Online simulation tools for orbit correction in the CBETA Machine

A. Nunez-delPrado, S. Peggs

September 2018

Collider Accelerator Department

Brookhaven National Laboratory

U.S. Department of Energy

USDOE Office of Science (SC), Nuclear Physics (NP) (SC-26)

Notice: This technical note has been authored by employees of Brookhaven Science Associates, LLC under Contract No.DE-SC0012704 with the U.S. Department of Energy. The publisher by accepting the technical note for publication acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this technical note, or allow others to do so, for United States Government purposes.
DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or any third party’s use or the results of such use of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof or its contractors or subcontractors. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
Control of beam trajectory is essential in particle accelerators. The trajectory of a beam is sensitive to various types of imperfections that cause it to deviate from its desired path along the accelerator. Orbit correction is hence fundamental for the control of beam trajectory. In this work we aim to demonstrate orbit correction using singular value decomposition techniques applied to distorted orbits generated in the CBETA-V Virtual Machine. In particular we present two concrete examples of such correction in different sections of the CBETA Fractional Arc Test layout. Our goal is to construct the backbone of an orbit correction software toolkit to be tested in the CBETA-V Virtual Machine and eventually implemented in the completed CBETA control system.

I. INTRODUCTION

A. CBETA Machine

The Cornell Brookhaven-National-Laboratory Energy-Recovery-Linac Test Accelerator (CBETA) is a particle accelerator being built at Cornell University that will serve to test several accelerator physics concepts relevant to development of the electron and Relativistic Heavy Ion Collider (eRHIC) at Brookhaven National Lab, as well as to the Energy Recovery Linac (ERL) community in general.

Linacs offer the advantage of producing beam whose quality is still roughly determined at the beam’s source, as opposed to rings, where the beam quality is determined by an equilibrium distribution set by the ring’s magnets. The drawback of a conventional linac approach includes their tendency to be wasteful due to their discarding of beams right after use and their creation of low average beam currents. The goal of CBETA, as an energy recovery linac (ERL), is to attempt to resolve both of these issues by: after the beam goes through the main accelerating cryomodule, a return loop utilizing a Non-scaling Fixed Field Alternating Gradient magnet design will direct the beam in a circular path back to the cryomodule where it will deliver its energy back to the SRF cavities, making the energy available for new beams.

In CBETA, this process starts in the in the DC Gun section, as seen in Fig. 1. By shining a laser onto a photocathode, electrons are ejected from the material and given an initial acceleration up to 300 keV. These electrons are then molded through the buncher section into an electron beam which then passes through the Injector Cryomodule (ICM) where the electron beam gains an initial 6MeV of energy due to the series of five two-cell superconducting radiofrequency (SRF) cavities. The beam is then sent through the Merger section in to the Main Linac Cryomodule (MLC), which consists of six SRF cavities, each of which can add up to an additional 6 MeV of energy to the beam. The beam is then passed into the low energy Splitter line which serves to provide the necessary matching into the first permanent magnet arc girder of the curved return loop comprised of Fixed Field Alternating Gradient (FFAG) magnets, shown in Fig. 2. The novelty of CBETA is that it will be the first energy recovery accelerator to control different-energy beams in the same structure relying on superconducting accelerating cavities and FFAG magnets in the return loop. To test the progress in construction made thus far, CBETA underwent the Fractional Arc Test (FAT) in the spring of 2018 with the purpose of testing a 42 MeV electron beam starting from the injector through the MLC all the way to the FA section.

