IMPROVING INJECTION AT DIAMOND USING DIMENSIONALITY REDUCTION TECHNIQUES AND BAYESIAN OPTIMISATION

S. Preston*[1,2,3], R. T. Fielder[2], N. Blaskovic Kraljevic[2], I. P. S. Martin[2,3], P. N. Burrows[3]
[1]Ada Lovelace Centre, Oxfordshire, UK
[2]Diamond Light Source, Oxfordshire, UK
[3]John Adams Institute for Accelerator Science, Oxfordshire, UK

Abstract

Injection efficiency into the Diamond Light Source Storage Ring (SR) is currently optimised by operators performing a grid scan over the final pair of corrector magnets inside the Booster-to-Storage Ring (BTS) transfer line. The phase advance between the pair is sufficient to provide good control over the position and angle of the beam as it enters the SR. However, the method is slow and the strengths of the corrector pair can approach power supply limits over time as machine conditions drift. We propose a Bayesian optimisation algorithm to optimise leading right-singular vector coefficients obtained from a Singular Value Decomposition (SVD) of the BTS response matrix over all corrector magnets, improving sample efficiency. We further transform the proposed coefficients through a non-linear map to restrict the set of solutions to a desirable range as determined by the user. We conclude by suggesting further refinements to the algorithm.

INTRODUCTION

Diamond Light Source is a third generation 3 GeV synchrotron light source facility. Due to timing jitter of the radio frequency (RF) waves in the accelerating structures relative to the beam, the shot-to-shot energy jitter in the LTB is much larger than in the BTS where the energy variation has been reduced by the radiation damping in the booster. This causes the stability of the electron beam trajectory to be different between the transfer lines. In addition, work is currently ongoing to measure the Twiss parameters accurately in the LTB [1], whereas those in the BTS are constrained by the closed orbit of the booster and are therefore known. For these reasons, recent work has focused on the BTS and adjusting correctors to improve injection efficiency and automate the process.

THEORY

Response Matrix

A response matrix quantifies the sensitivity of the beam position at each beam position monitor (BPM) to changes in corrector strengths [2] and it must be recalculated if the linear optics change. A distinction is made between a Trajectory Response Matrix (TRM) which pertains to a single pass lattice and an Orbit Response Matrix (ORM) for a multi-pass lattice. If the Twiss parameters are known to high accuracy it can be computed analytically by considering the position of the beam at a downstream BPM.

$$\vec{x_{b_i}} = \begin{pmatrix} x_{b_i} \\ x'_{b_i} \\ \frac{\delta p}{p_0} \end{pmatrix} = \begin{pmatrix} \mathrm{T}_{11} & \mathrm{T}_{12} & \mathrm{D} \\ \mathrm{T}_{21} & \mathrm{T}_{22} & \mathrm{D}' \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{c_j} \\ x'_{c_j} \\ \frac{\delta p}{p_0} \end{pmatrix} = \tilde{\mathrm{T}} \vec{x_{c_j}}. \quad (1)$$

where $\vec{x_{c_j}}$ is the augmented phase space coordinate at the jth corrector, $\delta p/p_0$ is the relative momentum offset at the corrector, D and D′ are the dispersion in the position and its derivative, respectively, at the ith BPM, $\tilde{\mathrm{T}}$ is the augmented Twiss transport matrix from the end of the corrector to the BPM, and $\vec{x_{b_i}}$ is the augmented phase space coordinate at the ith BPM. The beam position can be written out in full.

$$x_{b_i} = \mathrm{T}_{11} x_{c_j} + \mathrm{T}_{12}(x'_{c_j} + \theta) + \mathrm{D}\frac{\delta p}{p_0}. \quad (2)$$

where θ is an additional small change in the horizontal momentum from the jth corrector. The response is found by calculating the quotient in Eq. (3), where terms have been collected after expanding Eq. (2) at two different corrector strengths.

$$\frac{x^2_{b_i} - x^1_{b_i}}{\theta_2 - \theta_1} = \mathrm{T}_{12} + \frac{\delta p}{p_0}\frac{\mathrm{D}_2 - \mathrm{D}_1}{\theta_2 - \theta_1}. \quad (3)$$

where $x^n_{b_i}$ is the transverse beam position at the ith BPM after a θ_n corrector kick. D_{θ_n} is the dispersion as a function of the kick. Because p_0 represents the momentum of a reference particle, the ensemble average relative momentum offset for the entire beam $\langle \delta p/p_0 \rangle = 0$ if the beam is on-energy. In this instance, the equation reduces to the top-right entry of the 2×2 transport matrix.

$$\frac{x^2_{b_i} - x^1_{b_i}}{\theta_2 - \theta_1} = \mathrm{T}_{12}. \quad (4)$$

A further implication is that the dispersive term in Eq. (3) may not be negligible if the beam is off-energy or if there is significant energy jitter.

