Enhancing the MOGA optimization process at ALS-U with Machine Learning

Y. Lu, S.C. Leemann, C. Sun, M.P. Ehrlichman, T. Hellert, H. Nishimura, M. Venturini

Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

1. Introduction

- MOGA is the most commonly used algorithm in lattice optimization for ultra-high brightness storage rings.
- However, MOGA requires extensive runtime to arrive at a Pareto front due to its stochastic nature and costly evaluations of dynamic aperture (DA) and momentum aperture (MA) based on many-turn particle tracking.
- Since Machine Learning (ML) has proven its efficiency in building computational models to solve complex data-intensive problems compared to traditional statistical methods, we have applied ML techniques to speed up MOGA.

2. Machine Learning Approach

- We first pre-process training data acquired from prior simulations and use this data to obtain two well-trained models using the neural network (NN) depicted in Fig. 1.
- We then use these two NN models to replace DA/MA particle tracking in MOGA while the rest of the MOGA setup remains the same as in the original tracking-based MOGA (Tr-MOGA).
- We evaluate this ML-based MOGA (ML-MOGA) on a simpler 2-DoF problem and a more complex 11-DoF problem.

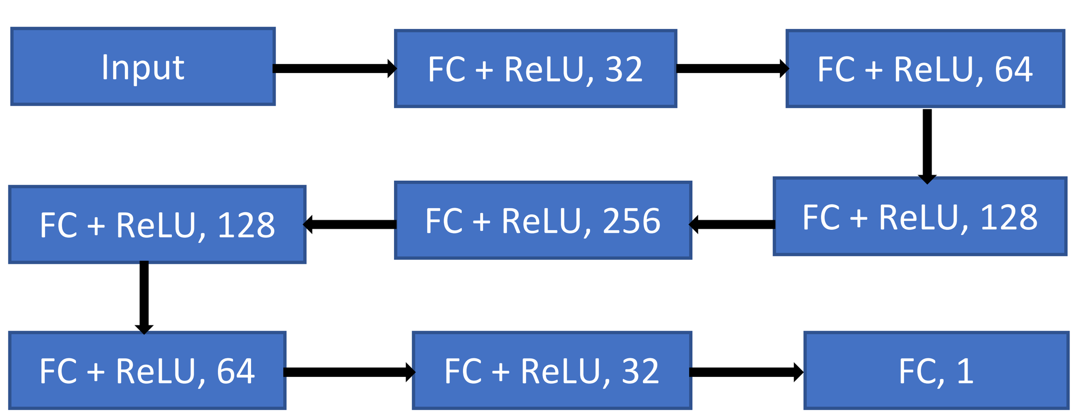

Figure 1: *8-layer fully-connected (FC) NN architecture for DA and MA prediction. Output dimension is indicated.*

3. Optimization in 2 DoF

- Lattice optimization at ALS-U (before introduction of reverse bending and high-field bends) consists of 11 DoF as well as several constraints and objectives. 9 DoF are linear (quadrupoles) while 2 DoF are nonlinear.
- In the 2 DoF study, 9 linear DoF were fixed, leaving only 2 harmonic sextupoles (SH1,SH2), for a data input size of 2.
- Random data which is uniformly distributed in input parameter space is used for training.