B. CBETA-V Virtual Machine

The CBETA-V Virtual Machine (VM) is an online model of the CBETA machine. It is a combination of open-source simulation packages written in Python and Fortran. It is based on BMAD, a subroutine library for charged-particle dynamics simulations in accelerators and storage rings [2], as well as TAO, an analysis simulation program based on BMAD which includes tools to solve problems such as designing lattices under specified constraints, simulating errors and changes in machine parameters, and simulate machine commissioning [3]. When the VM is started, it creates a copy of the CBETA EPICS control system, thereupon allowing one to control the elements of the simulated accelerator in the same way one would control the elements of the CBETA machine, i.e. the same command is used to change the strength of an element in the VM as it is to change it in the physical machine. The CBETA-V VM is of high importance to the CBETA project because of its ability to produce simulated physics data of real measurement scripts and to debug such scripts for use in either the virtual or real machine. During FAT, the virtual machine underwent preliminary testing. Several capabilities are yet to be added to the CBETA-V VM. This paper deals
Figure 1. Layout of the injector and linac sections of the FAT experiment. The beam begins in the gun section (A1), has initial acceleration to 6 MeV in the Injector Cryomodule (A2), and is then merged into the Main Linac Cryomodule (B1, D1) [1].

Figure 2. Layout of the post linac section of the FAT experiment. Beam exits the linac and enters the S1 splitter line at the bottom right and proceeds to the fractional FA arc (top right) [1].
II. ORBIT CORRECTION

It is well known in the field of accelerator physics that one of the most indispensable operations in a particle accelerator is orbit correction. Typically the ideal path for a beam in an accelerator is to travel through the longitudinal axis, i.e. the center of the pipe. In general errors in magnet production and placement cause the beam orbit to deviate from this ideal. Thus, an orbit correction capability is of primary importance to ensure adequate control of the beam orbit.

A. Response Matrix

The (linear) relationship between changes in current to corrector magnets and the changes in orbit position downstream caused by them can be represented by the corrector-to-BPM mapping called the response matrix. A system with $M$ BPMs and $N$ correctors generates an $M \times N$ dimensional response matrix

$$
R_{M \times N} = \begin{pmatrix}
    r_{11} & r_{12} & \cdots & r_{1N} \\
    r_{21} & r_{22} & \cdots & r_{2N} \\
    \vdots & \vdots & \ddots & \vdots \\
    r_{M1} & r_{M2} & \cdots & r_{MN}
\end{pmatrix}.
$$

Each element $r_{ij}$ in the response matrix represents the orbit shift diagnosed by the i-th BPM caused by the a unit current change in the j-th corrector magnet. Hence, all elements in the i-th row preceding the j-th column in that same row will be zero if the current given to j-th BPM was the only adjustment in the system. An orbit response matrix can either be calculated via simulation or measured.

B. Orbit Correction Theory

Given the importance of orbit correction, numerous orbit correction algorithms have been proposed over the years. One that is widely used is the technique of singular value decomposition (SVD). Suppose the desired orbit is represented by $r_d$, the position of a non-ideal orbit in the accelerator (as diagnosed by the BPMs) represented by $r$, and the change in current settings of the electromagnets that will shift the beam position from the uncorrected trajectory to the desired one by $\Delta I$. The equation to attain the desired trajectory is thus

$$
r_d = r + R \cdot \Delta I,
$$

where $R$ is the response matrix obtained from specified correctors and diagnostics. This equation however assumes previous knowledge of the change in corrector currents and finds the resulting change in beam position. Nevertheless, the way in which the practical problem is formulated is the reverse of that: knowing the desired orbit trajectory, we intend to find the necessary $\Delta I$ that will generate the best orbit shift that will bring $r_c$ as close as possible to $r_d$, i.e. that will minimize the difference between $r$ and $r_d$. If we define $\Delta r$ to be that difference,

$$
\Delta r = r_d - r,
$$

then orbit correction techniques essentially aim to invert the matrix techniques using SVD.

$$
\Delta r = R^{-1} \cdot r_d.
$$

Were the number of correctors $M$ the same as the number of diagnostics $N$, thus producing a square response matrix $R$, the next step would entail finding the inverse of $R$ and the procedure would be nearly spotless (assuming $R$ is nondegenerate). This, however, is usually not the case. Most of the times we deal with non-square response matrices, and these do not have an inverse, but rather a pseudo-inverse, which will produce the identity matrix when multiplied with $R$ depending on the order of their multiplication.

