14th International Particle Accelerator Conference,Venice, Italy JACoW Publishing
ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-WEPA026

ACTIVE DEEP LEARNING FOR NONLINEAR OPTICS DESIGN OF A VERTICAL FFA ACCELERATOR

A. Oeftiger*, GSI Helmholtz Centre, Darmstadt, Germany
A. Santamaria Garcia, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
J.-B. Lagrange, ISIS Department, RAL, STFC, UK
S. Hirlaender, Paris Lodron Universität Salzburg, Salzburg, Austria

Abstract

Vertical Fixed-Field Alternating Gradient (vFFA) accelerators exhibit particle orbits which move vertically during acceleration. This recently rediscovered circular accelerator type has several advantages over conventional ring accelerators, such as zero momentum compaction factor. At the same time, inherently non-planar orbits and a unique transverse coupling make controlling the beam dynamics a complex task. In general, betatron tune adjustment is crucial to avoid resonances, particularly when space charge effects are present. Due to highly nonlinear magnetic fields in the vFFA, it remains a challenging task to determine an optimal lattice design in terms of maximising the dynamic aperture. This contribution describes a deep learning based algorithm which strongly improves on regular grid scans and random search to find an optimal lattice: a surrogate model is built iteratively from simulations with varying lattice parameters to predict the dynamic aperture. The training of the model follows an active learning paradigm, which thus considerably reduces the number of samples needed from the computationally expensive simulations.

OVERVIEW

The concept of a vertical FFA (vertical Fixed Field alternating gradient Accelerator) [1] has recently gained popularity [2]. In this type of circular accelerator, the beams move vertically when accelerated as can be seen in Fig. 1. In contrast, orbits in the original FFA remain in the horizontal plane. Vertical FFAs are characterised by strong transverse coupling and highly nonlinear magnetic fields. The design of such lattices was so far carried out in a trial-and-error fashion [3], and it remains challenging to optimise them.

The goal of this study is to efficiently explore the lattice parameter space in order to identify a lattice design with maximum dynamic aperture (DA). In such higher-dimensional problems, an exhaustive grid scan can easily become impractical and a mere random selection too sparse to resolve the important regions. Data-driven approaches lend themselves as computationally economical and rewarding solutions. This paper guides the exploration by iterative supervised learning, which is also referred to as *active learning* [4]. Since many lattice configurations do not even yield a closed orbit (CO) around the accelerator, the domain of valid CO needs to be identified before searching for maximum DA.

The present study was staged in three steps. First, the lattice parameter space was randomly sampled to obtain an initial set of lattices which possess a CO. In a second step, a classification algorithm was trained on predicting CO existence for a given lattice configuration. Rejection sampling based on the predicted probability then significantly improved the efficiency of gathering more samples with a valid CO. The third and last stage qualifies the DA across the valid CO domain: an uncertainty-aware surrogate model was initially trained to predict the DA, and was later re-trained with additional simulations where predictions were most uncertain. With this approach, several interesting parameter regions could be identified where particular lattices yield a larger DA compared to previous studies based on grid scans.

MACHINE AND SIMULATION MODEL

The vertical beam excursion of a vFFA has several advantages. The momentum compaction factor is zero because of the vanishing radial dispersion function, which would be advantageous e.g. for acceleration at ultra-relativistic energies in a muon collider, a neutrino factory, or a relativistic cyclotron. It would also vertically separate the light rays of different radiation spectra in a light source. However, these benefits come at the cost of coupling the particle motion between the two transverse planes [5]. The vertical field gradient also implies a transversely oscillating closed orbit, making it challenging to derive essential accelerator parameters and optics other than by numerical modelling. In addition to this, the scaling condition requires magnetic fields that increase exponentially in the vertical direction z as $B_{x,y,z} \propto B_0 \exp(mz)$ to obtain zero-chromatic operation, leading to highly nonlinear magnetic fields. The DA is a critical parameter in high-power machines to avoid un-

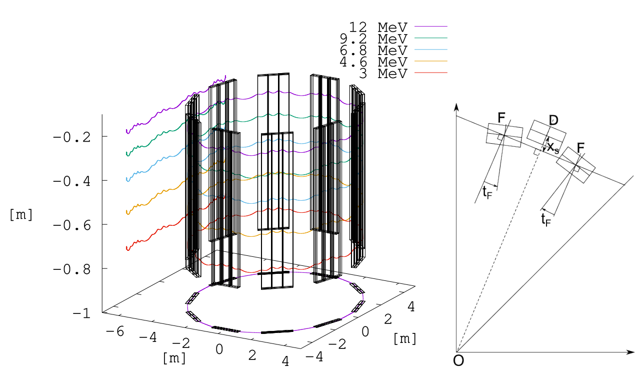

Figure 1: Closed orbit in a reference vFFA triplet lattice.

