EXTRACTING INFORMATION CONTENT WITHIN NOISY, SAMPLED PROFILE DATA FROM CHARGED PARTICLE BEAMS

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

Abstract

Profile diagnostic devices for Charged-particle beams diagnostic provide data sets describing the one-dimensional density distributions at particular locations. We explore requirements and methods computation of beam position and size from profile data. Typically these data require subjective, human, processing to extract meaningful results, which is inefficient and labor intensive. Our goal is to automate such computations, or at least streamline the process.

INTRODUCTION

Here we describe properties of profile data, which are called *projection data* in tomography applications. These data are typically obtained from particle-beam diagnostic devices such as wire scanner, laser strippers, or wire harps. Each device provides multiple "views", or projections, of the underlying particle beam distribution at specific beamline locations. We provide [*]a model for the data collection process which includes random noise components. Our goal is automating the estimation of beam parameters from profile data, explicitly the position μ and size σ. In turn, these values can be used for Twiss parameter estimation, transverse matching, and halo identification and mitigation. In order to implement any automation we must make real world considerations. Specifically, we consider information content, noise, and sampling theory.

Sampled Data and Noise

Let x represent a transverse beam axis. Then denote the set of measurements obtained from the diagnostic device as $\{m_k\}_{k=1}^{N}$. These ordered measurements correspond to a respective set of axis locations $\{x_k\}_{k=0}^{N-1}$. We take the axis locations to be equidistant so that $x_k = x_0 + kh, k = 0, \ldots, N-1$, where $h > 0$ is the (constant) step length and the quantity x_0 is an offset. This situation is represents a *sampled-data system*. Now say the projection of a beam distribution is represented by the (absolutely) continuous function $f(x)$. Our objective here is to reconstruct a reasonable approximation of f from the measurements $\{m_k\}$. Making the convenient definition

$$f_k \triangleq f(x_k)$$

then $\{f_k\}$ is the sampling of the profile *f*.

Gaussian Signal

A common practice used on the field is to assume that *f* has a Gaussian profile. (It can be shown in the limit of zero space charge the Gaussian is a stationary beam distribution [2].) For the general case, we will need four parameters to identify such a signal: amplitude *A*, mean $h\mu$, standard deviation $h\sigma$, and offset *B*. Note that we have normalized μ and σ by the step length *h* so that they are in units of 'number of samples'. The familiar continuous Gaussian form is

$$f(x; A, \mu, \sigma, B) = Ae^{-\frac{(x-h\mu)^2}{2h^2\sigma^2}} + B. \quad (1)$$

By looking at the sample locations $\{x_k\}$ we get the four-parameter, sampled form

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

Moments

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

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

where normalization constant Q is given by $Q \triangleq \int f dx$, the total mass. Because we are dealing with sampled data we can only approximate these values. The simplest form of approximation would be to replace the integration with a finite summation. We begin with a definition to simply the sequel

$$S_n(\bar{k}) \triangleq \sum_{k=0}^{N-1}(k-\bar{k})^n f_k \quad (3)$$

Then, the n^{th} (discrete) moment centered at $\bar{k}$ is given as $\langle (k-\bar{k})^n \rangle \triangleq S_n(\bar{k})/S_0(0)$. The conversion between $\langle (x-\bar{x})^n \rangle$ and $\langle (k-\bar{k})^n \rangle$ is given by the step length power h^n, that is, $\langle (x-\bar{x})^n \rangle \approx h^n \langle (k-\bar{k})^n \rangle$, where $\bar{x} = h\bar{k}$.

Another possible approach for computation of $\langle (x-\bar{x})^n \rangle$ would be if we were certain that *f* was band-limited by 1/2*h;* then a formula such as (4) below could be employed. However, the non-causal impulse response *sinc* would need to be replaced by some filter having causal response. The moment computation would be affected by the choice of filter. We may well wish to pursue this idea; however, we still have not addressed reconstruction of *f* in the presence of noisy measurements.

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

RECONSTRUCTION

Theoretically, if $f(x)$ is band-limited in frequency by the Nyquist rate $\pm 1/2h$, then the Shannon sampling theorem states that [1][3]

$$f(x) = \sum_{k=0}^{x/h} f_k \operatorname{sinc}\left(\frac{x}{h} - k\right), \qquad (4)$$

where $\operatorname{sinc} x \triangleq \sin x / x$ and we have taken $x_0 = 0$. If we only look at the points $x_l = lh, l = 0, \ldots, N-1$ then the above simplifies to the discrete convolution $f_l = \sum_{k=0}^{l} f_k \operatorname{sinc}(l - k)$. Because the *sinc* function is non-casual, the above convolution is impractical. Typically a finite impulse-response linear filter is used in lieu of the *sinc* response.