C. Singular Value Decomposition

One of the most common approaches to producing a 'pseudo-inverse' for the response matrix $R$ is through the well-known matrix decomposition method called Singular Value Decomposition (SVD). Using SVD, one can decompose an $M \times N$ matrix $R$ into a product of three matrices:

$$
R = U \cdot S \cdot V^T
$$

where $U$ and $V$ are both orthonormal matrices of sizes $M \times M$ and $N \times N$, respectively, and $U$ has columns that are left singular vectors and $V$ has rows that are
right singular vectors. $S$ is a diagonal matrix of the same size as $R$, i.e., $M \times N$, whose diagonal elements are non-negatives values called singular values, usually produced in descending order, and have the form
\[ S_{ij} = s_w \delta_{ij}, \]
where $1 \leq w < \min(M, N)$. The singular values correspond to the coupling efficiency between the first $w$ diagnostics and $w$ correctors.

To ‘invert’ $R$, as decomposed in Eq. (4), we notice that by definition of orthonormal matrix, the inverses of $U$ and $V^T$ are $U^T$ and $V$, respectively. The ‘inverse’ of $S$, is a $N \times M$ diagonal matrix whose elements are
\[ (S^{-1})_{ij} = t_w \delta_{ij}, \]
where, again, $1 \leq w < \min(M, N)$ and
\[ t_w = \begin{cases} 0 & s_n \leq \tau \\ \frac{1}{s_n} & \text{otherwise}. \end{cases} \]

The tolerance $\tau$ refers to the threshold for which all singular values in $S$ below it are zeroed out so that when dealing with $S^{-1}$, the inverse of these singular values do not result in diverging values [6].

The expression for $R^{-1}$ is thus
\[ R^{-1} = V \cdot S^{-1} \cdot U^T, \]
which allows for us to invert Eq. (4) to obtain the desired change in corrector strengths that will take the current beam position as close as possible to the desired position
\[ R^{-1} \cdot \Delta r = \Delta I. \]

It is very important to note a couple of points. One is that when applying orbit correction using SVD there is a free parameter that must be adjusted for, namely the tolerance. The tolerance must be chosen carefully: a threshold that is too small will result in an unsuitably high corrector strengths, whereas one that is too high will diminish the utility of a higher number of diagnostics and result in a mediocre solution. The second important point is to emphasize that the orbit correction procedure is iterative. Since the response matrix is only a linear approximation (accurate within a bounded offset displacement from the design orbit) of a response from the actual system, repeating the orbit correction procedure a number of times will indeed push the non-ideal orbit to converge to the desired corrected trajectory [4].

III. SIMULATIONS

We tested orbit correction using SVD using several combinations of BPMs and correctors/dipoles, with quadrupole offsets and scaling errors included. We also present a slight modification of orbit correction theory that will aid in finding the periodic orbit in the curved FA section of the FAT layout.

1. Correction in Injector and Merger section with Quadrupole Offsets and Errors

This example deals with the A3 Quad Telescope section, which is part of the Injector section, and the B1 Merger sections of the FAT layout (Refer to Fig. 1). We chose to apply orbit correction in the injector line because there had been no previous orbit correction attempts in this section and was thus seen as a natural test case. Nevertheless, since the method outlined in this paper is a general procedure, if it works in this section, it should also apply to the other sections of the accelerator.

Procedure - The simulation starts by producing the response matrix, which is determined from the set of diagnostics and correctors that will play a part in the orbit correction process. For this example, all available BPMs from the A1 Gun section to the S1 Splitter section were used as the set of diagnostics and the set of correctors consisted of all those in the A3 and B1 sections. A ‘perfect’ response matrix was thus obtained with all nominal settings, that is one that did not take into account errors on any of the virtual optical elements, and a tolerance of 0.4345 was manually chosen for use in forming the psuedo-inverse via SVD. To create randomized ‘bad’ orbits, the same set of correctors from the A3 and B1 sections were randomly set with currents ranging in the range of ±50 mA. Next, the quadrupoles in the A3 section were given a scaling error of 5% RMS, and transverse offsets of 100 µm. A total of 200 randomized ‘bad’ orbits were tested.