Roll errors on the magnets are assumed to be negligible. The TRM is shown in Fig. 1.

Singular Values as Trajectory Distortions

Stronger singular vectors are used in this work because the beam trajectory can be nudged with a smaller kick and thus

* shaun.preston@physics.ox.ac.uk

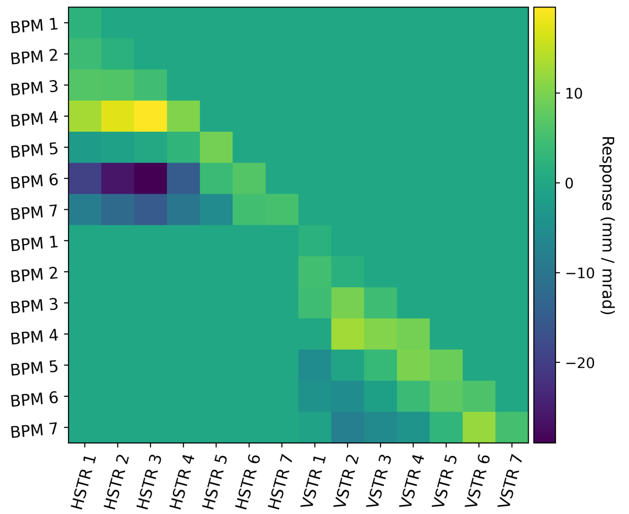

Figure 1: TRM of the Diamond BTS with nominal optics for an on-energy beam. It has been computed analytically using Eq. (4) where $T_{12} = \sqrt{\beta_{b_i}\beta_{c_j}} \sin\phi$ where β_{b_i} and β_{c_j} are the beta functions at the ith BPM and the jth corrector, respectively, and ϕ is the phase advance between them.

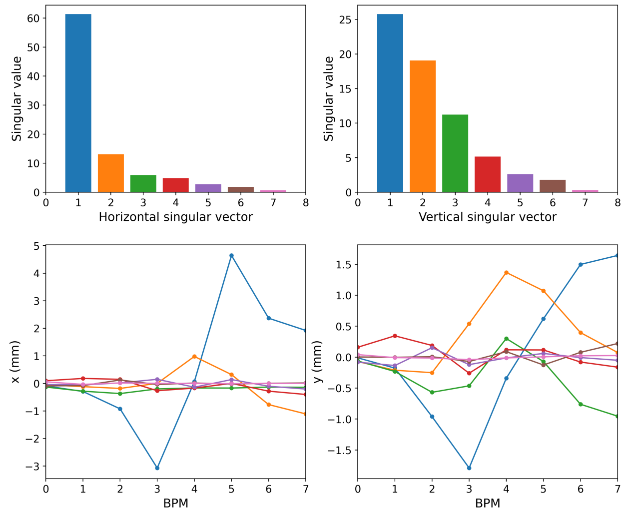

Figure 2: For each singular vector in V^T (top), the same reference beam has been tracked through the BTS model lattice. The position of the beam at the BPMs and at the very end was recorded (bottom). Singular values were scaled down by a factor 1×10^{-4} when setting corrector strengths to keep the beam within the aperture. The different lines in the bottom row correspond to different singular vectors. The singular values and singular vectors have been colour matched in each column.

less power from a power supply. This keeps correctors within their limits and helps the optimiser avoid getting stuck in flat regions where there is little to no change in the objective. A matrix $M \in \mathbb{R}^{m\times n}$ can be decomposed using SVD.

$$M = UV^T. \tag{5}$$

where $U \in \mathbb{R}^{m\times m}$ forms an orthonormal basis in BPM space, $\Sigma \in \mathbb{R}^{m\times n}$ is a diagonal matrix of singular values and $V \in \mathbb{R}^{n\times n}$ forms an orthonormal basis in corrector space. Furthermore, if the action, $\vec{\Phi}_n$, of the nth singular corrector vector is considered, the following is obtained.

