EXTRACTING INFORMATION CONTENT WITHIN NOISY, SAMPLED PROFILE DATA FROM CHARGED PARTICLE BEAMS*: PART II

C.K. Allen#, W. Blokland, S.M. Cousineau, J.D. Galambos, ORNL, Oak Ridge, TN 38371 USA

Abstract

This is a continuation of work in [1]. The objective is to design a robust procedure for automating the analysis of beam profile data. In particular, we wish to extract accurate values for the beam position and beam size from profile data sets. These values may be then used to estimate additional beam characteristics such as the Courant-Snyder parameters.

INTRODUCTION

Profile data are typically obtained from particle-beam diagnostic devices such as wire scanner, laser strippers, or wire harps. The beam distribution is projected upon various spatial axes the effect being marginalization of the beam distribution with all variables except the projection axis. We provide a model for the data collection process which includes random noise components. The goal is to design a robust, automated procedure for accurate estimation of beam position μ and RMS beam size σ. Additionally, we need a process for automatic identification of bad data sets (typically such a process requires human intervention, a tedious, time consuming, and expensive endeavor). This goal is the first requirement for automated procedures for Twiss parameter estimation, transverse matching, and halo identification and mitigation. To realize our current goal we must make real world considerations. Specifically, we consider information content, noise (randomness), and sampling theory. Information content was covered previously [1]; we briefly review sampling.

Sampling and Measurements

Let $f(x)$ represent the profile distribution where x represents some spatial beam axis. Take the axis sampling locations to be equidistant so that $x_k = kh, k = 0, \dots, N-1$, where $h > 0$ is the (constant) step length between measurements. Making the definition

$$f_k \triangleq f(x_k)$$

then $\{f_k\}$ is the sampling of the profile f. Now denote the set of profile measurements as $\{m_k\}_{k=0}^{N-1}$. These ordered measurements correspond, respectively, to the set of sampling locations $\{x_k\}_{k=0}^{N-1}$.

Moments

The n^{th} moment $\langle x^n \rangle$ of a distribution f is defined

$$\langle x^n \rangle \triangleq \frac{1}{Q} \int_{-\infty}^{+\infty} x^n f(x) dx, \tag{1}$$

where the constant $Q \triangleq \int f dx$ is the total beam charge.

Because we are dealing with sampled data we can only approximate these values. The beam position μ and RMS size σ are the first two moments of the distribution

$$\mu \triangleq \langle x \rangle,$$
$$\sigma \triangleq \langle (x-\mu)^2 \rangle^{\frac{1}{2}}. \tag{2}$$

Because we are working with sampled data we normalize these quantities by h. Specifically, the continuous approximations to beam position and RMS size are $h\mu$ and $h\sigma$, respectively.

We take two approaches for approximating μ and σ, 1) computing μ and σ directly from the measurements, and 2) fitting a Gaussian approximation for f from the measurements from which μ and σ are determined.

The Measurement Process

Each measurement m_k is composed of both the actual beam profile f_k plus a noise component W_k, where W_k is part of a random process $\{W_k\}$. This noise introduces indeterminacy. Henceforth we can only generalize in terms of probabilities and stochastic (or "random") processes. Denote by $\mathrm{E}[\cdot]$ the expectation operator of the random variable W_k, averaging over its ensemble. Then $\Omega_k \triangleq \mathrm{E}[W_k]$ is the mean and the quantity $V_k \triangleq \mathrm{E}[(W_k - \Omega_k)^2]^{1/2}$ is the standard deviation, or variance. Although the values of a random process are not deterministic but their statistics are, specifically, Ω_k and V_k can be measured through a calibration experiment. Most noise processes are modeled as random processes.

Take the process $\{W_k\}$ to be a Gaussian distributed white noise process with mean Ω and variance V, then $W_k = W$ for all k, and W is Gaussian distributed. We have

$$m_k = f_k + W, \qquad k = 0, \dots, N-1 \tag{3}$$

where

$$Pr(W = w | \Omega, V) = \frac{1}{\sqrt{2\pi} V} e^{-\frac{(w-\Omega)^2}{2V^2}}. \tag{4}$$

The notation $\Pr(M_k = m_k | \Omega, V) = \Pr(W = m_k - \Omega - f_k | \Omega, V)$ indicates the probability that the measurement (random variable) M_k at axial position x_k has value m_k, given that the noise has mean Ω and variance V.