Results - Fig. 3 depicts the orbit correction of 200 randomized orbits produced from random current changes given to the A3 and B1 correctors and random offsets and scaling errors to the quadrupole magnets. The four plots show results where the vertical axes represent the change in position in the corresponding transverse directions from the design values as a function of position along the accelerator. The lines in all four plots represent the orbits as simulated by TAO. The location of the BPMs are represented by the horizontal coordinates of the red dots, and their vertical coordinates represent the position readings of the beams. The left top and bottom plots depict the original uncorrected orbits. We can see that initial section of the orbits in the left plots are zeroed out since they correspond to the A1 and A2 sections of the lattice and the randomization of the orbits begins in A3. The right top and bottom plots depict the ‘corrected’ orbits in the corresponding transverse directions after iterating the SVD orbit correction process 10 times.

A measurement of the scale of the error of the orbit correction procedure in the injector and merger sections is shown in Fig. 4. Our metric of the error is the RMS
Figure 3. Orbit correction of 200 random trajectories simulated by CBETA-V with quadrupole offsets and quadrupole errors using A3 Injector and B1 Merger correctors. The two left plots show the transverse offset of the randomized beam orbit. The two right plots show the beam orbit after ten orbit correction iterations. The transverse offsets are the change in position in the corresponding transverse direction from the design value in the VM. The red dots represent the position of the diagnostics along the accelerator pipe as well as the diagnostics’ readings.

value of the BPM residuals, given by

$$\sigma_{Res} = \sqrt{\frac{\sum_{i=x,y} (r_i - r_{design,i})^2}{N}}. \quad (10)$$

The results suggest that the RMS values for each of the 200 tests converge after about 5 iterations and the orbit correction errors are at the micron and sub-micron scales.

Figure 4. Measurement correction error for the 200 simulated orbits in the simulation presented in Fig. 3 as a function of the number of iterations of orbit correction. The error metric used measures the offset of the diagnostic readings of an orbit from the design value, and it is calculated with Eq. (10).

2. Finding the Periodic Orbit in the FA section

In this example we show that by slightly modifying the orbit correction approach previously outlined, it is possible to find the periodic orbit in the FA 1st Arc Girder section of the FAT layout. The particularity that comes with applying orbit correction to this first part of the return loop of the final CBETA machine is that, unlike all the other sections previous to it, it is a curved trajectory with a repeating collection of magnets called cells. For this section it turns out it is possible to find a specific input trajectory into the the arc such that the beam follows a trajectory that will repeat itself in every cell, and hence making the diagnostics in each repeated cell read the same position. The desired beam position can be represented by the constant vector

$$C = \begin{pmatrix} C \\ C \\ \vdots \\ C \end{pmatrix} = C \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}. \quad (11)$$

In simulation, we know what this orbit position $C$ for the nominal 42 MeV beam energy. In contrast, when using the real CBETA machine we allow for a deviation from the design value of $C$. To try and simulate this scenario with the virtual machine, we aim to set up a situation where we do not know the value of this periodic trajectory $C$. Because the periodic orbit is a function of energy, a straightforward way of arriving at such situation is to simply change the energy of the beam. This can be achieved by changing the cavity voltages in the main accelerating cryomodule. After applying the change in
beam energy (which occurs at the MLC), it is necessary to scale the dipoles in the successive S1 Splitter to bring the beam back to the center of the pipe. However, scaling the dipoles for the beam to arrive at the FA section does not guarantee that it will go into the FA section with the right position and angle to get onto the periodic orbit.

It is important to emphasize that this problem is different from the one tackled in the first simulation. Previously, we knew what our desired orbit was (the nominal beam path), but in the FA section we do not. Hence, in this scenario not only do we need to find a vector of current changes that will steer onto the periodic trajectory, we must also identify what this periodic orbit is. We can make use of the fact that our desired orbit takes a special form, namely, that it is constant, and thus be able to solve this problem.