$$\vec{x} = (UV^T)\vec{\Phi}_n = U(\vec{\Phi}_n^T V)^T = U\, \sigma_{\Phi_n} 1_{\vec{\Phi}_n}. \tag{6}$$

where the transpose identity $(AB)^T = B^T A^T$ has been used to reorder the matrix-vector product, σ_{Φ_n} is the singular value of the nth singular vector, and $1_{\vec{\Phi}_n}$ is the basis vector corresponding to σ_{Φ_n}. Finally, it is noted that $1_{\vec{\Phi}_n}$ picks out the first column of U. If the magnitude of both sides is taken, the following expression is obtained.

$$\| \vec{x} \| = \sigma_{\Phi_n} \| U_{\cdot,n} \| = \sigma_{\Phi_n}. \tag{7}$$

where $U_{\cdot,n}$ is the nth column of U. This relationship between $\| \vec{x} \|$ and σ_{Φ_n} is made explicit in Fig. 2.

Constraints and Nonlinear Transformations

An optimiser can be configured to enforce soft or hard constraints on the values it considers. However, when the values are only indirectly used, a nonlinear transformation may be necessary to impose constraints. Consider a corrector vector chosen indirectly by an optimiser.

$$\vec{C_s} = V\vec{S}. \tag{8}$$

where $V \in \mathbb{R}^{n\times n}$ is the corrector sub-matrix obtained from performing a singular value decomposition on the full response matrix, $\text{TRM} \in \mathbb{R}^{m\times n}$, $\vec{S}$ is a column-vector of coefficients chosen by an optimiser and $\vec{C_s}$ is the resulting matrix-vector product. A nonlinear transformation followed by multiplication with a scale factor transforms the elements of $\vec{C_s}$ onto an arbitrary bounded domain.

$$\vec{C_*} = \overrightarrow{C_{low}} + 0.5(\vec{1} + \tanh(\vec{C_s})) \cdot \overrightarrow{C_{rng}}. \tag{9}$$

where $\overrightarrow{C_{low}}$ and $\overrightarrow{C_{rng}}$ are vectors of minimum power supply limits and power supply ranges, respectively, tanh is the nonlinear function that maps the real line onto the open and bounded interval (-1, 1), and $\vec{C_*}$ is the final corrector value. Multiplication is to be taken element-wise when calculating the second term on the RHS. The power supplies that feed the correctors in the transfer lines at Diamond have limits of ±5 Amps, which limits the magnet strengths to a few mrad.

METHODOLOGY

The Twiss parameters at the entrance of the BTS are $\alpha_x = -2.92$, $\alpha_y = 0.75$, $\beta_x = 12.13$m, $\beta_y = 2.94$m. Together with quadrupole strengths, the magnet layout and dipole parameters, the transfer matrix along the transfer line can be calculated, and hence the analytic response matrix. A Gaussian Process (GP) optimiser was configured in Xopt [3] with an Upper Confidence Bound (UCB) acquisition function, Radial Basis Function with Automatic Relevance Determination (RBF-ARD) kernel and a zero-mean prior. It was also configured to maximise the following objective.