DIRECT MOMENT COMPUTATION

Because we are dealing with sampled data we can only approximate the moments $\langle x^n \rangle$. The simplest form of approximation would be to replace the integration in (1) with a finite summation. We begin with a definition to simply the sequel:

* This work supported by SNS through UT-Battelle, LLC, under contract DE-AC05-00OR22725 for the U.S. DOE.
# Corresponding author allenck@ornl.gov

$$S_n(\bar{k}) \triangleq \sum_{k=0}^{N-1} (k-\bar{k})^n f_k \tag{5}$$

Then, the n^{th} discrete moment centered at $\bar{k}$ is given as

$$\langle (k-\bar{k})^n \rangle \triangleq \tilde{S}_n(\bar{k})/S_0(0) \tag{6}$$

where $\bar{k}$ is meant to identify the mean value of k.

Computations involving random processes require that we properly observe their statistics. Considering the direct moment calculations based upon Eq. (5) we define

$$\tilde{S}_n(\bar{k}) \triangleq \sum_{k=0}^{N-1} (k-\bar{k})^n (m_k - \Omega), \tag{7}$$

the measurement central summations. Then

$$\begin{aligned} \bar{Q} &\triangleq \tilde{S}_0(0), \\ \bar{\mu} &\triangleq \frac{\tilde{S}_1(0)}{\tilde{S}_0(0)}, \\ \bar{\sigma} &\triangleq \tilde{S}_2(\bar{\mu})/\tilde{S}_0(0), \end{aligned} \tag{8}$$

are the computed values of the beam charge, position, and RMS size, respectively. The expected values for these quantities follow from Eq. (7), it is straightforward to show

$$E[\tilde{S}_n(\bar{k})] = \sum_{k=0}^{N-1} (k-\bar{k})^n E[(m_k - \Omega)] = S_n(\bar{k}) \tag{9}$$

since $E[\cdot]$ is a linear operator. This result is exactly that which we need. However, consider the variance of $\tilde{S}_n(\bar{k})$

$$E\left[\left(\tilde{S}_n(\bar{k}) - S_n(\bar{k})^2\right)\right]^{\frac{1}{2}} = N_n(\bar{k})V, \tag{10}$$

where $N_n(\bar{k}) \triangleq \sum_{k=0}^{N-1} (k-\bar{k})^n$ is a form of the Riemann zeta function. This function can become prohibitively large with relatively moderate values of n and N [1]. Ironically, increasing sample count N provides more certainty in f but less in $\tilde{S}_n(\bar{k})$. Yet, we can use Eq. (10) to compute the variances of Eq. (8). For two random variables X and Y with means $\bar{x}$ and $\bar{y}$ and variances σ_x and σ_y, respectively, the random variable $Z \triangleq X/Y$ has mean $\bar{z} = \bar{x}/\bar{y}$ and variance $\sigma_z = \bar{z}\sqrt{\sigma_x^2/\bar{x}^2 + \sigma_y^2/\bar{y}^2}$ [3]. Thus, we have (noting $N_0(0) = N$)

$$\begin{aligned} Var[\bar{Q}] &= NV, \\ Var[\bar{\mu}] &= \bar{\mu}\Delta_1(\bar{Q},\bar{\mu})V, \\ Var[\bar{\sigma}] &= \bar{\sigma}\Delta_2(\bar{Q},\bar{\mu},\bar{\sigma}), \end{aligned} \tag{11}$$

where

$$\Delta_1(\bar{Q},\bar{\mu}) \triangleq \sqrt{N^2/\bar{Q}^2 + N_1^2(\bar{\mu})/\bar{\sigma}^2},$$

Thus, the values of $\Delta_1(\bar{Q},\bar{\mu})V$ and $\Delta_2(\bar{Q},\bar{\mu},\bar{\sigma})V$ can be used to determine the order of magnitude to which $\bar{\mu}$ and $\bar{\sigma}$ are accurate.