We can use Eq. (1) in a way that the desired orbit is the constant vector $C$,

$$C = \mathbf{r} + \mathbf{R} \cdot \Delta \mathbf{I}.$$  \hspace{1cm} (12)

We can rewrite this as

$$\mathbf{R} \cdot \Delta \mathbf{I} - C = -\mathbf{r}.$$  \hspace{1cm} (13)

In matrix form, this equation is represented by

$$
\begin{pmatrix}
  r_{11} & r_{12} & r_{13} & \cdots & r_{1N} \\
  r_{21} & r_{22} & r_{23} & \cdots & r_{2N} \\
  r_{31} & r_{32} & r_{33} & \cdots & r_{3N} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  r_{M1} & r_{M2} & r_{M3} & \cdots & r_{MN}
\end{pmatrix}
\begin{pmatrix}
  \Delta I_1 \\
  \Delta I_2 \\
  \vdots \\
  \Delta I_N
\end{pmatrix}
- 
\begin{pmatrix}
  C \\
  C \\
  \vdots \\
  C
\end{pmatrix}
= 
\begin{pmatrix}
  -r_1 \\
  -r_2 \\
  \vdots \\
  -r_M
\end{pmatrix}.
$$

As it is, this system seems to be unsolvable, given that there are two unknowns, $\Delta \mathbf{I}$ and $C$. However, it is possible to write this problem as only a tiny difference from the problem we have dealt with before. By expanding the response matrix $\mathbf{R}$ such that it contains an added column of a series of ones, and by multiplying this with a new vector identical to $\Delta \mathbf{I}$ except for an added last entry
of value \(-C\), we can obtain the same vector \(-r\),

\[
\begin{pmatrix}
  r_{11} & r_{12} & r_{13} & \ldots & r_{1N} & 1 \\
  r_{21} & r_{22} & & & r_{2N} & 1 \\
  r_{31} & & & & r_{3N} & 1 \\
  \vdots & & & & \vdots & \vdots \\
  r_{M1} & & & & r_{MN} & 1
\end{pmatrix}
\begin{pmatrix}
  \Delta I_1 \\
  \Delta I_2 \\
  \vdots \\
  \Delta I_N
\end{pmatrix}
= -
\begin{pmatrix}
  r_1 \\
  r_2 \\
  \vdots \\
  r_M
\end{pmatrix}.
\]

This step allows us to put all unknown values in one vector, and thus ending up with a matrix equation where we solve for the extended version of the \(\Delta I\) vector, which includes the constant position to be read from the diagnostics. From here on, the steps of matrix inversion through singular value decomposition are the same.

**Procedure** - As described in this subsection, in this simulation we alter our previous approach of producing a randomized orbit. Instead of randomly setting correctors or dipoles to a random value, we increase the energy of the beam from the nominal value of 42 MeV to 52 MeV by applying a 10 MeV increase in the last cavity of the MLC (RDICAV01). We set out to avoid beam loss in the S1 splitter section by autoscaling the dipoles in S1 accordingly from their design values. Once the beam arrives at the FA section entrance, it is imperative to ensure that it enters the section at the correct position and angle. To solve for these two variables, only two dipoles are needed for orbit correction. Thus, the last two dipoles in S1 are used to produce the response matrix for this simulation, along with the set of all diagnostics in FA, and the tolerance for the response matrix was manually set to 0. Orbit correction using SVD as presented on this page is applied 10 consecutive times.

We repeated this same procedure for 10 sample energies ranging from the minimum energy needed to get a periodic orbit in FA, 39 MeV, to the maximum output energy from the MLC, 59 MeV.

**Results** - Fig. 5 shows our finding of the periodic orbit for an arbitrary energy beam value of 52 MeV. It is worth mentioning that the design value for the beam position throughout the length of the accelerator is 0, except in the FA section, where we know in advance that the design beam position is at about 15 mm for the nominal beam energy of 42 MeV. The two plots in this figure display results where the vertical axes represent the difference in BPM position from the periodic orbit at 42 MeV as a function of position along the accelerator. The left plot shows the transverse offset of the uncorrected produced orbit at 52 MeV and the right one shows the transverse offsets after 10 correction iterations. In these plots the beam offset is zero until it encounters the two dipoles at the very end of the splitter section. The uncorrected orbit goes through the FA section in a seemingly random fashion, as displayed by the random BPM readings represented by the red dots in the left plot. In the right plot, we can see that after 10 iterations the BPM diagnostics in FA read nearly the same value, suggesting that the entering beam orbit has matched the periodic orbit trajectory at 52 MeV.