* a.oeftiger@gsi.de

ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-WEPA026

controlled beam loss, and numerical simulations have been conducted to investigate it in vFFAs in Ref. [3]. The layout of the vFFA lattice is similar in our study with a ring made of 10 FDF triplet cells with an average radius of 4.45 m. The rectangular F and D magnets have identical lengths of 50 cm. Preliminary hardware studies have shown that tanh fringe fields used previously [3] do not properly fit first magnet prototypes, so the fringe field model here is of the form

$$f(y) = \frac{1}{\pi}\left(\arctan\left(\frac{y - y_{\text{ent}}}{\lambda}\right) - \arctan\left(\frac{y - y_{\text{ex}}}{\lambda}\right)\right), \quad (1)$$

where y is the longitudinal coordinate, y_{en} and y_{ex} the effective entrance and exit boundaries of the magnet, respectively, and λ the characteristic fringe field extent.

Analogous to the previous study in Ref. [3], the following 5 lattice parameters are varied in this study: the field at the centre of the F magnet (B_{0f}) and D magnet (B_{0d}) at the reference altitude $z = 0$, the normalised field gradient m, the radial displacement between F and D magnet x_s, and the tilt angle t_f of F magnets with respect to the D magnet (cf. the sketch on the right of Fig. 1). The characteristic fringe field extent λ is fixed at 15 cm, and the extrapolation off the magnetic median plane is truncated at the 10$^{\text{th}}$ order. The neighbouring triplet cells are included in the simulation due to overlapping fringe fields in the long straight section. Simulations were carried out with the particle tracking code FixField [6].

DOMAIN OF VALID CLOSED ORBITS

Steps 1 and 2 of the study identify the domain of valid COs in the 5-dimensional parameter space. Due to the nonlinear potential, a lattice described by the 5-tuple $(B_{0f}, B_{0d}, m, x_s, t_f)$ might not feature a CO. In this case the tracked particles either turn backwards, get lost in the far-located collimators[1], or even get quasi-trapped[2]. Figure 1 shows CO projections for various particle momenta in a reference vFFA triplet lattice $(-1\,\text{T}, 1.15\,\text{T}, 1.31\,\text{m}^{-1}, 0, 0)$. The particle used in the present study is a 3 MeV proton. At the centre of the long straight section (the beginning of a triplet cell), the CO of the reference lattice is located radially (horizontally) at $x_{co} = 4.36$ m and vertically at $z_{co} = -0.73$ m. Owing to the symmetry, the corresponding transverse momenta vanish.

In step 1, an initial data set is generated via uniform random sampling of the 5D parameter space to generate 20 000 lattice configurations. Simulations computed valid COs for 429 lattices i.e., 2% success rate for random sampling. Step 2 involves a classification algorithm which is trained to predict the probability for a new random lattice configuration to possess a valid CO. Samples from the uniform random distribution can then be rejected based on probability limits. A new set of 10 000 lattices is readily generated using the rejection sampling, and simulations are run to determine the CO. With each iteration of 10 000 lattices, the classifier improves in prediction quality as the data set increases. The employed rejection strategy first gathered more data by excluding lattices below a 30% probability of having a CO, then resolved the domain boundary in finer detail by accepting lattices between 30% and 70%, and eventually focused on gathering more CO by accepting only lattices above a 90% probability. Two initial iterations used Gaussian Naive Bayes classification [7], resulting in ≈ 10% valid COs. However, the emerging rather complex domain shape in the 5D space was not well captured by the classifier. Changing to random forest classification [7] resulted in reproducing the domain shape in much more detail, which in turn tremendously improved the fraction of valid COs. Approximately 85% of valid COs were obtained when rejecting lattices below 90% of predicted probability.

Proceeding with the rejection sampling stage, a total of 170 000 lattices was simulated, with more than a quarter possessing a valid CO. Figure 2 presents an overview of the domain of valid CO (after step 3) along with a histogram of all simulated lattices in 2D projections of the 5D parameter space. The clear overlap of the highest number of COs with the highest number of launched simulations show that the classifier-based rejection sampling successfully guided the parameter space exploration to the most interesting regions.