GAUSSIAN PROFILES

It can be shown in the limit of zero space charge the Gaussian profile is a stationary beam distribution [2]. Here we make that assumption. We need three parameters to identify a Gaussian profile: amplitude A, mean μ, and standard deviation σ. The sampled Gaussian profile is

$$f_k(A,\mu,\sigma,B) = Ae^{-\frac{(k-\mu)^2}{2\sigma^2}}. \tag{12}$$

where, recall, μ and σ are normalized by step length h.

Chi-Squared Fitting

From Eq. (3) note again that $\Pr(M_k = m_k|\Omega,V)$ is equal to $\Pr(W = m_k - f_k|\Omega,V)$. From Eq. (4), and assuming that each measurement m_k is independent, the probability of obtaining all the measurements $\{m_k\}$ is then

$$Pr(\{m_k\}|A,\mu,\sigma) = \prod_{k=0}^{N-1} Pr(m_k|\Omega,V) = \frac{e^{-\frac{\chi^2}{2V^2}}}{\left(\sqrt{2\pi}V\right)^N} \tag{13}$$

where

$$\chi^2(A,\mu,\sigma|\Omega) \triangleq \sum_{k=0}^{N-1} [m_k - \Omega - f_k(A,\mu,\sigma)]^2. \tag{14}$$

We can use the above equations to determine the most probable values of (A,μ,σ) which produce measurements $\{m_k\}$; this condition occurs when $\chi^2(A,\mu,\sigma|\Omega)$ is a minimum. Notice that noise variance V does not explicitly occur in χ^2. Consequently, we can take Ω as an additional free parameter in the minimization to avoid the separate calibration experiment for Ω and V.

Bayesian Methods

Bayes'theorem states that

$$Pr(A,\mu,\sigma|\{m_k\},\Omega,V) \propto Pr(\{m_k\}|A,\mu,\sigma)\, Pr(A,\mu,\sigma)$$

We know the first factor in the right hand side by Eq. (13). The final factor above is the *prior distribution*, consisting of all the information we know about A,μ,σ prior to the measurements. Note that μ is independent but A and σ are correlated, their product proportional to beam charge Q. Thus $\Pr(A,\mu,\sigma) = \Pr(A,\sigma)\Pr(\mu)$.

For the prior there are several other quantities we can infer from the data and our familiarity with the experiment. We can use the most reliable deterministic quantities in Eq. (8) and Eq. (11) to form the prior distribution, namely $\bar{Q}$ and $\bar{\mu}$. (If we find them unusable by inspecting variances, we must return to χ^2 fitting.) The amplitude A is correlated to the span of the measurements, however, we can say no more; that is, A is uniformly distributed between the extremes. We have

$$Pr(Q|\Omega,V) = \frac{1}{\sqrt{2\pi}NV} e^{-\frac{(Q-\bar{Q})^2}{2N^2V^2}}$$

$$Pr(\mu|\Omega,V) = \frac{1}{\sqrt{2\pi}\Delta_1(\bar{Q},\bar{\mu})V} e^{-\frac{(\mu-\bar{\mu})^2}{2\Delta_1^2(\bar{Q},\bar{\mu})V^2}}$$

$$Pr(A|\Omega,V) = \begin{cases} \dfrac{1}{m_{max} - m_{min}} & A \in [m_{max} - m_{min}] \\ 0 & A \notin [m_{max} - m_{min}] \end{cases}$$

To find $\Pr(A,\sigma|\Omega,V)$ note that A and Q are independent so $\Pr(A,Q) = \Pr(A)\Pr(Q)$. The random variables A,σ and Q are related by $Q = \sqrt{2\pi}\sigma A$. Then $\Pr(A,\sigma) = \Pr(A)\Pr\big(Q(A,\sigma)\big)dQ/d\sigma$. The result,

$$Pr(A,\sigma) = \frac{A}{NV} e^{-\frac{\left(\sqrt{2\pi}A\sigma - \bar{Q}\right)^2}{2N^2V^2}}.$$