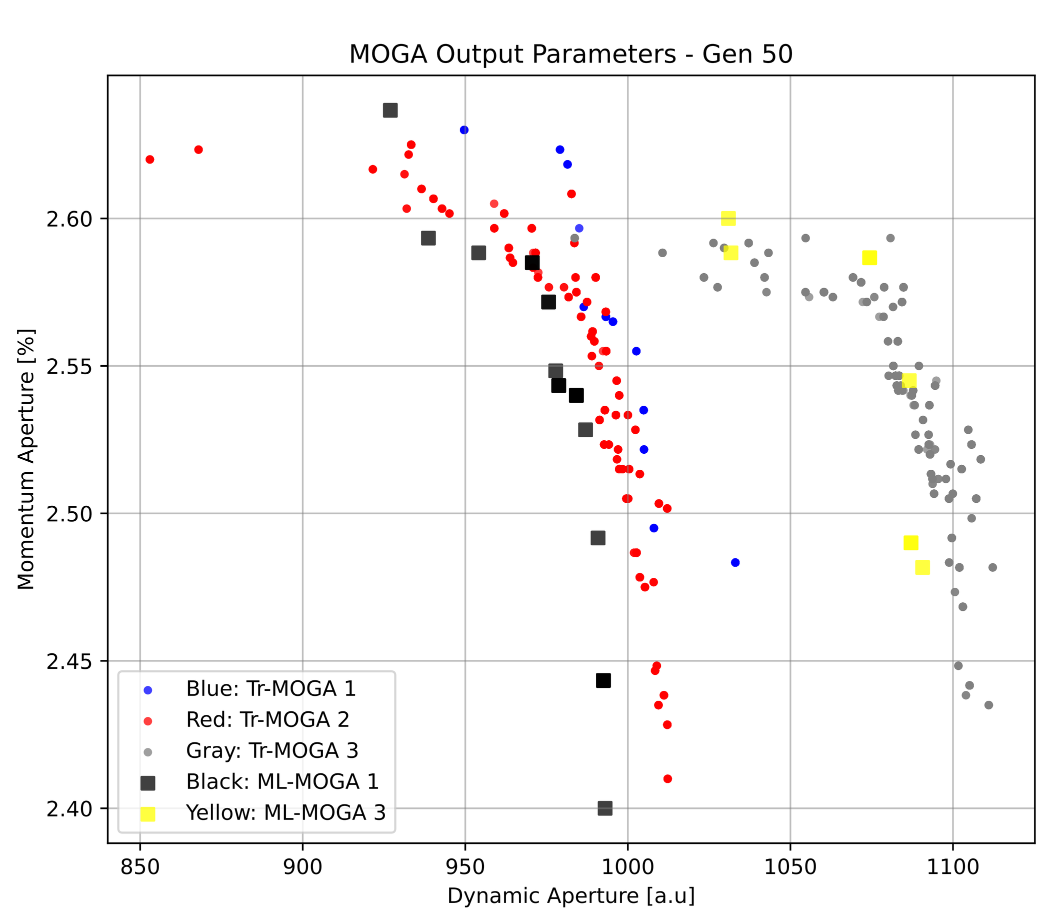

Figure 2: *Solution space comparison between Tr-MOGA runs and ML-MOGA runs with different random seeds.*