Measurement of the error scale of our finding of the periodic orbit corresponding to the 52 MeV energy beam is shown in Fig. 6. In this case our error metric makes use of the difference between FA BPM diagnostics:

\[
\sigma_{Res} = \sqrt{\frac{\sum_{i,j}(x_i - x_j)^2}{2N}}.
\]

This gives a sense of how much, on average, the BPM readings deviate from one another; consequently giving insight into how perfect the periodic orbit can be found. Fig. 5 shows us that the uncorrected orbit achieves periodicity to some level of accuracy in each of the first three iterations, and it seems to stop improving thereafter, suggesting that this uncorrected orbit will match into the corresponding periodic orbit by the third iteration.

Having tested our periodic orbit-finding technique with a random energy supported by CBETA, we proceed to show the results for this same procedure for 10 energies between 39 MeV to 59 MeV. In Fig. 7, both plots have been zoomed into the section of the orbits that gets modified by altering the energy of the beam. The different lines represent the simulated orbits produced from the 10 different energies. The left plot depicts the uncorrected orbits as somewhat chaotic in the curved FA section, whereas the right plot shows the corrected orbits after only 3 iterations to be rather periodic. The BPM readings in both plots align well with the simulated orbits, furthering agreeing with the realization of periodicity.

Fig. 8 depicts the same metric represented by Eq. (14) plotted versus the number of iterations. Because in this simulation no offsets or errors were added to any optical elements, the orbit correction procedure is complete in under 10 iterations. This occurs because orbit correction has indeed found the exact solution, i.e. a perfectly periodic orbit for a corresponding energy, and hence all BPM diagnostics in FA read the same value. Consequently, the residual function at that iteration is zero. In the log plot scale depicted in Fig. 8, this would become negative infinity, and is thus not depicted in the plot.

In Fig. 9 we plot the periodic orbit Shift, measured by

\[
\Delta x = x(E) - x(E_{\text{design}}),
\]

where \(x(E)\) is the orbit position diagnosed by the BPMs in FA (all of which read the same, or nearly the same value, per input beam energy) and \(x(E_{\text{design}})\) represents orbit position at the nominal beam energy of 42 MeV. We can see that the shift from the nominal position increases as you increase the input beam energy.
IV. CONCLUSION

Two cases were tested with the CBETA-V virtual machine using the current FAT lattice, and presented two important ones in this paper. The first one demonstrated orbit correction of two hundred randomized orbits using SVD along the A3 line in the Injector and the B1 Merger sections when taking quadrupole magnets offsets and scaling errors into account. After ten iterations, the error scale of the corrections was shown to be at the micron and sub-micron level. In the second case, a slightly modified version of the orbit correction procedure with SVD was used to find both the necessary current changes that will match the input beam to a periodic orbit, and the periodic orbit itself. This was tested at ten different energies and it was shown that, after ten iterations, the error for this simulation was also at the micron and sub-micron scale.

Thus far, this orbit correction toolkit has been successfully tested to work based on the FAT lattice implemented in the virtual machine. Work is currently being done to scale up this software in order to test it with the full 1-pass CBETA lattice. Once this is completed and all procedures are in place for quickly switching between lattices, the scaled-up version of this software shall be implemented in the future completed CBETA machine.

V. ACKNOWLEDGEMENTS

I would like to express my most sincere thanks to my mentor Colwyn Gulliford for his enormous help and guidance throughout this project as well as for his patience and words of encouragement. I would also like to express my gratitude to Adam Bartnik for his insights, wonderful collaboration, and delectable explaining style. Thanks to
both Colwyn and Adam, whose collaborative spirit were the main ingredients in creating such a pleasurable work environment. I would also like to thank Ivan Bazarov for helping organize the CLASSE REU program which has given me an amazing opportunity to grow as a researcher. Lastly, I would like to thank the National Science Foundation for making all this possible by funding this program with the award PHY-1757811 for the REU Site Accelerator Physics and Synchrotron Radiation Science at Cornell University.