We can get an appreciation for the step size h needed to accurately represent f by considering the Gaussian case (1). The Fourier transform of a Gaussian with standard deviation $h\sigma$ is also a Gaussian, with standard deviation $1/h\sigma$; specifically, with $\mu = B = 0$, the transform is

$$\hat{f}(\omega) \triangleq \mathfrak{F}[f](\omega) = A\sqrt{2\pi}\sigma e^{-\frac{\sigma^2\omega^2}{2}}. \qquad (5)$$

From the Nyquist criterion we know the largest frequency content ν_{max} in $\{f_k\}$ is $1/2h$ [4]. The percentage of power remaining in the sampling process is given by

$$P = \int_{-\frac{\pi}{h}}^{+\frac{\pi}{h}} \left|\hat{f}\right|^2(\omega)d\omega \Big/ \int_{-\infty}^{+\infty} \left|\hat{f}\right|^2(\omega)d\omega,$$
$$= \operatorname{erf}(\sigma/2),$$

where, recall, sigma is normalized by step length h.

Table 1: Signal power versus samples/sigma

σ	1	1.5	2.0	2.5	3.0
P	52.1%	71.1%	84.3%	92.3%	96.6%

Table 1 demonstrates that a standard deviation between two and three steps wide will preserve a reasonable amount of profile information.

NOISE

The presence of 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 a random variable, averaging over the ensemble. Then for any random variable R, $\bar{R} \triangleq \mathrm{E}[R]$ is the mean. The quantity $\sigma_R \triangleq \mathrm{E}[(R-\bar{R})^2]^{1/2}$ is the standard deviation, or variance. A random process is a random variable that varies in position, that is, $R = R_k$. The values of a random process are not deterministic, but their statistics are; notably $\bar{R} = \bar{R}_k$ and $\sigma_R = \sigma_{R,k}$. Most noise processes are modeled as random processes.

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

$$m_k = f_k + W, \qquad k = 0, \ldots, N-1 \qquad (6)$$

where

$$\Pr(W = w|B,V) = \frac{1}{\sqrt{2\pi V}} e^{-\frac{(w-B)^2}{2V^2}}. \qquad (7)$$

The notation $\Pr(M_k = m_k|B,V)$ 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 B and variance V.

Computations with Noise

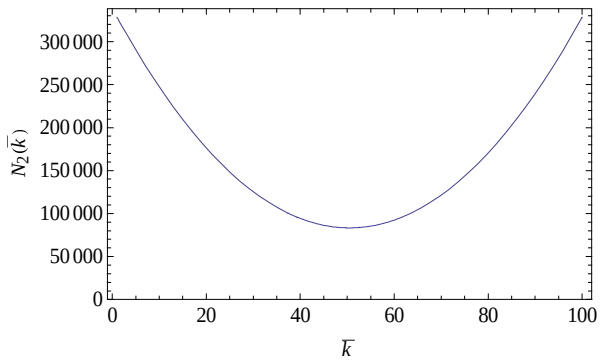

Figure 1: $N_k(\bar{k})$ for *n*=2 and *N*=100

Computations involving random processes require that we properly observe their statistics. Considering the direct moment calculations based upon (3) we define

$$\tilde{S}_n(\bar{k}) \triangleq \sum_{k=0}^{N-1} (k - \bar{k})^n (m_k - B), \qquad (8)$$

the weighted central summations of the measurements. The expected value of this computation is

$$E\left[\tilde{S}_n(\bar{k})\right] = \sum_{k=0}^{N-1} (k - \bar{k})^n \, E[(m_k - B)] = S_n(\bar{k}) \qquad (9)$$

since $E[\cdot]$ is a linear operator. The variance of $\tilde{S}_n(\bar{k})$ is a more involved computation, the result is

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

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 enormous with relatively moderate values of n and N, as demonstrated in

Figure 1. Ironically, as evident from the figure, increased sample count provides more certainty in f but less in $S_n(\bar{k})$. The conclusion is that we are unable to ascertain meaningful results from the approach of Eq. (8).

Bayesian Methods

A common technique is to compute the probability density function (i.e., p.d.f.) for a measurement process, then find the distribution $\{f_k\}$ that (locally) maximizes this p.d.f., a Bayesian method. A major drawback here is that we usually require the assumption of a given profile for *f* to be useful.