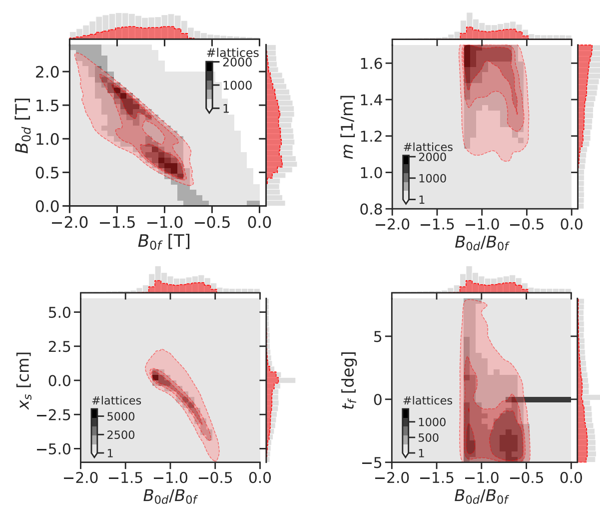

Figure 2: Distribution of identified closed orbits in 5D lattice parameter space. The number of lattices simulated per bin is shown in grey, where non-white bins contain at least one simulation. The red contours show the kernel density estimation for the distribution of valid closed orbits.

DYNAMIC APERTURE QUALIFICATION

Step 3 continues the exploration of lattice space by moving from fast CO simulations (≈ 5 min) to slow DA simulations (≈ 2 × 3 h). Ref. [3] establishes a useful approach to determine the dynamic aperture (DA) of a vFFA lattice in decoupled transverse space, (u, p_u, v, p_v): particles are tracked for 1000 lattice cells starting from an initial amplitude, once by scanning an initial u_i as $(u_i, 0, 0, 0)$ and likewise for v_i

[1] limiting to radial $2\,\text{m} < x < 6\,\text{m}$ and vertical $-10\,\text{m} < z < 10\,\text{m}$.
[2] i.e. reaching a maximum number of integration steps ≫ than for the CO.

ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-WEPA026

as $(0, 0, v_i, 0)$. The amplitude is considered unstable if the particle gets lost, turns backwards, or is quasi-trapped during the tracking. More than 1000 lattice periods are not required as the DA estimate remained almost identical. The present paper builds on this approach. For a given lattice configuration, a DA amplitude is identified using the divide-and-conquer algorithm separately for both decoupled planes, u_i and v_i. Thirteen recursions of interval bisection yield a $2^{-13} \approx 10^{-4}$ accuracy. This estimate is then confirmed by probing the interval from zero in ten equidistant steps to exclude the existence of smaller unstable amplitudes.

An initial DA data set is evaluated for the previously identified lattices with a valid CO. To continue exploring the lattice parameter space, a surrogate model is built to predict the two transverse DA values for a given 5D lattice configuration. The uncertainty-aware surrogate model class employed here is based on the *ensembling* of deep neural networks [8], which is well suited for large and growing data sets (where e.g. Gaussian processes do not scale well). In our case, five structurally equivalent pyramid-shaped networks with three hidden layers of 2048-1024-512 neurons were initialised with different random seeds and then trained independently on the data. The Adam algorithm [9] was used as optimiser and the mean squared error (MSE) as loss function. The standard deviation of the prediction for a given sample is taken as estimate for the epistemic uncertainty [4], which is used to guide the exploration. The devised algorithm iteratively follows these steps:

1. Train the model on the existing DA data;
2. Draw new rejection-sampled lattice configurations based on the trained classifier to ensure a valid CO;
3. Predict DA and epistemic uncertainty for new samples;
4. Rank new samples w.r.t. their epistemic uncertainty;
5. Run DA simulations for selected new samples.

Step 3 added three more iterations of 10 000 lattice samples each, which were selected as the top 25% quantile from the ranked samples. The RMS prediction error from each iteration is shown in Fig. 3 against the DA amplitude in decoupled u and v spaces. The underlying inference samples are indicated by the grey histogram. Only a fraction of 0.5% lattices does not feature a CO in this last iteration. From the total of 200 000 simulated lattices, 25% possess a valid CO. Figure 4 displays a maximum DA figure[3] in the 5D lattice parameter space projections. With a DA of 5.5 cm the reference lattice lies in the light green area, red indicates the location of the ten lattices with the largest DA up to 6.5 cm.

CONCLUSIVE REMARKS

The employed active learning approach tremendously improved the efficiency of finding closed orbits in the 5D lattice parameter space: from an initial 2% with random sampling to 85% using classifier-based rejection sampling, and finally

[3] sum of $u + v/2$ since the β-function in v is about twice as large as in u.

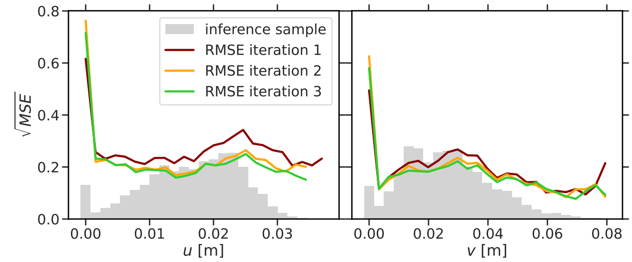

Figure 3: Training results after each active-learning iteration.