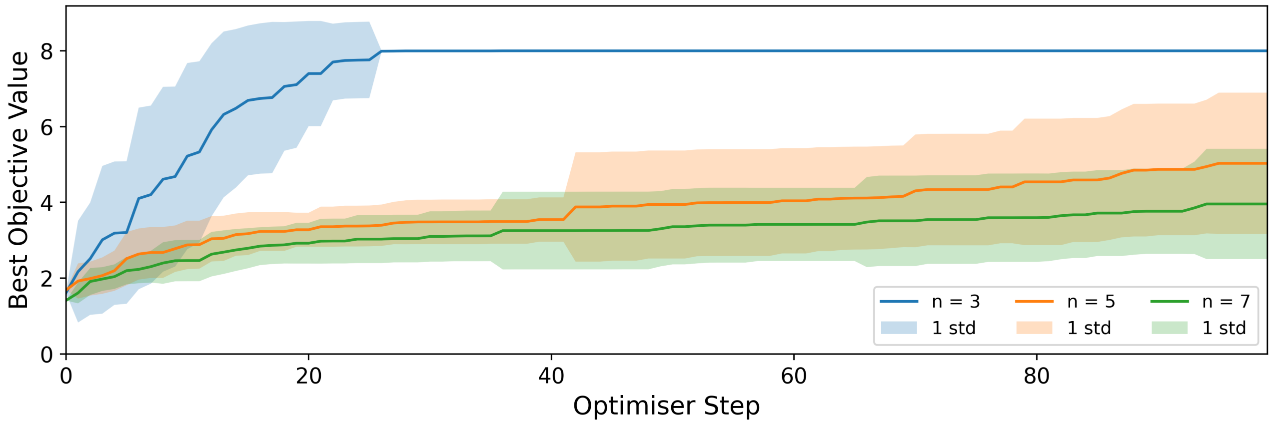

Figure 3: Studies looking at the average behaviour of optimisers using the top $n = 3$ (1H, 2V), $n = 5$ (2H, 3V) and $n = 7$ (3H, 4V) singular vectors. BPM charge values are normalised so that a wholly surviving beam reads a value of one, making the maximum possible score 8. Within each study, 20 identical optimisers were given the same poor starting point. No random samples beyond this are given. The shaded regions represent the 1σ spread of cumulative best scores.

$$f(\vec{C}_*) = \sum_i Q_{b_i} + Q_{\mathrm{DCCT}}. \quad (10)$$

where Q_{b_i} is the charge recorded by the ith BPM and Q_{DCCT} is the charge recorded inside the Booster.

Simulation

A lattice model of the BTS was created in the PyAT [4] particle tracking code. Rectangular apertures of $\pm$25mm were added between every element. For each separate optimiser run, a unique randomly initialised bunch was instantiated with 2,000 particles using the full sigma matrix at the start of the transfer line with horizontal and vertical emittance $\epsilon_{xy} = 2.6 \times 10^{-7}$ m rad. Since separate studies were performed to determine the performance of the optimiser with three, five and seven top singular vectors. In each study, the same optimiser was restarted with the same intentionally poor starting candidate. The variability in performance within a study hence comes from the stochasticity that appears when hyperparameters of the GP kernel are found by minimising the negative log likelihood of the model. In simulation the charge in the storage ring is substituted with the charge at the very end of the BTS.

RESULTS

In simulation, using the top $n = 3$ singular vectors yielded the largest improvement in sample efficiency, as seen in Fig. 3. The variability in the optimiser trajectories despite identical initial conditions comes from the random seeding of trial initial hyperparameter values when the negative log likelihood of the GP surrogate model is minimised. This step is necessary to shape the kernel in such a way that the error between a superposition of kernels fixed at the observed data points and the observed data points themselves is minimised. There was a steady but slow improvement in the cumulative best score for $n = 5$ and $n = 7$, however, the additional vectors required larger coefficients to be selected by the optimiser which may have exceeded the limits of [-5 mrad, 5 mrad] imposed in the simulation. Hence, they may not have been discoverable at all. The correctors suffer from hysteresis which means a fixed path must be taken through state-space when setting the magnet strength. A fast-hysteresis compensation scheme has been tested at Diamond where the correctors are ramped up to their maximum strength of 5 Amps, then ramped down to -5 Amps, and then up again to the target setpoint [1]. At each step, a two second timeout is observed to allow the magnetic field to settle. This is a compromise between reducing measurement times and experimental repeatability. The BPMs in the real machine have a refresh rate of 5Hz. Ten repeat measurement were taken and the average value was returned.

CONCLUSION AND FUTURE WORK

An optimiser has been used to determine optimal singular vector coefficients that maximise injection efficiency through a transfer line at Diamond Light Source in simulation. The impacts of a beam being off-energy on the response matrix measurement are highlighted and used to conclude that this method may not be suitable in its existing form to energetically unstable regions of an accelerator. To further generalise this method, a neural network prior mean for the optimiser shall be trained on a dataset of simulated trajectories through a transfer line at a variety of different optics settings. With the model parameters frozen, an additional linear layer will be added and trained on experimental data to learn a calibration factor.

REFERENCES

[1] N. Blaskovic Kraljevic *et al.*, “Characterisation and optimisation of the Diamond linac”, presented at the 17th International Particle Accelerator Conf. (IPAC’26), Deauville, France, May 2026, paperTHP2057, this conference.

[2] S. Y. Lee, *Accelerator Physics (Fourth Edition)*. World Scientific Publishing Company, 2018. doi:10.1142/11111

[3] R. Roussel *et al.*, “Xopt: A simplified framework for optimization of accelerator problems using advanced algorithms”, in *Proc. IPAC’23*, Venice, Italy, pp. 4847–4850, Sep. 2023. doi:10.18429/JACoW-IPAC2023-THPL164

[4] W. A. H. Rogers *et al.*, “pyAT: A Python Build of Accelerator Toolbox”, in *Proc. IPAC’17*, Copenhagen, Denmark, May 2017, pp. 3855–3857. doi:10.18429/JACoW-IPAC2017-THPAB060