# REINFORCEMENT LEARNING FOR LASER ALIGNMENT

*Matthew B. Schwab*

## MOTIVATION

Laser Plasma Accelerators (LPAs) demand stable laser operation due to the nonlinear nature of various acceleration mechanisms, e.g., LWFA, TNSA, etc. Reinforcement Learning (RL) is tested as a control technique to re-optimize the alignment of the laser's constituent amplifiers and subsystems.

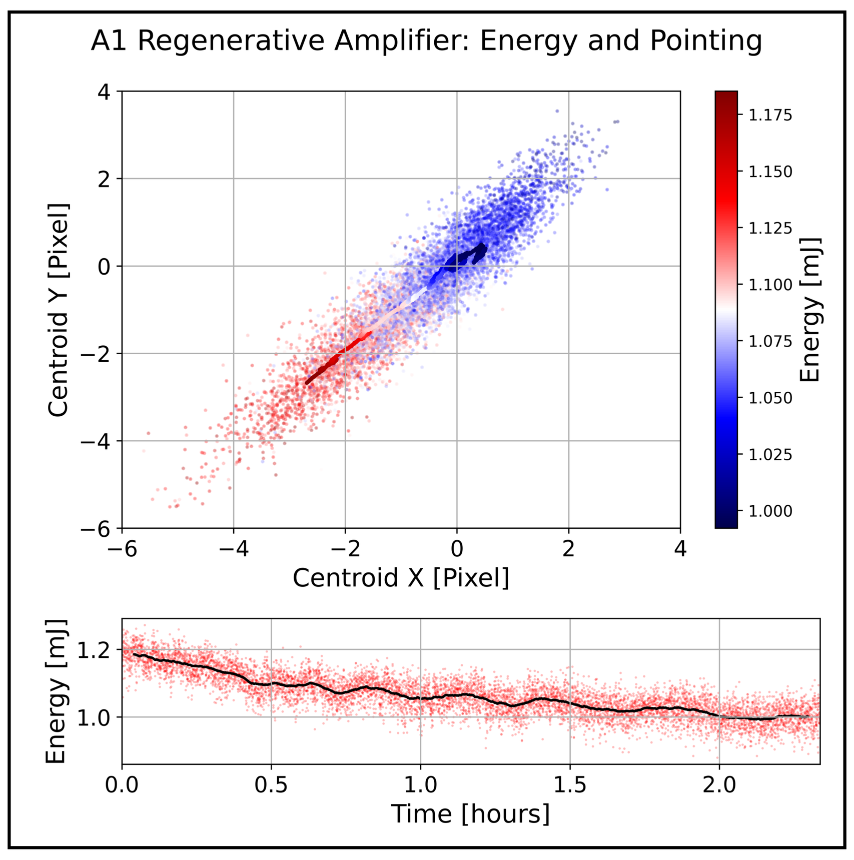



**Figure 1**
Exemplary energy and pointing measurement of a regenerative amplifier in the Polaris laser. (Top) color coded energy measurement correlates with a long-term drift in the laser's centroid beam position. (Bottom) laser's shot-to-shot energy (red) and average energy (black) plotted over ~2.5 hours.

As displayed in **Fig. 1**, the output energy of a laser amplifier can degrade over time-scales on the order of minutes to hours, while the shot-to-shot performance of the laser also varies (lower subplot). The long-term drift can be correlated with a change in the optical alignment of the system as seen in the upper subplot of the laser's centroid position.

If any amplifier in the laser chain exhibits such a long-term drift, the laser's overall performance can suffer severely. Automated realignment systems have several potential benefits compared to human intervention:

- Saved time due to faster alignment
- Saved data due to more stable laser performance
- More comparability due to objective realignment

Variation in laser performance over the course of a day is primarily caused by changes in temperature and humidity, as well as vibrational noise in the laboratory.

