ISBN: 978-3-95450-208-0 doi:10.18429/JACoW-IPAC2019-WEPTS072

APPLICATION OF BAYESIAN INFERENCE IN ACCELERATOR COMMISSIONING OF FRIB*

Yue Hao and Leo Neufcourt, FRIB/NSCL, MSU, East Lansing, MI, USA

Abstract

We report the application of the Bayesian Inference of the unknown parameters of the accelerator model using the FRIB commissioning data. The inference result not only indicates the value of the unknown parameter but also the confidence of adopting the value. The Bayesian approach provides an alternative method to understand the difference between the accelerator model and the hardware and may help to achieve ultimate beam parameters of FRIB.

INTRODUCTION

Modern accelerators equip with complex diagnostics system to probe the beam in both machine commissioning and operation. A significant amount of diagnostic data is recorded to understand the statistical properties of the bunched charged particles. The diagnostic data may be the first order moment (beam centroid) from beam position monitors (BPMs), second order moment (beam size) or sampling of the beam distribution from the projection on the varies types of the profile monitors. These data are the only clue to tune the control knobs to make the accelerator as the machine we designed.

FRIB accelerator delivers up to 400 KW heavy ion beams to the target to generate rare isotope for nuclear physics researches. It features various types of diagnostic devices to probe the beam position and beam profile at different locations. FRIB is now being commissioned by stage. In this paper, we use the measurement data of Front End of the FRIB accelerator [1], which is sketched in Figure 1.

As in all linac accelerator, the initial condition from a certain starting point largely determines the beam properties downstream and the machine performance. They have to be determined as the unknown parameter of the machine model from the starting point to the location of the diagnostic devices, with a set of measurement data.

Usually, a fitting routine is adapted to find the unknown parameters of the model using the measurement data. However, there are limitations to the fitting method. First, in most cases, it is not easy to get the uncertainties of the fitting results. The uncertainties of the unknown parameter are very important in machine optimization. Second, the fitting routine tends to lose its efficiency in high dimensional parameter space.

In this paper, we report an attempt of using Bayesian Inference to infer the unknown parameters in the model using measurement data from FRIB Front End accelerator. The Bayesian Inference algorithm is based on the well known

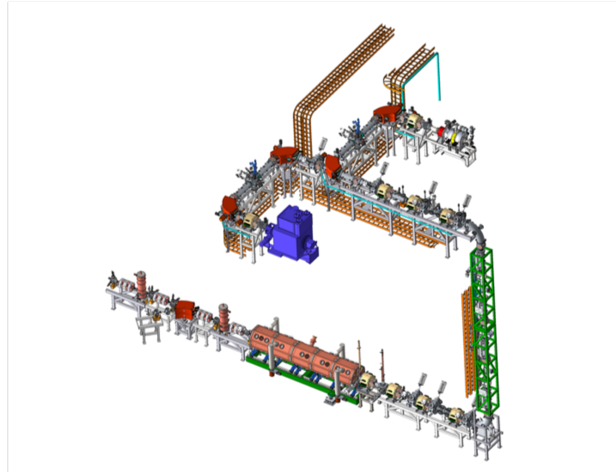

Figure 1: 3-D Drawing of Front End accelerator in FRIB.

Bayes' theorem,

$$P(H \mid M) = \frac{P(M \mid H) \cdot P(H)}{P(M)} \tag{1}$$

Here, H is the hypothesis, which represents the guess of the distribution of the unknown parameters. M is the set of measurements. The left-hand side of Eq. 1 is called posterior probability, the distribution of unknown parameter with given measurement results. On the right-hand side, $P(M \mid H)$ is the possibility of achieving some measurement result assuming the hypothesis H is valid, while the $P(H)$ is the prior guess of the distribution before the measurement is taken. The denominator $P(M)$ is the marginal probability of measurement M, which is very hard to compute in reality.

Since the direct evaluation of Eq. 1 is difficult, we adopt the Markov Chain Monte Carlo (MCMC) methods [2], to sample the posterior probability which is proportional to

$$P(H \mid M) \sim P(M \mid H) \cdot P(H) \tag{2}$$

without evaluating $P(M)$.

