>Good</td>
</tr>
<tr>
<td>Noisy data</td>
<td>Bad</td>
<td>marginal</td>
</tr>
<tr>
<td>Jittery data</td>
<td>Bad</td>
<td>Bad</td>
</tr>
</tbody>
</table>

• 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 $m_k$ 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.

We want to recover...

• Beam Position $\mu$
• Beam Size $\sigma$

This is a reasonable expectation.

The difficulty arises because we have so many data, and it’s noisy ….
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 $\mu$ be the beam centroid position (i.e., beam position)

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

where the $S_n$ are the sampled summations

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

- Let $\sigma$ 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

$$\widetilde{\sigma}(N) = h^2 \frac{\widetilde{S}_2(N)}{\widetilde{S}_0(N)} - h^2 \left[ \frac{\widetilde{S}_1(N)}{\widetilde{S}_0(N)} \right]^2 \quad \text{where} \quad \widetilde{S}_n(N) = \sum_{k=1}^{N} k^2 m_k$$
The Problem – Halo
What is the Beam Size?

Determining beam size can be a very subjective process

Direct calculation using $\{m_k\}$

Direct calculation with manual processing

Labor Intensive!

Gaussian fit result

1.516
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

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sigma</td>
<td>2.291</td>
<td>0.245</td>
</tr>
<tr>
<td>Amp.</td>
<td>0.112</td>
<td>0.008</td>
</tr>
<tr>
<td>Center</td>
<td>94.401</td>
<td>0.245</td>
</tr>
<tr>
<td>Offset</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>