## GYMNASIUM ENVIRONMENT

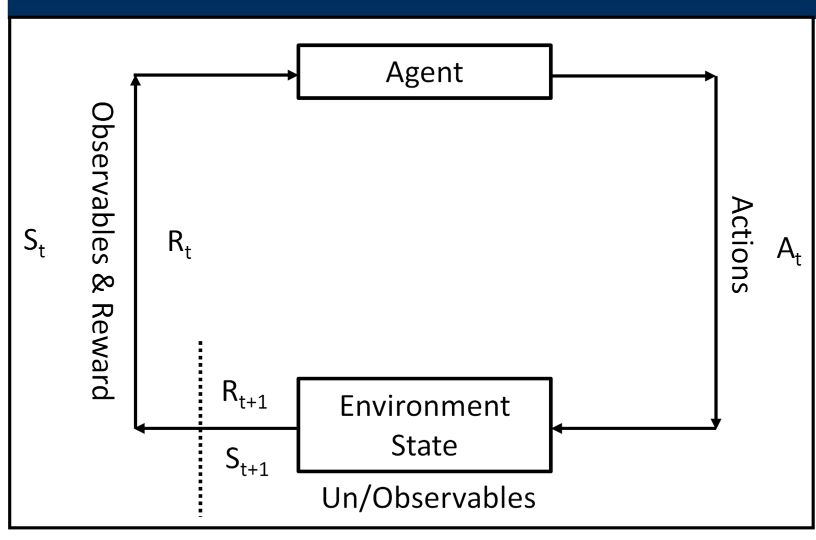



**Figure 2**
Iterative scheme following the framework of a Markov Decision Process (MDP) for sequential decision making.

The Python package *Gymnasium* allows users to code their environment; mapping the agent's actions, environment's state (observable and unobservable), and state change's reward (see **Fig. 2**) to their respective simulation or real-world counterparts.

A two mirror system is the most basic optical alignment setup. Each mirror is motorized in the horizontal and vertical directions and two imaging sensors, typically CCDs, are used to measure the position of the laser beam downstream of each mirror (see **Fig. 3**). The two mirror system can be both easily built with hardware and easily modelled in simulation.

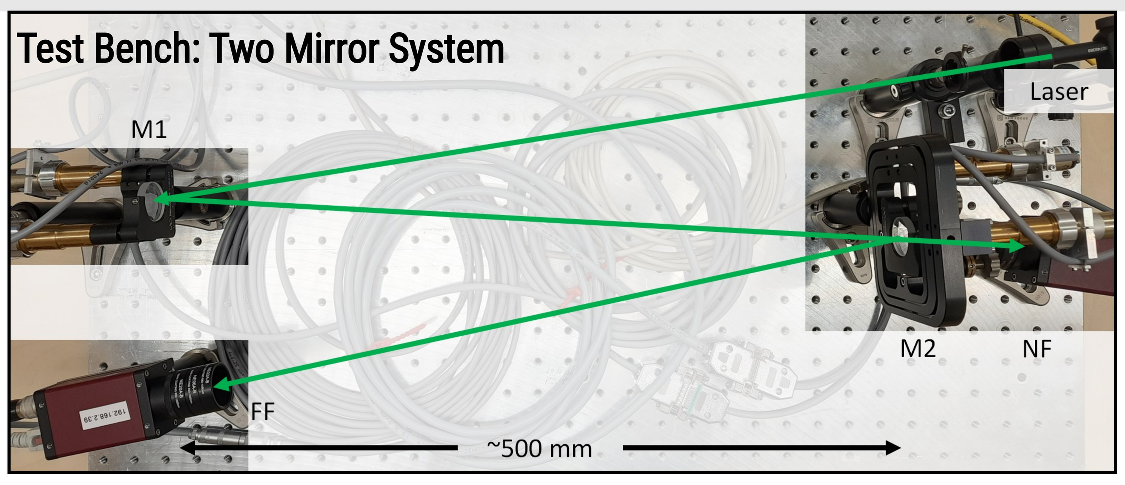



**Figure 3**
Typical two mirror alignment system. Mirrors M1 and M2 are each motorized to rotate around their local x– and y-axes. The nearfield (NF) and farfield (FF) sensors record the position of the laser at two planes downstream of the mirrors. Each sensor will have a reference position that represents the system's previous best alignment.