Putting it all together

$$Pr(A,\mu,\sigma|\{m_k\}) \propto \frac{\sqrt{2\pi}A}{\left(\sqrt{2\pi}V\right)^{N+2}} \frac{e^{-\frac{\chi^2}{2V^2} - \frac{\left(\sqrt{2\pi}A\sigma-\bar{Q}\right)^2}{2N^2V^2} - \frac{(\mu-\bar{\mu})^2}{2\Delta_1 V^2}}}{N\Delta_1}$$

Taking the logarithm of the above, multiplying through by $2V^2$ and ignoring constant terms yields

$$J \triangleq \ln A - \chi^2 - \frac{1}{N^2}\left(\sqrt{2\pi}\sigma A - \bar{Q}\right)^2 - \frac{1}{\Delta_1^2}(\mu - \bar{\mu})^2.$$

Since the logarithm is a monotonic function, the maximum of J is also a maximum of $\Pr(A,\mu,\sigma|\Omega,V)$. Maximizing J with respect to our parameters yields the most probable values of A,μ,σ.

EXAMPLES

The first example is a profile with obvious (asymmetric) halo, shown in Figure 1. The second curve in Figure 1 is the Gaussian profile obtained when applying the parameters $\bar{A},\bar{\mu},\bar{\sigma}$ computed with the direct method.

Table 1: Example #1 Beam Parameter Comparison

Method	A	μ	σ	Ω
Direct	0.130	61.1±0.279	3.12±1.45	0.3e-3
Bayes	0.153	62.5	2.30	0.3e-3
χ^2	0.153	61.5	2.18	7.0e-3
l_1 fit	0.181	61.1	3.31	0.0

Table 2: Example #2 Beam Parameter Comparison

Method	A	μ	σ	Ω
Direct	NaN	53.6±21.2	NaN±156	1.07e-3
Bayes	0.113	50.4	2.27	1.07e-3
χ^2	0.112	49.4	2.26	1.81e-3

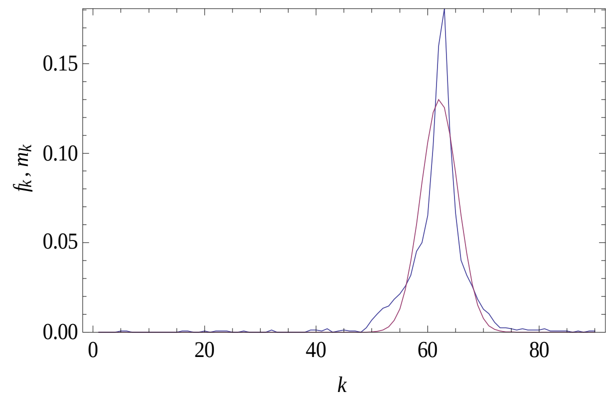

Figure 1: Example #1 profile with direct method Gaussian.

Table 1 provides a comparison for the different methods. In the table we have also included an l_1 fit to $\{m_k\}$ where the sum of the absolute values $|m_k - \Omega - f_k|$ is minimized. Interestingly enough this gives the same results as the direct method.

The second example is a pathological case involving extreme noise (not shown). Neither the direct method nor the l_1 fitting converge. 2 provides a comparison of the successful methods while demonstrating the failure of the direct method. Noise variances for the examples were 6.81×10^{-5} and 2.23×10^{-3}, respectively.

CONCLUSIONS

Accuracy for the direct computation of beam position and RMS size is quantified. When only the position and charge are accurate, they may be used in the Bayesian method. If both the position and sizes are inaccurate, then the χ^2 approach is always available. Additionally, χ^2 minimization does not require a calibration for noise mean and variance. Both χ^2 minimization and Bayesian methods are robust with noisy data. The use of double-Gaussian profile has been suggested and would appear appealing for separating core and halo.

REFERENCES

[1] C.K. Allen, W. Blokland, S.M. Cousineau, J.D. Galambos, High Brightness Workshop 2008, Nashville, TN, Aug. 25-29, 2008.

[2] M. Reiser, *Theory and Design of Charged Particle Beams* (Wiley-VCH, 2008).

[3] D.S. Sivia, *Data Analysis: A Bayesian Tutorial*, (Oxford Press, Oxford, 1996).