MODEL DESCRIPTION

We use the model of the FRIB Front End accelerator, which starts at the exit of the ion source. The initial 4-D linearized beam distribution is unknown, which can be represented in 4×4 matrix:

$$\begin{pmatrix} \langle x^2 \rangle & \langle xx' \rangle & \langle xy \rangle & \langle xy' \rangle \\ \langle x'x \rangle & \langle x'^2 \rangle & \langle x'y \rangle & \langle x'y' \rangle \\ \langle yx \rangle & \langle yx' \rangle & \langle y^2 \rangle & \langle yy' \rangle \\ \langle y'x \rangle & \langle y'x' \rangle & \langle y'y \rangle & \langle y'^2 \rangle \end{pmatrix}$$

* Work supported by the National Science Foundation under Cooperative Agreement PHY-1102511, the State of Michigan and Michigan State University.

ISBN: 978-3-95450-208-0 doi:10.18429/JACoW-IPAC2019-WEPTS072

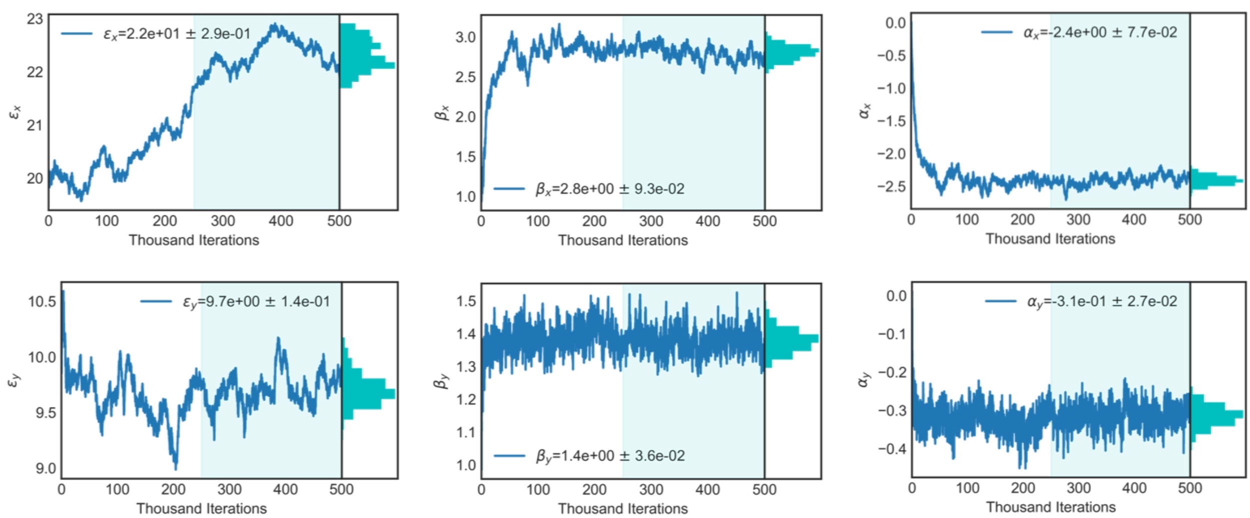

Figure 2: Inference Results of ϵ_x(up left), β_x(up middle), α_x(up right), ϵ_y(bottom left), β_y(bottom middle), α_y(bottom right).

the 16 matrix elements contain 10 degree of freedom. We can express 10 parameters (denote by θ) in the Twiss parameter form as

$$\theta = \left(\epsilon_x, \beta_x, \alpha_x, \epsilon_y, \beta_y, \alpha_y, c_{xy}, c_{xy'}, c_{x'y}, c_{x'y'}\right)$$

where $\epsilon_{x/y}$ are the rms emittance, $\beta_{x/y}$ are the beta functions, $\alpha_{x/y}$ are the alpha functions and 4 coupling factors represents the coupling between the horizontal and vertical plane.