## Gymnasium Env():

**Unobservables**: reset at start of each episode
- Start position of the laser
- Start direction of the laser
- Absolute mirror orientations

**Observables**:
- Nearfield position of the laser (NFx,NFy)
- Farfield position of the laser (FFx, FFy)
- Normalized time
- Motors' change of direction
- Laser energy

**Actions**:
- Number of steps driven by each motor

**Reward**:
- Measure of distance to NF and FF references
- Measure of laser energy

**Truncation**:
- Terminal state outside of MDP

**Termination**:
- Terminal state inside of MDP

## SIM2REAL GAP

Simulating the optical system for training is required due to the low repetition rate of the laser and low sample efficiency needed by „vanilla" RL algorithms. Several aspects of the system can be difficult to model correctly, yet are critical for success.

Mechanical backlash/hysteresis (see **Fig. 4**):
- Significant nonlinearity in the system
- Modelled as a Normal distribution at each step size
- Depends on the type of motors and mirror mounts
- Faulhaber stepper motors and U100 Newport mounts used
- Large backlash at step sizes < 100 steps is very problematic for fine optical alignment.

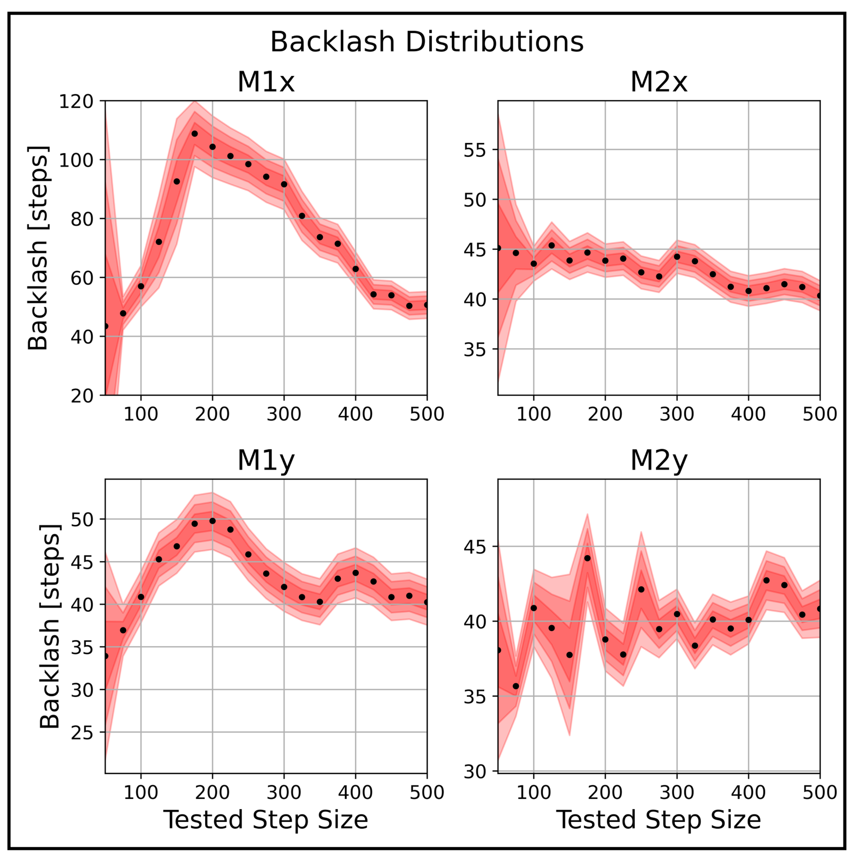
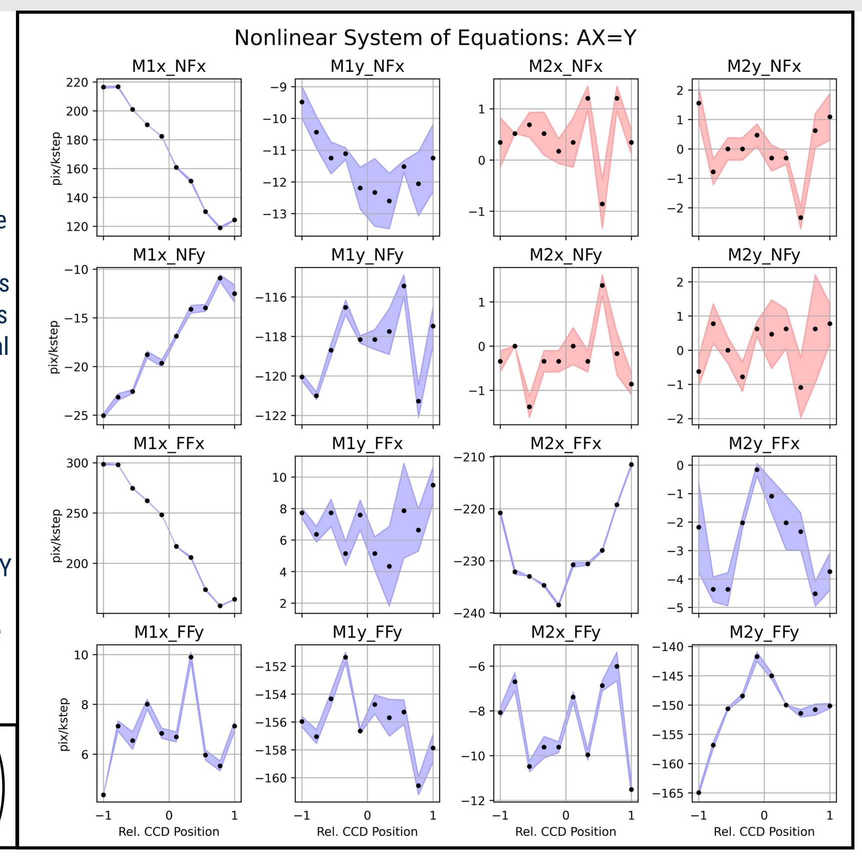