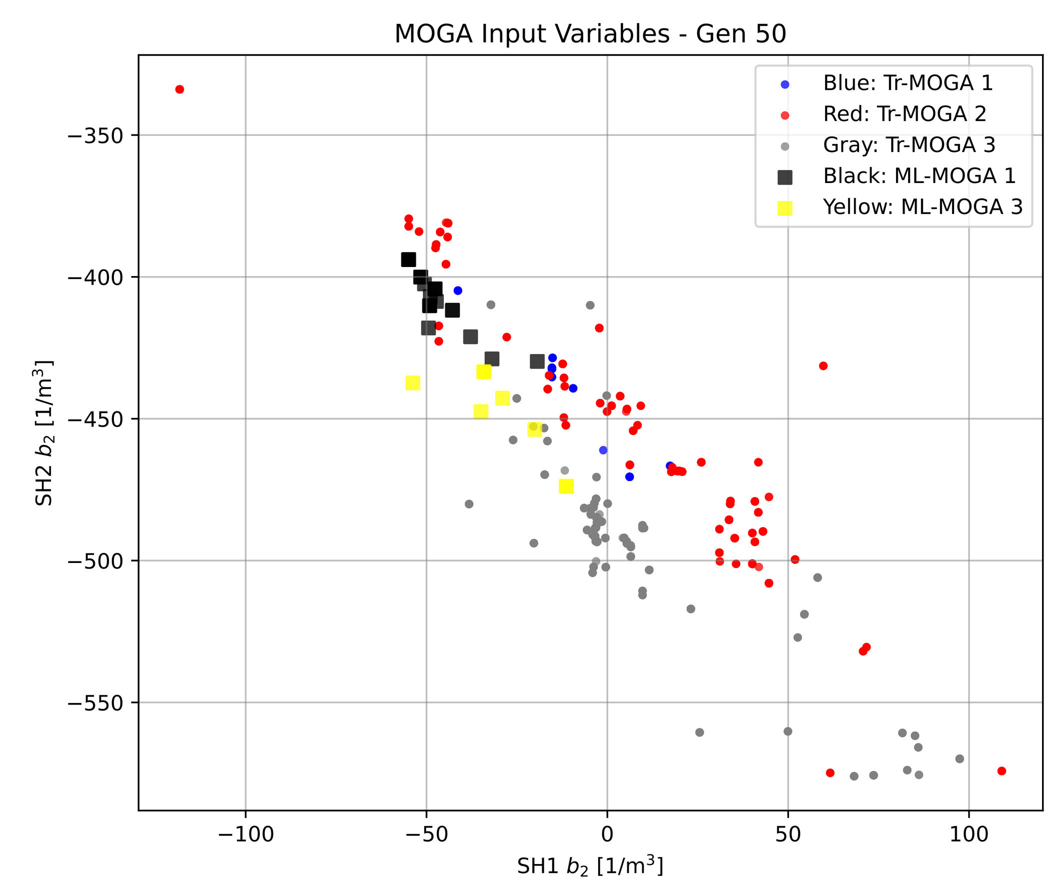

Figure 3: *Input space comparison between Tr-MOGA runs and ML-MOGA runs with different random seeds.*

- We also studied the effect of training data sampling size on the accuracy of our NN predictions.
- The root mean square error (RMSE) proves to be very robust, allowing us to reduce the sampling to as low as 20×20 without loss of fidelity.
- Consequently, we can arrive at similar ML-MOGA results with much smaller sets of training data, effectively reducing tracking effort to 400 samples for ML-MOGA vs. 2.5×10^5 for Tr-MOGA.

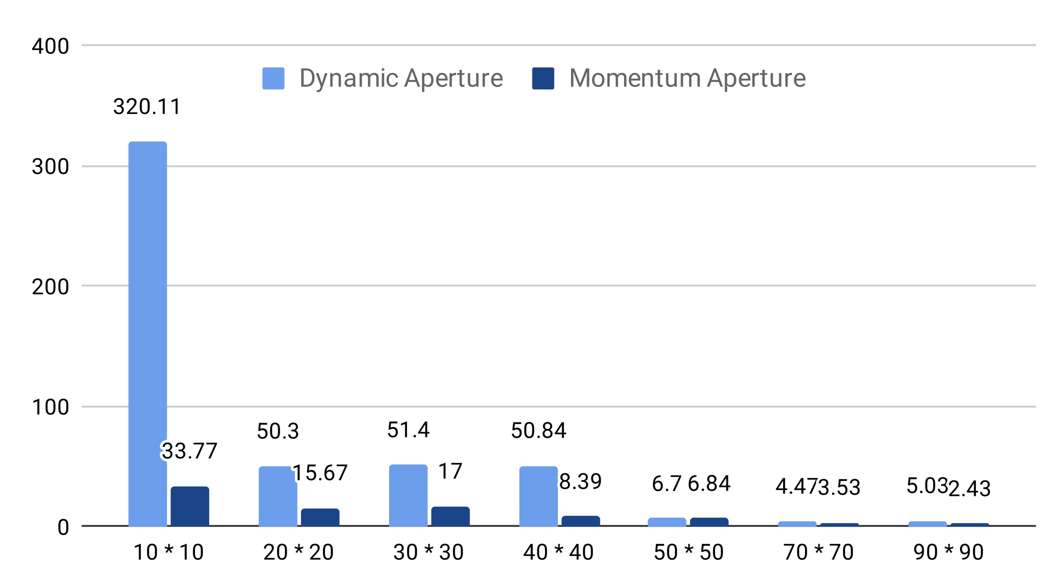

Figure 4: *RMSE increase in percent compared to the initial sampling size choice of 115×115.*