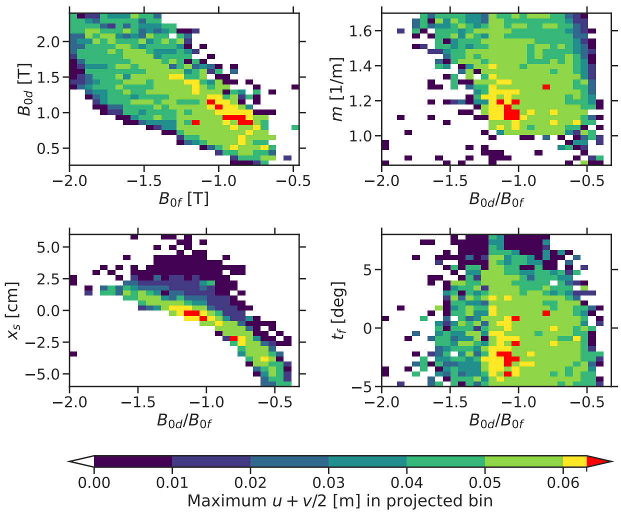

Figure 4: Simulation results for dynamic aperture (DA) in 5D lattice parameter space. Each panel plots the maximum DA per projected bin for the nearly 50 000 valid closed orbits. Red bins contain the top 10 lattices (overall maximum DA).

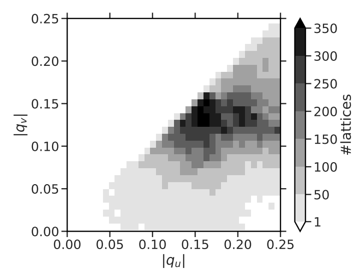

Figure 5: Histogram of stable lattices in tune space.

99.5% via ranking using the deep neural network ensemble. The gathered data set and established surrogate models link lattice parameter space to dynamic aperture as well as tunes (see Fig. 5). This result now allows to guide vFFA lattice design by choosing a target tune and then obtaining a lattice configuration with maximum dynamic aperture: a helpful tool to gain space toward resonances located close by in (decoupled) tune space, in particular in view of intensity limitations due to space-charge detuning.

ACKNOWLEDGEMENT

Shinji Machida provided inputs and valuable feedback.

ISBN: 978-3-95450-231-8 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2023-WEPA026

REFERENCES

[1] J. Teichmann, "Accelerators with a vertically increasing field," *J. Nucl. Energy Part C*, vol. 5, no. 1, p. 57, 1963. `doi:10.1088/0368-3281/5/1/312`

[2] S. Brooks, "Vertical orbit excursion fixed field alternating gradient accelerators," *Phys. Rev. ST Accel. Beams*, vol. 16, no. 8, p. 084 001, 2013. `doi:10.1103/PhysRevSTAB.16.084001`

[3] S. Machida, D. J. Kelliher, J.-B. Lagrange, and C. T. Rogers, "Optics design of vertical excursion fixed-field alternating gradient accelerators," *Phys. Rev. Accel. Beams*, vol. 24, p. 021 601, 2021. `doi:10.1103/PhysRevAccelBeams.24.021601`

[4] K. P. Murphy, *Probabilistic Machine Learning: Advanced Topics*. MIT Press, 2023. `http://probml.github.io/book2`

[5] M. Vanwelde, C. Hernalsteens, S. A. Bogacz, S. Machida, and N. Pauly, "Review of coupled betatron motion parametrizations and applications to strongly coupled lattices," *arXiv preprint*, 2022. `doi:10.48550/arXiv.2210.11866`

[6] J.-B. Lagrange, "The Particle Tracking Code Fixfield," in *Proc. IPAC'21*, Campinas, Brazil, May 2021, pp. 1905–1906. `doi:10.18429/JACoW-IPAC2021-TUPAB209`

[7] L. Buitinck *et al.*, "API design for machine learning software: Experiences from the scikit-learn project," in *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, 2013, pp. 108–122. `doi:10.48550/arXiv.1309.0238`

[8] B. Lakshminarayanan, A. Pritzel, and C. Blundell, "Simple and scalable predictive uncertainty estimation using deep ensembles," *Advances in neural information processing systems*, vol. 30, 2017. `doi:10.48550/arXiv.1612.01474`

[9] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint*, 2014. `doi:10.48550/arXiv.1412.6980`