**Figure 4**
Backlash distributions calibrated for four motor/mount pairs in the two mirror optical system. Step sizes below 50 steps could not be calibrated, because the resulting backlash was larger than the desired step size. This is a limiting factor for automating fine alignment and must be remedied.




$$A \begin{pmatrix} \Delta M1x \\ \Delta M1y \\ \Delta M2x \\ \Delta M2y \end{pmatrix} = \begin{pmatrix} \Delta NFx \\ \Delta NFy \\ \Delta FFx \\ \Delta FFy \end{pmatrix}$$

**Figure 5**
The alignment system comprises a nonlinear system of equations represented by **AX=Y**. The subplots in the figure represent the 16 elements of matrix **A**, whose values vary depending on several system variables. An operational set point is chosen to linearize the system. (Below) column vector X is the relative motion of each motor in steps and column vector Y is the relative motion of the laser in pixels on the NF and FF sensors

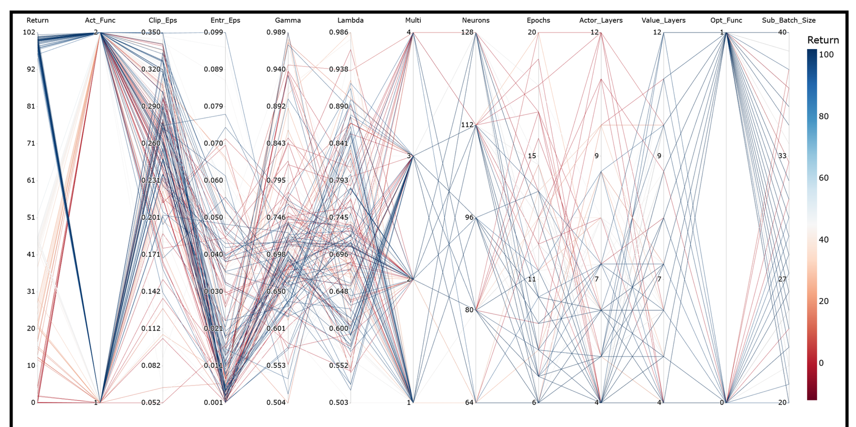
If a ray tracing approach proves too slow or complex, another approach is to use the hardware directly to calibrate the ray transfer matrix, or how the laser beam moves on each sensor depending on the motion of the motors. This is described in **Fig. 5** for the two mirror setup described previously.