4. Optimization in 11 DoF

- Unlike the 2-DoF problem, the 11-DoF optimization includes all 9 linear and 2 nonlinear variables. The first step for both Tr-MOGA and ML-MOGA is to find reasonable ranges for the 9 linear DoF (c_l).
- We then use only the first 10 generations of Tr-MOGA data (samples violating any constraints are filtered out) as training data to build NNs. Once the ML-MOGA run converges, we again perform one tracking run involving inputs (a_{jl}) from the final ML-MOGA generation for validation purposes. Finally, by combining this tracked generation with the previous training data, we can retrain the NNs allowing us to iterate this ML pipeline until it fully converges.
- To properly assess convergence among all MOGA runs, we introduce two Euclidean distance metrics for input and output space, respectively (M = MA, D = DA, m = gen. no.):

$$\delta_i(m) = \frac{1}{n(n-1)} \sum_{j=1}^{n} \sum_{k=1}^{n} \sqrt{\sum_{l=1}^{11} \left(\frac{a_{jl}^{(m)} - a_{kl}^{(m-1)}}{c_l} \right)^2}, \quad (1)$$

$$\delta_o(m) = \frac{1}{n} \sqrt{\sum_{j=1}^{n} \left(A_{mj}^2 + B_{mj}^2 + C_{mj}^2 \right)}, \quad \text{where}$$

$$A_{mj} = \frac{\varepsilon_{mj} - \varepsilon_0}{\varepsilon_0}, \quad B_{mj} = \frac{D_{mj} - D_0}{D_0}, \quad C_{mj} = \frac{M_{mj} - M_0}{M_0}. \quad (2)$$