To experimentally determine these 10 parameters, we change the voltages of three electric quadrupoles ($V = (V_1, V_2, V_3)$) and record the beam profile using a viewer downstream. The viewer gives the rms beam sizes in both transverse plane and the correlation, written as $\sigma = \left(\sigma_x, \sigma_y, \sigma_{xy}\right)$. Applying the Bayes' theorem, we get:

$$P\left(\theta \mid (\sigma, V)\right) \sim P\left((\sigma, V) \mid \theta\right) \cdot P\left(\theta\right) \tag{3}$$

To evaluate the likelihood function $P\left((\sigma, V) \mid \theta\right)$, we have to make further assumptions. With given machine parameter V and initial beam distribution θ, the model will predict the measurement result at the viewer, denoted as σ_{mod}. We assume that the real measurement σ_{mea} only differs from σ_{mod}, by a Gaussian random number $\delta\xi$, where ξ is normalized Gaussian random number and $\delta = \left(\delta_x, \delta_y, \delta_{xy}\right)$ are the rms random deviation between the model and the real measurement. Equivalently, we assume that there is no systematic error in the model. With this assumption, the likelihood function can be written as:

$$P\left((\sigma, V) \mid \theta, \delta\right) = \prod_i P\left((\sigma, V) \mid \theta, \delta\right)$$
$$\sim \prod_i \frac{\exp\left(-\left(\sigma_{\text{mea,i}} - \sigma_{\text{mod,i}}\right)^2 / 2\delta^2\right)}{\delta_x \delta_y \delta_{xy}}$$

where the subscript i denotes the i^{th} measurement using different machine setting V.

In the likelihood assumption, we introduce 3 more unknown parameter $\delta = \left(\delta_x, \delta_y, \delta_{xy}\right)$. They should be determined together with the 10 unknown parameter θ. Eq. 3 becomes:

$$P\left(\theta, \delta \mid (\sigma, V)\right) \sim P\left((\sigma, V) \mid \theta, \delta\right) \cdot P\left(\theta, \delta\right) \tag{4}$$

The prior distribution $P\left(\theta, \delta\right)$ should be determined by the prior knowledge of the parameters. Here we only impose the necessary limitations such as the emittance and beta functions should be positive. Otherwise, the parameters are assumed to have a uniform distribution.

INFERENCE RESULTS

We write MCMC python program to evaluate the posterior distribution $P\left(\theta, \delta \mid (\sigma, V)\right)$ using Eq. 4. The program adopt the Metropolis–Hastings algorithm to evaluate the MCMC. The algorithm is detailed as below steps:

1. Choose initial condition (0^{th} iteration) θ^0 and δ^0. The choice does not affect the inference result. However, a reasonable guess of the initial condition reduces the required iteration to reach equilibrium.

2. Evaluate the $\pi(i) = P\left((\sigma, V) \mid \theta^i, \delta^i\right) \cdot P\left(\theta^i, \delta^i\right)$ for the i^{th} iteration

3. Make Gaussian random walk centered at the value of θ^i and δ^i , with preset step size as the standard deviation, to get the new trial parameters θ^t and δ^t

4. Evaluate the $\pi(t) = P\left((\sigma, V) \mid \theta^t, \delta^t\right) \cdot P\left(\theta^t, \delta^t\right)$

5. Get a sample u from uniform random distribution $[0, 1]$

ISBN: 978-3-95450-208-0 doi:10.18429/JACoW-IPAC2019-WEPTS072

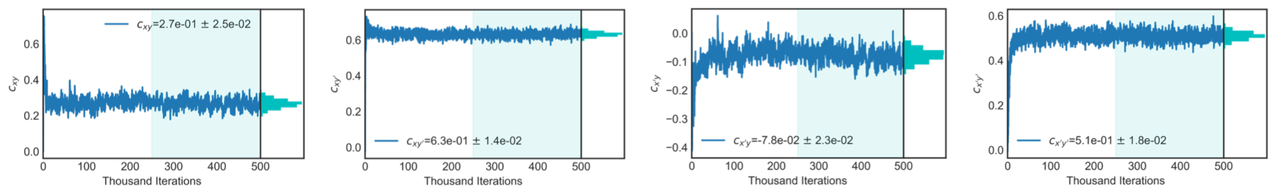

Figure 3: Inference Results of c_{xy}, $c_{xy'}$, $c_{x'y}$, $c_{x'y'}$ (from left to right).