Other complications when moving to hardware may include:
- Non-Gaussian laser profile
- Dirt or damages on sensors and optics
- Noise sources: vibrational, electrical, stray light, etc.
- Timing systems and synchronous event acquisition

## CURRENT RESULTS

Using the Proximal Policy Optimization (PPO) algorithm and the Optuna hyperparameter tuning package (**Fig. 6**), several agents have been trained that successfully realign a misaligned two mirror system to NF and FF references. The agent achieves near sub-pixel accuracy within 2-5 laser shots, which is faster than the average human. Oscillatory behavior (**Fig. 7**) can be currently attributed to poor motor/mount settings and calibration, as well as a lack of well defined consistency between normalized and scaled variables during training.

**Figure 6**
Hyperparameter (HP) tuning using a sampler and pruner is highly recommended compared to using simple grid search. Combined with a parallel coordinates plot, the results can be investigated.



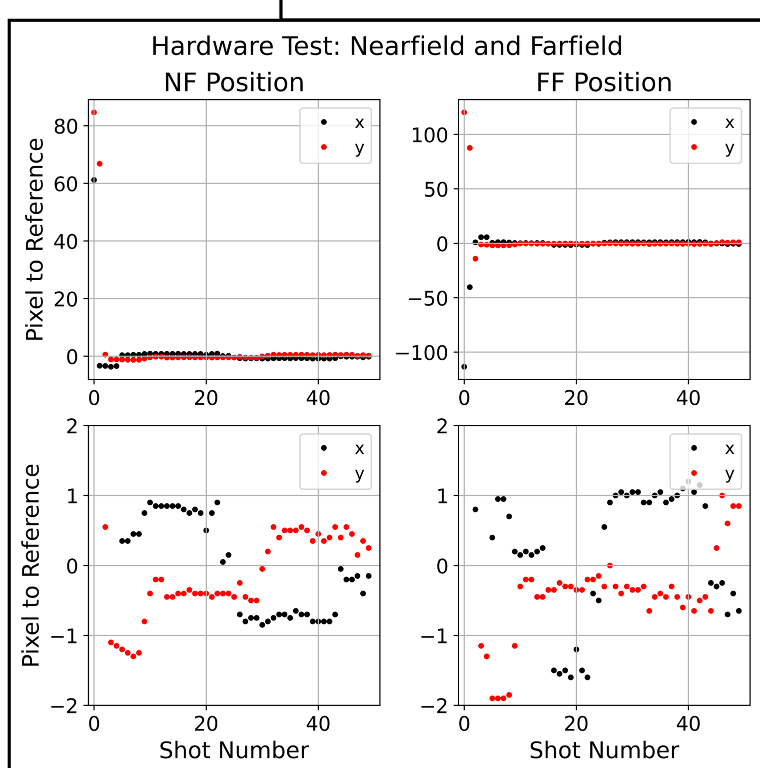



**Figure 7**
The agent has 50 laser shots (specific to the Polaris laser) to realign the two mirror optical setup. Rough alignment down to near sub-pixel accuracy is completed in 2-5 shots, faster than the average human. Improved accuracy should be achievable by increasing the angular resolution of the NF and FF diagnostics as well as improving the mirror/mount calibration.

## OUTLOOK

- Stopping conditions to prevent oscillatory behavior
- Domain randomization for energy optimization
- Improvement of motor/mount calibration and settings
- Improved angular resolution of NF/FF diagnostics
- Implementation on regenerative amplifier input and cavity

Relativistic Laser Physics, Institute for Optics and Quantum Electronics
and
Helmholtz-Institute Jena

FRIEDRICH-SCHILLER-UNIVERSITÄT JENA