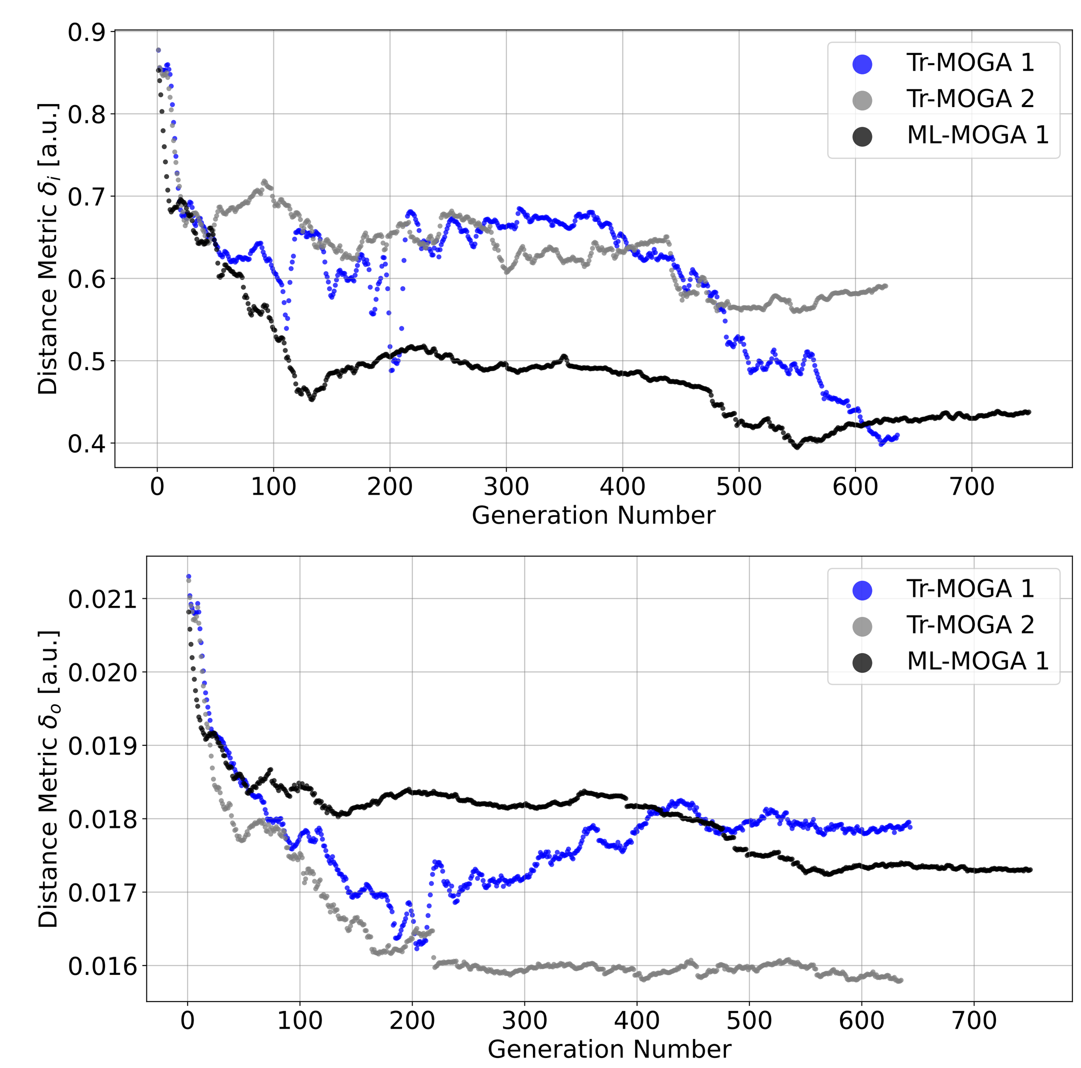

Figure 5: *Distance metric for input variables (top) and solution space (bottom) for two Tr-MOGA runs with different random seeds and one ML-MOGA run.*

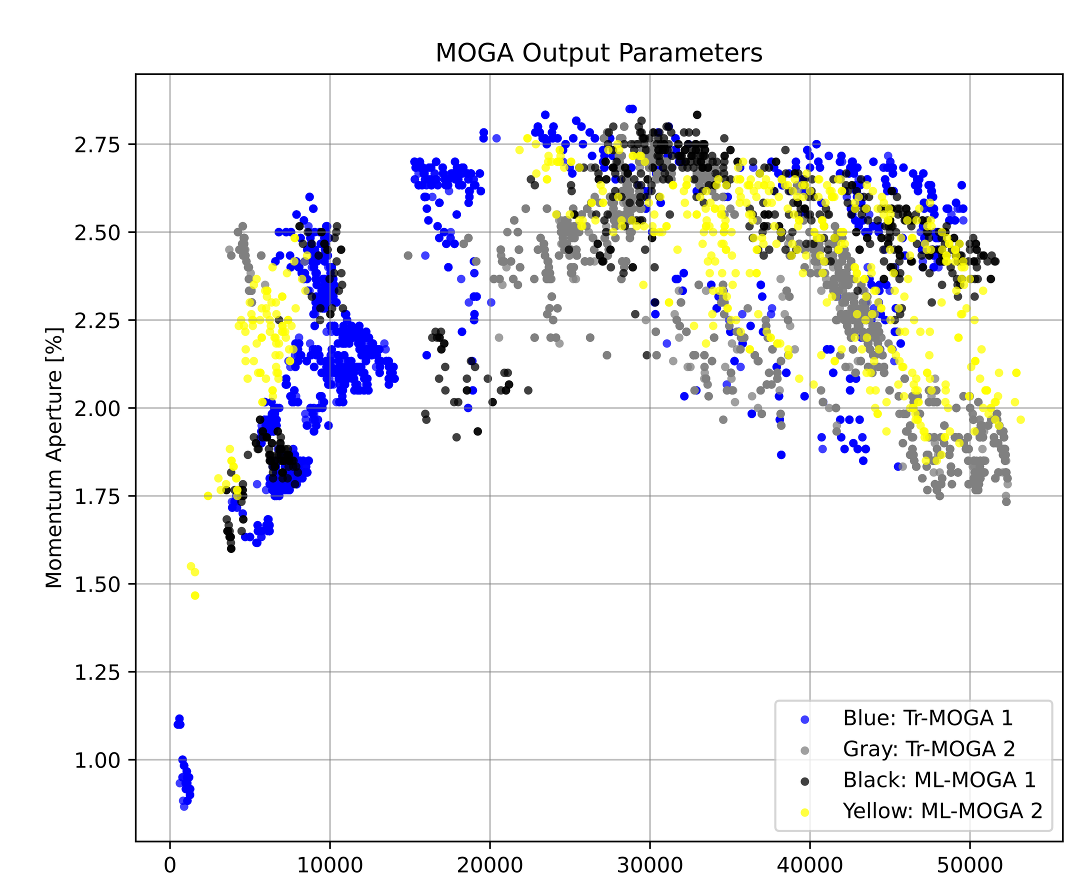

Figure 6: *Comparison in reduced solution space (emittance omitted for clarity) between Tr-MOGA and ML-MOGA runs with different random seeds.*

Acknowledgments

This research is funded by the US Department of Energy (BES & ASCR Programs), and supported by the Director of the Office of Science of the US Department of Energy under Contract No. DE-AC02-05CH11231.