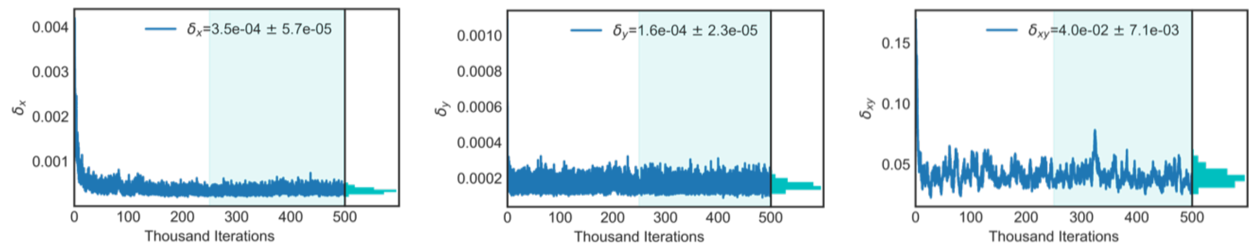

Figure 4: Inference Results of δ_x, δ_y, δ_{xy}(from left to right).

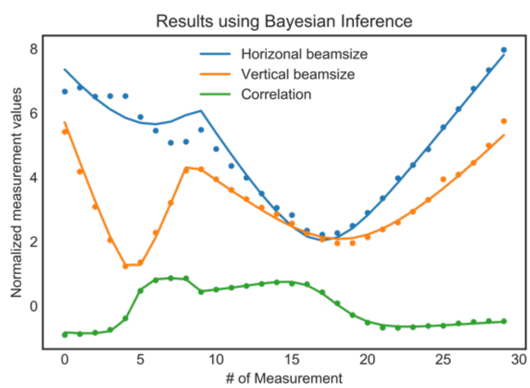

Figure 5: Comparison of the measurement results (dots) with the results from the model with inferred parameters. The blue, orange and green color represent the horizontal beam size, vertical beam size and correlation between the two directions.

6. If $u < \min\left(\frac{\pi(t)}{\pi(i)}, 1\right)$, the random walk is accepted, $\theta^{i+1} = \theta^t\ \delta^{i+1} = \delta^t$; otherwise $\theta^{i+1} = \theta^i \delta^{i+1} = \delta^i$

The step 2-6 is one iteration to get the distribution of unknown parameter θ and δ. We repeat the iteration and till the saturation of each parameter is reached. After saturation, we continue the iterations, so that we can achieve the stable distribution of each parameter. The Figure 2,3 and 4 show the iteration history of each parameter in θ and δ. In all figures, the x-axis is the iteration number and y-axis is the value of the parameter. Totally 500k iterations are calculated. The sample distributions of all parameters reach saturation except the horizontal emittance after 250k iterations. In each plot, the histogram is attached to the right of the figure to represent the stable distribution of the corresponding parameter.

By taking the average of the stable distribution of each parameter in θ and using them in the model, we achieve the expected measurement value and compare it with the experimental data in Figure 5. Generally, the inferred parameters predict the results very close to the measurement, except for the several horizontal beam size measurement. The discrepancy cannot be improved by choosing a better combination of the parameters. Such discrepancy may due to the missing physics in the model or the measurement error in some measurement. Since the measurement is achieved from a viewer which is not a precise diagnostics device, the measurement error has larger change to induce the discrepancy. Another indicator of such discrepancy is that the iteration plot of the horizontal emittance shows poor convergence at 500k iterations. We may resolve the unknown discrepancy by multiple measurements in the future.

Compare with the regular fitting method, Bayesian Inference provides the uncertainties of each parameter from the sampling of the distribution. This uncertainty information can be used in tuning the accelerator and allow us to trust these parameters with different confidence, which will prevent or reduce the overfitting problem. We will demonstrate this advantage in the future beam experiments.

SUMMARY

We use a set of simple measurement data in the FRIB Front End accelerator to demonstrate the effectiveness of the Bayesian Inference. The inference result is the distributions of the unknown parameters. The advantage of this approach is the capability of knowing how good does the parameter in the model to describe the data. This information may help to tune the machine without the over-fitting problem.

REFERENCES

[1] G. Pozdeyev, “FRIB Front End Construction and Commissioning”, in *Proc. 9th Int. Particle Accelerator*

ISBN: 978-3-95450-208-0 doi:10.18429/JACoW-IPAC2019-WEPTS072

Conf. (IPAC'18), Vancouver, Canada, Apr.-May 2018, pp. 58–62. doi:10.18429/JACoW-IPAC2018-MOZGBF1

[2] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications", in *Biometrika*, vol. 57, no. 1, pp. 97–109, Apr. 1970. doi:10.1093/biomet/57.1.97