Take *f* to be a Gaussian so that f_k is given by (2). Note that $\Pr(M_k = m_k|B,V)$ is the same as $\Pr(W = m_k - f_k|B,V)$, the probability that $W_k = M_k - f_k$ given B and V. From (7), assuming that each measurement is independent, the probability of obtaining all the measurements $\{m_k\}$ is

$$\Pr(\{m_k\}|B,V) = \prod_{k=0}^{N-1} \Pr(m_k|B,V) = \frac{1}{(2\pi)^{\frac{N}{2}}}\frac{1}{V^N} e^{-\frac{\chi^2}{V^2}}$$

where

$$\chi^2(A,\mu,\sigma,B) \triangleq \sum_{k=0}^{N-1} [m_k - f_k(A,\mu,\sigma) - B]^2. \quad (11)$$

We are interested in the most probable values of A, μ, σ, and B given the measurements $\{m_k\}$(the noise variance V contributes nothing). However, Bayes'theorem states that

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

The final factor above is called the *prior distribution*, consisting of all the information we know about A, μ, σ, and B prior to the measurements. Immediately note that μ and B are independent and uniformly distributed. Quantity B requires a calibration experiment and, unless we have prior knowledge of the scanner and beam position we can presume nothing about μ. The values of A and σ are related by the $Q = \sqrt{2\pi}\sigma A$, where Q is the beam charge. The measurement system cannot be guaranteed to produce the exact beam charge, the best we can assume is that $\Pr(A,\sigma) = \Pr(Q)$, a uniform distribution and $\Pr(A,\mu,\sigma,B) = \Pr(A,\sigma,)\Pr(\mu)\Pr(B) = const.$ The *most probable* distribution $\{f_k\}$ is then the *most likely* distribution based upon maximizing χ^2 in (11). This results in a simple least-squares fit to the data $\{m_k\}$.

Gaussian Fitting

Including B as a parameter in the fit eliminates the need for a calibration experiment. We have found that Gaussian fitting is robust concerning noise and the parameters A, μ, and σ, see Figure 2. However, as shown in Figure 3, Gaussian fitting fails to provide an accurate μ and σ for signals with significant (asymmetric) halo.

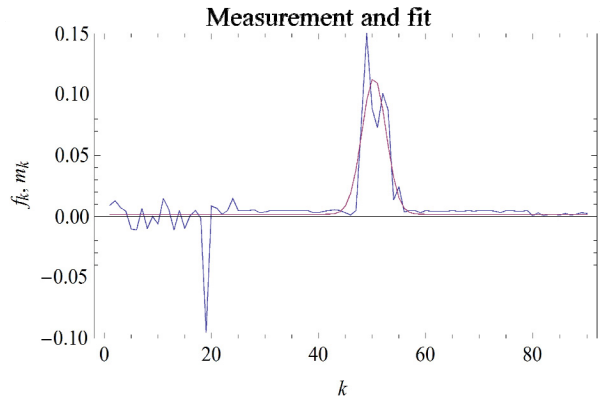

Figure 2: Gaussian fit for noisy signal

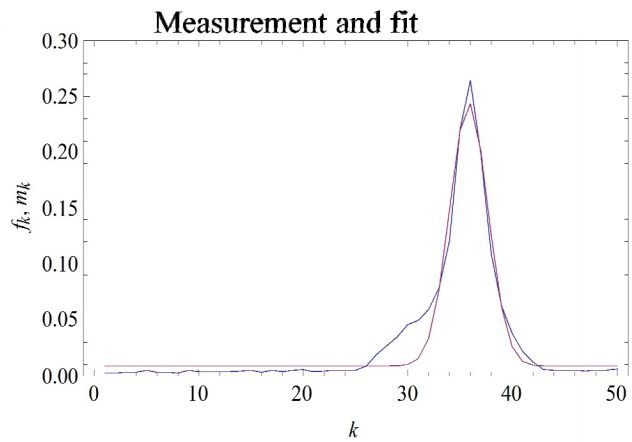

Figure 3: Gaussian fit for signal with halo

CONCLUSIONS

Referring to Table 1 the selection of sampling step is an important part in the accurate representation of the beam distribution. From the discussion on computations with noise, attempting to directly compute position and, more significantly, beam size is difficult (if not impossible). Gaussian fitting works well with noisy data but fails to provide accurate values when halo is present. The use of double-Gaussian fits has been suggested and would appear appealing for signals with symmetric halo components, this is currently being pursued.

REFERENCES

[1] H. Nyquist, "Certain topics in telegraph transmission theory", Trans. AIEE, vol. 47, pp. 617-644, Apr. 1928.

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

[3] C.E. Shannon, "Communication in the presence of noise", Proc. Institute of Radio Engineers, vol. 37, no.1, pp. 10-21, Jan. 1949.

[4] E.T. Whittaker, "On the Functions Which are Represented by the Expansions of the Interpolation Theory", Proc. Royal Soc. Edinburgh, Sec. A, vol.35, pp.181-194, 1915.